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

Lecture 16: Nielsen's theorem on pure state entanglement transformation

In this lecture we will consider *pure-state entanglement transformation*. The setting is as follows: Alice and Bob share a pure state $x \in \mathcal{X}_A \otimes \mathcal{X}_B$, and they would like to transform this state to another pure state $y \in \mathcal{Y}_A \otimes \mathcal{Y}_B$ by means of local operations and classical communication. This is possible for some choices of x and y and impossible for others, and what we would like is to have a condition on x and y that tells us precisely when it is possible. Nielsen's theorem, which we will prove in this lecture, provides such a condition.

Theorem 16.1 (Nielsen's theorem)**.** *Let $x \in \mathcal{X}_A \otimes \mathcal{X}_B$ and $y \in \mathcal{Y}_A \otimes \mathcal{Y}_B$ be unit vectors, for any choice of complex Euclidean spaces $\mathcal{X}_A$, $\mathcal{X}_B$, $\mathcal{Y}_A$, and $\mathcal{Y}_B$. There exists a channel $\Phi \in \mathrm{LOCC}(\mathcal{X}_A, \mathcal{Y}_A : \mathcal{X}_B, \mathcal{Y}_B)$ such that $\Phi(xx^*) = yy^*$ if and only if $\mathrm{Tr}_{\mathcal{X}_B}(xx^*) \prec \mathrm{Tr}_{\mathcal{Y}_B}(yy^*)$.*

It may be that $\mathcal{X}_A$ and $\mathcal{Y}_A$ do not have the same dimension, so the relationship

$$\mathrm{Tr}_{\mathcal{X}_B}(xx^*) \prec \mathrm{Tr}_{\mathcal{Y}_B}(yy^*)$$

requires an explanation. In general, given positive semidefinite operators $P \in \mathrm{Pos}(\mathcal{X})$ and $Q \in \mathrm{Pos}(\mathcal{Y})$, we define that $P \prec Q$ if and only if

$$VPV^* \prec WQW^* \tag{16.1}$$

for some choice of a complex Euclidean space $\mathcal{Z}$ and isometries $V \in \mathrm{U}(\mathcal{X}, \mathcal{Z})$ and $W \in \mathrm{U}(\mathcal{Y}, \mathcal{Z})$. If the above condition (16.1) holds for one such choice of $\mathcal{Z}$ and isometries V and W, it holds for all other possible choices of these objects. In particular, one may always take $\mathcal{Z}$ to have dimension equal to the larger of $\dim(\mathcal{X})$ and $\dim(\mathcal{Y})$.

In essence, this interpretation is analogous to padding vectors with zeroes, as is done when we wish to consider the majorization relation between vectors of nonnegative real numbers having possibly different dimensions. In the operator case, the isometries V and W embed the operators P and Q into a single space so that they may be related by our definition of majorization.

It will be helpful to note that if $P \in \mathrm{Pos}(\mathcal{X})$ and $Q \in \mathrm{Pos}(\mathcal{Y})$ are positive semidefinite operators, and $P \prec Q$, then it must hold that $\mathrm{rank}(P) \geq \mathrm{rank}(Q)$. One way to verify this claim is to examine the vectors of eigenvalues $\lambda(P)$ and $\lambda(Q)$, whose nonzero entries agree with $\lambda(VPV^*)$ and $\lambda(WQW^*)$ for any choice of isometries V and W, and to note that Theorem 13.2 implies that $\lambda(WQW^*)$ cannot possibly majorize $\lambda(VPV^*)$ if $\lambda(Q)$ has strictly more nonzero entries than $\lambda(P)$. An alternate way to verify the claim is to note that mixed unitary channels can never decrease the rank of any positive semidefinite operator. It follows from this observation that if $P \in \mathrm{Pd}(\mathcal{X})$ and $Q \in \mathrm{Pd}(\mathcal{Y})$ are positive definite operators, and $P \prec Q$, then $\dim(\mathcal{X}) \geq \dim(\mathcal{Y})$. The condition $P \prec Q$ is therefore equivalent to the existence of an isometry $W \in \mathrm{U}(\mathcal{Y}, \mathcal{X})$ such that $P \prec WQW^*$ in this case.

The remainder of this lecture will be devoted to proving Nielsen's theorem. For the sake of the proof, it will be helpful to make a simplifying assumption, which causes no loss of generality. The assumption is that these equalities hold:

$$\dim(\mathcal{X}_A) = \operatorname{rank}\left(\operatorname{Tr}_{\mathcal{X}_B}(xx^*)\right) = \dim(\mathcal{X}_B),$$
$$\dim(\mathcal{Y}_A) = \operatorname{rank}\left(\operatorname{Tr}_{\mathcal{Y}_B}(yy^*)\right) = \dim(\mathcal{Y}_B).$$

That we can make this assumption follow from a consideration of Schmidt decompositions of x and y:

$$x = \sum_{j=1}^{m} \sqrt{p_j}\, x_{A,j} \otimes x_{B,j} \quad \text{and} \quad y = \sum_{k=1}^{n} \sqrt{q_k}\, y_{A,k} \otimes y_{B,k},$$

where $p_1, \ldots, p_m > 0$ and $q_1, \ldots, q_n > 0$, so that $m = \operatorname{rank}\left(\operatorname{Tr}_{\mathcal{X}_B}(xx^*)\right)$ and $n = \operatorname{rank}\left(\operatorname{Tr}_{\mathcal{Y}_B}(yy^*)\right)$. By restricting $\mathcal{X}_A$ to $\operatorname{span}\{x_{A,1}, \ldots, x_{A,m}\}$, $\mathcal{X}_B$ to $\operatorname{span}\{x_{B,1}, \ldots, x_{B,m}\}$, $\mathcal{Y}_A$ to $\operatorname{span}\{y_{A,1}, \ldots, y_{A,n}\}$, and $\mathcal{Y}_B$ to $\operatorname{span}\{y_{B,1}, \ldots, y_{B,n}\}$, we have that the spaces $\mathcal{X}_A$, $\mathcal{X}_B$, $\mathcal{Y}_A$, and $\mathcal{Y}_B$ are only as large in dimension as they need to be to support the vectors x and y. The reason why this assumption causes no loss of generality is that neither the notion of an LOCC channel transforming xx^* to yy^*, nor the majorization relationship $\operatorname{Tr}_{\mathcal{X}_B}(xx^*) \prec \operatorname{Tr}_{\mathcal{Y}_B}(yy^*)$, is sensitive to the possibility that the ambient spaces in which xx^* and yy^* exist are larger than necessary to support x and y.

16.1 The easier implication: from mixed unitary channels to LOCC channels

We will begin with the easier implication of Nielsen's theorem, which states that the majorization relationship

$$\operatorname{Tr}_{\mathcal{X}_B}(xx^*) \prec \operatorname{Tr}_{\mathcal{Y}_B}(yy^*) \tag{16.2}$$

implies the existence of an LOCC channel mapping xx^* to yy^*. To prove the implication, let us begin by letting $X \in \mathrm{L}(\mathcal{X}_B, \mathcal{X}_A)$ and $Y \in \mathrm{L}(\mathcal{Y}_B, \mathcal{Y}_A)$ satisfy

$$x = \operatorname{vec}(X) \qquad \text{and} \qquad y = \operatorname{vec}(Y),$$

so that (16.2) is equivalent to $XX^* \prec YY^*$. The assumption $\dim(\mathcal{X}_A) = \operatorname{rank}\left(\operatorname{Tr}_{\mathcal{X}_B}(xx^*)\right) = \dim(\mathcal{X}_B)$ implies that XX^* is positive definite (and therefore X is invertible). Likewise, the assumption $\dim(\mathcal{Y}_A) = \operatorname{rank}\left(\operatorname{Tr}_{\mathcal{Y}_B}(yy^*)\right) = \dim(\mathcal{Y}_B)$ implies that YY^* is positive definite. It follows that

$$XX^* = \Psi(WYY^*W^*)$$

for some choice of an isometry $W \in \mathrm{U}(\mathcal{Y}_A, \mathcal{X}_A)$ and a mixed unitary channel $\Psi \in \mathrm{C}(\mathcal{X}_A)$. Let us write this channel as

$$\Psi(\rho) = \sum_{a \in \Sigma} p(a) U_a \rho U_a^*$$

for Σ being a finite and nonempty set, $p \in \mathbb{R}^{\Sigma}$ being a probability vector, and $\{U_a : a \in \Sigma\} \subset \mathrm{U}(\mathcal{X}_A)$ being a collection of unitary operators.

Next, define a channel $\Xi \in \mathrm{C}(\mathcal{X}_A \otimes \mathcal{X}_B, \mathcal{X}_A \otimes \mathcal{Y}_B)$ as

$$\Xi(\rho) = \sum_{a \in \Sigma} \left(U_a^* \otimes \overline{B_a}\right) \rho \left(U_a^* \otimes \overline{B_a}\right)^*$$

for each $\rho \in \mathrm{L}(\mathcal{X}_A \otimes \mathcal{X}_B)$, where $B_a \in \mathrm{L}(\mathcal{X}_B, \mathcal{Y}_B)$ is defined as

$$B_a = \sqrt{p(a)}\left(X^{-1} U_a W Y\right)^*$$

for each $a \in \Sigma$. It holds that

$$\sum_{a\in\Sigma} B_a^* B_a = \sum_{a\in\Sigma} p(a) X^{-1} U_a W Y Y^* W^* U_a^* (X^{-1})^* = X^{-1} \Psi\left(W Y Y^* W^*\right) (X^{-1})^* = \mathbb{1}_{\mathcal{X}_A},$$

and therefore

$$\sum_{a\in\Sigma} B_a^{\mathsf{T}} \overline{B_a} = \overline{\sum_{a\in\Sigma} B_a^* B_a} = \mathbb{1}_{\mathcal{X}_A}.$$

It follows that Ξ is trace-preserving, because

$$\sum_{a\in\Sigma} \left(U_a^* \otimes \overline{B_a}\right)^* \left(U_a^* \otimes \overline{B_a}\right) = \sum_{a\in\Sigma} \left(\mathbb{1}_{\mathcal{X}_A} \otimes B_a^{\mathsf{T}} \overline{B_a}\right) = \mathbb{1}_{\mathcal{X}_A} \otimes \mathbb{1}_{\mathcal{X}_A}.$$

The channel Ξ is, in fact, an LOCC channel. To implement it as an LOCC channel, Bob may first apply the local channel

$$\xi \mapsto \sum_{a\in\Sigma} E_{a,a} \otimes \overline{B_a} \xi B_a^{\mathsf{T}},$$

which has the form of a mapping from $\mathrm{L}(\mathcal{X}_B)$ to $\mathrm{L}(\mathcal{Z} \otimes \mathcal{Y}_B)$ for $\mathcal{Z} = \mathbb{C}^\Sigma$. He then sends the register Z corresponding to the space $\mathcal{Z}$ through a classical channel to Alice. Alice then performs the local channel given by

$$\sigma \mapsto \sum_{b\in\Sigma} \left(U_b^* \otimes e_b^*\right) \sigma \left(U_b^* \otimes e_b^*\right)^*,$$

which has the form of a mapping from $\mathrm{L}(\mathcal{X}_A \otimes \mathcal{Z})$ to $\mathrm{L}(\mathcal{X}_A)$. The composition of these three channels is given by Ξ, which shows that $\Xi \in \mathrm{LOCC}(\mathcal{X}_A, \mathcal{X}_A : \mathcal{X}_B, \mathcal{Y}_B)$ as claimed.

The channel Ξ almost satisfies the requirements of the theorem, for we have

$$\begin{aligned}
\Xi(xx^*) &= \sum_{a\in\Sigma} \left(U_a^* \otimes \overline{B_a}\right) \mathrm{vec}(X)\, \mathrm{vec}(X)^* \left(U_a^* \otimes \overline{B_a}\right)^* \\
&= \sum_{a\in\Sigma} \mathrm{vec}\left(U_a^* X B_a^*\right) \mathrm{vec}\left(U_a^* X B_a^*\right)^* \\
&= \sum_{a\in\Sigma} p(a)\, \mathrm{vec}\left(U_a^* X X^{-1} U_a W Y\right) \mathrm{vec}\left(U_a^* X X^{-1} U_a W Y\right)^* \\
&= \mathrm{vec}(WY)\, \mathrm{vec}(WY)^* \\
&= (W \otimes \mathbb{1}_{\mathcal{Y}_B}) yy^* (W \otimes \mathbb{1}_{\mathcal{Y}_B})^*.
\end{aligned}$$

That is, Ξ transforms xx^* to yy^*, followed by the isometry W being applied to Alice's space $\mathcal{Y}_A$, embedding it in $\mathcal{X}_A$. To "undo" this embedding, Alice may apply the channel

$$\xi \mapsto W^* \xi W + \langle \mathbb{1}_{\mathcal{Y}_A} - WW^*, \xi \rangle\, \sigma \tag{16.3}$$

to her portion of the state $(W \otimes \mathbb{1}_{\mathcal{Y}_B}) yy^* (W \otimes \mathbb{1}_{\mathcal{Y}_B})^*$, where $\sigma \in \mathrm{D}(\mathcal{Y}_A)$ is an arbitrary density matrix that has no influence on the proof. Letting $\Phi \in \mathrm{C}(\mathcal{X}_A \otimes \mathcal{X}_B, \mathcal{Y}_A \otimes \mathcal{Y}_B)$ be the channel that results from composing (16.3) with Ξ, we have that Φ is an LOCC channel and satisfies $\Phi(xx^*) = yy^*$ as required.

16.2 The harder implication: from LOCC channels to mixed unitary channels

The reverse implication, from the one proved in the previous section, states that if $\Phi(xx^*) = yy^*$ for an LOCC channel $\Phi \in \mathrm{C}(\mathcal{X}_A, \mathcal{Y}_A : \mathcal{X}_B, \mathcal{Y}_B)$, then $\mathrm{Tr}_{\mathcal{X}_B}(xx^*) \prec \mathrm{Tr}_{\mathcal{Y}_B}(yy^*)$. The main difficulty in proving this fact is that our proof must account for all possible LOCC channels, which do not admit a simple mathematical characterization (so far as anyone knows). For instance, a given LOCC channel could potentially require a composition of 1,000,000 channels that alternate between product channels and classical communication channels, possibly without any shorter composition yielding the same channel.

However, in the situation that we only care about the action of a given LOCC channel on a single pure state—such as the state xx^* being considered in the context of the implication we are trying to prove—LOCC channels can always be reduced to a very simple form. To describe this form, let us begin by defining a restricted class of LOCC channels, acting on the space of operators $\mathrm{L}(\mathcal{Z}_A \otimes \mathcal{Z}_B)$ for any fixed choice of complex Euclidean spaces $\mathcal{Z}_A$ and $\mathcal{Z}_B$, as follows.

1. A channel $\Phi \in \mathrm{C}(\mathcal{Z}_A \otimes \mathcal{Z}_B)$ will be said to be an *A→B channel* if there exists a finite and nonempty set Σ, a collection of operators $\{A_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{Z}_A)$ satisfying the constraint

   $$\sum_{a \in \Sigma} A_a^* A_a = \mathbb{1}_{\mathcal{Z}_A},$$

   and a collection of unitary operators $\{U_a : a \in \Sigma\} \subset \mathrm{U}(\mathcal{Z}_B)$ such that

   $$\Phi(\rho) = \sum_{a \in \Sigma} (A_a \otimes U_a) \rho (A_a \otimes U_a)^*.$$

   One imagines that such an operation represents the situation where Alice performs a non-destructive measurement represented by the collection $\{A_a : a \in \Sigma\}$, transmits the result to Bob, and Bob applies a unitary channel to his system that depends on Alice's measurement result.

2. A channel $\Phi \in \mathrm{C}(\mathcal{Z}_A \otimes \mathcal{Z}_B)$ will be said to be a *B→A channel* if there exists a finite and nonempty set Σ, a collection of operators $\{B_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{Z}_B)$ satisfying the constraint

   $$\sum_{a \in \Sigma} B_a^* B_a = \mathbb{1}_{\mathcal{Z}_B},$$

   and a collection of unitary operators $\{V_a : a \in \Sigma\} \subset \mathrm{U}(\mathcal{Z}_A)$ such that

   $$\Phi(\rho) = \sum_{a \in \Sigma} (V_a \otimes B_a) \rho (V_a \otimes B_a)^*.$$

   Such a channel is analogous to an A→B channel, but where the roles of Alice and Bob are reversed. (The channel constructed in the previous section had this basic form, although the operators $\{B_a\}$ were not necessarily square in that case.)

3. Finally, a channel $\Phi \in \mathrm{C}(\mathcal{Z}_A \otimes \mathcal{Z}_B)$ will be said to be a *restricted LOCC channel* if it is a composition of A→B and $B \to$A channels.

It should be noted that the terms *A→B channel*, *B→A channel*, and *restricted LOCC channel* are being used for the sake of this proof only: they are not standard terms, and will not be used elsewhere in the course.

It is not difficult to see that every restricted LOCC channel is an LOCC channel, using a similar argument to the one showing that the channel Ξ from the previous section was indeed an LOCC channel. As the following theorem shows, restricted LOCC channels turn out to be as powerful as general LOCC channels, provided they are free to act on sufficiently large spaces.

Theorem 16.2. *Suppose $\Phi \in \mathrm{LOCC}(\mathcal{X}_A, \mathcal{Y}_A : \mathcal{X}_B, \mathcal{Y}_B)$ is an LOCC channel. There exist complex Euclidean spaces $\mathcal{Z}_A$ and $\mathcal{Z}_B$, linear isometries*

$$V_A \in \mathrm{U}(\mathcal{X}_A, \mathcal{Z}_A), \quad W_A \in \mathrm{U}(\mathcal{Y}_A, \mathcal{Z}_A), \quad V_B \in \mathrm{U}(\mathcal{X}_B, \mathcal{Z}_B), \quad W_B \in \mathrm{U}(\mathcal{Y}_B, \mathcal{Z}_B),$$

and a restricted LOCC channel $\Psi \in \mathrm{C}(\mathcal{Z}_A \otimes \mathcal{Z}_B)$ such that

$$(W_A \otimes W_B)\Phi(\rho)(W_A \otimes W_B)^* = \Psi\left((V_A \otimes V_B)\rho(V_A \otimes V_B)^*\right) \tag{16.4}$$

for all $\rho \in \mathrm{L}(\mathcal{X}_A \otimes \mathcal{X}_B)$.

Remark 16.3. Before we prove this theorem, let us consider what it is saying. Alice's input space $\mathcal{X}_A$ and output space $\mathcal{Y}_A$ may have different dimensions, but we want to view these two spaces as being embedded in a single space $\mathcal{Z}_A$. The isometries V_A and W_A describe these embeddings. Likewise, V_B and W_B describe the embeddings of Bob's input and output spaces $\mathcal{X}_B$ and $\mathcal{Y}_B$ in a single space $\mathcal{Z}_B$. The above equation (16.4) simply means that Ψ correctly represents Φ in terms of these embeddings: Alice and Bob could either embed the input $\rho \in \mathrm{L}(\mathcal{X}_A \otimes \mathcal{X}_B)$ in $\mathrm{L}(\mathcal{Z}_A \otimes \mathcal{Z}_B)$ as $(V_A \otimes V_B)\rho(V_A \otimes V_B)^*$, and then apply Ψ; or they could first perform Φ and then embed the output $\Phi(\rho)$ into $\mathrm{L}(\mathcal{Z}_A \otimes \mathcal{Z}_B)$ as $(W_A \otimes W_B)\Phi(\rho)(W_A \otimes W_B)^*$. The equation (16.4) means that they obtain the same thing either way.

Proof. Let us suppose that Φ is a composition of mappings

$$\Phi = \Phi_{n-1} \cdots \Phi_1,$$

where each mapping Φ_k takes the form

$$\Phi_k \in \mathrm{LOCC}\left(\mathcal{X}_A^k, \mathcal{X}_A^{k+1} : \mathcal{X}_B^k, \mathcal{X}_B^{k+1}\right),$$

and is either a local operation for Alice, a local operation for Bob, a classical communication from Alice to Bob, or a classical communication from Bob to Alice. Here we assume

$$\mathcal{X}_A^1 = \mathcal{X}_A, \quad \mathcal{X}_B^1 = \mathcal{X}_B, \quad \mathcal{X}_A^n = \mathcal{Y}_A, \quad \text{and} \quad \mathcal{X}_B^n = \mathcal{Y}_B;$$

the remaining spaces are arbitrary, so long as they have forms that are appropriate to the choices $\Phi_1, \ldots, \Phi_{n-1}$. For instance, if Φ_k is a local operation for Alice, then $\mathcal{X}_B^k = \mathcal{X}_B^{k+1}$, while if Φ_k is a classical communication from Alice to Bob, then $\mathcal{X}_A^k = \mathcal{X}_A^{k+1} \otimes \mathcal{W}_k$ and $\mathcal{X}_B^{k+1} = \mathcal{X}_B^k \otimes \mathcal{W}_k$ for $\mathcal{W}_k$ representing the system that stores the classical information communicated from Alice to Bob. There is no loss of generality in assuming that every such $\mathcal{W}_k$ takes the form $\mathcal{W}_k = \mathbb{C}^\Gamma$ for some fixed finite and non-empty set Γ, chosen to be large enough to account for any one of the message transmissions among the mappings $\Phi_1, \ldots, \Phi_{n-1}$.

We will take

$$\mathcal{Z}_A = \mathcal{X}_A^1 \oplus \cdots \oplus \mathcal{X}_A^n \qquad \text{and} \qquad \mathcal{Z}_B = \mathcal{X}_B^1 \oplus \cdots \oplus \mathcal{X}_B^n.$$

These spaces will generally not have minimal dimension among the possible choices that would work for the proof, but they are convenient choices that allow for a simple presentation of the

proof. Let us also define isometries $V_{A,k} \in \mathrm{U}\left(\mathcal{X}_A^k, \mathcal{Z}_A\right)$ and $V_{B,k} \in \mathrm{U}\left(\mathcal{X}_B^k, \mathcal{Z}_B\right)$ to be the most straightforward ways of embedding $\mathcal{X}_A^k$ into $\mathcal{Z}_A$ and $\mathcal{X}_B^k$ into $\mathcal{Z}_B$, i.e.,

$$V_{A,k}x_k = \underbrace{0 \oplus \cdots \oplus 0}_{k-1 \text{ times}} \oplus x_k \oplus \underbrace{0 \oplus \cdots \oplus 0}_{n-k \text{ times}}$$

for every choice of $k = 1, \ldots, n$ and $x_1 \in \mathcal{X}_A^1, \ldots, x_n \in \mathcal{X}_A^n$, and likewise for $V_{B,1}, \ldots, V_{B,n}$.

Suppose Φ_k is a local operation for Alice. This means that there exist a collection of operators

$$\{A_{k,a} \,:\, a \in \Sigma\} \subset \mathrm{L}\left(\mathcal{X}_A^k, \mathcal{X}_A^{k+1}\right)$$

such that

$$\sum_{a \in \Sigma} A_{k,a}^* A_{k,a} = \mathbb{1}_{\mathcal{X}_A^k}$$

and

$$\Phi_k(\rho) = \sum_{a \in \Sigma} \left(A_{k,a} \otimes \mathbb{1}_{\mathcal{X}_B^k, \mathcal{X}_B^{k+1}}\right) \rho \left(A_{k,a} \otimes \mathbb{1}_{\mathcal{X}_B^k, \mathcal{X}_B^{k+1}}\right)^*.$$

This expression refers to the identity mapping from $\mathcal{X}_B^k$ to $\mathcal{X}_B^{k+1}$, which makes sense if we keep in mind that these spaces are equal (given that Φ_k is a local operation for Alice). We wish to extend this mapping to an A$\to$B channel on $\mathrm{L}\left(\mathcal{Z}_A \otimes \mathcal{Z}_B\right)$. Let

$$\{B_{k,b} \,:\, b \in \Delta\} \subset \mathrm{L}\left(\mathcal{X}_A^{k+1}, \mathcal{X}_A^k\right)$$

be an arbitrary collection of operators for which

$$\sum_{b \in \Delta} B_{k,b}^* B_{k,b} = \mathbb{1}_{\mathcal{X}_A^{k+1}},$$

and define $C_{k,a,b} \in \mathrm{L}\left(\mathcal{Z}_A\right)$ as

$$C_{k,a,b} = \begin{pmatrix} \mathbb{1}_{\mathcal{X}_A^1} & & & & & & & & \\ & \mathbb{1}_{\mathcal{X}_A^2} & & & & & & & \\ & & \ddots & & & & & & \\ & & & 0 & B_{k,b} & & & & \\ & & & A_{k,a} & 0 & & & & \\ & & & & & \mathbb{1}_{\mathcal{X}_A^{k+2}} & & & \\ & & & & & & \ddots & & \\ & & & & & & & \mathbb{1}_{\mathcal{X}_A^n} \end{pmatrix}$$

for each $a \in \Sigma$ and $b \in \Delta$, define $U_k \in \mathrm{U}\left(\mathcal{Z}_B\right)$ as

$$U_k = \begin{pmatrix} \mathbb{1}_{\mathcal{X}_A^1} & & & & & & & & \\ & \mathbb{1}_{\mathcal{X}_A^2} & & & & & & & \\ & & \ddots & & & & & & \\ & & & 0 & \mathbb{1}_{\mathcal{X}_B^{k+1}, \mathcal{X}_B^k} & & & & \\ & & & \mathbb{1}_{\mathcal{X}_B^k, \mathcal{X}_B^{k+1}} & 0 & & & & \\ & & & & & \mathbb{1}_{\mathcal{X}_A^{k+2}} & & & \\ & & & & & & \ddots & & \\ & & & & & & & \mathbb{1}_{\mathcal{X}_A^n} \end{pmatrix},$$

and define

$$\Xi_k(\sigma) = \sum_{\substack{a\in\Sigma\\ b\in\Delta}} \left(C_{k,a,b}\otimes U_k\right)\rho\left(C_{k,a,b}\otimes U_k\right)^*.$$

(In both of the matrices above, empty entries are to be understood as containing zero operators of the appropriate dimensions.) It holds that Ξ_k is an A→B channel, and it may be verified that

$$\Xi_k\left((V_{A,k}\otimes V_{B,k})\rho(V_{A,k}\otimes V_{B,k})^*\right) = (V_{A,k+1}\otimes V_{B,k+1})\Phi_k(\rho)(V_{A,k+1}\otimes V_{B,k+1})^* \tag{16.5}$$

for every $\rho\in \mathrm{L}\left(\mathcal{X}_A^k\otimes\mathcal{X}_B^k\right)$.

In case Φ_k is a local operation for Bob rather than Alice, we define Ξ_k to be a $B\to A$ channel through a similar process, where the roles of Alice and Bob are reversed. The equality (16.5) holds in this case through similar reasoning.

Now suppose that Φ_k is a classical message transmission from Alice to Bob. As stated above, we assume that $\mathcal{X}_A^k = \mathcal{X}_A^{k+1}\otimes\mathbb{C}^\Gamma$ and $\mathcal{X}_B^{k+1} = \mathcal{X}_B^k\otimes\mathbb{C}^\Gamma$. Define

$$C_{k,a} = \begin{pmatrix}
\mathbb{1}_{\mathcal{X}_A^1} & & & & & & & & \\
& \mathbb{1}_{\mathcal{X}_A^2} & & & & & & & \\
& & \ddots & & & & & & \\
& & & 0 & \mathbb{1}_{\mathcal{X}_A^{k+1}}\otimes e_a & & & & \\
& & & \mathbb{1}_{\mathcal{X}_A^{k+1}}\otimes e_a^* & 0 & & & & \\
& & & & & \mathbb{1}_{\mathcal{X}_A^{k+2}} & & & \\
& & & & & & \ddots & & \\
& & & & & & & \mathbb{1}_{\mathcal{X}_A^n}
\end{pmatrix}$$

for each $a\in\Gamma$, define $U_{k,a}\in \mathrm{U}\left(\mathcal{Z}_B\right)$ as

$$U_{k,a} = \begin{pmatrix}
\mathbb{1}_{\mathcal{X}_A^1} & & & & & & & & \\
& \mathbb{1}_{\mathcal{X}_A^2} & & & & & & & \\
& & \ddots & & & & & & \\
& & & 0 & \mathbb{1}_{\mathcal{X}_B^{k+1}}\otimes e_a^* & & & & \\
& & & \mathbb{1}_{\mathcal{X}_B^k}\otimes e_a & \mathbb{1}_{\mathcal{X}_B^{k+1}}\otimes \Pi_a & & & & \\
& & & & & \mathbb{1}_{\mathcal{X}_A^{k+2}} & & & \\
& & & & & & \ddots & & \\
& & & & & & & \mathbb{1}_{\mathcal{X}_A^n}
\end{pmatrix},$$

where

$$\Pi_a = \sum_{\substack{b\in\Gamma\\ b\neq a}} E_{b,b},$$

and define

$$\Xi_k(\sigma) = \sum_{a\in\Gamma}\left(C_{k,a}\otimes U_{k,a}\right)\rho\left(C_{k,a}\otimes U_{k,a}\right)^*.$$

It may be checked that each $U_{k,a}$ is unitary and that $\sum_{a\in\Gamma} C_{k,a}^* C_{k,a} = \mathbb{1}_{\mathcal{Z}_A}$. Thus, Ξ_k is an A→B channel, and once again it may be verified that (16.5) holds for every $\rho\in \mathrm{L}\left(\mathcal{X}_A^k\otimes\mathcal{X}_B^k\right)$. A similar

process is used to define a $B \to A$ channel Ξ_k obeying the equation (16.5) in case Φ_k is a message transmission from Bob to Alice.

By making use of (16.5) iteratively, we find that

$$(\Xi_{n-1} \cdots \Xi_1) \left((V_{A,1} \otimes V_{B,1}) \rho (V_{A,1} \otimes V_{B,1})^* \right) = (V_{A,n} \otimes V_{B,n})(\Phi_{n-1} \cdots \Phi_1)(\rho)(V_{A,n} \otimes V_{B,n})^*.$$

Setting $V_A = V_{A,1}$, $V_B = V_{B,1}$, $W_A = V_{A,n}$, $W_B = V_{A,n}$, and recalling that $\mathcal{Y}_A = \mathcal{X}_A^n$ and $\mathcal{Y}_B = \mathcal{X}_B^n$, we have that $\Psi = \Xi_n \cdots \Xi_1$ is a restricted LOCC channel satisfying the requirements of the theorem. □

Next, we observe that restricted LOCC channels can be "collapsed" to a single A→B or B→A channel, assuming their action on a single know pure state is the only concern.

Lemma 16.4. *For any choice of complex Euclidean spaces $\mathcal{Z}_A$ and $\mathcal{Z}_B$ having equal dimension, every restricted LOCC channel $\Phi \in \mathrm{C}(\mathcal{Z}_A \otimes \mathcal{Z}_B)$, and every vector $x \in \mathcal{Z}_A \otimes \mathcal{Z}_B$, the following statements hold.*

1. *There exists an $A \to B$ channel $\Psi \in \mathrm{C}(\mathcal{Z}_A \otimes \mathcal{Z}_B)$ such that $\Psi(xx^*) = \Phi(xx^*)$.*
2. *There exists a $B \to A$ channel $\Psi \in \mathrm{C}(\mathcal{Z}_A \otimes \mathcal{Z}_B)$ such that $\Psi(xx^*) = \Phi(xx^*)$.*

Proof. The idea of the proof is to show that A→B and B→A channels can be interchanged for fixed pure-state inputs, which allows any restricted LOCC channel to be collapsed to a single A→B or B→A channel by applying the interchanges recursively, and noting that A→B channels and $B \to A$ channels are (separately) both closed under composition.

Suppose that $\{A_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{Z}_A)$ is a collection of operators for which $\sum_{a \in \Sigma} A_a^* A = \mathbb{1}_{\mathcal{Z}_A}$, $\{U_a : a \in \Sigma\} \subset \mathrm{U}(\mathcal{Z}_B)$ is a collection of unitary operators, and

$$\Xi(\rho) = \sum_{a \in \Sigma} (A_a \otimes U_a) \rho (A_a \otimes U_a)^*$$

is the A→B channel that is described by these operators. Let $X \in \mathrm{L}(\mathcal{Z}_B, \mathcal{Z}_A)$ satisfy $\mathrm{vec}(X) = x$. It holds that

$$\Xi(xx^*) = \Xi(\mathrm{vec}(X)\,\mathrm{vec}(X)^*) = \sum_{a \in \Sigma} \mathrm{vec}\left(A_a X U_a^{\mathsf{T}}\right) \mathrm{vec}\left(A_a X U_a^{\mathsf{T}}\right)^*.$$

Our goal is to find a collection of operators $\{B_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{Z}_B)$ satisfying $\sum_{a \in \Sigma} B_a^* B_a = \mathbb{1}_{\mathcal{Z}_B}$ and a collection of unitary operators $\{V_a : a \in \Sigma\} \subset \mathrm{U}(\mathcal{Z}_A)$ such that

$$V_a X B_a^{\mathsf{T}} = A_a X U_a^{\mathsf{T}}$$

for all $a \in \Sigma$. If such a collection of operators is found, then we will have that

$$\begin{aligned} \sum_{a \in \Sigma} (V_a \otimes B_a)\,\mathrm{vec}(X)\,\mathrm{vec}(X)^* (V_a \otimes B_a)^* &= \sum_{a \in \Sigma} \mathrm{vec}(V_a X B_a^{\mathsf{T}})\,\mathrm{vec}(V_a X B_a^{\mathsf{T}})^* \\ &= \sum_{a \in \Sigma} \mathrm{vec}(A_a X U_a^{\mathsf{T}})\,\mathrm{vec}(A_a X U_a^{\mathsf{T}})^* = \sum_{a \in \Sigma} (A_a \otimes U_a)\,\mathrm{vec}(X)\,\mathrm{vec}(X)^* (A_a \otimes U_a)^*, \end{aligned}$$

so that $\Xi(uu^*) = \Lambda(uu^*)$ for Λ being the B→A channel defined by

$$\Lambda(\rho) = \sum_{a \in \Sigma} (V_a \otimes B_a) \rho (V_a \otimes B_a)^*.$$

Choose a unitary operator $U \in \mathrm{U}(\mathcal{Z}_A, \mathcal{Z}_B)$ such that $XU \in \mathrm{Pos}(\mathcal{Z}_A)$. Such a U can be found by considering a singular value decomposition of X. Also, for each $a \in \Sigma$, choose a unitary operator $W_a \in \mathrm{U}(\mathcal{Z}_A, \mathcal{Z}_B)$ such that

$$A_a X U_a^{\mathsf{T}} W_a \in \mathrm{Pos}(\mathcal{Z}_A).$$

We have that

$$A_a X U_a^{\mathsf{T}} W_a = (A_a X U_a^{\mathsf{T}} W_a)^* = (A_a (XU) U^* U_a^{\mathsf{T}} W_a)^* = W_a^* \overline{U_a} U (XU) A_a^*,$$

so that

$$A_a X U_a^{\mathsf{T}} = W_a^* \overline{U_a} U X U A_a^* W_a^*.$$

Define

$$V_a = W_a^* \overline{U_a} U \qquad \text{and} \qquad B_a = (U A_a^* W_a^*)^{\mathsf{T}}$$

for each $a \in \Sigma$. Each V_a is unitary and it can be checked that

$$\sum_{a \in \Sigma} B_a^* B_a = \mathbb{1}_{\mathcal{Z}_B}.$$

We have $V_a X B_a^{\mathsf{T}} = A_a X U_a^{\mathsf{T}}$ as required.

We have therefore proved that for every A→B channel $\Xi \in \mathrm{C}(\mathcal{Z}_A \otimes \mathcal{Z}_B)$, there exists a B→A channel $\Lambda \in \mathrm{C}(\mathcal{Z}_A \otimes \mathcal{Z}_B)$ such that $\Xi(xx^*) = \Lambda(xx^*)$. A symmetric argument shows that for every B→A channel Ξ, there exists an A→B channel Λ such that $\Lambda(xx^*) = \Xi(xx^*)$.

Finally, notice that the composition of any two A→B channels is also an A→B channel, and likewise for B→A channels. Therefore, by applying the above arguments repeatedly for the A→B and B→A channels from which Φ is composed, we find that there exists an A→B channel Ψ such that $\Psi(uu^*) = \Phi(uu^*)$, and likewise for Ψ being a B→A channel. □

We are now prepared to finish the proof of Nielsen's theorem. We assume that there exists an LOCC channel Φ mapping xx^* to yy^*. By Theorem 16.2 and Lemma 16.4, we have that there is no loss of generality in assuming $x, y \in \mathcal{Z}_A \otimes \mathcal{Z}_B$ for $\mathcal{Z}_A$ and $\mathcal{Z}_B$ having equal dimension, and moreover that $\Phi \in \mathrm{C}(\mathcal{Z}_A \otimes \mathcal{Z}_B)$ is a B→A channel. Write

$$\Phi(\rho) = \sum_{a \in \Sigma} (V_a \otimes B_a) \rho (V_a \otimes B_a)^*,$$

for $\{B_a : a \in \Sigma\}$ satisfying $\sum_{a \in \Sigma} B_a^* B_a = \mathbb{1}_{\mathcal{Z}_B}$ and $\{V_a : a \in \Sigma\}$ being a collection of unitary operators on $\mathcal{Z}_A$.

Let $X, Y \in \mathrm{L}(\mathcal{Z}_B, \mathcal{Z}_A)$ satisfy $x = \mathrm{vec}(X)$ and $y = \mathrm{vec}(Y)$, so that

$$\Phi(\mathrm{vec}(X)\,\mathrm{vec}(X)^*) = \sum_{a \in \Sigma} \mathrm{vec}(V_a X B_a^{\mathsf{T}})\,\mathrm{vec}(V_a X B_a^{\mathsf{T}})^* = \mathrm{vec}(Y)\,\mathrm{vec}(Y)^*.$$

This implies that

$$V_a X B_a^{\mathsf{T}} = \alpha_a Y$$

and therefore

$$X B_a^{\mathsf{T}} = \alpha_a V_a^* Y$$

for each $a \in \Sigma$, where $\{\alpha_a : a \in \Sigma\}$ is a collection of complex numbers. We now have

$$\sum_{a \in \Sigma} |\alpha_a|^2 V_a^* Y Y^* V_a = \sum_{a \in \Sigma} X B_a^{\mathsf{T}} \overline{B_a} X^* = X X^*.$$

Taking the trace of both sides of this equation reveals that $\sum_{a\in\Sigma} |\alpha_a|^2 = 1$. It has therefore been shown that there exists a mixed unitary channel mapping YY^* to XX^*. It therefore holds that $XX^* \prec YY^*$ (or, equivalently, $\mathrm{Tr}_{\mathcal{Z}_B}(xx^*) \prec \mathrm{Tr}_{\mathcal{Z}_B}(yy^*)$) as required.