Chapter 6

Bipartite entanglement

Entanglement is a fundamental quantum information theoretic concept. It is considered by many to be a quintessential characteristic that distinguishes quantum systems from their classical counterparts. Informally speaking, a state of a collection of registers $\mathsf{X}_1, \ldots, \mathsf{X}_n$ is said to be *entangled* when it is not possible to specify the correlations that exist among the registers in classical terms. When it is possible to describe these correlations in classical terms, the registers are said to be in a *separable* state. Entanglement among two or more registers is therefore synonymous with a lack of separability.

This chapter introduces notions associated with bipartite entanglement, in which correlations between precisely two registers (or two collections of registers) are considered. Topics to be discussed include the property of separability, which is applicable not only to states but also to channels and measurements; aspects of entanglement manipulation and quantification; and a discussion of operational phenomena associated with entanglement, including teleportation, dense coding, and non-classical correlations among measurements on separated systems.

6.1 Separability

This section introduces the notion of separability, which is applicable to states, channels, and measurements on bipartite systems. It is possible to define a multipartite variant of this concepts, but only bipartite separability is considered in this book.

6.1.1 Separable operators and states

The property of separability for operators acting on bipartite tensor product spaces is defined as follows.

Definition 6.1. For any choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, the set $\mathrm{Sep}(\mathcal{X}:\mathcal{Y})$ is defined as the set containing all positive semidefinite operators $R \in \mathrm{Pos}(\mathcal{X}\otimes\mathcal{Y})$ for which there exists an alphabet Σ and two collections of positive semidefinite operators,

$$\{P_a \,:\, a\in\Sigma\} \subset \mathrm{Pos}(\mathcal{X}) \quad \text{and} \quad \{Q_a \,:\, a\in\Sigma\} \subset \mathrm{Pos}(\mathcal{Y}), \tag{6.1}$$

such that

$$R = \sum_{a\in\Sigma} P_a \otimes Q_a. \tag{6.2}$$

Elements of the set $\mathrm{Sep}(\mathcal{X}:\mathcal{Y})$ are called *separable operators*.

Remark 6.2. As the previous definition reflects, separability is defined with respect to a particular tensor product structure of the underlying complex Euclidean space of a given operator. When the term *separable operator* is used, one must therefore make this tensor product structure known (if it is not implicit). For instance, an operator $R \in \mathrm{Pos}(\mathcal{X}\otimes\mathcal{Y}\otimes\mathcal{Z})$ may be an element of $\mathrm{Sep}(\mathcal{X}:\mathcal{Y}\otimes\mathcal{Z})$ but not $\mathrm{Sep}(\mathcal{X}\otimes\mathcal{Y}:\mathcal{Z})$.

By restricting the definition above to density operators, one obtains a definition of *separable states*.

Definition 6.3. Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces. One defines

$$\mathrm{SepD}(\mathcal{X}:\mathcal{Y}) = \mathrm{Sep}(\mathcal{X}:\mathcal{Y}) \cap \mathrm{D}(\mathcal{X}\otimes\mathcal{Y}). \tag{6.3}$$

Elements of the set $\mathrm{SepD}(\mathcal{X}:\mathcal{Y})$ are referred to as *separable states* (or *separable density operators*).

Convex properties of separable operators and states

The sets $\mathrm{Sep}(\mathcal{X}:\mathcal{Y})$ and $\mathrm{SepD}(\mathcal{X}:\mathcal{Y})$ possess various properties relating to convexity, a few of which will now be observed.

Proposition 6.4. *For every choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, the set $\mathrm{SepD}(\mathcal{X}:\mathcal{Y})$ is convex, and the set $\mathrm{Sep}(\mathcal{X}:\mathcal{Y})$ is a convex cone.*

Proof. It will first be proved that $\mathrm{Sep}(\mathcal{X}:\mathcal{Y})$ is a convex cone. It suffices to prove that $\mathrm{Sep}(\mathcal{X}:\mathcal{Y})$ is closed under addition as well as multiplication by any nonnegative real number. To this end, assume that $R_0, R_1 \in \mathrm{Sep}(\mathcal{X}:\mathcal{Y})$ are separable operators and $\lambda \geq 0$ is a nonnegative real number. One may write

$$R_0 = \sum_{a\in\Sigma_0} P_a \otimes Q_a \quad \text{and} \quad R_1 = \sum_{a\in\Sigma_1} P_a \otimes Q_a \tag{6.4}$$

for disjoint alphabets Σ_0 and Σ_1, where

$$\{P_a \,:\, a \in \Sigma_0 \cup \Sigma_1\} \subset \mathrm{Pos}(\mathcal{X}) \;\text{ and }\; \{Q_a \,:\, a \in \Sigma_0 \cup \Sigma_1\} \subset \mathrm{Pos}(\mathcal{Y}) \tag{6.5}$$

are collections of positive semidefinite operators. It holds that

$$R_0 + R_1 = \sum_{a\in\Sigma_0\cup\Sigma_1} P_a \otimes Q_a, \tag{6.6}$$

and therefore $R_0 + R_1 \in \mathrm{Sep}(\mathcal{X}:\mathcal{Y})$. Moreover, it holds that

$$\lambda R_0 = \sum_{a\in\Sigma_0} (\lambda P_a) \otimes Q_a. \tag{6.7}$$

As $\lambda P \in \mathrm{Pos}(\mathcal{X})$ for every positive semidefinite operator $P \in \mathrm{Pos}(\mathcal{X})$, it follows that $\lambda R_0 \in \mathrm{Sep}(\mathcal{X}:\mathcal{Y})$.

The fact that $\mathrm{SepD}(\mathcal{X}:\mathcal{Y})$ is convex follows from the fact that it is equal to the intersection of two convex sets, $\mathrm{Sep}(\mathcal{X}:\mathcal{Y})$ and $\mathrm{D}(\mathcal{X}\otimes\mathcal{Y})$. □

The next proposition, when combined with the previous one, implies that $\mathrm{Sep}(\mathcal{X}:\mathcal{Y})$ is equal to the cone generated by $\mathrm{SepD}(\mathcal{X}:\mathcal{Y})$.

Proposition 6.5. *Let $\mathcal{Z}$ be a complex Euclidean space, let $\mathcal{A} \subseteq \mathrm{Pos}(\mathcal{Z})$ be a cone, and assume that $\mathcal{B} = \mathcal{A} \cap \mathrm{D}(\mathcal{Z})$ is nonempty. It holds that*

$$\mathcal{A} = \big\{\lambda\rho \,:\, \lambda \geq 0,\ \rho \in \mathcal{B}\big\}. \tag{6.8}$$

Proof. One may write

$$\mathrm{cone}(\mathcal{B}) = \big\{\lambda\rho \,:\, \lambda \geq 0,\ \rho \in \mathcal{B}\big\} \tag{6.9}$$

for brevity. Suppose first that $\rho \in \mathcal{B}$ and $\lambda \geq 0$. It follows that $\lambda\rho \in \mathcal{A}$ by virtue of the fact that $\mathcal{B} \subseteq \mathcal{A}$ and $\mathcal{A}$ is a cone. Therefore $\mathrm{cone}(\mathcal{B}) \subseteq \mathcal{A}$.

Now suppose that $P \in \mathcal{A}$. If $P = 0$, then one has that $P = \lambda\rho$ for $\lambda = 0$ and $\rho \in \mathcal{B}$ being chosen arbitrarily. If $P \neq 0$, then consider the density operator $\rho = P/\operatorname{Tr}(P)$. It holds that $\rho \in \mathcal{A}$ because $1/\operatorname{Tr}(P) > 0$ and $\mathcal{A}$ is a cone, and therefore $\rho \in \mathcal{B}$. As $P = \lambda\rho$ for $\lambda = \operatorname{Tr}(P) > 0$, it follows that $P \in \mathrm{cone}(\mathcal{B})$. Therefore, $\mathcal{A} \subseteq \mathrm{cone}(\mathcal{B})$, which completes the proof. □

Two equivalent ways of specifying separable states are provided by the next proposition, which is a straightforward consequence of the spectral theorem.

Proposition 6.6. *Let $\xi \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ be a density operator, for complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$. The following statements are equivalent:*

1. $\xi \in \mathrm{SepD}(\mathcal{X} : \mathcal{Y})$.
2. *There exists an alphabet Σ, collections of states $\{\rho_a : a \in \Sigma\} \subseteq \mathrm{D}(\mathcal{X})$ and $\{\sigma_a : a \in \Sigma\} \subseteq \mathrm{D}(\mathcal{Y})$, and a probability vector $p \in \mathcal{P}(\Sigma)$, such that*

$$\xi = \sum_{a \in \Sigma} p(a)\, \rho_a \otimes \sigma_a. \tag{6.10}$$

3. *There exists an alphabet Σ, collections of unit vectors $\{x_a : a \in \Sigma\} \subset \mathcal{X}$ and $\{y_a : a \in \Sigma\} \subset \mathcal{Y}$, and a probability vector $p \in \mathcal{P}(\Sigma)$, such that*

$$\xi = \sum_{a \in \Sigma} p(a)\, x_a x_a^* \otimes y_a y_a^*. \tag{6.11}$$

Proof. The third statement trivially implies the second, and it is immediate that the second statement implies the first, as $\mathrm{SepD}(\mathcal{X} : \mathcal{Y})$ is convex and $\rho_a \otimes \sigma_a \in \mathrm{SepD}(\mathcal{X} : \mathcal{Y})$ for each $a \in \Sigma$. It remains to prove that the first statement implies the third.

Let $\xi \in \mathrm{SepD}(\mathcal{X} : \mathcal{Y})$. As $\xi \in \mathrm{Sep}(\mathcal{X} : \mathcal{Y})$, one may write

$$\xi = \sum_{b \in \Gamma} P_b \otimes Q_b \tag{6.12}$$

for some choice of an alphabet Γ and collections $\{P_b : b \in \Gamma\} \subset \mathrm{Pos}(\mathcal{X})$ and $\{Q_b : b \in \Gamma\} \subset \mathrm{Pos}(\mathcal{Y})$ of positive semidefinite operators. Let $n = \dim(\mathcal{X})$, let $m = \dim(\mathcal{X})$, and consider spectral decompositions of these operators as follows:

$$P_b = \sum_{j=1}^{n} \lambda_j(P_b) u_{b,j} u_{b,j}^* \quad \text{and} \quad Q_b = \sum_{k=1}^{m} \lambda_k(Q_b) v_{b,k} v_{b,k}^*, \tag{6.13}$$

for each $b \in \Gamma$. Define $\Sigma = \Gamma \times \{1, \ldots, n\} \times \{1, \ldots, m\}$, and define

$$\begin{aligned} p((b,j,k)) &= \lambda_j(P_b)\lambda_k(Q_b), \\ x_{(b,j,k)} &= u_{b,j}, \\ y_{(b,j,k)} &= v_{b,k}, \end{aligned} \tag{6.14}$$

for every $(b, j, k) \in \Sigma$. A straightforward computation reveals that

$$\sum_{a \in \Sigma} p(a)\, x_a x_a^* \otimes y_a y_a^* = \sum_{b \in \Gamma} P_b \otimes Q_b = \xi. \tag{6.15}$$

Moreover, each value $p(a)$ is nonnegative, and because

$$\sum_{a \in \Sigma} p(a) = \mathrm{Tr}(\xi) = 1, \tag{6.16}$$

it follows that p is a probability vector. It has therefore been proved that statement 1 implies statement 3. □

By the equivalence of the first and second statements in the previous proposition, it holds that a given separable state $\xi \in \mathrm{SepD}(\mathcal{X} : \mathcal{Y})$ represents a classical probability distribution over independent quantum states of a pair of registers (X, Y); and in this sense the possible states of the registers X and Y, when considered in isolation, are classically correlated.

For a separable state $\xi \in \mathrm{SepD}(\mathcal{X} : \mathcal{Y})$, the expression (6.11) is generally not unique—there may be many inequivalent ways that ξ can be expressed in this form. It is important to observe that an expression of this form cannot necessarily be obtained directly from a spectral decomposition of ξ. Indeed, for some choices of $\xi \in \mathrm{SepD}(\mathcal{X} : \mathcal{Y})$ it may hold that every expression of ξ in the form (6.11) requires that Σ has cardinality strictly larger than $\mathrm{rank}(\xi)$. An upper bound on the size of the alphabet Σ required for an expression of the form (6.11) to exist may, however, be obtained from Carathéodory's theorem (Theorem 1.9).

Proposition 6.7. *Let $\xi \in \mathrm{SepD}(\mathcal{X} : \mathcal{Y})$ for $\mathcal{X}$ and $\mathcal{Y}$ being complex Euclidean spaces. There exists an alphabet Σ such that $|\Sigma| \leq \mathrm{rank}(\xi)^2$, two collections of unit vectors $\{x_a : a \in \Sigma\} \subset \mathcal{X}$ and $\{y_a : a \in \Sigma\} \subset \mathcal{Y}$, and a probability vector $p \in \mathcal{P}(\Sigma)$ such that*

$$\xi = \sum_{a \in \Sigma} p(a)\, x_a x_a^* \otimes y_a y_a^*. \tag{6.17}$$

Proof. By Proposition 6.6 it holds that

$$\mathrm{SepD}(\mathcal{X} : \mathcal{Y}) = \mathrm{conv}\big\{ xx^* \otimes yy^* : x \in \mathcal{S}(\mathcal{X}),\, y \in \mathcal{S}(\mathcal{Y}) \big\}, \tag{6.18}$$

from which it follows that ξ is contained in the set

$$\mathrm{conv}\big\{ xx^* \otimes yy^* : x \in \mathcal{S}(\mathcal{X}),\, y \in \mathcal{S}(\mathcal{Y}),\, \mathrm{im}(xx^* \otimes yy^*) \subseteq \mathrm{im}(\xi) \big\}. \tag{6.19}$$

Every density operator $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ satisfying $\mathrm{im}(\rho) \subseteq \mathrm{im}(\xi)$ is contained in the real affine subspace

$$\big\{H \in \mathrm{Herm}(\mathcal{X} \otimes \mathcal{Y}) \,:\, \mathrm{im}(H) \subseteq \mathrm{im}(\xi),\, \mathrm{Tr}(H) = 1\big\} \tag{6.20}$$

of dimension $\mathrm{rank}(\xi)^2 - 1$, and therefore the proposition follows directly from Carathéodory's theorem. □

By combining the previous proposition with Proposition 6.5, one obtains the following corollary.

Corollary 6.8. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces and let $R \in \mathrm{Sep}(\mathcal{X} : \mathcal{Y})$ be a nonzero operator. There exists an alphabet Σ such that $|\Sigma| \leq \mathrm{rank}(R)^2$, along with two collections of vectors $\{x_a \,:\, a \in \Sigma\} \subset \mathcal{X}$ and $\{y_a \,:\, a \in \Sigma\} \subset \mathcal{Y}$, such that*

$$R = \sum_{a \in \Sigma} x_a x_a^* \otimes y_a y_a^*. \tag{6.21}$$

The last observation to be made about separable operators and states in this subsection is the following proposition, which establishes a basic topological property of the sets $\mathrm{Sep}(\mathcal{X} : \mathcal{Y})$ and $\mathrm{SepD}(\mathcal{X} : \mathcal{Y})$.

Proposition 6.9. *For every choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, the set $\mathrm{SepD}(\mathcal{X} : \mathcal{Y})$ is compact and the set $\mathrm{Sep}(\mathcal{X} : \mathcal{Y})$ is closed.*

Proof. The unit spheres $\mathcal{S}(\mathcal{X})$ and $\mathcal{S}(\mathcal{Y})$ are compact, which implies that their Cartesian product $\mathcal{S}(\mathcal{X}) \times \mathcal{S}(\mathcal{Y})$ is also compact. The function

$$\phi : \mathcal{S}(\mathcal{X}) \times \mathcal{S}(\mathcal{Y}) \to \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y}) : (x, y) \mapsto xx^* \otimes yy^* \tag{6.22}$$

is continuous, and therefore the set

$$\phi(\mathcal{S}(\mathcal{X}) \times \mathcal{S}(\mathcal{Y})) = \big\{xx^* \otimes yy^* \,:\, x \in \mathcal{S}(\mathcal{X}),\, y \in \mathcal{S}(\mathcal{Y})\big\} \tag{6.23}$$

is compact. Because the convex hull of a compact set is necessarily compact, it follows that $\mathrm{SepD}(\mathcal{X} : \mathcal{Y})$ is compact.

As $\mathrm{SepD}(\mathcal{X} : \mathcal{Y})$ is compact, and does not include 0, the cone it generates must be closed, and therefore $\mathrm{Sep}(\mathcal{X} : \mathcal{Y})$ is closed. □

The Horodecki criterion

The next theorem provides an alternative characterization of separability, demonstrating that the property of separability for operators has a close connection with the property of positivity for maps.

Theorem 6.10 (Horodecki criterion)**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces and let $R \in \operatorname{Pos}(\mathcal{X} \otimes \mathcal{Y})$ be a positive semidefinite operator. The following three statements are equivalent:*

1. $R \in \operatorname{Sep}(\mathcal{X} : \mathcal{Y})$.
2. *For every choice of a complex Euclidean space $\mathcal{Z}$ and a positive map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Z})$ it holds that*
$$\big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(R) \in \operatorname{Pos}(\mathcal{Z} \otimes \mathcal{Y}). \tag{6.24}$$
3. *For every positive and unital map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$, it holds that*
$$\big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(R) \in \operatorname{Pos}(\mathcal{Y} \otimes \mathcal{Y}). \tag{6.25}$$

Proof. Suppose first that $R \in \operatorname{Sep}(\mathcal{X} : \mathcal{Y})$, so that

$$R = \sum_{a \in \Sigma} P_a \otimes Q_a \tag{6.26}$$

for some choice of an alphabet Σ and collections $\{P_a : a \in \Sigma\} \subset \operatorname{Pos}(\mathcal{X})$ and $\{Q_a : a \in \Sigma\} \subset \operatorname{Pos}(\mathcal{Y})$. For every complex Euclidean space $\mathcal{Z}$ and every positive map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Z})$ it holds that

$$\big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(R) = \sum_{a \in \Sigma} \Phi(P_a) \otimes Q_a \in \operatorname{Pos}(\mathcal{Z} \otimes \mathcal{Y}), \tag{6.27}$$

by virtue of the fact that $\Phi(P_a)$ is a positive semidefinite operator for each $a \in \Sigma$. Statement 1 therefore implies statement 2.

Statement 2 trivially implies statement 3.

Finally, the fact that statement 3 implies statement 1 will be proved in the contrapositive form. To this end, assume $R \in \operatorname{Pos}(\mathcal{X} \otimes \mathcal{Y})$ is not a separable operator. As $\operatorname{Sep}(\mathcal{X} : \mathcal{Y})$ is a closed, convex cone within the real vector space $\operatorname{Herm}(\mathcal{X} \otimes \mathcal{Y})$, the hyperplane separation theorem (Theorem 1.11) implies that there must exist a Hermitian operator $H \in \operatorname{Herm}(\mathcal{X} \otimes \mathcal{Y})$ such that

$\langle H, R \rangle < 0$ and $\langle H, S \rangle \geq 0$ for every $S \in \mathrm{Sep}(\mathcal{X} : \mathcal{Y})$. The operator H will be used to define a positive and unital map $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ for which

$$\big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(R) \notin \mathrm{Pos}(\mathcal{Y} \otimes \mathcal{Y}). \tag{6.28}$$

First, let $\Psi \in \mathrm{T}(\mathcal{Y}, \mathcal{X})$ be the unique map for which $J(\Psi) = H$, choose $\varepsilon > 0$ to be a sufficiently small positive real number so that the inequality

$$\langle H, R \rangle + \varepsilon \operatorname{Tr}(R) < 0 \tag{6.29}$$

is satisfied, and define $\Xi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ as

$$\Xi(X) = \Psi^*(X) + \varepsilon \operatorname{Tr}(X) \mathbb{1}_{\mathcal{Y}} \tag{6.30}$$

for every $X \in \mathrm{L}(\mathcal{X})$. For arbitrarily chosen positive semidefinite operators $P \in \mathrm{Pos}(\mathcal{X})$ and $Q \in \mathrm{Pos}(\mathcal{Y})$, it is the case that

$$P \otimes \overline{Q} \in \mathrm{Sep}(\mathcal{X} : \mathcal{Y}), \tag{6.31}$$

and therefore

$$0 \leq \big\langle H, P \otimes \overline{Q} \big\rangle = \big\langle P \otimes \overline{Q}, J(\Psi) \big\rangle = \big\langle P, \Psi(Q) \big\rangle. \tag{6.32}$$

The fact that this inequality holds for every choice of $P \in \mathrm{Pos}(\mathcal{X})$ and $Q \in \mathrm{Pos}(\mathcal{Y})$ implies that $\Psi(Q) \in \mathrm{Pos}(\mathcal{X})$ for every choice of $Q \in \mathrm{Pos}(\mathcal{Y})$, and therefore Ψ is a positive map. It follows from Proposition 2.17 that Ψ^* is a positive map as well. For every nonzero positive semidefinite operator $P \in \mathrm{Pos}(\mathcal{X})$, the operator $\Xi(P)$ is therefore equal to a positive semidefinite operator $\Psi^*(P)$ plus a positive multiple of the identity operator.

Now let $A = \Xi(\mathbb{1}_{\mathcal{X}})$, which is necessarily a positive definite operator, and define $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ as

$$\Phi(X) = A^{-\frac{1}{2}} \Xi(X) A^{-\frac{1}{2}} \tag{6.33}$$

for every $X \in \mathrm{L}(\mathcal{X})$. It remains to verify that Φ is indeed a positive and unital map for which (6.28) holds. The positivity of Φ follows from the fact that Ξ is positive, and it holds that

$$\Phi(\mathbb{1}_{\mathcal{X}}) = A^{-\frac{1}{2}} \Xi(\mathbb{1}_{\mathcal{X}}) A^{-\frac{1}{2}} = A^{-\frac{1}{2}} A A^{-\frac{1}{2}} = \mathbb{1}_{\mathcal{Y}}, \tag{6.34}$$

establishing that Φ is unital. Finally, through the following computation one may verify that the operator $\big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(R)$ is not positive semidefinite:

$$\begin{aligned}
&\Big\langle \operatorname{vec}\big(\sqrt{A}\big) \operatorname{vec}\big(\sqrt{A}\big)^*, \big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(R) \Big\rangle \\
&\qquad = \Big\langle \operatorname{vec}\big(\mathbb{1}_{\mathcal{Y}}\big) \operatorname{vec}\big(\mathbb{1}_{\mathcal{Y}}\big)^*, \big(\Xi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(R) \Big\rangle \\
&\qquad = \big\langle J(\Xi^*), R \big\rangle \\
&\qquad = \big\langle J(\Psi) + \varepsilon \mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{Y}}, R \big\rangle \\
&\qquad = \langle H, R \rangle + \varepsilon \operatorname{Tr}(R) \\
&\qquad < 0.
\end{aligned} \tag{6.35}$$

This completes the proof. □

One immediate application of Theorem 6.10 is that it provides a method for proving that certain positive semidefinite operators are not separable. The following example demonstrates this method for two families of states known as *Werner states* and *isotropic states*.

Example 6.11. Let Σ be an alphabet, and let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces of the form $\mathcal{X} = \mathbb{C}^{\Sigma}$ and $\mathcal{Y} = \mathbb{C}^{\Sigma}$. The *swap operator* $W \in \mathrm{L}(\mathcal{X} \otimes \mathcal{Y})$ is the unique operator satisfying

$$W(x \otimes y) = y \otimes x \tag{6.36}$$

for all vectors $x, y \in \mathbb{C}^{\Sigma}$. Equivalently, this operator is given by

$$W = \sum_{a,b \in \Sigma} E_{a,b} \otimes E_{b,a}. \tag{6.37}$$

The operator W is both unitary and Hermitian, having eigenvalues 1 and -1. The eigenspace of W corresponding to the eigenvalue 1 is spanned by the orthonormal collection

$$\left\{ \frac{e_a \otimes e_b + e_b \otimes e_a}{\sqrt{2}} \,:\, a, b \in \Sigma,\, a < b \right\} \cup \big\{ e_a \otimes e_a \,:\, a \in \Sigma \big\}, \tag{6.38}$$

where it has been assumed that a total ordering of the alphabet Σ has been fixed, while the eigenspace corresponding to the eigenvalue -1 is spanned by the orthonormal collection

$$\left\{ \frac{e_a \otimes e_b - e_b \otimes e_a}{\sqrt{2}} \,:\, a, b \in \Sigma,\, a < b \right\}. \tag{6.39}$$

Let $n = |\Sigma|$, and define projection operators Δ_0, Δ_1, Π_0, $\Pi_1 \in \mathrm{Proj}(\mathcal{X} \otimes \mathcal{Y})$ as follows:

$$\Delta_0 = \frac{1}{n} \sum_{a,b \in \Sigma} E_{a,b} \otimes E_{a,b}\,, \quad \Pi_0 = \frac{1}{2}\mathbb{1} \otimes \mathbb{1} + \frac{1}{2}W\,, \tag{6.40}$$

$$\Delta_1 = \mathbb{1} \otimes \mathbb{1} - \Delta_0\,, \qquad \Pi_1 = \mathbb{1} \otimes \mathbb{1} - \Pi_0\,. \tag{6.41}$$

That these operators are indeed projection operators follows from the fact that they are Hermitian and square to themselves. Alternatively, one may observe that $\Delta_0 = uu^*$ is the projection onto the one-dimensional subspace of $\mathcal{X} \otimes \mathcal{Y}$ spanned by the unit vector

$$u = \frac{1}{\sqrt{n}} \sum_{a \in \Sigma} e_a \otimes e_a, \tag{6.42}$$

Δ_1 is the projection onto the orthogonal complement of this subspace, and Π_0 and Π_1 are the projection operators onto the subspaces spanned by the collections (6.38) and (6.39), respectively. (The images of Π_0 and Π_1 are also known as the *symmetric* and *antisymmetric* subspaces of $\mathbb{C}^\Sigma \otimes \mathbb{C}^\Sigma$, and are considered in greater detail and generality in Chapter 7.) It holds that

$$\begin{aligned} &\mathrm{rank}(\Delta_0) = 1, && \mathrm{rank}(\Pi_0) = \binom{n+1}{2}, \\ &\mathrm{rank}(\Delta_1) = n^2 - 1, && \mathrm{rank}(\Pi_1) = \binom{n}{2}. \end{aligned} \tag{6.43}$$

States of the form

$$\lambda \Delta_0 + (1 - \lambda) \frac{\Delta_1}{n^2 - 1} \tag{6.44}$$

are known as *isotropic states*, and states of the form

$$\lambda \frac{\Pi_0}{\binom{n+1}{2}} + (1 - \lambda) \frac{\Pi_1}{\binom{n}{2}} \tag{6.45}$$

are known as *Werner states* (for $\lambda \in [0, 1]$ in both cases).

Now, let $\mathrm{T} \in \mathrm{T}(\mathcal{X})$ denote the transpose mapping, defined by the action $\mathrm{T}(X) = X^{\mathsf{T}}$ for all $X \in \mathrm{L}(\mathcal{X})$. The mapping T is a positive map. Using the observation that

$$(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(\Delta_0) = \frac{1}{n} W, \tag{6.46}$$

which may be verified directly, as well as $\mathrm{T}(\mathbb{1}_{\mathcal{X}}) = \mathbb{1}_{\mathcal{X}}$ and $\mathrm{T}^2 = \mathbb{1}_{\mathrm{L}(\mathcal{X})}$, the following relations may be obtained:

$$(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(\Delta_0) = \frac{1}{n}\Pi_0 - \frac{1}{n}\Pi_1 , \tag{6.47}$$

$$(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(\Delta_1) = \frac{n-1}{n}\Pi_0 + \frac{n+1}{n}\Pi_1 , \tag{6.48}$$

$$(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(\Pi_0) = \frac{n+1}{2}\Delta_0 + \frac{1}{2}\Delta_1 , \tag{6.49}$$

$$(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(\Pi_1) = -\frac{n-1}{2}\Delta_0 + \frac{1}{2}\Delta_1 . \tag{6.50}$$

For $\lambda \in [0,1]$, the equations

$$\begin{aligned} &(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})\left(\lambda\Delta_0 + (1-\lambda)\frac{\Delta_1}{n^2-1}\right) \\ &\qquad = \left(\frac{1+\lambda n}{2}\right)\frac{\Pi_0}{\binom{n+1}{2}} + \left(\frac{1-\lambda n}{2}\right)\frac{\Pi_1}{\binom{n}{2}} \end{aligned} \tag{6.51}$$

and

$$\begin{aligned} &(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})\left(\lambda\frac{\Pi_0}{\binom{n+1}{2}} + (1-\lambda)\frac{\Pi_1}{\binom{n}{2}}\right) \\ &\qquad = \left(\frac{2\lambda-1}{n}\right)\Delta_0 + \left(1 - \frac{2\lambda-1}{n}\right)\frac{\Delta_1}{n^2-1} \end{aligned} \tag{6.52}$$

are implied. It therefore holds that the isotropic state (6.44) is entangled (i.e., not separable) for $\lambda \in (1/n, 1]$, while the Werner state (6.45) is entangled for $\lambda \in [0, 1/2)$.[1]

A separable neighborhood of the identity operator

By means of the Horodecki criterion (Theorem 6.10), it may be proved that there exists a neighborhood of the identity operator $\mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{Y}}$, for any choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, in which *every* positive semidefinite operator is separable. Consequently, every density operator $\mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ that is sufficiently close to the completely mixed state is separable. In order to

[1] It does indeed hold that the isotropic state (6.44) is separable for $\lambda \in [0, 1/n]$ and the Werner state (6.45) is separable for $\lambda \in [1/2, 1]$. These facts are proved in Chapter 7 (q.v. Example 7.26).

prove this fact, which is stated in more precise terms in Theorem 6.14 below, the following lemma will be used.

Lemma 6.12. *Let Σ be an alphabet, let $\mathcal{X}$ and $\mathcal{Y} = \mathbb{C}^{\Sigma}$ be complex Euclidean spaces, and let $\{A_{a,b} : a, b \in \Sigma\} \subset \mathrm{L}(\mathcal{X})$ be a collection of operators. For $A \in \mathrm{L}(\mathcal{X} \otimes \mathcal{Y})$ being the operator defined as*

$$A = \sum_{a,b \in \Sigma} A_{a,b} \otimes E_{a,b}\,, \tag{6.53}$$

one has that

$$\|A\|^2 \leq \sum_{a,b \in \Sigma} \|A_{a,b}\|^2. \tag{6.54}$$

Proof. For each $a \in \Sigma$, define an operator $B_a \in \mathrm{L}(\mathcal{X} \otimes \mathcal{Y})$ as

$$B_a = \sum_{b \in \Sigma} A_{a,b} \otimes E_{a,b}. \tag{6.55}$$

By expanding the product $B_a B_a^*$ and applying the triangle inequality, the multiplicativity of the spectral norm under tensor products, and the spectral norm identity (1.173), one finds that

$$\|B_a B_a^*\| = \left\| \sum_{b \in \Sigma} A_{a,b} A_{a,b}^* \otimes E_{a,a} \right\| \leq \sum_{b \in \Sigma} \|A_{a,b} A_{a,b}^*\| = \sum_{b \in \Sigma} \|A_{a,b}\|^2. \tag{6.56}$$

It holds that

$$A^* A = \sum_{a \in \Sigma} B_a^* B_a, \tag{6.57}$$

and therefore, by (6.56) together with the triangle inequality and spectral norm identity,

$$\|A\|^2 = \|A^* A\| \leq \sum_{a \in \Sigma} \|B_a^* B_a\| \leq \sum_{a,b \in \Sigma} \|A_{a,b}\|^2, \tag{6.58}$$

as required. □

In addition, the following theorem (which is equivalent to Theorem 3.42) will be needed.

Theorem 6.13. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces and let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be a positive and unital map. It holds that*

$$\|\Phi(X)\| \leq \|X\| \tag{6.59}$$

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

Proof. By the assumption that Φ is positive and unital, Proposition 2.17 and Theorem 2.26 imply that Φ^* is positive and trace-preserving. For operators $X \in \mathrm{L}(\mathcal{X})$ and $Y \in \mathrm{L}(\mathcal{Y})$, one therefore has

$$\begin{aligned} \big|\langle Y, \Phi(X)\rangle\big| = \big|\langle \Phi^*(Y), X\rangle\big| &\leq \|X\|\big\|\Phi^*(Y)\big\|_1 \\ &\leq \|X\|\|Y\|_1\|\Phi^*\|_1 = \|X\|\|Y\|_1, \end{aligned} \tag{6.60}$$

where the final equality follows by Corollary 3.43 (to Theorem 3.42). By maximizing over all operators $Y \in \mathrm{L}(\mathcal{Y})$ that satisfy $\|Y\|_1 \leq 1$, one finds that $\|\Phi(X)\| \leq \|X\|$ for every $X \in \mathrm{L}(\mathcal{X})$, as required. □

Theorem 6.14. *Let $H \in \mathrm{Herm}(\mathcal{X} \otimes \mathcal{Y})$ be a Hermitian operator satisfying $\|H\|_2 \leq 1$, for complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$. It holds that*

$$\mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{Y}} - H \in \mathrm{Sep}(\mathcal{X} : \mathcal{Y}). \tag{6.61}$$

Proof. Let $\Phi \in \mathrm{T}(\mathcal{X}, \mathcal{Y})$ be an arbitrarily chosen positive and unital map. Let Σ be the alphabet for which $\mathcal{Y} = \mathbb{C}^{\Sigma}$, and write

$$H = \sum_{a,b \in \Sigma} H_{a,b} \otimes E_{a,b}. \tag{6.62}$$

It holds that

$$\big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(H) = \sum_{a,b \in \Sigma} \Phi(H_{a,b}) \otimes E_{a,b}, \tag{6.63}$$

and therefore

$$\begin{aligned} \big\|\big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(H)\big\|^2 &\leq \sum_{a,b \in \Sigma} \|\Phi(H_{a,b})\|^2 \\ &\leq \sum_{a,b \in \Sigma} \|H_{a,b}\|^2 \leq \sum_{a,b \in \Sigma} \|H_{a,b}\|_2^2 = \|H\|_2^2 \leq 1. \end{aligned} \tag{6.64}$$

(The first inequality is implied by Lemma 6.12, and the second inequality is implied by Theorem 6.13.) The positivity of Φ implies that $(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(H)$ is Hermitian, and therefore $(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(H) \leq \mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{Y}}$. It follows that

$$\big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(\mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{Y}} - H) = \mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{Y}} - (\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(H) \geq 0. \tag{6.65}$$

Because (6.65) holds for all positive and unital maps Φ, one concludes from Theorem 6.10 that $\mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{Y}} - H$ is separable. □

Bipartite operator entanglement rank

Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, and consider the collection of all positive semidefinite operators $R \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$ for which there exists an alphabet Σ and a collection of operators $\{X_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{Y}, \mathcal{X})$ such that

$$R = \sum_{a \in \Sigma} \mathrm{vec}(X_a)\,\mathrm{vec}(X_a)^* \tag{6.66}$$

and $\mathrm{rank}(X_a) \leq 1$ for each $a \in \Sigma$. It holds that an operator $X \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ has rank at most 1 if and only if there exist vectors $u \in \mathcal{X}$ and $v \in \mathcal{Y}$ such that $\mathrm{vec}(X) = u \otimes v$, and from this observation it follows that the collection of operators R just described coincides with $\mathrm{Sep}(\mathcal{X} : \mathcal{Y})$.

It is useful to generalize this notion, allowing for arbitrary upper-bounds on the rank of the operators $\{X_a : a \in \Sigma\}$, along the lines of the following definition.

Definition 6.15. Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces and let $r \geq 1$ be a positive integer. The set $\mathrm{Ent}_r(\mathcal{X} : \mathcal{Y})$ is defined to be the set of all operators $R \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$ for which there exists an alphabet Σ and a collection of operators

$$\{X_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{Y}, \mathcal{X}) \tag{6.67}$$

satisfying $\mathrm{rank}(X_a) \leq r$ for each $a \in \Sigma$, such that

$$R = \sum_{a \in \Sigma} \mathrm{vec}(X_a)\,\mathrm{vec}(X_a)^*. \tag{6.68}$$

An element $R \in \mathrm{Ent}_r(\mathcal{X} : \mathcal{Y})$ is said to have *entanglement rank* bounded by r. The *entanglement rank* of $R \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$, with respect to the bipartition between $\mathcal{X}$ and $\mathcal{Y}$, is the minimum value of $r \geq 1$ such that $R \in \mathrm{Ent}_r(\mathcal{X} : \mathcal{Y})$.

As indicated above, it holds that

$$\mathrm{Sep}(\mathcal{X} : \mathcal{Y}) = \mathrm{Ent}_1(\mathcal{X} : \mathcal{Y}), \tag{6.69}$$

and from Definition 6.15 it is immediate that

$$\mathrm{Ent}_{r-1}(\mathcal{X} : \mathcal{Y}) \subseteq \mathrm{Ent}_r(\mathcal{X} : \mathcal{Y}) \tag{6.70}$$

for every integer $r \geq 2$.

The containment (6.70) is proper, provided $r \leq \min\{\dim(\mathcal{X}), \dim(\mathcal{Y})\}$. To see that this is so, consider any operator $Y \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ having rank equal to r, and suppose that

$$\operatorname{vec}(Y) \operatorname{vec}(Y)^* = \sum_{a \in \Sigma} \operatorname{vec}(X_a) \operatorname{vec}(X_a)^* \tag{6.71}$$

for some collection of operators $\{X_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{Y}, \mathcal{X})$. As the operator represented by this equation has rank equal to 1, it must hold that $X_a = \alpha_a Y$ for each $a \in \Sigma$, for $\{\alpha_a : a \in \Sigma\}$ being a collection of complex numbers satisfying

$$\sum_{a \in \Sigma} |\alpha_a|^2 = 1. \tag{6.72}$$

It is therefore not possible that (6.71) holds when each operator X_a has rank strictly smaller than r, and therefore

$$\operatorname{vec}(Y) \operatorname{vec}(Y)^* \notin \operatorname{Ent}_{r-1}(\mathcal{X} : \mathcal{Y}). \tag{6.73}$$

It is immediate, on the other hand, that $\operatorname{vec}(Y) \operatorname{vec}(Y)^* \in \operatorname{Ent}_r(\mathcal{X} : \mathcal{Y})$.

Finally, one may observe that

$$\operatorname{Ent}_n(\mathcal{X} : \mathcal{Y}) = \operatorname{Pos}(\mathcal{X} \otimes \mathcal{Y}) \tag{6.74}$$

for $n \geq \min\{\dim(\mathcal{X}), \dim(\mathcal{Y})\}$, as every operator $A \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ has rank bounded by n in this case.

The following simple proposition concerning entanglement rank will be useful in the subsequent sections of this chapter.

Proposition 6.16. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let $Y \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$, and assume that $\|Y\| \leq 1$. For every positive integer r and every operator*

$$P \in \operatorname{Ent}_r(\mathcal{X} : \mathcal{Y}) \tag{6.75}$$

having entanglement rank bounded by r, it holds that

$$\left\langle \operatorname{vec}(Y) \operatorname{vec}(Y)^*, P \right\rangle \leq r \operatorname{Tr}(P). \tag{6.76}$$

Proof. Under the assumption that P has entanglement rank bounded by r, one may write

$$P = \sum_{a \in \Sigma} \operatorname{vec}(X_a) \operatorname{vec}(X_a)^* \tag{6.77}$$

for some alphabet Σ and a collection of operators $\{X_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{Y}, \mathcal{X})$ for which $\operatorname{rank}(X_a) \leq r$ for every $a \in \Sigma$. For every operator $X \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$, one has

$$\big|\langle Y, X\rangle\big|^2 \leq \|X\|_1^2 \leq \operatorname{rank}(X)\,\|X\|_2^2, \tag{6.78}$$

so that evaluating the inner product in the statement of the proposition yields

$$\begin{aligned}\big\langle \operatorname{vec}(Y)\operatorname{vec}(Y)^*, P\big\rangle &= \sum_{a\in\Sigma} \big|\langle Y, X_a\rangle\big|^2 \\ &\leq \sum_{a\in\Sigma} \operatorname{rank}(X_a)\big\|X_a\big\|_2^2 \leq r\sum_{a\in\Sigma}\big\|X_a\big\|_2^2 = r\operatorname{Tr}(P),\end{aligned} \tag{6.79}$$

as required. □

Example 6.17. Let Σ be an alphabet, let $n = |\Sigma|$, let $\mathcal{X} = \mathbb{C}^\Sigma$ and $\mathcal{Y} = \mathbb{C}^\Sigma$, and define a density operator $\tau \in \mathrm{D}(\mathcal{X}\otimes\mathcal{Y})$ as

$$\tau = \frac{1}{n}\sum_{a,b\in\Sigma} E_{a,b}\otimes E_{a,b}. \tag{6.80}$$

The density operator τ, which coincides with the isotropic state Δ_0 defined in Example 6.11, is the canonical example of a maximally entangled state with respect to the spaces $\mathcal{X}$ and $\mathcal{Y}$. One may observe that

$$\tau = \frac{1}{n}\operatorname{vec}(\mathbb{1})\operatorname{vec}(\mathbb{1})^* \tag{6.81}$$

for $\mathbb{1}$ denoting the identity operator on $\mathbb{C}^\Sigma$, which may be viewed as an element of the set $\mathrm{L}(\mathcal{Y}, \mathcal{X})$ in the most straightforward way.

For every positive integer r and every density operator

$$\rho \in \mathrm{D}(\mathcal{X}\otimes\mathcal{Y}) \cap \mathrm{Ent}_r(\mathcal{X}:\mathcal{Y}) \tag{6.82}$$

having entanglement rank bounded by r, Proposition 6.16 implies that

$$\langle \tau, \rho\rangle = \frac{1}{n}\big\langle \operatorname{vec}(\mathbb{1})\operatorname{vec}(\mathbb{1})^*, \rho\big\rangle \leq \frac{r}{n}. \tag{6.83}$$

One therefore has that every state of bounded entanglement rank must have a proportionately small inner product with the state τ.

6.1.2 Separable maps and the LOCC paradigm

Separable maps are defined in an analogous way to separable operators, reflecting the natural correspondence between completely positive maps and positive semidefinite operators. The resulting notion of separability for maps, including channels, is algebraic in nature; and it cannot be said that it is directly motivated from a physical or operational viewpoint.

This notion of separability for channels is, however, closely connected to the more operationally motivated notion of channels implementable by *local operations and classical communication* (or *LOCC* for short). An LOCC channel is a channel that can be implemented by two individuals whose local actions are unrestricted (corresponding to arbitrary channels or measurements), but whose communications with one another are restricted to be classical. This paradigm provides a foundation from which properties of entanglement are commonly studied, particularly in settings in which entanglement is viewed as a resource for information processing.

Separable map and channels

As suggested above, the notion of separability for maps is defined in an analogous way to separability for operators. The following definition states this in more precise terms.

Definition 6.18. Let $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, and $\mathcal{W}$ be complex Euclidean spaces. The set

$$\text{SepCP}(\mathcal{X}, \mathcal{Z} : \mathcal{Y}, \mathcal{W}) \tag{6.84}$$

is defined as the set of all completely positive maps of the form

$$\Xi \in \text{CP}(\mathcal{X} \otimes \mathcal{Y}, \mathcal{Z} \otimes \mathcal{W}) \tag{6.85}$$

for which there exists an alphabet Σ and collections of completely positive maps

$$\{\Phi_a : a \in \Sigma\} \subset \text{CP}(\mathcal{X}, \mathcal{Z}) \quad \text{and} \quad \{\Psi_a : a \in \Sigma\} \subset \text{CP}(\mathcal{Y}, \mathcal{W}) \tag{6.86}$$

such that

$$\Xi = \sum_{a \in \Sigma} \Phi_a \otimes \Psi_a. \tag{6.87}$$

Elements of the set $\text{SepCP}(\mathcal{X}, \mathcal{Z} : \mathcal{Y}, \mathcal{W})$ are called *separable maps*.

As the following simple proposition states, separable maps are precisely those completely positive maps having Kraus representations in which the individual Kraus operators are tensor products of operators. A direct proof of this proposition is obtained by considering Kraus representations of the maps Φ_a and Ψ_a in Definition 6.18, along the same lines as the proof of Proposition 6.6.

Proposition 6.19. *Let $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, and $\mathcal{W}$ be complex Euclidean spaces and let*

$$\Xi \in \mathrm{CP}(\mathcal{X} \otimes \mathcal{Y}, \mathcal{Z} \otimes \mathcal{W}) \tag{6.88}$$

be a completely positive map. It holds that $\Phi \in \mathrm{SepCP}(\mathcal{X}, \mathcal{Z} : \mathcal{Y}, \mathcal{W})$ if and only if there exists an alphabet Σ and collections of operators

$$\{A_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{X}, \mathcal{Z}) \quad \text{and} \quad \{B_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{Y}, \mathcal{W}) \tag{6.89}$$

such that

$$\Xi(X) = \sum_{a \in \Sigma} (A_a \otimes B_a) X (A_a \otimes B_a)^* \tag{6.90}$$

for every operator $X \in \mathrm{L}(\mathcal{X} \otimes \mathcal{Y})$.

Another straightforward proposition regarding separable maps is the following proposition, which implies that the set of all separable maps is closed under composition. Like the previous proposition, it may be verified directly.

Proposition 6.20. *Let $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, $\mathcal{W}$, $\mathcal{U}$, and $\mathcal{V}$ be complex Euclidean spaces, and suppose that Φ and Ψ are separable maps of the form*

$$\Phi \in \mathrm{SepCP}(\mathcal{X}, \mathcal{U} : \mathcal{Y}, \mathcal{V}) \quad \text{and} \quad \Psi \in \mathrm{SepCP}(\mathcal{U}, \mathcal{Z} : \mathcal{V}, \mathcal{W}). \tag{6.91}$$

It holds that the composition $\Psi\Phi$ is separable:

$$\Psi\Phi \in \mathrm{SepCP}(\mathcal{X}, \mathcal{Z} : \mathcal{Y}, \mathcal{W}). \tag{6.92}$$

Similar to the analogous case for states, one defines the set of separable channels by simply restricting the definition of separability for completely positive maps to channels.

Definition 6.21. For complex Euclidean spaces $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, and $\mathcal{W}$, one defines

$$\begin{aligned} &\mathrm{SepC}(\mathcal{X}, \mathcal{Z} : \mathcal{Y}, \mathcal{W}) \\ &\quad = \mathrm{SepCP}(\mathcal{X}, \mathcal{Z} : \mathcal{Y}, \mathcal{W}) \cap \mathrm{C}(\mathcal{X} \otimes \mathcal{Y}, \mathcal{Z} \otimes \mathcal{W}). \end{aligned} \tag{6.93}$$

Elements of the set $\mathrm{SepC}(\mathcal{X}, \mathcal{Z} : \mathcal{Y}, \mathcal{W})$ are referred to as *separable channels*.

It should be noted that, unlike the analogous case of states, separable channels need not be equal to convex combinations of product channels, as the following example illustrates.

Example 6.22. Let $\Sigma = \{0,1\}$ denote the binary alphabet, let $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, and $\mathcal{W}$ all be equal to $\mathbb{C}^{\Sigma}$, and define a channel $\Xi \in \mathrm{C}(\mathcal{X} \otimes \mathcal{Y}, \mathcal{Z} \otimes \mathcal{W})$ by the equation

$$\Xi(E_{a,b} \otimes E_{c,d}) = \begin{cases} E_{a,a} \otimes E_{a,a} & \text{if } a = b \text{ and } c = d \\ 0 & \text{if } a \neq b \text{ or } c \neq d, \end{cases} \tag{6.94}$$

holding for all $a, b, c, d \in \Sigma$. It is the case that Ξ is a separable channel, meaning that $\Xi \in \mathrm{SepC}(\mathcal{X}, \mathcal{Z} : \mathcal{Y}, \mathcal{W})$. Indeed, one may write

$$\Xi = \Phi_0 \otimes \Psi_0 + \Phi_1 \otimes \Psi_1 \tag{6.95}$$

for completely positive maps defined as follows:

$$\begin{aligned} \Phi_0(X) &= \langle E_{0,0}, X \rangle E_{0,0}, \quad \Psi_0(X) = \mathrm{Tr}(X) E_{0,0}, \\ \Phi_1(X) &= \langle E_{1,1}, X \rangle E_{1,1}, \quad \Psi_1(X) = \mathrm{Tr}(X) E_{1,1}, \end{aligned} \tag{6.96}$$

for every $X \in \mathrm{L}\big(\mathbb{C}^{\Sigma}\big)$.

It is not possible, however, to express the channel Ξ in the form

$$\Xi = \sum_{a \in \Gamma} p(a) \Phi_a \otimes \Psi_a \tag{6.97}$$

for any choice of an alphabet Γ, a probability vector $p \in \mathcal{P}(\Gamma)$, and two collections of channels

$$\{\Phi_a : a \in \Gamma\} \subseteq \mathrm{C}(\mathcal{X}, \mathcal{Z}) \quad \text{and} \quad \{\Psi_a : a \in \Gamma\} \subseteq \mathrm{C}(\mathcal{Y}, \mathcal{W}). \tag{6.98}$$

To verify this claim, consider the fact that

$$\Xi(E_{0,0} \otimes \rho) = E_{0,0} \otimes E_{0,0} \quad \text{and} \quad \Xi(E_{1,1} \otimes \rho) = E_{1,1} \otimes E_{1,1} \tag{6.99}$$

for every density operator $\rho \in \mathrm{D}(\mathcal{Y})$. If it were the case that (6.97) were true for each Φ_a and Ψ_a being a channel, then one would necessarily have

$$\sum_{a \in \Gamma} p(a) \Phi_a(E_{0,0}) \otimes \Psi_a(\rho) = E_{0,0} \otimes E_{0,0}, \tag{6.100}$$

and therefore, by tracing over the space $\mathcal{Z}$,

$$\sum_{a\in\Sigma} p(a)\Psi_a(\rho) = E_{0,0} \tag{6.101}$$

for every $\rho \in \mathrm{D}(\mathcal{Y})$. By similar reasoning, it would simultaneously hold that

$$\sum_{a\in\Sigma} p(a)\Phi_a(E_{1,1}) \otimes \Psi_a(\rho) = E_{1,1} \otimes E_{1,1}, \tag{6.102}$$

and therefore

$$\sum_{a\in\Sigma} p(a)\Psi_a(\rho) = E_{1,1} \tag{6.103}$$

for every $\rho \in \mathrm{D}(\mathcal{Y})$. The equations (6.101) and (6.103) are in contradiction, implying that Ξ is not equal to a convex combination of product channels.

Intuitively speaking, the situation represented by the previous example is quite simple. Channels that can be expressed as a convex combination of product channels correspond to transformations that may be implemented by means of *local operations and shared randomness*—no communication is needed to implement them, and such channels do not allow for a direct causal relationship to hold among the input and output systems across the bipartition with respect to which separability is considered. The channel Ξ, on the other hand, induces a direct causal relationship of this form.

As the following proposition states, a given completely positive map is separable if and only if its Choi representation is separable, with respect to the natural bipartition of the tensor product space over which it is defined.

Proposition 6.23. *Let $\Xi \in \mathrm{CP}(\mathcal{X}\otimes\mathcal{Y}, \mathcal{Z}\otimes\mathcal{W})$ be a completely positive map, for complex Euclidean spaces $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, and $\mathcal{W}$, and define an isometry*

$$V \in \mathrm{U}(\mathcal{Z}\otimes\mathcal{W}\otimes\mathcal{X}\otimes\mathcal{Y},\ \mathcal{Z}\otimes\mathcal{X}\otimes\mathcal{W}\otimes\mathcal{Y}) \tag{6.104}$$

by the equation

$$V\,\mathrm{vec}(A\otimes B) = \mathrm{vec}(A)\otimes\mathrm{vec}(B) \tag{6.105}$$

holding for all operators $A \in \mathrm{L}(\mathcal{X},\mathcal{Z})$ and $B \in \mathrm{L}(\mathcal{Y},\mathcal{W})$. It holds that

$$\Xi \in \mathrm{SepCP}(\mathcal{X},\mathcal{Z}:\mathcal{Y},\mathcal{W}) \tag{6.106}$$

if and only if

$$VJ(\Xi)V^* \in \mathrm{Sep}(\mathcal{Z}\otimes\mathcal{X}:\mathcal{W}\otimes\mathcal{Y}). \tag{6.107}$$

Proof. Assume first that Ξ is a separable map. By Proposition 6.19, there must exist an alphabet Σ and two collections of operators,

$$\{A_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{X}, \mathcal{Z}) \quad \text{and} \quad \{B_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{Y}, \mathcal{W}), \tag{6.108}$$

such that

$$\Xi(X) = \sum_{a \in \Sigma} (A_a \otimes B_a) X (A_a \otimes B_a)^* \tag{6.109}$$

for every operator $X \in \mathrm{L}(\mathcal{X} \otimes \mathcal{Y})$. The Choi representation of Ξ is therefore given by

$$J(\Xi) = \sum_{a \in \Sigma} \operatorname{vec}(A_a \otimes B_a) \operatorname{vec}(A_a \otimes B_a)^*, \tag{6.110}$$

so that

$$VJ(\Xi)V^* = \sum_{a \in \Sigma} \operatorname{vec}(A_a) \operatorname{vec}(A_a)^* \otimes \operatorname{vec}(B_a) \operatorname{vec}(B_a)^*, \tag{6.111}$$

which is evidently contained in $\operatorname{Sep}(\mathcal{Z} \otimes \mathcal{X} : \mathcal{W} \otimes \mathcal{Y})$.

Conversely, if $VJ(\Xi)V^*$ is separable, then it must be possible to express this operator in the form (6.111) for some choice of an alphabet Σ and two collections of operators as in (6.108). It therefore follows that (6.110) is a Choi representation of Ξ, so that (6.109) holds for all $X \in \mathrm{L}(\mathcal{X} \otimes \mathcal{Y})$. The map Ξ is therefore separable, which completes the proof. □

Remark 6.24. The isometry V defined in Proposition 6.23 may alternatively be defined by the action

$$V(z \otimes w \otimes x \otimes y) = z \otimes x \otimes w \otimes y, \tag{6.112}$$

for every choice of vectors $x \in \mathcal{X}$, $y \in \mathcal{Y}$, $z \in \mathcal{Z}$, and $w \in \mathcal{W}$. In words, this isometry represents a permutation of tensor factors, allowing a relationship concerning separability with respect to a particular bipartition to be stated precisely.

It is not uncommon in the theory of quantum information literature that statements of this nature are made without an explicit mention of such an isometry. This can sometimes simplify expressions and generally does not lead to any confusion—the isometry can usually be taken as being implicit, particularly in cases when the underlying complex Euclidean spaces have distinct names. In the interest of clarity and formality, however, this book will always represent such permutations of tensor factors explicitly.

Separable channels are not capable of creating entanglement; when a separable channel is applied to a separable state, the output is necessarily another separable state. More generally, separable maps cannot cause an increase in entanglement rank, as the following theorem establishes.

Theorem 6.25. *Let $\Xi \in \mathrm{SepCP}(\mathcal{X}, \mathcal{Z} : \mathcal{Y}, \mathcal{W})$ be a separable map, for complex Euclidean spaces $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, and $\mathcal{W}$. For every positive integer r and every operator $P \in \mathrm{Ent}_r(\mathcal{X} : \mathcal{Y})$, it holds that $\Xi(P) \in \mathrm{Ent}_r(\mathcal{Z} : \mathcal{W})$.*

Proof. For an operator $P \in \mathrm{Ent}_r(\mathcal{X} : \mathcal{Y})$ having entanglement rank bounded by r, there must exist an alphabet Γ and a collection of operators

$$\{X_b \,:\, b \in \Gamma\} \subset \mathrm{L}(\mathcal{Y}, \mathcal{X}), \tag{6.113}$$

satisfying $\mathrm{rank}(X_b) \leq r$ for every $b \in \Gamma$, such that

$$P = \sum_{b \in \Gamma} \mathrm{vec}(X_b)\,\mathrm{vec}(X_b)^*. \tag{6.114}$$

By Proposition 6.19, it follows that

$$\begin{aligned} \Xi(P) &= \sum_{a \in \Sigma} \sum_{b \in \Gamma} (A_a \otimes B_a)\,\mathrm{vec}(X_b)\,\mathrm{vec}(X_b)^* (A_a \otimes B_a)^* \\ &= \sum_{a \in \Sigma} \sum_{b \in \Gamma} \mathrm{vec}\big(A_a X_b B_a^{\mathsf{T}}\big)\,\mathrm{vec}\big(A_a X_b B_a^{\mathsf{T}}\big)^* \end{aligned} \tag{6.115}$$

for some choice of an alphabet Σ and two collections of operators

$$\{A_a \,:\, a \in \Sigma\} \subset \mathrm{L}(\mathcal{X}, \mathcal{Z}) \quad \text{and} \quad \{B_a \,:\, a \in \Sigma\} \subset \mathrm{L}(\mathcal{Y}, \mathcal{W}). \tag{6.116}$$

For every $a \in \Sigma$ and $b \in \Gamma$, it holds that

$$\mathrm{rank}\big(A_a X_b B_a^{\mathsf{T}}\big) \leq \mathrm{rank}(X_b) \leq r, \tag{6.117}$$

and therefore $\Xi(P) \in \mathrm{Ent}_r(\mathcal{Z} : \mathcal{W})$, as required. □

Corollary 6.26. *Let $\Xi \in \mathrm{SepCP}(\mathcal{X}, \mathcal{Z} : \mathcal{Y}, \mathcal{W})$ be a separable map, for complex Euclidean spaces $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, and $\mathcal{W}$. For every separable operator $P \in \mathrm{Sep}(\mathcal{X} : \mathcal{Y})$, it holds that $\Xi(P)$ is also separable: $\Xi(P) \in \mathrm{Sep}(\mathcal{Z} : \mathcal{W})$.*

LOCC channels

As was stated at the beginning of the present subsection, LOCC channels represent transformations of quantum states that may be implemented by two individuals that communicate with one another classically and perform quantum channels and measurements on registers they hold locally.

For instance, one individual may apply a combination of channels and measurements to a collection of registers in their possession and transmit the measurement outcomes to the other individual. Upon receiving this transmission, the other individual may apply channels and measurements depending on the communicated measurement outcomes to registers in their possession. In general, LOCC channels may represent the cumulative effect of composing any finite[2] number of transformations of this sort.

The following definition formalizes this notion. Naturally, it is possible to generalize this definition to three or more individuals, although this will not be done in this book.

Definition 6.27. Let $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, and $\mathcal{W}$ be complex Euclidean spaces and let

$$\Xi \in \mathrm{C}(\mathcal{X} \otimes \mathcal{Y}, \mathcal{Z} \otimes \mathcal{W}) \tag{6.118}$$

be a channel. The channel Ξ is an *LOCC channel* under these conditions:

1. The channel Ξ is a *one-way right LOCC channel* if and only if there exists an alphabet Σ and a collection

   $$\{\Phi_a : a \in \Sigma\} \subset \mathrm{CP}(\mathcal{X}, \mathcal{Z}) \tag{6.119}$$

   of completely positive maps satisfying

   $$\sum_{a \in \Sigma} \Phi_a \in \mathrm{C}(\mathcal{X}, \mathcal{Z}), \tag{6.120}$$

   along with a collection

   $$\{\Psi_a : a \in \Sigma\} \subseteq \mathrm{C}(\mathcal{Y}, \mathcal{W}) \tag{6.121}$$

   of channels, such that

   $$\Xi = \sum_{a \in \Sigma} \Phi_a \otimes \Psi_a. \tag{6.122}$$

[2] One may consider variants of the definition that allow for an unbounded number of classical transmissions that terminate with probability 1 according to a chosen stopping rule. Only the finite case is considered in this book for simplicity.

2. The channel Ξ is a *one-way left LOCC channel* if and only if there exists an alphabet Σ and a collection

$$\{\Psi_a : a \in \Sigma\} \subset \mathrm{CP}(\mathcal{Y}, \mathcal{W}) \tag{6.123}$$

of completely positive maps satisfying

$$\sum_{a \in \Sigma} \Psi_a \in \mathrm{C}(\mathcal{Y}, \mathcal{W}), \tag{6.124}$$

along with a collection

$$\{\Phi_a : a \in \Sigma\} \subseteq \mathrm{C}(\mathcal{X}, \mathcal{Z}) \tag{6.125}$$

of channels, such that (6.122) holds.

3. The channel Ξ is an *LOCC channel* if and only if it is equal to a finite composition of one-way left and one-way right LOCC channels. That is, either Ξ is a one-way left LOCC channel, a one-way right LOCC channel, or there exists an integer $m \geq 2$, complex Euclidean spaces $\mathcal{U}_1, \ldots, \mathcal{U}_{m-1}$ and $\mathcal{V}_1, \ldots, \mathcal{V}_{m-1}$, and channels

$$\begin{aligned} \Xi_1 &\in \mathrm{C}(\mathcal{X} \otimes \mathcal{Y}, \mathcal{U}_1 \otimes \mathcal{V}_1), \\ \Xi_2 &\in \mathrm{C}(\mathcal{U}_1 \otimes \mathcal{V}_1, \mathcal{U}_2 \otimes \mathcal{V}_2), \\ &\vdots \\ \Xi_m &\in \mathrm{C}(\mathcal{U}_{m-1} \otimes \mathcal{V}_{m-1}, \mathcal{Z} \otimes \mathcal{W}), \end{aligned} \tag{6.126}$$

each of which is either a one-way left LOCC channel or a one-way right LOCC channel, such that Ξ is equal to the composition $\Xi = \Xi_m \cdots \Xi_1$.

The collection of all such LOCC channels is denoted $\mathrm{LOCC}(\mathcal{X}, \mathcal{Z} : \mathcal{Y}, \mathcal{W})$.

Remark 6.28. In the definition above, one-way left and one-way right LOCC channels represent channels that can be implemented by local operations and one-way classical communication. In both cases, the channel Ξ may be viewed as having resulted from actions performed by two individuals, Alice and Bob. Alice begins with a register X and Bob begins with Y, and as a result of their actions these registers are transformed into Z and W, respectively.

In the case of a one-way right LOCC channel Ξ, the communication is from Alice to Bob (moving to the *right*, assuming Alice is on the left and

Bob is on the right), with the alphabet Σ representing the set of possible classical messages that may be transmitted. Alice's actions are described by a collection of completely positive maps

$$\{\Phi_a : a \in \Sigma\} \subset \mathrm{CP}(\mathcal{X}, \mathcal{Z}) \tag{6.127}$$

that satisfies the constraint

$$\sum_{a \in \Sigma} \Phi_a \in \mathrm{C}(\mathcal{X}, \mathcal{Z}). \tag{6.128}$$

In essence, this collection specifies a quantum instrument (q.v. Section 2.3.2). Assuming the classical communication is represented by a classical register V having associated complex Euclidean space $\mathcal{V} = \mathbb{C}^{\Sigma}$, Alice's action would be described by the channel $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Z} \otimes \mathcal{V})$ defined by

$$\Phi(X) = \sum_{a \in \Sigma} \Phi_a(X) \otimes E_{a,a} \tag{6.129}$$

for all $X \in \mathrm{L}(\mathcal{X})$. The register V is sent to Bob, who observes its classical state (or, equivalently, measures V with respect to the standard basis) and transforms his register Y into W according to the channel $\Psi_a \in \mathrm{C}(\mathcal{Y}, \mathcal{W})$, for $a \in \Sigma$ being the classical state of V that was observed. Assuming that the register V is discarded after Bob applies the appropriate channel, the combined actions of Alice and Bob are described by Ξ.

For a one-way left LOCC channel Ξ, the situation is similar, with the roles of Alice and Bob switched.

It is apparent from Definition 6.27, together with the fact that separable channels are closed under composition (Proposition 6.20), that every LOCC channel is a separable channel.

Proposition 6.29. *For every choice of complex Euclidean spaces $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, and $\mathcal{W}$, it holds that*

$$\mathrm{LOCC}(\mathcal{X}, \mathcal{Z} : \mathcal{Y}, \mathcal{W}) \subseteq \mathrm{SepC}(\mathcal{X}, \mathcal{Z} : \mathcal{Y}, \mathcal{W}). \tag{6.130}$$

6.1.3 Separable and LOCC measurements

As was explained in Section 2.3.1, one may associate a quantum-to-classical channel with each measurement, with the classical output of the channel representing the outcome of the measurement. Through an identification of this sort, the notions of separable and LOCC channels may be extended to measurements.

Definitions of separable and LOCC measurements

The following definition of separable and LOCC measurements refers to an association of quantum-to-classical channels with measurements that has been adapted to a bipartite setting.

Definition 6.30. Let Σ be an alphabet, let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, and let $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$ be a measurement. Define complex Euclidean spaces $\mathcal{Z} = \mathbb{C}^{\Sigma}$ and $\mathcal{W} = \mathbb{C}^{\Sigma}$, and define a channel

$$\Phi_{\mu} \in \mathrm{C}(\mathcal{X} \otimes \mathcal{Y}, \mathcal{Z} \otimes \mathcal{W}) \tag{6.131}$$

as

$$\Phi_{\mu}(X) = \sum_{a \in \Sigma} \langle \mu(a), X \rangle \, E_{a,a} \otimes E_{a,a} \tag{6.132}$$

for every $X \in \mathrm{L}(\mathcal{X} \otimes \mathcal{Y})$. The measurement μ is a *separable measurement* if and only if

$$\Phi_{\mu} \in \mathrm{SepC}(\mathcal{X}, \mathcal{Z} : \mathcal{Y}, \mathcal{W}), \tag{6.133}$$

and μ is an *LOCC measurement* if and only if

$$\Phi_{\mu} \in \mathrm{LOCC}(\mathcal{X}, \mathcal{Z} : \mathcal{Y}, \mathcal{W}). \tag{6.134}$$

For a given measurement μ, the channel Φ_{μ} specified in Definition 6.30 is similar to the quantum-to-classical channel one would normally associate with μ, except that two copies of the measurement outcome are produced rather than one. In a bipartite setting, this is natural way of associating a quantum-to-classical channel with a measurement. If this measurement is performed on a pair of registers (X, Y) by two individuals, Alice and Bob, where it is assumed that Alice holds X and Bob holds Y, the channel Φ_{μ} represents the measurement μ under the assumption that both individuals learn the measurement outcome after the measurement is performed.

One alternative to Definition 6.30 is to replace the channel Φ_{μ} by the quantum-to-classical channel that would ordinarily be associated with the measurement μ, along with a specification of which side of the bipartition the measurement outcome is to fall (requiring this channel to be separable or LOCC, as in the stated definition). In essence, with respect to a situation in which Alice and Bob are performing the measurement μ as suggested above, such a definition specifies which of the two individuals obtains the

measurement outcome. This alternative creates an artificial asymmetry in the definition, but is equivalent to Definition 6.30.

With respect to Definition 6.30, the separability of a given measurement is equivalent to the constraint that each measurement operator is separable, as the following proposition states.

Proposition 6.31. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let Σ be an alphabet, and let μ be a measurement of the form $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$. It holds that μ is a separable measurement if and only if $\mu(a) \in \mathrm{Sep}(\mathcal{X} : \mathcal{Y})$ for every $a \in \Sigma$.*

Proof. Consider the Choi representation of the mapping Φ_μ, as specified in Definition 6.30, which is given by

$$J(\Phi_\mu) = \sum_{a \in \Sigma} E_{a,a} \otimes E_{a,a} \otimes \overline{\mu(a)}. \tag{6.135}$$

Along similar lines to the statement of Proposition 6.23, let

$$V \in \mathrm{U}(\mathcal{Z} \otimes \mathcal{W} \otimes \mathcal{X} \otimes \mathcal{Y},\, \mathcal{Z} \otimes \mathcal{X} \otimes \mathcal{W} \otimes \mathcal{Y}) \tag{6.136}$$

be the isometry defined by the equation

$$V \operatorname{vec}(A \otimes B) = \operatorname{vec}(A) \otimes \operatorname{vec}(B) \tag{6.137}$$

holding for all operators $A \in \mathrm{L}(\mathcal{X}, \mathcal{Z})$ and $B \in \mathrm{L}(\mathcal{Y}, \mathcal{W})$. If it is the case that $\mu(a) \in \mathrm{Sep}(\mathcal{X} \otimes \mathcal{Y})$ for every $a \in \Sigma$, then it follows directly that

$$VJ(\Phi_\mu)V^* \in \mathrm{Sep}(\mathcal{Z} \otimes \mathcal{X} : \mathcal{W} \otimes \mathcal{Y}), \tag{6.138}$$

which implies that μ is a separable measurement by Proposition 6.23.

Now suppose that μ is a separable measurement, so that (6.138) holds. Define a mapping $\Xi_a \in \mathrm{T}(\mathcal{Z} \otimes \mathcal{X} \otimes \mathcal{W} \otimes \mathcal{Y}, \mathcal{X} \otimes \mathcal{Y})$, for each $a \in \Sigma$, as

$$\Xi_a(X) = \big((e_a^* \otimes \mathbb{1}_\mathcal{X}) \otimes (e_a^* \otimes \mathbb{1}_\mathcal{Y})\big) X \big((e_a \otimes \mathbb{1}_\mathcal{X}) \otimes (e_a \otimes \mathbb{1}_\mathcal{Y})\big) \tag{6.139}$$

for all $X \in \mathrm{L}(\mathcal{Z} \otimes \mathcal{X} \otimes \mathcal{W} \otimes \mathcal{Y})$. It is evident from this definition that Ξ_a is a separable mapping for each $a \in \Sigma$, meaning

$$\Xi_a \in \mathrm{SepCP}(\mathcal{Z} \otimes \mathcal{X}, \mathcal{X} : \mathcal{W} \otimes \mathcal{Y}, \mathcal{Y}). \tag{6.140}$$

It holds that

$$\overline{\mu(a)} = \Xi_a\big(VJ(\Phi_\mu)V^*\big) \tag{6.141}$$

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

$$\overline{\mu(a)} \in \text{Sep}(\mathcal{X} : \mathcal{Y}) \tag{6.142}$$

by Corollary 6.26. This is equivalent to $\mu(a) \in \text{Sep}(\mathcal{X} : \mathcal{Y})$ for each $a \in \Sigma$, as the entry-wise complex conjugate of every separable operator is evidently separable, which completes the proof. □

For two complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, along with an alphabet Σ, it is the case that the set of all separable measurements of the form

$$\mu : \Sigma \to \text{Pos}(\mathcal{X} \otimes \mathcal{Y}) \tag{6.143}$$

is a proper subset of the set of all measurements of the same form (aside from the trivial cases in which one of $\dim(\mathcal{X})$, $\dim(\mathcal{Y})$, or $|\Sigma|$ equals 1). As every LOCC channel is separable, it follows that every LOCC measurement is a separable measurement.

One-way LOCC measurements

An interesting restricted type of LOCC measurement is one in which only *one-way communication* is permitted. The following definition formalizes this type of measurement.

Definition 6.32. Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let Σ be an alphabet, and let

$$\mu : \Sigma \to \text{Pos}(\mathcal{X} \otimes \mathcal{Y}) \tag{6.144}$$

be a measurement. The measurement μ is a *one-way LOCC measurement* if and only if either of the following two conditions is met:

1. There exists an alphabet Γ and a measurement $\nu : \Gamma \to \text{Pos}(\mathcal{X})$, along with a measurement $\pi_b : \Sigma \to \text{Pos}(\mathcal{Y})$ for each $b \in \Gamma$, such that the equation

   $$\mu(a) = \sum_{b \in \Gamma} \nu(b) \otimes \pi_b(a) \tag{6.145}$$

   holds for every $a \in \Sigma$. In this case the measurement μ is said to be an *one-way right LOCC measurement*.

2. There exists an alphabet Γ and a measurement $\nu : \Gamma \to \mathrm{Pos}(\mathcal{Y})$, along with a measurement $\pi_b : \Sigma \to \mathrm{Pos}(\mathcal{X})$ for each $b \in \Gamma$, such that the equation
$$\mu(a) = \sum_{b \in \Gamma} \pi_b(a) \otimes \nu(b) \tag{6.146}$$
holds for every $a \in \Sigma$. In this case the measurement μ is said to be a *one-way left LOCC measurement*.

Limitations on state discrimination by separable measurements

One may consider the problem of state discrimination, as was discussed in Chapter 3, in which measurements are restricted to be separable or LOCC measurements. Many examples of sets of orthogonal pure states that cannot be distinguished without error by separable or LOCC measurements are known. The following theorem provides one class of examples, and implies that there exist relatively small sets of orthogonal pure states having this characteristic.

Theorem 6.33. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces of equal dimension n, let*
$$\{U_1, \ldots, U_m\} \in \mathrm{U}(\mathcal{Y}, \mathcal{X}) \tag{6.147}$$
be a set of pairwise orthogonal isometries, meaning that $\langle U_j, U_k \rangle = 0$ for $j \neq k$, and let $u_k \in \mathcal{X} \otimes \mathcal{Y}$ be the vector defined as
$$u_k = \frac{1}{\sqrt{n}} \mathrm{vec}(U_k) \tag{6.148}$$
for each $k \in \{1, \ldots, m\}$. For every separable measurement of the form
$$\mu : \{1, \ldots, m\} \to \mathrm{Sep}(\mathcal{X} : \mathcal{Y}) \tag{6.149}$$
it holds that
$$\sum_{k=1}^{m} \langle \mu(k), u_k u_k^* \rangle \leq n. \tag{6.150}$$

Proof. Under the assumption that μ is a separable measurement, one may write
$$\mu(k) = \sum_{a \in \Sigma} P_{k,a} \otimes Q_{k,a} \tag{6.151}$$

for each $k \in \{1, \ldots, m\}$, for some choice of an alphabet Σ and being some alphabet and collections

$$\begin{aligned} &\{P_{k,a} \,:\, k \in \{1, \ldots, m\},\, a \in \Sigma\} \subset \mathrm{Pos}(\mathcal{X}) \\ &\{Q_{k,a} \,:\, k \in \{1, \ldots, m\},\, a \in \Sigma\} \subset \mathrm{Pos}(\mathcal{Y}) \end{aligned} \tag{6.152}$$

of positive semidefinite operators. (There is no generality lost in using the same alphabet Σ in the expression (6.151) for each choice of k, as one is free to choose Σ to be as large as is needed, and to set $P_{k,a} = 0$ or $Q_{k,a}$ for some choices of k and a as necessary.) It holds that

$$\begin{aligned} \langle \mu(k), \mathrm{vec}(U_k)\,\mathrm{vec}(U_k)^* \rangle &= \sum_{a \in \Sigma} \left\langle U_k, P_{k,a} U_k Q_{k,a}^{\mathsf{T}} \right\rangle \\ &\leq \sum_{a \in \Sigma} \left\| P_{k,a} U_k Q_{k,a}^{\mathsf{T}} \right\|_1 \leq \sum_{a \in \Sigma} \left\| P_{k,a} \right\|_1 \left\| U_k Q_{k,a}^{\mathsf{T}} \right\|_1 \\ &= \sum_{a \in \Sigma} \mathrm{Tr}(P_{k,a})\,\mathrm{Tr}(Q_{k,a}) = \mathrm{Tr}(\mu(k)), \end{aligned} \tag{6.153}$$

and therefore

$$\sum_{k=1}^{m} \langle \mu(k), \mathrm{vec}(U_k)\,\mathrm{vec}(U_k)^* \rangle \leq \sum_{k=1}^{m} \mathrm{Tr}(\mu(k)) = n^2. \tag{6.154}$$

The theorem follows by dividing both sides of this inequality by n. □

For any set of pure states $\{u_1, \ldots, u_m\}$ as described by this theorem, for which $m > n$, one therefore has that

$$\frac{1}{m} \sum_{k=1}^{m} \langle \mu(k), u_k u_k^* \rangle \leq \frac{n}{m} < 1. \tag{6.155}$$

Consequently, for one of these m states being uniformly selected at random, any separable measurement that aims to discriminate these states must err with probability strictly greater than 0.

LOCC discrimination of any pair of orthogonal pure states

Although Theorem 6.33 establishes that there exist relatively small sets of orthonormal pure states that cannot be perfectly discriminated by separable measurements, the same cannot be said about *pairs* of orthonormal pure states. Indeed, every pair of orthonormal pure states can be discriminated without error by a one-way LOCC measurement. The following lemma is used to prove this fact.

Lemma 6.34. *Let $\mathcal{X}$ be a complex Euclidean space, let $X \in \mathrm{L}(\mathcal{X})$ be an operator satisfying $\mathrm{Tr}(X) = 0$, and let $n = \dim(\mathcal{X})$. There exists an orthonormal basis $\{x_1, \ldots, x_n\}$ of $\mathcal{X}$ such that $x_k^* X x_k = 0$ for all $k \in \{1, \ldots, n\}$.*

Proof. The proof is by induction on n. The base case $n = 1$ is immediate, so it will be assumed that $n \geq 2$ for the rest of the proof. It will also be assumed that $\mathcal{X} = \mathbb{C}^n$, which causes no loss of generality.

For every integer $k \in \{1, \ldots, n\}$, it holds that $\lambda_k(X) \in \mathcal{N}(X)$, where $\mathcal{N}(X)$ denotes the numerical range of X. By the Toeplitz–Hausdorff theorem (Theorem 3.56), the numerical range is convex, and therefore

$$0 = \frac{1}{n}\mathrm{Tr}(X) = \frac{1}{n}\sum_{k=1}^{n} \lambda_k(X) \in \mathcal{N}(X). \tag{6.156}$$

By the definition of the numerical range, there must therefore exist a unit vector $x_n \in \mathcal{X}$ such that $x_n^* X x_n = 0$.

Let $V \in \mathrm{U}(\mathbb{C}^{n-1}, \mathbb{C}^n)$ be any isometry that satisfies $x_n \perp \mathrm{im}(V)$, which is equivalent to $VV^* = \mathbb{1} - x_n x_n^*$. It holds that

$$\mathrm{Tr}(V^* X V) = \mathrm{Tr}((\mathbb{1} - x_n x_n^*)X) = \mathrm{Tr}(X) - x_n^* X x_n = 0. \tag{6.157}$$

As $V^* X V \in \mathrm{L}(\mathbb{C}^{n-1})$, the hypothesis of induction implies that there exist an orthonormal basis $\{u_1, \ldots, u_{n-1}\}$ of $\mathbb{C}^{n-1}$ such that

$$u_k^*(V^* X V)u_k = 0 \tag{6.158}$$

for all $k \in \{1, \ldots, n-1\}$. Define $x_k = Vu_k$ for each $k \in \{1, \ldots, n-1\}$, and observe that $\{x_1, \ldots, x_{n-1}\}$ is an orthonormal set, with each element x_k of this set satisfying $x_k^* X x_k = 0$. As V is an isometry and $x_n \perp \mathrm{im}(X)$, it follows that $\{x_1, \ldots, x_n\}$ is an orthonormal basis of $\mathcal{X}$ having the property stated by the lemma. □

Theorem 6.35. *Let $u_0, u_1 \in \mathcal{X} \otimes \mathcal{Y}$ be orthogonal unit vectors, for $\mathcal{X}$ and $\mathcal{Y}$ being complex Euclidean spaces. There exists a one-way LOCC measurement*

$$\mu : \{0, 1\} \to \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y}) \tag{6.159}$$

such that

$$\langle \mu(0), u_0 u_0^* \rangle = 1 = \langle \mu(1), u_1 u_1^* \rangle. \tag{6.160}$$

Proof. Let $n = \dim(\mathcal{Y})$ and let $A_0, A_1 \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ be the unique operators satisfying $u_0 = \mathrm{vec}(A_0)$ and $u_1 = \mathrm{vec}(A_1)$. The orthogonality of the vectors u_0 and u_1 is equivalent to the condition $\mathrm{Tr}(A_0^* A_1) = 0$. By Lemma 6.34, there exists an orthonormal basis $\{x_1, \ldots, x_n\}$ of $\mathcal{Y}$ with the property that $x_k^* A_0^* A_1 x_k = 0$, which is equivalent to the condition that

$$\left\langle A_0 x_k x_k^* A_0^*, A_1 x_k x_k^* A_1^* \right\rangle = 0, \tag{6.161}$$

for every $k \in \{1, \ldots, n\}$.

Define a measurement $\nu : \{1, \ldots, n\} \to \mathrm{Pos}(\mathcal{Y})$ as

$$\nu(k) = \overline{x_k} x_k^{\mathsf{T}} \tag{6.162}$$

for each $k \in \{1, \ldots, n\}$. By the equation (6.161), one has that there must exist a measurement $\pi_k : \{0, 1\} \to \mathrm{Pos}(\mathcal{X})$, for each $k \in \{1, \ldots, n\}$, such that

$$\left\langle \pi_k(0), A_1 x_k x_k^* A_1^* \right\rangle = 0 = \left\langle \pi_k(1), A_0 x_k x_k^* A_0^* \right\rangle. \tag{6.163}$$

Finally, define $\mu : \{0, 1\} \to \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$ as

$$\mu(a) = \sum_{k=1}^{n} \pi_k(a) \otimes \nu(k) \tag{6.164}$$

for each $a \in \{0, 1\}$, which is a one-way measurement with respect to the second condition of Definition 6.32. It holds that

$$\begin{aligned} \left\langle \mu(0), u_1 u_1^* \right\rangle &= \sum_{k=1}^{n} \left\langle \pi_k(0), (\mathbb{1} \otimes x_k^{\mathsf{T}}) \,\mathrm{vec}(A_1)\,\mathrm{vec}(A_1)^* (\mathbb{1} \otimes \overline{x_k}) \right\rangle \\ &= \sum_{k=1}^{n} \left\langle \pi_k(0), A_1 x_k x_k^* A_1^* \right\rangle = 0, \end{aligned} \tag{6.165}$$

and through a similar calculation one finds that $\left\langle \mu(1), u_0 u_0^* \right\rangle = 0$, which completes the proof. □

Remark 6.36. The preceding proof may be adapted in a straightforward way to prove that there exists a one-way LOCC measurement respecting the first condition of Definition 6.32, as opposed to the second, that satisfies the requirements of the theorem.

6.2 Manipulation of entanglement

As presented in the previous section, entanglement is defined as a lack of separability—for two complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, a bipartite state $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ that is not contained in the set $\mathrm{SepD}(\mathcal{X} : \mathcal{Y})$ is entangled with respect to the bipartition between $\mathcal{X}$ and $\mathcal{Y}$. This definition is qualitative, in the sense that it does not provide a measure of how much entanglement is present in a given state or suggest how two entangled states might or might not relate to one another. The present section discusses such notions, and develops basic concepts and techniques relating to quantitative aspects of entanglement.

6.2.1 Entanglement transformation

The next theorem establishes a necessary and sufficient condition under which two individuals may transform one pure state into another by means of local operations and classical communication. The condition concerns the reductions of the initial and final pure states to one of the two individuals, requiring that the reduction of the initial state is majorized by the reduction of the final state. This condition is not only equivalent to the existence of an LOCC (or even a separable) channel transforming the initial state to the final state, but also implies that the transformation can be accomplished with one-way classical communication, from either of the two individuals to the other. The theorem offers a tool through which two fundamental ways of quantifying how much entanglement exists in a given state, called the *entanglement cost* and the *distillable entanglement*, may be analyzed for pure states.

Theorem 6.37 (Nielsen's theorem)**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces and let $u, v \in \mathcal{X} \otimes \mathcal{Y}$ be unit vectors. The following statements are equivalent:*

1. $\mathrm{Tr}_{\mathcal{Y}}(uu^*) \prec \mathrm{Tr}_{\mathcal{Y}}(vv^*)$.
2. *There exists an alphabet Σ, a collection of operators $\{B_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{Y})$ satisfying*
$$\sum_{a \in \Sigma} B_a^* B_a = \mathbb{1}_{\mathcal{Y}}, \tag{6.166}$$
and a collection of unitary operators $\{U_a : a \in \Sigma\} \subset \mathrm{U}(\mathcal{X})$ such that
$$vv^* = \sum_{a \in \Sigma} (U_a \otimes B_a) uu^* (U_a \otimes B_a)^*. \tag{6.167}$$

3. *There exists an alphabet* Σ*, a collection of operators* $\{A_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{X})$ *satisfying*

$$\sum_{a\in\Sigma} A_a^* A_a = \mathbb{1}_{\mathcal{X}}, \tag{6.168}$$

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

$$vv^* = \sum_{a\in\Sigma} (A_a \otimes V_a) uu^* (A_a \otimes V_a)^*. \tag{6.169}$$

4. *There exists a separable channel*[3] $\Phi \in \mathrm{SepC}(\mathcal{X} : \mathcal{Y})$ *such that* $vv^* = \Phi(uu^*)$.

Proof. Let $X, Y \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ be the unique operators for which $u = \mathrm{vec}(X)$ and $v = \mathrm{vec}(Y)$, and let

$$X = \sum_{k=1}^{r} s_k x_k y_k^* \tag{6.170}$$

be a singular value decomposition of X, for $r = \mathrm{rank}(X)$.

Assume first that statement 1 holds, which is equivalent to $XX^* \prec YY^*$. There must therefore exist an alphabet Σ, a probability vector $p \in \mathcal{P}(\Sigma)$, and a collection of unitary operators $\{W_a : a \in \Sigma\} \subset \mathrm{U}(\mathcal{X})$ such that

$$XX^* = \sum_{a\in\Sigma} p(a) W_a YY^* W_a^*. \tag{6.171}$$

Let $\mathcal{Z} = \mathbb{C}^{\Sigma}$ and define an operator $Z \in \mathrm{L}(\mathcal{Y} \otimes \mathcal{Z}, \mathcal{X})$ as

$$Z = \sum_{a\in\Sigma} \sqrt{p(a)}\, W_a Y \otimes e_a^*. \tag{6.172}$$

It holds that

$$ZZ^* = \sum_{a\in\Sigma} p(a) W_a YY^* W_a^* = XX^*, \tag{6.173}$$

and therefore Z and X agree on their singular values, and on the possible choices for their left singular vectors. It follows that one may write

$$Z = \sum_{k=1}^{r} s_k x_k w_k^* \tag{6.174}$$

for $\{w_1, \ldots, w_r\} \subset \mathcal{Y} \otimes \mathcal{Z}$ being an orthonormal collection of vectors. Let $V \in \mathrm{U}(\mathcal{Y}, \mathcal{Y} \otimes \mathcal{Z})$ be an isometry for which $Vy_k = w_k$ for all $k \in \{1, \ldots, r\}$, so that $XV^* = Z$.

[3] As one may expect, the notation $\mathrm{SepC}(\mathcal{X} : \mathcal{Y})$ is a shorthand for $\mathrm{SepC}(\mathcal{X}, \mathcal{X} : \mathcal{Y}, \mathcal{Y})$.

Now, define operators

$$U_a = W_a^* \quad \text{and} \quad B_a = (\mathbb{1}_{\mathcal{Y}} \otimes e_a^*)\overline{V} \tag{6.175}$$

for each $a \in \Sigma$. As V is an isometry, so too is $\overline{V}$, and therefore

$$\sum_{a\in\Sigma} B_a^* B_a = \sum_{a\in\Sigma} V^{\mathsf{T}}(\mathbb{1}_{\mathcal{Y}} \otimes E_{a,a})\overline{V} = V^{\mathsf{T}}\overline{V} = \mathbb{1}_{\mathcal{Y}}. \tag{6.176}$$

It holds that

$$W_a^* X B_a^{\mathsf{T}} = W_a^* X V^*(\mathbb{1}_{\mathcal{Y}} \otimes e_a) = W_a^* Z(\mathbb{1}_{\mathcal{Y}} \otimes e_a) = \sqrt{p(a)}\,Y \tag{6.177}$$

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

$$\begin{aligned}
&\sum_{a\in\Sigma} (U_a \otimes B_a) uu^* (U_a \otimes B_a)^* \\
&\qquad = \sum_{a\in\Sigma} \operatorname{vec}\bigl(W_a^* X B_a^{\mathsf{T}}\bigr) \operatorname{vec}\bigl(W_a^* X B_a^{\mathsf{T}}\bigr)^* \\
&\qquad = \sum_{a\in\Sigma} p(a) \operatorname{vec}(Y) \operatorname{vec}(Y)^* \\
&\qquad = vv^*.
\end{aligned} \tag{6.178}$$

It has been established that statement 1 implies statement 2.

The fact that statement 1 implies statement 3 is established by a similar argument with the roles of $\mathcal{X}$ and $\mathcal{Y}$ exchanged, along with the observation that $\operatorname{Tr}_{\mathcal{Y}}(uu^*) \prec \operatorname{Tr}_{\mathcal{Y}}(vv^*)$ is equivalent to $\operatorname{Tr}_{\mathcal{X}}(uu^*) \prec \operatorname{Tr}_{\mathcal{X}}(vv^*)$.

Statements 2 and 3 each imply statement 4 directly, as the mappings defined by the actions

$$\begin{aligned}
uu^* &\mapsto \sum_{a\in\Sigma} (U_a \otimes B_a) uu^* (U_a \otimes B_a)^*, \\
uu^* &\mapsto \sum_{a\in\Sigma} (A_a \otimes V_a) uu^* (A_a \otimes V_a)^*
\end{aligned} \tag{6.179}$$

are both separable channels.

Finally, assume statement 4 holds, letting $\Phi \in \operatorname{SepC}(\mathcal{X} : \mathcal{Y})$ be a fixed separable channel for which $\Phi(uu^*) = vv^*$. It will be proved that

$$\lambda(XX^*) \prec \lambda(YY^*); \tag{6.180}$$

by Theorem 4.33, this relation is equivalent to $XX^* \prec YY^*$, which in turn is equivalent to statement 1. Let $n = \dim(\mathcal{X})$, and observe that

$$\sum_{k=1}^{n} \lambda_k(XX^*) = \mathrm{Tr}(XX^*) = 1 = \mathrm{Tr}(YY^*) = \sum_{k=1}^{n} \lambda_k(YY^*), \tag{6.181}$$

by the assumption that u and v are unit vectors. By Theorem 4.30, one finds that the relation (6.180) will therefore follow from the inequality

$$\sum_{k=m}^{n} \lambda_k(YY^*) \leq \sum_{k=m}^{n} \lambda_k(XX^*) \tag{6.182}$$

holding for every choice of $m \in \{1, \ldots, n\}$.

By the separability of the channel Φ, there must exist an alphabet Σ and two collections of operators

$$\{A_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{X}) \quad \text{and} \quad \{B_a : a \in \Sigma\} \subset \mathrm{L}(\mathcal{Y}), \tag{6.183}$$

with $\{A_a \otimes B_a : a \in \Sigma\}$ being a set of Kraus operators of Φ, for which

$$vv^* = \sum_{a \in \Sigma} (A_a \otimes B_a) uu^* (A_a \otimes B_a)^*. \tag{6.184}$$

As vv^* is a rank-one operator, it follows that there must exist a probability vector $p \in \mathcal{P}(\Sigma)$ such that

$$(A_a \otimes B_a) uu^* (A_a \otimes B_a)^* = p(a) vv^*, \tag{6.185}$$

which is equivalent to

$$\mathrm{vec}\big(A_a X B_a^{\mathsf{T}}\big)\, \mathrm{vec}\big(A_a X B_a^{\mathsf{T}}\big)^* = p(a)\, \mathrm{vec}(Y)\, \mathrm{vec}(Y)^*, \tag{6.186}$$

for each $a \in \Sigma$. By taking the partial trace over $\mathcal{Y}$, it follows that

$$A_a X B_a^{\mathsf{T}} \overline{B_a} X^* A_a^* = p(a) YY^* \tag{6.187}$$

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

$$\sum_{k=m}^{n} \lambda_k\big(YY^*\big) = \sum_{k=m}^{n} \sum_{a \in \Sigma} \lambda_k\Big(A_a X B_a^{\mathsf{T}} \overline{B_a} X^* A_a^*\Big) \tag{6.188}$$

for each $m \in \{1, \ldots, n\}$.

Next, for each choice of $a \in \Sigma$ and $m \in \{1, \ldots, n\}$, let $\Pi_{a,m} \in \mathrm{Proj}(\mathcal{X})$ be the projection operator onto the orthogonal complement of the subspace of $\mathcal{X}$ spanned by the set $\{A_a x_1, \ldots, A_a x_{m-1}\}$, where one is to assume $x_k = 0$ for $k > r$. By the definition of these projection operators, it is evident that

$$\left\langle \Pi_{m,a}, A_a X B_a^{\mathsf{T}} \overline{B_a} X^* A_a^* \right\rangle = \left\langle \Pi_{m,a}, A_a X_m B_a^{\mathsf{T}} \overline{B_a} X_m^* A_a^* \right\rangle \tag{6.189}$$

for every $a \in \Sigma$ and $m \in \{1, \ldots, n\}$, where

$$X_m = \sum_{k=m}^{r} s_k x_k y_k^*, \tag{6.190}$$

and one is to interpret that $X_m = 0$ for $m > r$. Because each operator $\Pi_{m,a}$ is a projection, and the operator $A_a X_m B_a^{\mathsf{T}} \overline{B_a} X_m^* A_a^*$ is positive semidefinite, it follows that

$$\left\langle \Pi_{m,a}, A_a X_m B_a^{\mathsf{T}} \overline{B_a} X_m^* A_a^* \right\rangle \leq \mathrm{Tr}\Big(A_a X_m B_a^{\mathsf{T}} \overline{B_a} X_m^* A_a^* \Big). \tag{6.191}$$

Using the fact that Φ is a channel, and therefore preserves trace, one finds that

$$\begin{aligned} \sum_{a \in \Sigma} & \mathrm{Tr}\Big(A_a X_m B_a^{\mathsf{T}} \overline{B_a} X_m^* A_a^* \Big) \\ &= \mathrm{Tr}\big(\Phi(\mathrm{vec}(X_m)\,\mathrm{vec}(X_m)^*) \big) \\ &= \mathrm{Tr}\big(\mathrm{vec}(X_m)\,\mathrm{vec}(X_m)^* \big) \\ &= \mathrm{Tr}(X_m X_m^*) \\ &= \sum_{k=m}^{n} \lambda_k(XX^*) \end{aligned} \tag{6.192}$$

for each $m \in \{1, \ldots, n\}$.

Finally, as it necessarily holds that $\mathrm{rank}(\Pi_{a,m}) \geq n - m + 1$ for every $a \in \Sigma$ and $m \in \{1, \ldots, n\}$, it follows that

$$\left\langle \Pi_{m,a}, A_a X B_a^{\mathsf{T}} \overline{B_a} X^* A_a^* \right\rangle \geq \sum_{k=m}^{n} \lambda_k\big(A_a X B_a^{\mathsf{T}} \overline{B_a} X^* A_a^* \big). \tag{6.193}$$

By combining (6.188), (6.189), (6.191), (6.192), and (6.193), one finds that

$$\sum_{k=m}^{n} \lambda_k\big(YY^*\big) \leq \sum_{k=m}^{n} \lambda_k(XX^*), \tag{6.194}$$

which establishes (6.180), and therefore completes the proof. □

Theorem 6.37 implies the following corollary, characterizing the pure state transformations from one tensor product space to a possibly different tensor product space that may be realized by LOCC channels.

Corollary 6.38. *Let $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, and $\mathcal{W}$ be complex Euclidean spaces and let $x \in \mathcal{X} \otimes \mathcal{Y}$ and $y \in \mathcal{Z} \otimes \mathcal{W}$ be unit vectors. The following statements are equivalent:*

1. *For $\rho = \mathrm{Tr}_{\mathcal{Y}}(xx^*)$, $\sigma = \mathrm{Tr}_{\mathcal{W}}(yy^*)$, and $r = \min\{\mathrm{rank}(\rho), \mathrm{rank}(\sigma)\}$, it holds that*
$$\lambda_1(\rho) + \cdots + \lambda_m(\rho) \leq \lambda_1(\sigma) + \cdots + \lambda_m(\sigma) \tag{6.195}$$
for every $m \in \{1, \ldots, r\}$.
2. *There exists a one-way right LOCC channel $\Phi \in \mathrm{LOCC}(\mathcal{X}, \mathcal{Z} : \mathcal{Y}, \mathcal{W})$ for which it holds that $\Phi(xx^*) = yy^*$.*
3. *There exists a one-way left LOCC channel $\Phi \in \mathrm{LOCC}(\mathcal{X}, \mathcal{Z} : \mathcal{Y}, \mathcal{W})$ for which it holds that $\Phi(xx^*) = yy^*$.*
4. *There exists a separable channel $\Phi \in \mathrm{SepC}(\mathcal{X}, \mathcal{Z} : \mathcal{Y}, \mathcal{W})$ for which it holds that $\Phi(xx^*) = yy^*$.*

Proof. Define four linear isometries, $A_0 \in \mathrm{U}(\mathcal{X}, \mathcal{X} \oplus \mathcal{Z})$, $B_0 \in \mathrm{U}(\mathcal{Y}, \mathcal{Y} \oplus \mathcal{W})$, $A_1 \in \mathrm{U}(\mathcal{Z}, \mathcal{X} \oplus \mathcal{Z})$, and $B_1 \in \mathrm{U}(\mathcal{W}, \mathcal{Y} \oplus \mathcal{W})$, as follows:

$$\begin{aligned} A_0 x &= x \oplus 0, \quad & A_1 z &= 0 \oplus z, \\ B_0 y &= y \oplus 0, \quad & B_1 w &= 0 \oplus w, \end{aligned} \tag{6.196}$$

for every choice of vectors $x \in \mathcal{X}$, $y \in \mathcal{Y}$, $z \in \mathcal{Z}$, and $w \in \mathcal{W}$. Also define four channels, $\Psi_0 \in \mathrm{C}(\mathcal{X} \oplus \mathcal{Z}, \mathcal{X})$, $\Lambda_0 \in \mathrm{C}(\mathcal{Y} \oplus \mathcal{W}, \mathcal{Y})$, $\Psi_1 \in \mathrm{C}(\mathcal{X} \oplus \mathcal{Z}, \mathcal{Z})$, and $\Lambda_1 \in \mathrm{C}(\mathcal{Y} \oplus \mathcal{W}, \mathcal{W})$, as

$$\begin{aligned} \Psi_0(X) &= A_0^* X A_0 + \left\langle \mathbb{1}_{\mathcal{X}\oplus\mathcal{Z}} - A_0 A_0^*, X \right\rangle \tau_0, \\ \Lambda_0(Y) &= B_0^* X B_0 + \left\langle \mathbb{1}_{\mathcal{X}\oplus\mathcal{Z}} - B_0 B_0^*, X \right\rangle \xi_0, \\ \Psi_1(X) &= A_1^* X A_1 + \left\langle \mathbb{1}_{\mathcal{X}\oplus\mathcal{Z}} - A_1 A_1^*, X \right\rangle \tau_1, \\ \Lambda_1(Y) &= B_1^* X B_1 + \left\langle \mathbb{1}_{\mathcal{X}\oplus\mathcal{Z}} - B_1 B_1^*, X \right\rangle \xi_1, \end{aligned} \tag{6.197}$$

for all $X \in \mathrm{L}(\mathcal{X} \oplus \mathcal{Z})$ and $Y \in \mathrm{L}(\mathcal{Y} \oplus \mathcal{W})$, where $\tau_0 \in \mathrm{D}(\mathcal{X})$, $\xi_0 \in \mathrm{D}(\mathcal{Y})$, $\tau_1 \in \mathrm{D}(\mathcal{Z})$, and $\xi_1 \in \mathrm{D}(\mathcal{W})$ are fixed, but otherwise arbitrarily selected, density operators.

Assume first that statement 1 holds. One concludes that

$$A_0 \rho A_0^* \prec A_1 \sigma A_1^*, \tag{6.198}$$

and therefore the four equivalent statements of Theorem 6.37 hold for the vectors

$$u = (A_0 \otimes B_0)x \quad \text{and} \quad v = (A_1 \otimes B_1)y. \tag{6.199}$$

There must therefore exist a one-way right LOCC channel Ξ, of the form specified in the statement of Theorem 6.37, such that $\Xi(uu^*) = vv^*$. Define $\Phi \in \mathrm{C}(\mathcal{X} \otimes \mathcal{Y}, \mathcal{Z} \otimes \mathcal{W})$ as

$$\Phi(X) = \big((\Psi_1 \otimes \Lambda_1)\Xi\big)\big((A_0 \otimes B_0)X(A_0 \otimes B_0)^*\big) \tag{6.200}$$

for every $X \in \mathrm{L}(\mathcal{X} \otimes \mathcal{Y})$. It holds that Φ is a one-way right LOCC channel satisfying $\Phi(xx^*) = yy^*$, and therefore statement 1 implies statement 2. The fact that statement 1 implies statement 3 is similar.

Statements 2 and 3 trivially imply that statement 4 holds.

Finally, assume statement 4 holds. Define a channel Ξ as

$$\Xi(Z) = (A_1 \otimes B_1)\big(\Phi(\Psi_0 \otimes \Lambda_0)\big)(X)(A_1 \otimes B_1)^* \tag{6.201}$$

for all $X \in \mathrm{L}((\mathcal{X} \oplus \mathcal{Z}) \otimes (\mathcal{Y} \oplus \mathcal{W}))$. The channel Ξ is separable and satisfies

$$\Xi(uu^*) = vv^* \tag{6.202}$$

for vectors u and v as in (6.199). The four equivalent statements listed in Theorem 6.37 therefore hold for u and v, which implies

$$\begin{aligned} &\mathrm{Tr}_{\mathcal{Y}\oplus\mathcal{W}}\big((A_0 \otimes B_0)xx^*(A_0 \otimes B_0)^*\big) \\ &\qquad \prec \mathrm{Tr}_{\mathcal{Y}\oplus\mathcal{W}}\big((A_1 \otimes B_1)yy^*(A_1 \otimes B_1)^*\big). \end{aligned} \tag{6.203}$$

This relation is equivalent to

$$A_0 \rho A_0^* \prec A_1 \sigma A_1^*, \tag{6.204}$$

which implies that statement 1 holds, and completes the proof. □

6.2.2 Distillable entanglement and entanglement cost

Suppose $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ is a state, for $\mathcal{X}$ and $\mathcal{Y}$ being complex Euclidean spaces. There are various ways in which one may quantify the amount of entanglement that is present in ρ, with respect to the bipartition between $\mathcal{X}$ and $\mathcal{Y}$. The *distillable entanglement* and *entanglement cost* represent two such measures. The distillable entanglement concerns the rate at which copies of the state ρ can be converted into copies of the maximally entangled two-qubit state

$$\tau = \frac{1}{2} \sum_{a,b \in \{0,1\}} E_{a,b} \otimes E_{a,b} \tag{6.205}$$

with high accuracy by means of an LOCC channel. The entanglement cost refers to the reverse process; it is the rate at which approximate copies of ρ may be produced from copies of τ by an LOCC channel. In both cases, it is the asymptotic behavior of these processes, as the number of copies of each state grows, that is taken as the measure of entanglement.

For every bipartite state, the distillable entanglement is upper-bounded by the entanglement cost, with the two measures coinciding for pure states. In general, however, the two quantities may differ, with the entanglement cost being strictly larger than the distillable entanglement in some cases.

Notation related to distillable entanglement and entanglement cost

The following notation will be useful when discussing both the distillable entanglement and entanglement cost of a bipartite state $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$.

For each $n \in \mathbb{N}$, which represents a number of copies of ρ that are to be manipulated for either of the two measures, complex Euclidean spaces $\mathcal{X}_1, \ldots, \mathcal{X}_n$ and $\mathcal{Y}_1, \ldots, \mathcal{Y}_n$ will represent isomorphic copies of $\mathcal{X}$ and $\mathcal{Y}$, respectively, with respect to which the individual copies of ρ are assumed to be defined. Thus, n distinct copies of the state ρ may be represented by the operator

$$\rho^{\otimes n} \in \mathrm{D}\big((\mathcal{X}_1 \otimes \mathcal{Y}_1) \otimes \cdots \otimes (\mathcal{X}_n \otimes \mathcal{Y}_n)\big). \tag{6.206}$$

When it is convenient, the notations

$$\mathcal{X}^{\otimes n} = \mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n \quad \text{and} \quad \mathcal{Y}^{\otimes n} = \mathcal{Y}_1 \otimes \cdots \otimes \mathcal{Y}_n \tag{6.207}$$

will be used, as well as

$$(\mathcal{X} \otimes \mathcal{Y})^{\otimes n} = (\mathcal{X}_1 \otimes \mathcal{Y}_1) \otimes \cdots \otimes (\mathcal{X}_n \otimes \mathcal{Y}_n). \tag{6.208}$$

One may define an isometry

$$U_n \in \mathrm{U}\big((\mathcal{X} \otimes \mathcal{Y})^{\otimes n}, \mathcal{X}^{\otimes n} \otimes \mathcal{Y}^{\otimes n}\big) \tag{6.209}$$

by the action

$$U_n\big(\mathrm{vec}(A_1) \otimes \cdots \otimes \mathrm{vec}(A_n)\big) = \mathrm{vec}(A_1 \otimes \cdots \otimes A_n) \tag{6.210}$$

for all operators $A_1, \ldots, A_n \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$. Equivalently, U_n is defined by the action

$$\begin{aligned} &U_n((x_1 \otimes y_1) \otimes \cdots \otimes (x_n \otimes y_n)) \\ &\quad = (x_1 \otimes \cdots \otimes x_n) \otimes (y_1 \otimes \cdots \otimes y_n) \end{aligned} \tag{6.211}$$

for all vectors $x_1, \ldots, x_n \in \mathcal{X}$ and $y_1, \ldots, y_n \in \mathcal{Y}$. This isometry has the effect of re-ordering the tensor factors of the space $(\mathcal{X} \otimes \mathcal{Y})^{\otimes n}$ so that it takes the form of a bipartite tensor product space $\mathcal{X}^{\otimes n} \otimes \mathcal{Y}^{\otimes n}$ that allows for notions concerning entanglement and separability to be conveniently stated .

The binary alphabet will be denoted $\Gamma = \{0, 1\}$, and the state τ defined above is to be considered as an element of the set $\mathrm{D}(\mathcal{Z} \otimes \mathcal{W})$, for $\mathcal{Z} = \mathbb{C}^{\Gamma}$ and $\mathcal{W} = \mathbb{C}^{\Gamma}$. Along the same lines as the convention described above, complex Euclidean spaces $\mathcal{Z}_1, \ldots, \mathcal{Z}_m$ and $\mathcal{W}_1, \ldots, \mathcal{W}_m$, for each $m \in \mathbb{N}$, will denote isomorphic copies of $\mathcal{Z}$ and $\mathcal{W}$, over which distinct copies of the state τ are defined.

Also similar to above, one may define an isometry

$$V_m \in \mathrm{U}\big((\mathcal{Z} \otimes \mathcal{W})^{\otimes m}, \mathcal{Z}^{\otimes m} \otimes \mathcal{W}^{\otimes m}\big) \tag{6.212}$$

playing an analogous role to the isometry U_n, but for the spaces $\mathcal{Z}_1, \ldots, \mathcal{Z}_m$ and $\mathcal{W}_1, \ldots, \mathcal{W}_m$. This isometry is defined by the action

$$V_m\big(\mathrm{vec}(B_1) \otimes \cdots \otimes \mathrm{vec}(B_m)\big) = \mathrm{vec}(B_1 \otimes \cdots \otimes B_m) \tag{6.213}$$

for all operators $B_1, \ldots, B_m \in \mathrm{L}(\mathcal{W}, \mathcal{Z})$. Equivalently, V_m is defined by the action

$$\begin{aligned} &V_m\big((z_1 \otimes w_1) \otimes \cdots \otimes (z_m \otimes w_m)\big) \\ &\quad = (z_1 \otimes \cdots \otimes z_m) \otimes (w_1 \otimes \cdots \otimes w_m) \end{aligned} \tag{6.214}$$

for all vectors $z_1, \ldots, z_m \in \mathcal{Z}$ and $w_1, \ldots, w_m \in \mathcal{W}$.

Definitions of distillable entanglement and entanglement cost

With respect to the notation introduced above, the *distillable entanglement* and *entanglement cost* are defined as follows.

Definition 6.39. Let X and Y be registers with associated complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, and let $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ be a state of (X, Y). With respect to the state ρ, the *distillable entanglement* $\mathrm{E}_{\mathrm{D}}(\mathsf{X} : \mathsf{Y})$ of the pair (X, Y) is defined to be the supremum over all nonnegative real numbers $\alpha \geq 0$ for which the following statement holds: there exists a sequence of LOCC channels $\{\Psi_n : n \in \mathbb{N}\}$, where

$$\Psi_n \in \mathrm{LOCC}\big(\mathcal{X}^{\otimes n}, \mathcal{Z}^{\otimes m} : \mathcal{Y}^{\otimes n}, \mathcal{W}^{\otimes m}\big) \tag{6.215}$$

for $m = \lfloor \alpha n \rfloor$, such that

$$\lim_{n\to\infty} \mathrm{F}\Big(V_m \tau^{\otimes m} V_m^*, \Psi_n\big(U_n \rho^{\otimes n} U_n^*\big)\Big) = 1. \tag{6.216}$$

Definition 6.40. Let X and Y be registers with associated complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, and let $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ be a state of (X, Y). With respect to the state ρ, the *entanglement cost* $\mathrm{E}_{\mathrm{C}}(\mathsf{X} : \mathsf{Y})$ of the pair (X, Y) is defined to be the infimum over all nonnegative real numbers $\alpha \geq 0$ for which the following statement holds: there exists a sequence of LOCC channels $\{\Phi_n : n \in \mathbb{N}\}$, where

$$\Phi_n \in \mathrm{LOCC}\big(\mathcal{Z}^{\otimes m}, \mathcal{X}^{\otimes n} : \mathcal{W}^{\otimes m}, \mathcal{Y}^{\otimes n}\big) \tag{6.217}$$

for $m = \lfloor \alpha n \rfloor$, such that

$$\lim_{n\to\infty} \mathrm{F}\Big(U_n \rho^{\otimes n} U_n^*, \Phi_n\big(V_m \tau^{\otimes m} V_m^*\big)\Big) = 1. \tag{6.218}$$

It is intuitive that the entanglement cost should be at least as large as the distillable entanglement, for any choice of $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$, for otherwise one could repeatedly distill copies of the state τ from copies of a given state ρ, use them to produce more copies of ρ, and repeat this process indefinitely, eventually producing any desired number of copies of τ from a finite number of copies of ρ. Such an "entanglement factory" must surely not be possible through local operations and classical communication alone. The following proposition confirms this intuition.

Proposition 6.41. *Let* X *and* Y *be registers. With respect to every state of the pair* (X, Y) *it holds that* $\mathrm{E}_{\mathrm{D}}(\mathsf{X} : \mathsf{Y}) \leq \mathrm{E}_{\mathrm{C}}(\mathsf{X} : \mathsf{Y})$.

Proof. Suppose that n, m, and k are nonnegative integers with $k > m$, and

$$\begin{aligned} \Phi_n &\in \mathrm{LOCC}\big(\mathcal{Z}^{\otimes m}, \mathcal{X}^{\otimes n} : \mathcal{W}^{\otimes m}, \mathcal{Y}^{\otimes n}\big) \\ \Psi_n &\in \mathrm{LOCC}\big(\mathcal{X}^{\otimes n}, \mathcal{Z}^{\otimes k} : \mathcal{Y}^{\otimes n}, \mathcal{W}^{\otimes k}\big) \end{aligned} \tag{6.219}$$

are LOCC channels. It holds that

$$V_m \tau^{\otimes m} V_m^* \in \mathrm{Ent}_{2^m}\big(\mathcal{Z}^{\otimes m} : \mathcal{W}^{\otimes m}\big), \tag{6.220}$$

and therefore, given that the composition $\Psi_n \Phi_n$ is an LOCC (and therefore separable) channel, Theorem 6.25 implies that

$$(\Psi_n \Phi_n)\big(V_m \tau^{\otimes m} V_m^*\big) \in \mathrm{Ent}_{2^m}\big(\mathcal{Z}^{\otimes k} : \mathcal{W}^{\otimes k}\big). \tag{6.221}$$

It follows by Proposition 6.16 that

$$\begin{aligned} &\mathrm{F}\Big((\Psi_n \Phi_n)\big(V_m \tau^{\otimes m} V_m^*\big), V_k \tau^{\otimes k} V_k^*\Big)^2 \\ &\quad = \Big\langle (\Psi_n \Phi_n)\big(V_m \tau^{\otimes m} V_m^*\big), V_k \tau^{\otimes k} V_k^* \Big\rangle \leq 2^{m-k} \leq \frac{1}{2}. \end{aligned} \tag{6.222}$$

Now, let $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ be any state of the pair (X, Y), and suppose α and β are nonnegative real numbers satisfying the definitions of entanglement cost and distillable entanglement, respectively, for the state ρ. For all $\varepsilon > 0$, there must therefore exist a sufficiently large value of $n \in \mathbb{N}$ such that, for $m = \lfloor \alpha n \rfloor$ and $k = \lfloor \beta m \rfloor$, there exist LOCC channels of the form (6.219) for which the following bounds hold:

$$\begin{aligned} \mathrm{F}\Big(\Phi_n\big(V_m \tau^{\otimes m} V_m^*\big), U_n \rho^{\otimes n} U_n^*\Big) &> 1 - \varepsilon, \\ \mathrm{F}\Big(\Psi_n\big(U_n \rho^{\otimes n} V_n^*\big), V_k \tau^{\otimes k} V_k^*\Big) &> 1 - \varepsilon. \end{aligned} \tag{6.223}$$

Therefore, by Theorem 3.32, one may conclude that

$$\mathrm{F}\Big((\Psi_n \Phi_n)\big(V_m \tau^{\otimes m} V_m^*\big), V_k \tau^{\otimes k} V_k^*\Big) > 1 - 4\varepsilon. \tag{6.224}$$

Taking $\varepsilon < 1/8$, one finds from the bound above that $k \leq m$ (for sufficiently large n), and therefore $\alpha \geq \beta$, from it follows that $\mathrm{E}_{\mathrm{D}}(\mathsf{X} : \mathsf{Y}) \leq \mathrm{E}_{\mathrm{C}}(\mathsf{X} : \mathsf{Y})$. □

Pure state entanglement

The next theorem demonstrates that the entanglement cost and distillable entanglement are equal for bipartite pure states; in both cases, the value of these measures agrees with the von Neumann entropy of the states obtained by restricting the given pure state to either part of its bipartition.

Theorem 6.42. *Let* X *and* Y *be registers. With respect to every pure state of the pair* (X, Y)*, one has*

$$\mathrm{E_D}(\mathsf{X}:\mathsf{Y}) = \mathrm{H}(\mathsf{X}) = \mathrm{H}(\mathsf{Y}) = \mathrm{E_C}(\mathsf{X}:\mathsf{Y}). \tag{6.225}$$

Proof. Let $u \in \mathcal{X} \otimes \mathcal{Y}$ be a unit vector, and consider the pure state uu^* of the pair (X, Y). By means of the Schmidt decomposition, one may write

$$u = \sum_{a \in \Sigma} \sqrt{p(a)}\, x_a \otimes y_a \tag{6.226}$$

for some choice of an alphabet Σ, a probability vector $p \in \mathcal{P}(\Sigma)$, and two orthonormal collections $\{x_a : a \in \Sigma\} \subset \mathcal{X}$ and $\{y_a : a \in \Sigma\} \subset \mathcal{Y}$. It holds that

$$\mathrm{Tr}_{\mathcal{Y}}(uu^*) = \sum_{a \in \Sigma} p(a) x_a x_a^* \quad \text{and} \quad \mathrm{Tr}_{\mathcal{X}}(uu^*) = \sum_{a \in \Sigma} p(a) y_a y_a^*, \tag{6.227}$$

which implies that $\mathrm{H}(\mathsf{X}) = \mathrm{H}(p) = \mathrm{H}(\mathsf{Y})$.

Next, recall that, for every choice of $\varepsilon > 0$ and $n \in \mathbb{N}$, the set of ε-typical strings $T_{n,\varepsilon}$ with respect to p contains those strings $a_1 \cdots a_n \in \Sigma^n$ for which

$$2^{-n(\mathrm{H}(p)+\varepsilon)} < p(a_1) \cdots p(a_n) < 2^{-n(\mathrm{H}(p)-\varepsilon)}. \tag{6.228}$$

With this set in mind, one may define a vector $v_{n,\varepsilon} \in \mathcal{X}^{\otimes n} \otimes \mathcal{Y}^{\otimes n}$, for every choice of $\varepsilon > 0$ and $n \in \mathbb{N}$, as

$$v_{n,\varepsilon} = \sum_{a_1 \cdots a_n \in T_{n,\varepsilon}} \sqrt{p(a_1) \cdots p(a_n)}\, x_{a_1 \cdots a_n} \otimes y_{a_1 \cdots a_n}, \tag{6.229}$$

where the shorthand notations

$$x_{a_1 \cdots a_n} = x_{a_1} \otimes \cdots \otimes x_{a_n} \quad \text{and} \quad y_{a_1 \cdots a_n} = y_{a_1} \otimes \cdots \otimes y_{a_n} \tag{6.230}$$

have been used for the sake of brevity. Also define a normalized version of the vector $v_{n,\varepsilon}$ as

$$w_{n,\varepsilon} = \frac{v_{n,\varepsilon}}{\| v_{n,\varepsilon} \|}. \tag{6.231}$$

Observe that

$$2^{-n(\mathrm{H}(p)+\varepsilon)} < \lambda_k\Big(\mathrm{Tr}_{\mathcal{Y}^{\otimes n}}\big(v_{n,\varepsilon}v_{n,\varepsilon}^*\big)\Big) < 2^{-n(\mathrm{H}(p)-\varepsilon)}, \tag{6.232}$$

and therefore

$$\frac{2^{-n(\mathrm{H}(p)+\varepsilon)}}{\|v_{n,\varepsilon}\|^2} < \lambda_k\Big(\mathrm{Tr}_{\mathcal{Y}^{\otimes n}}\big(w_{n,\varepsilon}w_{n,\varepsilon}^*\big)\Big) < \frac{2^{-n(\mathrm{H}(p)-\varepsilon)}}{\|v_{n,\varepsilon}\|^2}, \tag{6.233}$$

for $k = 1, \ldots, |T_{n,\varepsilon}|$, while the remaining eigenvalues are zero in both cases.

Now, consider the entanglement cost of the pair (X, Y) with respect to the state uu^*. Let α be any real number such that $\alpha > \mathrm{H}(p)$, let $\varepsilon > 0$ be sufficiently small so that $\alpha > \mathrm{H}(p) + 2\varepsilon$, and consider any choice of $n > 1/\varepsilon$. Denoting $m = \lfloor \alpha n \rfloor$, it follows that $m \geq n(\mathrm{H}(p) + \varepsilon)$. Moreover, it holds that

$$\lambda_k\Big(\mathrm{Tr}_{\mathcal{W}^{\otimes m}}\big(V_m\tau^{\otimes m}V_m^*\big)\Big) = 2^{-m} \tag{6.234}$$

for $k = 1, \ldots, 2^m$. As

$$2^{-m} \leq 2^{-n(\mathrm{H}(p)+\varepsilon)} \leq \frac{2^{-n(\mathrm{H}(p)+\varepsilon)}}{\|v_{n,\varepsilon}\|^2}, \tag{6.235}$$

it follows that

$$\sum_{j=1}^{k} \lambda_j\Big(\mathrm{Tr}_{\mathcal{W}^{\otimes m}}\big(V_m\tau^{\otimes m}V_m^*\big)\Big) \leq \sum_{j=1}^{k} \lambda_j\Big(\mathrm{Tr}_{\mathcal{Y}^{\otimes n}}\big(w_{n,\varepsilon}w_{n,\varepsilon}^*\big)\Big) \tag{6.236}$$

for every $k \in \{1, \ldots, 2^m\}$. It follows by Corollary 6.38 to Nielsen's theorem (Theorem 6.37) that there exists an LOCC channel

$$\Phi_n \in \mathrm{LOCC}\big(\mathcal{Z}^{\otimes m}, \mathcal{X}^{\otimes n} : \mathcal{W}^{\otimes m}, \mathcal{Y}^{\otimes n}\big) \tag{6.237}$$

such that

$$\Phi_n(V_m\tau^{\otimes m}V_m^*) = w_{n,\varepsilon}w_{n,\varepsilon}^*. \tag{6.238}$$

As

$$\mathrm{F}\Big(U_n(uu^*)^{\otimes n}U_n^*, w_{n,\varepsilon}w_{n,\varepsilon}^*\Big)^2 = \sum_{a_1\cdots a_n \in T_{n,\varepsilon}} p(a_1)\cdots p(a_n), \tag{6.239}$$

which approaches 1 in the limit as n approaches infinity, it follows that $\mathrm{E_C}(\mathsf{X}:\mathsf{Y}) \leq \alpha$, and therefore $\mathrm{E_C}(\mathsf{X}:\mathsf{Y}) \leq \mathrm{H}(p)$.

Next, consider the distillable entanglement of (X,Y) with respect to the state uu^*. If $\mathrm{H}(p) = 0$, then there is nothing to prove, as the distillable entanglement is trivially nonnegative, so it will be assumed hereafter that $\mathrm{H}(p) > 0$. Let α be a real number such that $\alpha < \mathrm{H}(p)$, and let $\varepsilon \in (0,1)$ be sufficiently small so that $\alpha < \mathrm{H}(p) - 2\varepsilon$. For all but finitely many values of $n \in \mathbb{N}$, one has that $-\varepsilon n < \log(1-\varepsilon)$, from which it follows that

$$m = \lfloor \alpha n \rfloor \leq n(\mathrm{H}(p) - \varepsilon) + \log(1-\varepsilon), \tag{6.240}$$

and therefore

$$\frac{2^{-n(\mathrm{H}(p)-\varepsilon)}}{1-\varepsilon} \leq 2^{-m}. \tag{6.241}$$

As the quantity

$$\|v_{n,\varepsilon}\|^2 = \sum_{a_1\cdots a_n \in T_{n,\varepsilon}} p(a_1)\cdots p(a_n) \tag{6.242}$$

approaches 1 in the limit as n approaches infinity, it follows that

$$\frac{2^{-n(\mathrm{H}(p)-\varepsilon)}}{\|v_{n,\varepsilon}\|^2} \leq 2^{-m} \tag{6.243}$$

for all but finitely many choices of $n \in \mathbb{N}$.

Now, consider any choice of n for which (6.243) holds (where $m = \lfloor \alpha n \rfloor$ as usual). One therefore has

$$\sum_{j=1}^{k} \lambda_j\Big(\mathrm{Tr}_{\mathcal{Y}^{\otimes n}}\big(w_{n,\varepsilon} w_{n,\varepsilon}^*\big)\Big) \leq \sum_{j=1}^{k} \lambda_j\Big(\mathrm{Tr}_{\mathcal{W}^{\otimes m}}\big(V_m \tau^{\otimes m} V_m^*\big)\Big) \tag{6.244}$$

for every $k \in \{1,\ldots,2^m\}$. Again using Corollary 6.38, one has that there must exist an LOCC channel

$$\Phi_n \in \mathrm{LOCC}\big(\mathcal{X}^{\otimes n}, \mathcal{Z}^{\otimes m} : \mathcal{Y}^{\otimes n}, \mathcal{W}^{\otimes m}\big) \tag{6.245}$$

such that

$$\Phi_n\big(w_{n,\varepsilon} w_{n,\varepsilon}^*\big) = V_m \tau^{\otimes m} V_m^*. \tag{6.246}$$

Making use of the monotonicity of the fidelity function under the action of

a channel (Theorem 3.30), one finds that

$$\begin{aligned}
&\mathrm{F}\Big(\Phi_n\big(U_n(uu^*)^{\otimes n}U_n^*\big), V_m\tau^{\otimes m}V_m^*\Big)^2 \\
&\qquad = \mathrm{F}\Big(\Phi_n\big(U_n(uu^*)^{\otimes n}U_n^*\big), \Phi_n(w_{n,\varepsilon}w_{n,\varepsilon}^*)\Big)^2 \\
&\qquad \geq \mathrm{F}\Big(U_n(uu^*)^{\otimes n}U_n^*\,,\, w_{n,\varepsilon}w_{n,\varepsilon}^*\Big)^2 \\
&\qquad = \sum_{a_1\cdots a_n\in T_{n,\varepsilon}} p(a_1)\cdots p(a_n).
\end{aligned} \tag{6.247}$$

The quantity on the right-hand-side of this inequality approaches 1 in the limit as n approaches infinity, from which it follows that $\mathrm{E}_{\mathrm{D}}(\mathsf{X}:\mathsf{Y}) \geq \alpha$, and therefore $\mathrm{E}_{\mathrm{D}}(\mathsf{X}:\mathsf{Y}) \geq \mathrm{H}(p)$.

It has been proved that

$$\mathrm{E}_{\mathrm{C}}(\mathsf{X}:\mathsf{Y}) \leq \mathrm{H}(p) \leq \mathrm{E}_{\mathrm{D}}(\mathsf{X}:\mathsf{Y}). \tag{6.248}$$

The inequality $\mathrm{E}_{\mathrm{D}}(\mathsf{X}:\mathsf{Y}) \leq \mathrm{E}_{\mathrm{C}}(\mathsf{X}:\mathsf{Y})$ holds by Proposition 6.41, so the proof is complete. □

Remark 6.43. For a given unit vector $u \in \mathcal{X}\otimes\mathcal{Y}$, for complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, the quantity in (6.225) is known as the *entanglement entropy* of the pure state uu^*.

6.2.3 Bound entanglement and partial transposition

Informally speaking, Theorem 6.42 implies that all pure state entanglement is equivalent in the bipartite setting. A bipartite pure state is entangled if and only if it has positive entanglement entropy. Moreover, given any two entangled pure states, one necessarily has that an approximate conversion between a large number of copies of the first state to the second state is possible through the use of an LOCC channel, at a rate determined by the ratio of the entanglement entropies of the two states.

The situation is more complex for mixed states. One respect in which this is so is that there exist entangled states having no distillable entanglement. The entanglement in such states, which is referred to as *bound entanglement*, can never be converted into pure state entanglement through the use of an LOCC channel. The fact that states of this sort exist may be proved through the use of properties of the transpose mapping.

The partial transpose and separability

For any complex Euclidean space $\mathcal{X}$, the transpose mapping $\mathrm{T} \in \mathrm{T}(\mathcal{X})$ is defined as

$$\mathrm{T}(X) = X^{\mathsf{T}} \tag{6.249}$$

for all $X \in \mathrm{L}(\mathcal{X})$. As this is a positive map, it follows by the Horodecki criterion (Theorem 6.10) that

$$\big(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(R) \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y}) \tag{6.250}$$

for every separable operator $R \in \mathrm{Sep}(\mathcal{X} : \mathcal{Y})$. If $P \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$ is a positive semidefinite operator for which

$$\big(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(P) \notin \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y}), \tag{6.251}$$

then one may therefore conclude that P is not separable.

The converse of this statement does not hold in general. Given a positive semidefinite operator $P \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$ for which

$$\big(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(P) \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y}), \tag{6.252}$$

one may not conclude that P is separable; an example of a non-separable operator possessing the property (6.252) is described below.

It is the case, however, that an operator $P \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$ satisfying the condition (6.252) is highly constrained, in some sense, with respect to the way it is entangled. With this idea in mind, one defines the sets of PPT operators and PPT states (short for *positive partial transpose* operators and states) as follows.

Definition 6.44. For any choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, the set $\mathrm{PPT}(\mathcal{X} : \mathcal{Y})$ is defined as the set of all operators $P \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$ that satisfy

$$\big(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(P) \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y}). \tag{6.253}$$

Elements of the set $\mathrm{PPT}(\mathcal{X} : \mathcal{Y})$ are called *PPT operators*, while elements of the set $\mathrm{PPT}(\mathcal{X} : \mathcal{Y}) \cap \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ are called *PPT states*.

Unextendable product sets and non-separable PPT operators

One method by which non-separable PPT operators may be constructed involves the notion of an *unextendable product set*. For complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, an orthonormal collection of vectors of the form

$$\mathcal{A} = \{u_1 \otimes v_1, \ldots, u_m \otimes v_m\}, \tag{6.254}$$

for unit vectors $u_1, \ldots, u_m \in \mathcal{X}$ and $v_1, \ldots, v_m \in \mathcal{Y}$, is an *unextendable product set* if two properties hold:

1. $\mathcal{A}$ spans a proper subspace of $\mathcal{X} \otimes \mathcal{Y}$. (Equivalently, $m < \dim(\mathcal{X} \otimes \mathcal{Y})$.)
2. For every choice of vectors $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ satisfying $x \otimes y \perp \mathcal{A}$, it must hold that $x \otimes y = 0$.

Example 6.45. Define unit vectors $u_1, \ldots, u_5 \in \mathcal{X}$ and $v_1, \ldots, v_5 \in \mathcal{Y}$, for $\mathcal{X} = \mathbb{C}^{\mathbb{Z}_3}$ and $\mathcal{Y} = \mathbb{C}^{\mathbb{Z}_3}$, as follows:

$$\begin{aligned}
u_1 &= e_0\,, & v_1 &= \frac{1}{\sqrt{2}}(e_0 - e_1)\,,\\
u_2 &= e_2\,, & v_2 &= \frac{1}{\sqrt{2}}(e_1 - e_2)\,,\\
u_3 &= \frac{1}{\sqrt{2}}(e_0 - e_1)\,, & v_3 &= e_2\,,\\
u_4 &= \frac{1}{\sqrt{2}}(e_1 - e_2)\,, & v_4 &= e_0\,,\\
u_5 &= \frac{1}{\sqrt{3}}(e_0 + e_1 + e_2)\,, & v_5 &= \frac{1}{\sqrt{3}}(e_0 + e_1 + e_2).
\end{aligned} \tag{6.255}$$

It therefore holds that

$$\begin{aligned}
u_1 \otimes v_1 &= \frac{1}{\sqrt{2}}\, e_0 \otimes (e_0 - e_1),\\
u_2 \otimes v_2 &= \frac{1}{\sqrt{2}}\, e_2 \otimes (e_1 - e_2),\\
u_3 \otimes v_3 &= \frac{1}{\sqrt{2}}\, (e_0 - e_1) \otimes e_2,\\
u_4 \otimes v_4 &= \frac{1}{\sqrt{2}}\, (e_1 - e_2) \otimes e_0,\\
u_5 \otimes v_5 &= \frac{1}{3}\, (e_0 + e_1 + e_2) \otimes (e_0 + e_1 + e_2).
\end{aligned} \tag{6.256}$$

The set $\{u_1 \otimes v_1, \ldots, u_5 \otimes v_5\}$ is orthonormal by inspection. If $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ satisfy

$$\langle x \otimes y, u_k \otimes v_k \rangle = \langle x, u_k \rangle \langle y, v_k \rangle = 0 \tag{6.257}$$

for $k = 1, \ldots, 5$, then one must have $\langle x, u_k \rangle = 0$ for at least 3 distinct choices of $k \in \{1, \ldots, 5\}$ or $\langle y, v_k \rangle = 0$ for at least 3 distinct choices of $k \in \{1, \ldots, 5\}$. As every 3 distinct choices of u_k span all of $\mathcal{X}$ and every 3 distinct choices of v_k span all of $\mathcal{Y}$, it follows that $x \otimes y = 0$. The set $\{u_1 \otimes v_1, \ldots, u_5 \otimes v_5\}$ is therefore an unextendable product set.

The projection onto the subspace orthogonal to an unextendable product set must be both PPT and entangled, as the following theorem states.

Theorem 6.46. *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, let*

$$\{u_1 \otimes v_1, \ldots, u_m \otimes v_m\} \tag{6.258}$$

be an unextendable product set in $\mathcal{X} \otimes \mathcal{Y}$, and define

$$\Pi = \sum_{k=1}^{m} u_k u_k^* \otimes v_k v_k^*. \tag{6.259}$$

It holds that

$$\mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{Y}} - \Pi \in \mathrm{PPT}(\mathcal{X} : \mathcal{Y}) \setminus \mathrm{Sep}(\mathcal{X} : \mathcal{Y}). \tag{6.260}$$

Proof. From the assumption that $\{u_1 \otimes v_1, \ldots, u_m \otimes v_m\}$ is an orthonormal set, one may conclude that $\{\overline{u_1} \otimes v_1, \ldots, \overline{u_k} \otimes v_k\}$ is an orthonormal set as well. It follows that

$$\big(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(\Pi) = \sum_{k=1}^{m} \overline{u_k} u_k^{\mathsf{T}} \otimes v_k v_k^* \tag{6.261}$$

is a projection operator, and therefore

$$\big(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(\Pi) \leq \mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{Y}}. \tag{6.262}$$

As

$$\big(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(\mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{Y}}) = \mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{Y}}, \tag{6.263}$$

one obtains the inclusion

$$\big(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(\mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{Y}} - \Pi) \in \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y}). \tag{6.264}$$

Now, toward a contradiction, assume that

$$\mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{Y}} - \Pi \in \operatorname{Sep}(\mathcal{X} : \mathcal{Y}), \tag{6.265}$$

which implies that

$$\mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{Y}} - \Pi = \sum_{a \in \Sigma} x_a x_a^* \otimes y_a y_a^* \tag{6.266}$$

for some choice of an alphabet Σ and collections $\{x_a : a \in \Sigma\} \subset \mathcal{X}$ and $\{y_a : a \in \Sigma\} \subset \mathcal{Y}$. It holds that

$$\begin{aligned} &\sum_{k=1}^{m} \sum_{a \in \Sigma} |\langle x_a \otimes y_a, u_k \otimes v_k \rangle|^2 \\ &\quad = \sum_{k=1}^{m} (u_k \otimes v_k)^* (\mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{Y}} - \Pi)(u_k \otimes v_k) = 0, \end{aligned} \tag{6.267}$$

and therefore $\langle x_a \otimes y_a, u_k \otimes v_k \rangle = 0$ for every $a \in \Sigma$ and $k \in \{1, \ldots, m\}$. By the assumption that $\{u_1 \otimes v_1, \ldots, u_m \otimes v_m\}$ is an unextendable product set, it follows that $x_a \otimes y_a = 0$ for every $a \in \Sigma$, and therefore

$$\mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{Y}} - \Pi = 0. \tag{6.268}$$

This, however, is in contradiction with the assumption $m < \dim(\mathcal{X} \otimes \mathcal{Y})$. It follows that

$$\mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{Y}} - \Pi \notin \operatorname{Sep}(\mathcal{X} : \mathcal{Y}), \tag{6.269}$$

which completes the proof. □

PPT states have no distillable entanglement

PPT states may not always be separable, but they exhibit similar properties to separable states in some respects. One such respect is that their overlap with every maximally entangled state is small. The next proposition, which is reminiscent of Proposition 6.16, is representative of this fact.

Proposition 6.47. *Let $Y \in \mathrm{L}(\mathcal{Y}, \mathcal{X})$ be an operator satisfying $\|Y\| \leq 1$, for $\mathcal{X}$ and $\mathcal{Y}$ being complex Euclidean spaces. For every operator $P \in \mathrm{PPT}(\mathcal{X} : \mathcal{Y})$ it holds that*

$$\big\langle \operatorname{vec}(Y) \operatorname{vec}(Y)^*, P \big\rangle \leq \operatorname{Tr}(P). \tag{6.270}$$

Proof. Observe that one may write

$$\operatorname{vec}(Y) = \left(\mathbb{1}_{\mathcal{X}} \otimes Y^{\mathsf{T}}\right) \operatorname{vec}(\mathbb{1}_{\mathcal{X}}), \tag{6.271}$$

which implies that

$$(\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(\operatorname{vec}(Y) \operatorname{vec}(Y)^*) = \left(\mathbb{1}_{\mathcal{X}} \otimes Y^{\mathsf{T}}\right) W \left(\mathbb{1}_{\mathcal{X}} \otimes Y^{\mathsf{T}}\right)^* \tag{6.272}$$

for $W \in \mathrm{U}(\mathcal{X} \otimes \mathcal{X})$ denoting the swap operator on $\mathcal{X} \otimes \mathcal{X}$. It holds that

$$\begin{aligned} &\big\langle \operatorname{vec}(Y) \operatorname{vec}(Y)^*, P \big\rangle \\ &\quad = \big\langle (\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(\operatorname{vec}(Y) \operatorname{vec}(Y)^*), (\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(P) \big\rangle \\ &\quad \leq \big\| (\mathrm{T} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Y})})(P) \big\|_1 \\ &\quad = \operatorname{Tr}(P); \end{aligned} \tag{6.273}$$

the first equality follows from the fact that the transpose mapping is its own adjoint and inverse, the inequality follows from the fact that the operator (6.272) has spectral norm at most 1, and the second equality follows from the assumption that $P \in \operatorname{PPT}(\mathcal{X} : \mathcal{Y})$ along with the observation that the transpose mapping preserves trace. □

Example 6.48. Similar to Example 6.17, let Σ be an alphabet, let $n = |\Sigma|$, and let $\mathcal{X} = \mathbb{C}^{\Sigma}$ and $\mathcal{Y} = \mathbb{C}^{\Sigma}$. Define a density operator $\tau \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ as

$$\tau = \frac{1}{n} \sum_{a,b \in \Sigma} E_{a,b} \otimes E_{a,b} = \frac{1}{n} \operatorname{vec}(\mathbb{1}) \operatorname{vec}(\mathbb{1})^*, \tag{6.274}$$

where $\mathbb{1}$ denotes the identity operator on $\mathbb{C}^{\Sigma}$, which may be viewed as an element of the set $\mathrm{L}(\mathcal{Y}, \mathcal{X})$. For every PPT density operator

$$\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y}) \cap \operatorname{PPT}(\mathcal{X} : \mathcal{Y}), \tag{6.275}$$

it holds that

$$\langle \tau, \rho \rangle = \frac{1}{n} \big\langle \operatorname{vec}(\mathbb{1}) \operatorname{vec}(\mathbb{1})^*, \rho \big\rangle \leq \frac{1}{n} \tag{6.276}$$

by Proposition 6.47. Thus, with respect to their overlap with the maximally entangled state τ, one has that PPT operators are bounded in a similar way to separable operators.

Proposition 6.47, when combined with the following proposition stating that separable maps (and therefore LOCC channels) map PPT operators to PPT operators, leads to a proof that PPT states have distillable entanglement equal to zero.

Proposition 6.49. *Let $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, and $\mathcal{W}$ be complex Euclidean spaces, and suppose that $P \in \mathrm{PPT}(\mathcal{X}:\mathcal{Y})$ is a PPT operator and $\Phi \in \mathrm{SepCP}(\mathcal{X},\mathcal{Z}:\mathcal{Y},\mathcal{W})$ is a separable map. It holds that $\Phi(P) \in \mathrm{PPT}(\mathcal{Z}:\mathcal{W})$.*

Proof. For any choice of operators $A \in \mathrm{L}(\mathcal{X},\mathcal{Z})$ and $B \in \mathrm{L}(\mathcal{Y},\mathcal{W})$, it holds that

$$\begin{aligned}&\big(\mathrm{T}\otimes\mathbb{1}_{\mathrm{L}(\mathcal{W})}\big)\big((A\otimes B)P(A\otimes B)^*\big)\\ &\quad=\big(\overline{A}\otimes B\big)\big(\mathrm{T}\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(P)\big(\overline{A}\otimes B\big)^*\in\mathrm{Pos}\big(\mathcal{Z}\otimes\mathcal{W}\big),\end{aligned}\tag{6.277}$$

by virtue of the fact that $P \in \mathrm{PPT}(\mathcal{X}:\mathcal{Y})$. As Φ is separable, one has

$$\Phi(X) = \sum_{a\in\Sigma}(A_a\otimes B_a)X(A_a\otimes B_a)^*\tag{6.278}$$

for all $X \in \mathrm{L}(\mathcal{X}\otimes\mathcal{Y})$, for some choice of an alphabet Σ and collections of operators

$$\{A_a : a\in\Sigma\}\subset\mathrm{L}(\mathcal{X},\mathcal{Z}) \;\text{ and }\; \{B_a : a\in\Sigma\}\subset\mathrm{L}(\mathcal{Y},\mathcal{W}).\tag{6.279}$$

Consequently, one has that

$$\big(\mathrm{T}\otimes\mathbb{1}_{\mathrm{L}(\mathcal{W})}\big)(\Phi(P)) = \sum_{a\in\Sigma}\big(\overline{A_a}\otimes B_a\big)\big(\mathrm{T}\otimes\mathbb{1}_{\mathrm{L}(\mathcal{Y})}\big)(P)\big(\overline{A_a}\otimes B_a\big)^*\tag{6.280}$$

is a positive semidefinite operator, and therefore $\Phi(P) \in \mathrm{PPT}\big(\mathcal{Z}:\mathcal{W}\big)$, as required. □

Theorem 6.50. *Let X and Y be registers and consider a PPT state*

$$\rho\in\mathrm{PPT}(\mathcal{X}:\mathcal{Y})\cap\mathrm{D}(\mathcal{X}\otimes\mathcal{Y})\tag{6.281}$$

of the pair (X,Y). With respect to the state ρ, it holds that $\mathrm{E}_{\mathrm{D}}(\mathsf{X}:\mathsf{Y}) = 0$.

Proof. Let $\Gamma = \{0,1\}$, let $\mathcal{Z} = \mathbb{C}^{\Gamma}$ and $\mathcal{W} = \mathbb{C}^{\Gamma}$, and let $\tau \in \mathrm{D}(\mathcal{Z}\otimes\mathcal{W})$ be defined as

$$\tau = \frac{1}{2}\sum_{a,b\in\Gamma}E_{a,b}\otimes E_{a,b}.\tag{6.282}$$

Let n and m be arbitrary positive integers, consider any LOCC channel

$$\Phi\in\mathrm{LOCC}\big(\mathcal{X}^{\otimes n},\mathcal{Z}^{\otimes m}:\mathcal{Y}^{\otimes n},\mathcal{W}^{\otimes m}\big),\tag{6.283}$$

and recall the operators U_n and V_m as defined by (6.211) and (6.213). It holds that

$$U_n \rho^{\otimes n} U_n^* \in \mathrm{PPT}\big(\mathcal{X}^{\otimes n} : \mathcal{Y}^{\otimes n}\big), \tag{6.284}$$

and therefore

$$\Phi(V_m \rho^{\otimes n} V_m^*) \in \mathrm{PPT}\big(\mathcal{Z}^{\otimes m} : \mathcal{W}^{\otimes m}\big) \tag{6.285}$$

by Proposition 6.49. One may therefore conclude from Proposition 6.47 that

$$\mathrm{F}\big(V_m \tau^{\otimes m} V_m^*, \Phi(U_n \rho^{\otimes n} U_n)\big) \leq 2^{-\frac{m}{2}} \leq \frac{1}{\sqrt{2}}. \tag{6.286}$$

It follows that $\mathrm{E}_{\mathrm{D}}(\mathsf{X} : \mathsf{Y}) = 0$. □

6.3 Phenomena associated with entanglement

This section discusses two notions—teleportation (together with the related notion of dense coding) and non-classical correlations—that are generally associated with entanglement, and serve as representatives of the sorts of operational effects that entanglement may induce.

6.3.1 Teleportation and dense coding

In the setting of quantum information, *teleportation* has traditionally referred to a protocol by which a single-qubit quantum channel is implemented through the use of a maximally entangled pair of qubits combined with two classical bits of communication. Informally speaking, teleportation suggests the transformation

1 pair of maximally entangled qubits
+ 2 bits of classical communication
$\rightarrow$ 1 qubit of quantum communication.

Teleportation is often associated with the *dense coding* protocol, which offers a complementary trade-off between resources; dense coding traditionally refers to a protocol by which a two-bit classical channel is implemented through the use of a maximally entangled pair of qubits and a single-qubit quantum channel. In this case, the suggested transformation is

1 pair of maximally entangled qubits
+ 1 qubit of quantum communication
$\rightarrow$ 2 bits of classical communication.

In both cases, the maximally entangled pair of qubits is consumed by the conversion between two classical bits and one qubit of communication; in essence, the entangled pair of qubits functions as a resource allowing for this conversion.

In the discussion that follows, teleportation and dense coding will be considered in greater generality. The traditional protocols suggested above will emerge as specific instances of more general classes of protocols.

Teleportation

Consider the following scenario in which two individuals, Alice and Bob, aim to implement an ideal quantum channel through the combined use of entanglement and classical communication.

Scenario 6.51 (Teleportation)**.** Alice holds a register X and Bob holds Y. Both registers have the same classical state set Σ, and the state of the pair (X, Y) is given by the maximally entangled state

$$\tau = \frac{1}{|\Sigma|} \sum_{b,c \in \Sigma} E_{b,c} \otimes E_{b,c}. \tag{6.287}$$

Alice obtains a new register Z, whose classical state set is also Σ, and she wishes to transmit Z to Bob. Alice and Bob attempt to accomplish this task using classical communication together with the shared entangled state τ, by means of a protocol as follows:

1. Alice performs a measurement $\mu : \Gamma \rightarrow \mathrm{Pos}(\mathcal{Z} \otimes \mathcal{X})$ on the pair (Z, X), where Γ is an arbitrarily chosen alphabet, and sends the outcome $a \in \Gamma$ of this measurement to Bob.
2. For $\{\Psi_a : a \in \Gamma\} \subseteq \mathrm{C}(\mathcal{Y}, \mathcal{Z})$ being a collection of channels indexed by Γ, Bob applies the channel Ψ_a to Y, for whichever symbol $a \in \Gamma$ was sent to him by Alice, transforming this register into a new register Z.

An analysis reveals that this protocol accomplishes the task at hand for a suitable choice for Alice's measurement and Bob's collection of channels.

Remark 6.52. One may consider more general scenarios along similar lines to Scenario 6.51. For instance, X, Y, and Z might not share the same classical state set, the initial state of the pair (X, Y) might be initialized to a different state than τ, and Alice and Bob might aim to implement a channel different from the identity channel. The discussion that follows, however, will focus on the setting described in Scenario 6.51 in the interest of simplicity.

For a given choice of Alice's measurement μ and Bob's collection of channels $\{\Psi_a : a \in \Gamma\}$, the channel $\Phi \in \mathrm{C}(\mathcal{Z})$ that is implemented by the protocol described in Scenario 6.51 may be expressed as

$$\Phi(Z) = \frac{1}{|\Sigma|} \sum_{a \in \Gamma} \sum_{b,c \in \Sigma} \big\langle \mu(a), Z \otimes E_{b,c} \big\rangle \Psi_a(E_{b,c}) \tag{6.288}$$

for all $Z \in \mathrm{L}(\mathcal{Z})$. The following theorem provides a characterization of those measurements and collections of channels for which the channel Φ is equal to the identity channel, which represents an ideal transmission of quantum information from Alice to Bob. (The statement of the theorem includes the assumption that none of the measurement operators of μ are identically zero, as this allows for a cleaner statement of the characterization.)

Theorem 6.53. *Let Σ and Γ be alphabets, let $\mathcal{X} = \mathbb{C}^{\Sigma}$, $\mathcal{Y} = \mathbb{C}^{\Sigma}$, and $\mathcal{Z} = \mathbb{C}^{\Sigma}$ be complex Euclidean spaces, let $\mu : \Gamma \to \mathrm{Pos}(\mathcal{Z} \otimes \mathcal{X})$ be a measurement such that $\mu(a) \neq 0$ for every $a \in \Gamma$, and let $\{\Psi_a : a \in \Gamma\} \subseteq \mathrm{C}(\mathcal{Y}, \mathcal{Z})$ be a collection of channels. The following two statements are equivalent:*

1. *It holds that*
$$Z = \frac{1}{|\Sigma|} \sum_{a \in \Gamma} \sum_{b,c \in \Sigma} \big\langle \mu(a), Z \otimes E_{b,c} \big\rangle \Psi_a(E_{b,c}) \tag{6.289}$$
for every $Z \in \mathrm{L}(\mathcal{Z})$.
2. *There exists a collection $\{U_a : a \in \Gamma\} \subset \mathrm{U}(\mathbb{C}^{\Sigma})$ of unitary operators and a probability vector $p \in \mathcal{P}(\Gamma)$ such that*
$$\mu(a) = p(a)|\Sigma| \operatorname{vec}(U_a) \operatorname{vec}(U_a)^* \quad \text{and} \quad \Psi_a(Y) = U_a Y U_a^* \tag{6.290}$$
for every choice of $a \in \Gamma$ and $Y \in \mathrm{L}(\mathcal{Y})$.

The proof of Theorem 6.53 will make use of the following proposition, which establishes that a channel of the form $\Phi \in \mathrm{C}(\mathcal{X})$, for any complex Euclidean space $\mathcal{X}$, has a completely positive inverse only if Φ is a unitary channel.

Proposition 6.54. *Let $\mathcal{X}$ be a complex Euclidean space, let $\Phi \in \mathrm{C}(\mathcal{X})$ be a channel, and let $\Psi \in \mathrm{CP}(\mathcal{X})$ be a completely positive map for which $\Phi\Psi = \mathbb{1}_{\mathrm{L}(\mathcal{X})}$. There exists a unitary operator $U \in \mathrm{U}(\mathcal{X})$ such that*

$$\Phi(X) = U^* X U \quad \text{and} \quad \Psi(X) = U X U^* \tag{6.291}$$

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

Proof. As Ψ is completely positive, and evidently nonzero, its Choi operator $J(\Psi)$ is a nonzero positive semidefinite operator. By the spectral theorem (Corollary 1.4), it is therefore possible to write

$$J(\Psi) = \sum_{k=1}^{r} \operatorname{vec}(A_k) \operatorname{vec}(A_k)^* \tag{6.292}$$

for $r = \operatorname{rank}(J(\Psi))$ and $\{A_1, \ldots, A_r\} \subset \mathrm{L}(\mathcal{X})$ being an orthogonal collection of nonzero operators. Consequently, one has

$$\begin{aligned} &\sum_{k=1}^{r} \bigl(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\bigr)\bigl(\operatorname{vec}(A_k) \operatorname{vec}(A_k)^*\bigr) \\ &\quad = \bigl(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\bigr)(J(\Psi)) = J(\Phi\Psi) = \operatorname{vec}(\mathbb{1}_{\mathcal{X}}) \operatorname{vec}(\mathbb{1}_{\mathcal{X}})^*. \end{aligned} \tag{6.293}$$

As $\operatorname{vec}(\mathbb{1}_{\mathcal{X}}) \operatorname{vec}(\mathbb{1}_{\mathcal{X}})^*$ has rank equal to one, and each operator

$$\bigl(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\bigr)\bigl(\operatorname{vec}(A_k) \operatorname{vec}(A_k)^*\bigr) \tag{6.294}$$

is positive semidefinite (by the complete positivity of Φ), it follows that there must exist a probability vector $(p_1, \ldots, p_r)$ such that

$$\bigl(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\bigr)\bigl(\operatorname{vec}(A_k) \operatorname{vec}(A_k)^*\bigr) = p_k \operatorname{vec}(\mathbb{1}_{\mathcal{X}}) \operatorname{vec}(\mathbb{1}_{\mathcal{X}})^* \tag{6.295}$$

for each $k \in \{1, \ldots, r\}$. Because Φ preserves trace, it follows that

$$\bigl(A_k^* A_k\bigr)^{\mathsf{T}} = \bigl(\mathrm{Tr} \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\bigr)\bigl(\operatorname{vec}(A_k) \operatorname{vec}(A_k)^*\bigr) = p_k \mathbb{1}_{\mathcal{X}}, \tag{6.296}$$

and therefore $A_k = \sqrt{p_k} U_k$ for some choice of a unitary operator $U_k \in \mathrm{U}(\mathcal{X})$, for each $k \in \{1, \ldots, r\}$. This implies that

$$\begin{aligned} \bigl(\mathbb{1}_{\mathcal{X}} \otimes U_k^{\mathsf{T}}\bigr) J(\Phi) \bigl(\mathbb{1}_{\mathcal{X}} \otimes U_k^{\mathsf{T}}\bigr)^* &= \bigl(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})}\bigr)\bigl(\operatorname{vec}(U_k) \operatorname{vec}(U_k)^*\bigr) \\ &= \operatorname{vec}(\mathbb{1}_{\mathcal{X}}) \operatorname{vec}(\mathbb{1}_{\mathcal{X}})^*, \end{aligned} \tag{6.297}$$

and therefore

$$J(\Phi) = \operatorname{vec}(U_k^*) \operatorname{vec}(U_k^*)^*, \tag{6.298}$$

again for each $k \in \{1, \ldots, r\}$. As $\{A_1, \ldots, A_r\}$ is a collection of nonzero, orthogonal operators, and is therefore linearly independent, one concludes that $r = 1$ and $p_1 = 1$; and by setting $U = U_1$ the proposition is proved. □

Proof of Theorem 6.53. Assume first that statement 1 holds. For each $a \in \Gamma$, define a map $\Xi_a \in \mathrm{T}(\mathbb{C}^\Sigma)$ as

$$\Xi_a(Z) = \frac{1}{|\Sigma|} \sum_{b,c \in \Sigma} \langle \mu(a), Z \otimes E_{b,c} \rangle \, E_{b,c} \tag{6.299}$$

for all $Z \in \mathrm{L}(\mathbb{C}^\Sigma)$. The Choi operator of Ξ_a is given by

$$\mathrm{J}(\Xi_a) = \frac{1}{|\Sigma|} W \overline{\mu(a)} W, \tag{6.300}$$

for $W \in \mathrm{U}(\mathbb{C}^\Sigma \otimes \mathbb{C}^\Sigma)$ denoting the swap operator. As $J(\Xi_a)$ is positive semidefinite for each $a \in \Gamma$, it follows that Ξ_a is completely positive, and moreover is nonzero by the assumption that $\mu(a)$ is nonzero. Statement 1 may now be expressed as

$$\sum_{a \in \Gamma} \Psi_a \Xi_a = \mathbb{1}_{\mathrm{L}(\mathcal{Z})}, \tag{6.301}$$

which is equivalent to

$$\sum_{a \in \Gamma} \mathrm{J}(\Psi_a \Xi_a) = \operatorname{vec}(\mathbb{1}_{\mathcal{Z}}) \operatorname{vec}(\mathbb{1}_{\mathcal{Z}})^*. \tag{6.302}$$

As the composition $\Psi_a \Xi_a$ is necessarily completely positive and nonzero for each $a \in \Gamma$, and the operator $\operatorname{vec}(\mathbb{1}_{\mathcal{Z}}) \operatorname{vec}(\mathbb{1}_{\mathcal{Z}})^*$ has rank equal to 1, it follows that there must exist a probability vector $p \in \mathcal{P}(\Gamma)$ such that

$$\mathrm{J}(\Psi_a \Xi_a) = p(a) \operatorname{vec}(\mathbb{1}_{\mathcal{Z}}) \operatorname{vec}(\mathbb{1}_{\mathcal{Z}})^* \tag{6.303}$$

for each $a \in \Gamma$. Consequently,

$$\frac{(\Psi_a \Xi_a)(Z)}{p(a)} = Z \tag{6.304}$$

for every $Y \in \mathrm{L}(\mathcal{Z})$. By Proposition 6.54, there must exist a collection of unitary operators $\{U_a : a \in \Sigma\} \subset \mathrm{U}(\mathbb{C}^\Sigma)$ such that

$$\Psi_a(Z) = U_a Z U_a^* \quad \text{and} \quad \frac{1}{p(a)}\Xi_a(Z) = U_a^* Z U_a \tag{6.305}$$

for every $a \in \Gamma$ and $Z \in \mathrm{L}(\mathbb{C}^\Sigma)$. Thus,

$$\frac{1}{|\Sigma|} W \overline{\mu(a)} W = \mathrm{J}(\Xi_a) = p(a)\,\mathrm{vec}(U_a^*)\,\mathrm{vec}(U_a^*)^*, \tag{6.306}$$

and because $W\,\mathrm{vec}(A) = \mathrm{vec}(A^\mathsf{T})$ for every $A \in \mathrm{L}(\mathbb{C}^\Sigma)$, one therefore has

$$\mu(a) = p(a)\,|\Sigma|\,\mathrm{vec}(U_a)\,\mathrm{vec}(U_a)^* \tag{6.307}$$

for each $a \in \Gamma$. Statement 1 therefore implies statement 2.

Now assume statement 2 holds. As μ is assumed to be a measurement, it must be the case that

$$|\Sigma| \sum_{a\in\Gamma} p(a)\,\mathrm{vec}(U_a)\,\mathrm{vec}(U_a)^* = \mathbb{1}_\Sigma \otimes \mathbb{1}_\Sigma, \tag{6.308}$$

and therefore

$$\sum_{a\in\Gamma} p(a)\,\mathrm{vec}(U_a)\,\mathrm{vec}(U_a)^* = \frac{1}{|\Sigma|}\mathbb{1}_\Sigma \otimes \mathbb{1}_\Sigma. \tag{6.309}$$

The operator represented by the equation (6.309) coincides with the Choi operator $\mathrm{J}(\Omega)$ of the completely depolarizing channel $\Omega \in \mathrm{C}(\mathbb{C}^\Sigma)$. It follows that one may write

$$\Omega(X) = \sum_{a\in\Gamma} p(a) U_a X U_a^* \tag{6.310}$$

for every $X \in \mathrm{L}(\mathbb{C}^\Sigma)$. As the natural representation $\mathrm{K}(\Omega)$ of the completely depolarizing channel is equal to the operator τ, one has that

$$\sum_{a\in\Gamma} p(a)\, U_a \otimes \overline{U_a} = \mathrm{K}(\Omega) = \tau \tag{6.311}$$

by Proposition 2.20, and because τ is invariant under taking the entry-wise complex conjugate it follows that

$$\sum_{a\in\Gamma} p(a)\, \overline{U_a} \otimes U_a = \tau. \tag{6.312}$$

Now consider the channel $\Phi \in \mathrm{C}(\mathbb{C}^{\Sigma})$ defined by

$$\Phi(Z) = \frac{1}{|\Sigma|} \sum_{a \in \Gamma} \sum_{b,c \in \Sigma} \langle \mu(a), Z \otimes E_{b,c} \rangle \, \Psi_a(E_{b,c}) \tag{6.313}$$

for every $Z \in \mathrm{L}(\mathcal{Z})$. Making use of the expression (6.312), one may write

$$\Phi(Z) = \sum_{a,b \in \Gamma} p(b) \big\langle \mu(a), Z \otimes \overline{U_b} \big\rangle \, \Psi_a(U_b) \tag{6.314}$$

for every $Z \in \mathrm{L}(\mathcal{Z})$. By substituting according to (6.290), one obtains

$$\begin{aligned} \Phi(Z) &= |\Sigma| \sum_{a,b \in \Gamma} p(a) p(b) \operatorname{vec}(U_a)^* \big(Z \otimes \overline{U_b}\big) \operatorname{vec}(U_a) \, U_a U_b U_a^* \\ &= |\Sigma| \sum_{a,b \in \Gamma} p(a) p(b) \big\langle U_a U_b U_a^*, Z \big\rangle \, U_a U_b U_a^*. \end{aligned} \tag{6.315}$$

The natural representation $K(\Phi)$ of the channel Φ is therefore given by

$$\begin{aligned} & |\Sigma| \sum_{a,b \in \Gamma} p(a) p(b) \operatorname{vec}(U_a U_b U_a^*) \operatorname{vec}(U_a U_b U_a^*)^* \\ &= \sum_{a \in \Gamma} p(a) \big(U_a \otimes \overline{U_a}\big) \left(|\Sigma| \sum_{b \in \Gamma} p(b) \operatorname{vec}(U_b) \operatorname{vec}(U_b)^* \right) \big(U_a \otimes \overline{U_a}\big)^* \\ &= \mathbb{1}_{\Sigma} \otimes \mathbb{1}_{\Sigma}, \end{aligned} \tag{6.316}$$

where the last equality has made use of (6.309). It follows that Φ is equal to the identity channel, and therefore statement 2 implies statement 1. □

Theorem 6.53 implies that every mixed-unitary representation of the completely depolarizing channel gives rise to a teleportation protocol, as the following corollary makes precise.

Corollary 6.55. *Let Σ and Γ be alphabets, let $\{U_a : a \in \Gamma\} \subset \mathrm{U}(\mathbb{C}^{\Sigma})$ be a collection of unitary operators, let $p \in \mathcal{P}(\Gamma)$ be a probability vector, and assume that*

$$\Omega(X) = \sum_{a \in \Gamma} p(a) U_a X U_a^* \tag{6.317}$$

for every $X \in \mathrm{L}(\mathbb{C}^{\Sigma})$, where $\Omega \in \mathrm{C}(\mathbb{C}^{\Sigma})$ denotes the completely depolarizing channel with respect to the space $\mathbb{C}^{\Sigma}$. For $\mu : \Gamma \to \mathrm{Pos}(\mathbb{C}^{\Sigma} \otimes \mathbb{C}^{\Sigma})$ defined as

$$\mu(a) = p(a) |\Sigma| \operatorname{vec}(U_a) \operatorname{vec}(U_a)^* \tag{6.318}$$

for each $a \in \Gamma$, one has that μ is a measurement, and moreover

$$X = \frac{1}{|\Sigma|} \sum_{a \in \Gamma} \sum_{b,c \in \Sigma} \langle \mu(a), X \otimes E_{b,c} \rangle \, U_a E_{b,c} U_a^* \tag{6.319}$$

for all $X \in \mathrm{L}(\mathbb{C}^\Sigma)$.

Proof. There is no loss of generality in assuming that $p(a) \neq 0$ for every $a \in \Gamma$, for otherwise one could define an alphabet $\Gamma_0 = \{a \in \Gamma : p(a) \neq 0\}$, verify that the corollary holds in this case, and observe that the statement of the corollary is equivalent when Γ is replaced by Γ_0 in this way.

It is evident that μ is a measurement, as each $\mu(a)$ is positive semidefinite and it holds that

$$\sum_{a \in \Gamma} \mu(a) = \sum_{a \in \Gamma} p(a) |\Sigma| \operatorname{vec}(U_a) \operatorname{vec}(U_a)^* = |\Sigma| J(\Omega) = \mathbb{1}_\Sigma \otimes \mathbb{1}_\Sigma. \tag{6.320}$$

By defining $\Psi_a(X) = U_a X U_a^*$ for every $X \in \mathrm{L}(\mathbb{C}^\Sigma)$ and $a \in \Gamma$, one has that statement 2 of Theorem 6.53 is satisfied. This implies that statement 1 of that theorem holds, which is equivalent to (6.319), and therefore completes the proof. □

Example 6.56. Let $\Sigma = \{0, 1\}$ denote the binary alphabet and let $\Gamma = \Sigma \times \Sigma$. Elements of Γ will be viewed as binary strings of length 2 for convenience. Define $p \in \mathcal{P}(\Gamma)$ as $p(00) = p(01) = p(10) = p(11) = 1/4$ and define unitary operators $U_{00}, U_{01}, U_{10}, U_{11} \in \mathrm{U}(\mathbb{C}^\Sigma)$ as follows:

$$\begin{aligned} U_{00} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad U_{01} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \\ U_{10} &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad U_{11} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \end{aligned} \tag{6.321}$$

The operators $U_{00}, U_{01}, U_{10}, U_{11}$ coincide with the discrete Weyl operators acting on the space $\mathbb{C}^\Sigma$, and (as explained in Section 4.1.2) provide a mixed-unitary realization of the completely depolarizing channel $\Omega \in \mathrm{C}(\mathbb{C}^\Sigma)$:

$$\frac{1}{4} \sum_{a,b \in \Sigma} U_{ab} X U_{ab}^* = \frac{\operatorname{Tr}(X)}{2} \mathbb{1} \tag{6.322}$$

for every $X \in \mathrm{L}(\mathbb{C}^{\Sigma})$. Consequently, by taking $\mu : \Gamma \to \mathrm{Pos}(\mathbb{C}^{\Sigma} \otimes \mathbb{C}^{\Sigma})$ to be the measurement defined as

$$\begin{aligned}
\mu(00) &= \frac{\mathrm{vec}(U_{00})\,\mathrm{vec}(U_{00})^*}{2} = u_{00}u_{00}^*, \\
\mu(01) &= \frac{\mathrm{vec}(U_{01})\,\mathrm{vec}(U_{01})^*}{2} = u_{01}u_{01}^*, \\
\mu(10) &= \frac{\mathrm{vec}(U_{10})\,\mathrm{vec}(U_{10})^*}{2} = u_{10}u_{10}^*, \\
\mu(11) &= \frac{\mathrm{vec}(U_{11})\,\mathrm{vec}(U_{11})^*}{2} = u_{11}u_{11}^*,
\end{aligned} \tag{6.323}$$

for

$$\begin{aligned}
u_{00} &= \frac{e_{00} + e_{11}}{\sqrt{2}}, \qquad & u_{01} &= \frac{e_{00} - e_{11}}{\sqrt{2}}, \\
u_{10} &= \frac{e_{01} + e_{10}}{\sqrt{2}}, \qquad & u_{11} &= \frac{e_{01} - e_{10}}{\sqrt{2}},
\end{aligned} \tag{6.324}$$

and setting

$$\Psi_{ab}(X) = U_{ab}XU_{ab}^* \tag{6.325}$$

for each $X \in \mathrm{L}(\mathbb{C}^{\Sigma})$ and $a, b \in \Sigma$, one obtains a teleportation protocol as described in Scenario 6.51. Indeed, the resulting protocol is equivalent to the traditional notion of teleportation in which an ideal single-qubit channel is implemented using a maximally entangled pair of qubits along with two classical bits of communication.

Example 6.57. The previous example may be generalized in the following way. Let $\Sigma = \mathbb{Z}_n$ for any positive integer n, let $\Gamma = \Sigma \times \Sigma$, and let the collection

$$\{U_{ab} \,:\, a, b \in \Sigma\} \subset \mathrm{U}(\mathbb{C}^{\Sigma}) \tag{6.326}$$

of unitary operators be in correspondence with the discrete Weyl operators acting on $\mathbb{C}^{\Sigma}$. By taking $\mu : \Gamma \to \mathrm{Pos}(\mathbb{C}^{\Sigma} \otimes \mathbb{C}^{\Sigma})$ to be the measurement defined as

$$\mu(ab) = \frac{\mathrm{vec}(U_{ab})\,\mathrm{vec}(U_{ab})^*}{n} \tag{6.327}$$

for each $a, b \in \Sigma$, and setting

$$\Psi_{ab}(X) = U_{ab}XU_{ab}^* \tag{6.328}$$

for each $X \in \mathrm{L}(\mathbb{C}^{\Sigma})$, one again obtains a teleportation protocol as described in Scenario 6.51.

In the teleportation protocols described in the previous two examples, the number of distinct classical symbols that must be tramsmitted is equal to the square of the number of classical states in the quantum system that is teleported. This is optimal, as the following corollary states.

Corollary 6.58. *Suppose that Σ and Γ are alphabets, $\mu : \Gamma \to \mathrm{Pos}(\mathbb{C}^{\Sigma} \otimes \mathbb{C}^{\Sigma})$ is a measurement, and $\{\Psi_a : a \in \Gamma\} \subseteq \mathrm{C}(\mathbb{C}^{\Sigma})$ is a collection of channels such that*

$$X = \frac{1}{|\Sigma|} \sum_{a \in \Gamma} \sum_{b,c \in \Sigma} \langle \mu(a), X \otimes E_{b,c} \rangle \Psi_a(E_{b,c}) \tag{6.329}$$

for every $X \in \mathrm{L}(\mathbb{C}^{\Sigma})$. It holds that $|\Gamma| \geq |\Sigma|^2$.

Proof. By Theorem 6.53, it follows that

$$\mu(a) = p(a)|\Sigma| \operatorname{vec}(U_a) \operatorname{vec}(U_a)^* \tag{6.330}$$

for each $a \in \Gamma$, for some choice of a probability vector $p \in \mathcal{P}(\Gamma)$ and a collection of unitary operators $\{U_a : a \in \Gamma\} \subset \mathrm{U}(\mathbb{C}^{\Sigma})$. Each operator $\mu(a)$ has rank at most one, while

$$\sum_{a \in \Gamma} \mu(a) = \mathbb{1}_{\Sigma} \otimes \mathbb{1}_{\Sigma} \tag{6.331}$$

has rank $|\Sigma|^2$. It follows that $|\Gamma| \geq |\Sigma|^2$ as required. □

Dense coding

Along similar lines to the discussion of teleportation above, a scenario in which Alice and Bob aim to implement an ideal classical channel through shared entanglement and quantum communication may be considered.

Scenario 6.59 (Dense coding)**.** Alice holds a register X and Bob holds Y. Both registers have the same classical state set Σ, and the state of the pair (X, Y) is given by the maximally entangled state

$$\tau = \frac{1}{|\Sigma|} \sum_{b,c \in \Sigma} E_{b,c} \otimes E_{b,c}. \tag{6.332}$$

Assume that Alice obtains a classical register Z whose state set is given by some alphabet Γ. She wishes to transmit the classical state $a \in \Gamma$ of Z to Bob by means of a protocol as follows:

1. For $\{\Psi_a : a \in \Gamma\} \subset \mathrm{C}(\mathcal{X})$ being a collection of channels indexed by Γ, Alice performs Ψ_a on X and sends this register to Bob.
2. Bob applies a measurement $\mu : \Gamma \to \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$ on the pair (X, Y), and interprets the resulting measurement outcome $b \in \Gamma$ as the outcome of Alice's transmission.

It is not surprising that protocols of this sort exist that function as desired—when Γ is no larger than Σ, the task is trivially accomplished. What is more interesting is that there are protocols of this form that work perfectly in the case that Γ is as large as $\Sigma \times \Sigma$.

The following proposition establishes that a dense coding protocol may be derived from an arbitrary mixed-unitary realization of the completely depolarizing channel, provided the unitary operators are drawn uniformly from a set indexed by $\Sigma \times \Sigma$.

Proposition 6.60. *Let Σ be an alphabet, let $\Gamma = \Sigma \times \Sigma$, let $\mathcal{X} = \mathbb{C}^{\Sigma}$, and let*

$$\tau = \frac{1}{|\Sigma|} \sum_{c,d \in \Sigma} E_{c,d} \otimes E_{c,d}. \tag{6.333}$$

Assume $\{U_a : a \in \Gamma\} \subset \mathrm{U}(\mathcal{X})$ is a collection of unitary operators such that

$$\Omega(X) = \frac{1}{|\Sigma|^2} \sum_{a \in \Gamma} U_a X U_a^* \tag{6.334}$$

for all $X \in \mathrm{L}(\mathcal{X})$, where $\Omega \in \mathrm{C}(\mathcal{X})$ is the completely depolarizing channel with respect to the space $\mathcal{X}$. For $\{\Psi_a : a \in \Gamma\} \subseteq \mathrm{C}(\mathcal{X})$ being a collection of channels defined as

$$\Psi_a(X) = U_a X U_a^* \tag{6.335}$$

for each $a \in \Gamma$ and $X \in \mathrm{L}(\mathcal{X})$, and for $\mu : \Gamma \to \mathrm{Pos}(\mathcal{X} \otimes \mathcal{X})$ being defined as

$$\mu(a) = \frac{\mathrm{vec}(U_a)\,\mathrm{vec}(U_a)^*}{|\Sigma|} \tag{6.336}$$

for each $a \in \Gamma$, it holds that μ is a measurement and

$$\big\langle \mu(a), (\Psi_b \otimes \mathbb{1}_{\mathrm{L}(\mathcal{X})})(\tau) \big\rangle = \begin{cases} 1 & \text{if } a = b \\ 0 & \text{if } a \neq b \end{cases} \tag{6.337}$$

for all $a, b \in \Gamma$.

Proof. It holds that

$$\sum_{a\in\Gamma}\mu(a) = \frac{1}{|\Sigma|}\sum_{a\in\Gamma}\operatorname{vec}(U_a)\operatorname{vec}(U_a)^* = |\Sigma|J(\Omega) = \mathbb{1}_{\mathcal{X}}\otimes\mathbb{1}_{\mathcal{X}}. \tag{6.338}$$

As each operator $\mu(a)$ is evidently positive semidefinite, it follows that μ is a measurement. For each $a\in\Gamma$, one has

$$\begin{aligned}&\big\langle \mu(a), (\Psi_a\otimes\mathbb{1}_{\mathrm{L}(\mathcal{X})})(\tau)\big\rangle\\ &\quad= \frac{1}{|\Sigma|^2}\big\langle \operatorname{vec}(U_a)\operatorname{vec}(U_a)^*, \operatorname{vec}(U_a)\operatorname{vec}(U_a)^*\big\rangle = 1.\end{aligned} \tag{6.339}$$

Because $(\Psi_b\otimes\mathbb{1}_{\mathrm{L}(\mathcal{X})})(\tau)$ is a density operator for each $b\in\Gamma$, it follows that

$$\big\langle \mu(a), (\Psi_b\otimes\mathbb{1}_{\mathrm{L}(\mathcal{X})})(\tau)\big\rangle = 0 \tag{6.340}$$

for $a\neq b$, which completes the proof. □

Example 6.61. As in Example 6.56, let $\Sigma=\{0,1\}$, let $\Gamma=\Sigma\times\Sigma$, and define unitary operators $U_{00},U_{01},U_{10},U_{11}\in\mathrm{U}(\mathbb{C}^\Sigma)$ as follows:

$$\begin{aligned}U_{00} &= \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \quad U_{01} = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix},\\ U_{10} &= \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \quad U_{11} = \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}.\end{aligned} \tag{6.341}$$

As the operators $U_{00},U_{01},U_{10},U_{11}$ provide a mixed-unitary realization of the completely depolarizing channel, by taking $\mu:\Gamma\to\operatorname{Pos}(\mathbb{C}^\Sigma\otimes\mathbb{C}^\Sigma)$ to be the measurement defined as

$$\begin{aligned}\mu(00) &= \frac{\operatorname{vec}(U_{00})\operatorname{vec}(U_{00})^*}{2}, \quad \mu(01) = \frac{\operatorname{vec}(U_{01})\operatorname{vec}(U_{01})^*}{2},\\ \mu(10) &= \frac{\operatorname{vec}(U_{10})\operatorname{vec}(U_{10})^*}{2}, \quad \mu(11) = \frac{\operatorname{vec}(U_{11})\operatorname{vec}(U_{11})^*}{2},\end{aligned} \tag{6.342}$$

and setting $\Psi_{ab}(X) = U_{ab}XU_{ab}^*$ for each $X\in\mathrm{L}(\mathbb{C}^\Sigma)$, as in Example 6.56, one obtains a dense coding protocol as described in Scenario 6.59. The resulting protocol is equivalent to the traditional notion of dense coding in which an ideal two-bit classical channel is implemented using a maximally entangled pair of qubits along with one qubit of communication.

In analogy to the more general type of teleportation protocol described previously, one may consider the capabilities of dense coding protocols for arbitrary choices of an alphabet Γ, as opposed to $\Gamma = \Sigma \times \Sigma$. In particular, suppose Alice's channels are given by the collection $\{\Psi_a : a \in \Gamma\}$, for an arbitrary alphabet Γ, and that the symbol $a \in \Gamma$ Alice wishes to send to Bob is randomly selected according to a probability vector $p \in \mathcal{P}(\Gamma)$. The state of the pair (X, Y) prior to Bob's measurement is described by the ensemble $\eta : \Gamma \to \mathrm{Pos}(\mathcal{X} \otimes \mathcal{X})$ defined as

$$\eta(a) = \frac{p(a)}{|\Sigma|} \sum_{b,c \in \Sigma} \Psi_a(E_{b,c}) \otimes E_{b,c} \tag{6.343}$$

for all $a \in \Gamma$. The following theorem provides a characterization of when the Holevo information $\chi(\eta)$ of this ensemble attains its maximum possible value, which is $2\log(|\Sigma|)$.

Theorem 6.62. *Let Σ and Γ be alphabets, let $p \in \mathcal{P}(\Gamma)$ be a probability vector such that $p(a) \neq 0$ for all $a \in \Gamma$, and let $\{\Psi_a : a \in \Gamma\} \subseteq \mathrm{C}(\mathbb{C}^\Sigma)$ be a collection of channels. The following two statements are equivalent:*

1. *For the ensemble $\eta : \Gamma \to \mathrm{Pos}(\mathbb{C}^\Sigma \otimes \mathbb{C}^\Sigma)$ defined as*

$$\eta(a) = \frac{p(a)}{|\Sigma|} \sum_{b,c \in \Sigma} \Psi_a(E_{b,c}) \otimes E_{b,c} \tag{6.344}$$

   for all $a \in \Gamma$, one has that $\chi(\eta) = 2\log(|\Sigma|)$.

2. *There exists a collection $\{U_a : a \in \Gamma\} \subset \mathrm{U}(\mathbb{C}^\Sigma)$ of unitary operators with*

$$\Omega(Y) = \sum_{a \in \Gamma} p(a) U_a Y U_a^* \tag{6.345}$$

   for all $Y \in \mathrm{L}(\mathbb{C}^\Sigma)$, where $\Omega \in \mathrm{C}(\mathbb{C}^\Sigma)$ denotes the completely depolarizing channel with respect to the space $\mathbb{C}^\Sigma$, such that $\Psi_a(X) = U_a X U_a^*$ for every choice of $a \in \Gamma$ and $X \in \mathrm{L}(\mathbb{C}^\Sigma)$.

Proof. The Holevo information of the ensemble η defined by (6.344) is

$$\begin{aligned} \chi(\eta) = \mathrm{H}\left(\sum_{a \in \Gamma} \frac{p(a)}{|\Sigma|} \sum_{b,c \in \Sigma} \Psi_a(E_{b,c}) \otimes E_{b,c} \right) \\ - \sum_{a \in \Gamma} p(a)\, \mathrm{H}\left(\frac{1}{|\Sigma|} \sum_{b,c \in \Sigma} \Psi_a(E_{b,c}) \otimes E_{b,c} \right), \end{aligned} \tag{6.346}$$

which may alternatively be written as

$$\chi(\eta) = \mathrm{H}\left(\sum_{a\in\Gamma} p(a)\frac{\mathrm{J}(\Psi_a)}{|\Sigma|}\right) - \sum_{a\in\Gamma} p(a)\,\mathrm{H}\left(\frac{\mathrm{J}(\Psi_a)}{|\Sigma|}\right). \tag{6.347}$$

Under the assumption that $\chi(\eta) = 2\log(|\Sigma|)$, it must hold that

$$\mathrm{H}\left(\sum_{a\in\Gamma} p(a)\frac{\mathrm{J}(\Psi_a)}{|\Sigma|}\right) = 2\log(|\Sigma|) \quad \text{and} \quad \mathrm{H}\left(\frac{\mathrm{J}(\Psi_a)}{|\Sigma|}\right) = 0 \tag{6.348}$$

for each $a \in \Gamma$. The rank of $J(\Psi_a)$ is therefore equal to 1 for each $a \in \Sigma$, and as each Ψ_a is a channel it follows that there must exist a collection of unitary operators $\{U_a : a \in \Gamma\} \subset \mathrm{U}(\mathbb{C}^\Sigma)$ such that $\Psi_a(X) = U_a X U_a^*$ for each $X \in \mathrm{L}(\mathbb{C}^\Sigma)$ and each $a \in \Gamma$. The left equation of (6.348) is equivalent to

$$\sum_{a\in\Gamma} p(a)\frac{\mathrm{J}(\Psi_a)}{|\Sigma|} = \frac{\mathbb{1}\otimes\mathbb{1}}{|\Sigma|^2}, \tag{6.349}$$

which implies

$$\sum_{a\in\Gamma} p(a)\,\mathrm{vec}(U_a)\,\mathrm{vec}(U_a)^* = \frac{\mathbb{1}\otimes\mathbb{1}}{|\Sigma|} = \mathrm{J}(\Omega), \tag{6.350}$$

and therefore

$$\sum_{a\in\Gamma} p(a) U_a Y U_a^* = \Omega(Y) \tag{6.351}$$

for all $Y \in \mathrm{L}(\mathbb{C}^\Sigma)$. Statement 1 therefore implies statement 2.

Under the assumption that statement 2 holds, the Holevo information of η may be calculated directly:

$$\begin{aligned}
\chi(\eta) &= \mathrm{H}\left(\sum_{a\in\Gamma}\frac{p(a)}{|\Sigma|}\sum_{b,c\in\Sigma}\Psi_a(E_{b,c})\otimes E_{b,c}\right) \\
&\qquad - \sum_{a\in\Gamma} p(a)\,\mathrm{H}\left(\frac{1}{|\Sigma|}\sum_{b,c\in\Sigma}\Psi_a(E_{b,c})\otimes E_{b,c}\right) \\
&= \mathrm{H}\left(\frac{\mathbb{1}\otimes\mathbb{1}}{|\Sigma|^2}\right) - \sum_{a\in\Gamma} p(a)\,\mathrm{H}\left(\frac{\mathrm{vec}(U_a)\,\mathrm{vec}(U_a)^*}{|\Sigma|}\right) \\
&= 2\log(|\Sigma|).
\end{aligned} \tag{6.352}$$

Statement 2 therefore implies statement 1, which completes the proof. □

6.3.2 Non-classical correlations

The definition of entanglement, as the absence of separability, is not directly represented by an observable physical phenomenon. There is, however, a fundamental connection between entanglement and the correlations that may exist among the outcomes of measurements performed on two or more separate parts of a physical system. It is helpful to refer to the following scenario when considering this connection.

Scenario 6.63. Two individuals, Alice and Bob, share a compound register (X, Y), with Alice holding X and Bob holding Y. Two events simultaneously occur:

1. Alice receives an input symbol, drawn from a fixed alphabet Σ_{A}, and she must produce an output symbol from a fixed alphabet Γ_{A}.
2. Bob receives an input symbol, drawn from a fixed alphabet Σ_{B}, and he must produce an output symbol from a fixed alphabet Γ_{B}.

Alice and Bob cannot communicate with one another at any point after they have received their input symbols. The output symbols they produce may, in general, be probabilistic, possibly resulting from measurements made on whichever one of the registers X or Y is in the possession of the individual performing the measurement.

The discussion that follows is primarily concerned with the collections of output distributions that may be produced by Alice and Bob, as described in the scenario above, through measurements on a shared entangled state, as compared with the correlations that may result from the initial state of (X, Y) being separable.

Correlation operators

The output distributions produced by Alice and Bob in a particular instance of Scenario 6.63, ranging over all pairs of input symbols, may collectively be described by a single operator

$$C \in \mathrm{L}\big(\mathbb{R}^{\Sigma_{\mathrm{A}} \times \Gamma_{\mathrm{A}}}, \mathbb{R}^{\Sigma_{\mathrm{B}} \times \Gamma_{\mathrm{B}}}\big), \tag{6.353}$$

defined so that $C((a,c),(b,d))$ is the probability that Alice and Bob output $(c,d) \in \Gamma_{\mathrm{A}} \times \Gamma_{\mathrm{B}}$, assuming they are given the input pair $(a,b) \in \Sigma_{\mathrm{A}} \times \Sigma_{\mathrm{B}}$.

Such an operator must satisfy certain constraints. For instance, to carry the interpretation that C represents a collection of probability distributions, each entry must be a nonnegative real number, and it must hold that

$$\sum_{(c,d)\in\Gamma_{\mathrm{A}}\times\Gamma_{\mathrm{B}}} C\big((a,c),(b,d)\big) = 1 \tag{6.354}$$

for every pair $(a,b) \in \Sigma_{\mathrm{A}} \times \Sigma_{\mathrm{B}}$. Additional constraints are imposed by the assumption that Alice and Bob are separated and cannot communicate.

Definition 6.64. Let Σ_{A}, Σ_{B}, Γ_{A}, and Γ_{B} be alphabets, and let

$$C \in \mathrm{L}\big(\mathbb{R}^{\Sigma_{\mathrm{A}}\times\Gamma_{\mathrm{A}}}, \mathbb{R}^{\Sigma_{\mathrm{B}}\times\Gamma_{\mathrm{B}}}\big) \tag{6.355}$$

be an operator.

1. It is said that C is a *deterministic correlation operator* if and only if it takes the form

   $$C = \sum_{(a,b)\in\Sigma_{\mathrm{A}}\times\Sigma_{\mathrm{B}}} E_{a,b} \otimes E_{f(a),g(b)}, \tag{6.356}$$

   or equivalently

   $$C\big((a,c),(b,d)\big) = \begin{cases} 1 & \text{if } c = f(a) \text{ and } d = g(b) \\ 0 & \text{otherwise,} \end{cases} \tag{6.357}$$

   for some choice of functions $f : \Sigma_{\mathrm{A}} \to \Gamma_{\mathrm{A}}$ and $g : \Sigma_{\mathrm{B}} \to \Gamma_{\mathrm{B}}$. It is said that C is a *probabilistic correlation operator* if and only if C is equal to a convex combination of deterministic correlation operators.

2. The operator C is a *quantum correlation operator* if there exist complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, a state $\rho \in \mathrm{D}(\mathcal{X}\otimes\mathcal{Y})$, and two collections of measurements $\{\mu_a : a \in \Sigma_{\mathrm{A}}\}$ and $\{\nu_b : b \in \Sigma_{\mathrm{B}}\}$, taking the form

   $$\mu_a : \Gamma_{\mathrm{A}} \to \mathrm{Pos}(\mathcal{X}) \quad \text{and} \quad \nu_b : \Gamma_{\mathrm{B}} \to \mathrm{Pos}(\mathcal{Y}), \tag{6.358}$$

   such that

   $$C\big((a,c),(b,d)\big) = \big\langle \mu_a(c) \otimes \nu_b(d), \rho \big\rangle \tag{6.359}$$

   for every $a \in \Sigma_{\mathrm{A}}$, $b \in \Sigma_{\mathrm{B}}$, $c \in \Gamma_{\mathrm{A}}$, and $d \in \Gamma_{\mathrm{B}}$.

$$\mu_0(0) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \mu_0(1) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$
$$\mu_1(0) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \qquad \mu_1(1) = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix},$$
$$\nu_0(0) = \begin{pmatrix} \frac{2+\sqrt{2}}{4} & \frac{\sqrt{2}}{4} \\ \frac{\sqrt{2}}{4} & \frac{2-\sqrt{2}}{4} \end{pmatrix}, \qquad \nu_0(1) = \begin{pmatrix} \frac{2-\sqrt{2}}{4} & -\frac{\sqrt{2}}{4} \\ -\frac{\sqrt{2}}{4} & \frac{2+\sqrt{2}}{4} \end{pmatrix},$$
$$\nu_1(0) = \begin{pmatrix} \frac{2+\sqrt{2}}{4} & -\frac{\sqrt{2}}{4} \\ -\frac{\sqrt{2}}{4} & \frac{2-\sqrt{2}}{4} \end{pmatrix}, \qquad \nu_1(1) = \begin{pmatrix} \frac{2-\sqrt{2}}{4} & \frac{\sqrt{2}}{4} \\ \frac{\sqrt{2}}{4} & \frac{2+\sqrt{2}}{4} \end{pmatrix}.$$

Figure 6.1: Matrix representations of the measurement operators described in Example 6.65.

Example 6.65. Let Σ_A, Σ_B, Γ_A, and Γ_B all be equal to the binary alphabet $\Sigma = \{0,1\}$, let $\mathcal{X} = \mathbb{C}^\Sigma$ and $\mathcal{Y} = \mathbb{C}^\Sigma$, define $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ to be the maximally entangled state

$$\rho = \frac{1}{2} \sum_{a,b \in \Sigma} E_{a,b} \otimes E_{a,b}, \tag{6.360}$$

and define measurements $\mu_0, \mu_1 : \Gamma_A \to \mathrm{Pos}(\mathcal{X})$ and $\nu_0, \nu_1 : \Gamma_B \to \mathrm{Pos}(\mathcal{Y})$ as

$$\begin{aligned} \mu_0(0) &= \Pi_0, & \mu_0(1) &= \Pi_{\pi/2}, \\ \mu_1(0) &= \Pi_{\pi/4}, & \mu_1(1) &= \Pi_{3\pi/4}, \\ \nu_0(0) &= \Pi_{\pi/8}, & \nu_0(1) &= \Pi_{5\pi/8}, \\ \nu_1(0) &= \Pi_{7\pi/8}, & \nu_1(1) &= \Pi_{3\pi/8}, \end{aligned} \tag{6.361}$$

for

$$\Pi_\theta = \begin{pmatrix} \cos^2(\theta) & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin^2(\theta) \end{pmatrix}. \tag{6.362}$$

Equivalently, these measurement operators are as described in Figure 6.1.

For this choice of ρ, and because each of the measurement operators above have real number entries, it holds that

$$\big\langle \mu_a(c) \otimes \nu_b(d), \rho \big\rangle = \frac{1}{2} \big\langle \mu_a(c), \nu_b(d) \big\rangle \tag{6.363}$$

for each $a \in \Sigma_A$, $b \in \Sigma_B$, $c \in \Gamma_A$, and $d \in \Gamma_B$. A calculation reveals that the quantum correlation operator defined by (6.359) is given by

$$C = \begin{pmatrix} \frac{2+\sqrt{2}}{8} & \frac{2-\sqrt{2}}{8} & \frac{2+\sqrt{2}}{8} & \frac{2-\sqrt{2}}{8} \\ \frac{2-\sqrt{2}}{8} & \frac{2+\sqrt{2}}{8} & \frac{2-\sqrt{2}}{8} & \frac{2+\sqrt{2}}{8} \\ \frac{2+\sqrt{2}}{8} & \frac{2-\sqrt{2}}{8} & \frac{2-\sqrt{2}}{8} & \frac{2+\sqrt{2}}{8} \\ \frac{2-\sqrt{2}}{8} & \frac{2+\sqrt{2}}{8} & \frac{2+\sqrt{2}}{8} & \frac{2-\sqrt{2}}{8} \end{pmatrix}. \tag{6.364}$$

It will be demonstrated shortly that the operator C is not a probabilistic correlation operator.

Example 6.66. Let Σ_A, Σ_B, Γ_A, and Γ_B all be equal to the binary alphabet $\Sigma = \{0, 1\}$. There are 16 deterministic correlation operators, which are in correspondence with the 16 possible pairs of functions (f, g) having the form $f : \Sigma_A \to \Gamma_A$ and $g : \Sigma_B \to \Gamma_B$. As matrices, these operators are as described in Figure 6.2.

Bell inequalities

By its definition, the set of all probabilistic correlation operators of the form

$$C \in \mathrm{L}\big(\mathbb{R}^{\Sigma_A \times \Gamma_A}, \mathbb{R}^{\Sigma_B \times \Gamma_B}\big) \tag{6.365}$$

is convex. Indeed, this set is given by the convex hull of a finite set, as there are finitely many deterministic correlation operators of the same form. From this fact it follows that the set of all probabilistic correlation operators of the form (6.365) is compact. Therefore, by the separating hyperplane theorem (Theorem 1.11), if an operator

$$D \in \mathrm{L}\big(\mathbb{R}^{\Sigma_A \times \Gamma_A}, \mathbb{R}^{\Sigma_B \times \Gamma_B}\big) \tag{6.366}$$

is not a probabilistic correlation operator, there must exist an operator

$$K \in \mathrm{L}\big(\mathbb{R}^{\Sigma_A \times \Gamma_A}, \mathbb{R}^{\Sigma_B \times \Gamma_B}\big) \tag{6.367}$$

and a real number α such that

$$\langle K, D \rangle > \alpha \quad \text{and} \quad \langle K, C \rangle \leq \alpha \tag{6.368}$$

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}.$$

Figure 6.2: Matrix representations of the correlation operators described in Example 6.66.

for all probabilistic correlation operators C of the form (6.365).

For a fixed choice of an operator K and a real number α, the inequality $\langle K, C \rangle \leq \alpha$ is traditionally called a *Bell inequality*, assuming it is satisfied for every probabilistic correlation operator C of the form (6.365). When this is the case, the inequality $\langle K, D \rangle > \alpha$ is called a *Bell inequality violation* if it holds for some choice of a quantum correlation operator D.

The illustration of a Bell inequality violation can provide a convenient way to demonstrate that certain correlation operators are not probabilistic, as the following example illustrates.

Example 6.67 (Clauser–Horn–Shimony–Holt inequality)**.** Let Σ_A, Σ_B, Γ_A, and Γ_B all be equal to the binary alphabet $\Sigma = \{0, 1\}$, and define

$$K \in \mathrm{L}\big(\mathbb{R}^{\Sigma_A \times \Gamma_A}, \mathbb{R}^{\Sigma_B \times \Gamma_B}\big) \tag{6.369}$$

as

$$K = \begin{pmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{pmatrix}. \tag{6.370}$$

For every deterministic correlation operator

$$C \in \mathrm{L}\big(\mathbb{R}^{\Sigma_{\mathrm{A}} \times \Gamma_{\mathrm{A}}}, \mathbb{R}^{\Sigma_{\mathrm{B}} \times \Gamma_{\mathrm{B}}}\big) \tag{6.371}$$

it holds that

$$\langle K, C \rangle \leq 2, \tag{6.372}$$

which may be verified by an inspection of the 16 deterministic correlation operators in Example 6.66. It follows by linearity that the same inequality holds for C being any probabilistic correlation operator. On the other hand, the quantum correlation operator

$$D = \frac{1}{8} \begin{pmatrix} 2+\sqrt{2} & 2-\sqrt{2} & 2+\sqrt{2} & 2-\sqrt{2} \\ 2-\sqrt{2} & 2+\sqrt{2} & 2-\sqrt{2} & 2+\sqrt{2} \\ 2+\sqrt{2} & 2-\sqrt{2} & 2-\sqrt{2} & 2+\sqrt{2} \\ 2-\sqrt{2} & 2+\sqrt{2} & 2+\sqrt{2} & 2-\sqrt{2} \end{pmatrix} \tag{6.373}$$

described in Example 6.65 satisfies

$$\langle K, D \rangle = 2\sqrt{2}. \tag{6.374}$$

This demonstrates that D is not a probabilistic correlation operator.

Correlations among binary-valued measurements

For a given choice of alphabets Σ_{A}, Σ_{B}, Γ_{A}, and Γ_{B}, and an operator

$$K \in \mathrm{L}\big(\mathbb{R}^{\Sigma_{\mathrm{A}} \times \Gamma_{\mathrm{A}}}, \mathbb{R}^{\Sigma_{\mathrm{B}} \times \Gamma_{\mathrm{B}}}\big), \tag{6.375}$$

it is evidently quite difficult in some cases to determine the supremum value of $\langle K, C \rangle$, optimized over all quantum correlation operators of the form

$$C \in \mathrm{L}\big(\mathbb{R}^{\Sigma_{\mathrm{A}} \times \Gamma_{\mathrm{A}}}, \mathbb{R}^{\Sigma_{\mathrm{B}} \times \Gamma_{\mathrm{B}}}\big). \tag{6.376}$$

There is, however, an interesting class of operators K for which this problem is solvable. This is the class for which the output alphabets Γ_{A} and Γ_{B} are both equal to the binary alphabet $\Sigma = \{0, 1\}$, and furthermore the operator K takes the form

$$K = M \otimes \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \tag{6.377}$$

for some choice of an operator

$$M \in \mathrm{L}\big(\mathbb{R}^{\Sigma_{\mathrm{B}}}, \mathbb{R}^{\Sigma_{\mathrm{A}}}\big). \tag{6.378}$$

Operators of the form (6.377) have a simple interpretation when considered in the context of Bell inequalities and violations—they effectively assign the value $M(a, b)$ to the event that Alice and Bob output equal binary-valued answers, and the value $-M(a, b)$ to the event that their outputs differ, for each possible question pair (a, b).

The following theorem, known as Tsirelson's theorem, provides the basis for a solution to the problem under consideration.

Theorem 6.68 (Tsirelson's theorem)**.** *Let $X \in \mathrm{L}\big(\mathbb{R}^{\Sigma_{\mathrm{B}}}, \mathbb{R}^{\Sigma_{\mathrm{A}}}\big)$ be an operator, for alphabets Σ_{A} and Σ_{B}. The following statements are equivalent:*

1. *There exist complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, a state $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$, and two collections $\{A_a : a \in \Sigma_{\mathrm{A}}\} \subset \mathrm{Herm}(\mathcal{X})$ and $\{B_b : b \in \Sigma_{\mathrm{B}}\} \subset \mathrm{Herm}(\mathcal{Y})$ of operators satisfying $\|A_a\| \leq 1$, $\|B_b\| \leq 1$, such that*

$$X(a, b) = \big\langle A_a \otimes B_b, \rho \big\rangle \tag{6.379}$$

   for every $a \in \Sigma_{\mathrm{A}}$ and $b \in \Sigma_{\mathrm{B}}$.

2. *Statement 1 holds under the additional requirement that, for some choice of an alphabet Γ, one has $\mathcal{X} = \mathbb{C}^{\Gamma}$, $\mathcal{Y} = \mathbb{C}^{\Gamma}$, and*

$$\rho = \frac{1}{|\Gamma|} \sum_{c,d \in \Gamma} E_{c,d} \otimes E_{c,d}, \tag{6.380}$$

   and furthermore that the operators in the collections

$$\{A_a : a \in \Sigma_{\mathrm{A}}\} \quad \text{and} \quad \{B_b : b \in \Sigma_{\mathrm{B}}\} \tag{6.381}$$

   are unitary (in addition to being Hermitian).

3. *There exist operators*

$$P \in \mathrm{Pos}\big(\mathbb{C}^{\Sigma_{\mathrm{A}}}\big) \quad \textit{and} \quad Q \in \mathrm{Pos}\big(\mathbb{C}^{\Sigma_{\mathrm{B}}}\big), \tag{6.382}$$

with $P(a,a) = 1$ *and* $Q(b,b) = 1$ *for every* $a \in \Sigma_{\mathrm{A}}$ *and* $b \in \Sigma_{\mathrm{B}}$*, such that*

$$\begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \in \mathrm{Pos}\big(\mathbb{C}^{\Sigma_{\mathrm{A}}} \oplus \mathbb{C}^{\Sigma_{\mathrm{B}}}\big). \tag{6.383}$$

4. *There exist two collections* $\{u_a : a \in \Sigma_{\mathrm{A}}\}$ *and* $\{v_b : b \in \Sigma_{\mathrm{B}}\}$ *of unit vectors, with elements drawn from the space* $\mathbb{R}^{\Sigma_{\mathrm{A}}} \oplus \mathbb{R}^{\Sigma_{\mathrm{B}}}$*, such that*

$$X(a,b) = \langle u_a, v_b \rangle \tag{6.384}$$

for every $a \in \Sigma_{\mathrm{A}}$ *and* $b \in \Sigma_{\mathrm{B}}$*.*

The proof of this theorem will make use of a collection of unitary and Hermitian operators known as *Weyl–Brauer operators*.

Definition 6.69. Let m be a positive integer, let $\Gamma = \{0,1\}$ denote the binary alphabet, and let $\mathcal{Z} = \mathbb{C}^{\Gamma}$. The *Weyl–Brauer operators* $V_0, \ldots, V_{2m} \in \mathrm{L}\big(\mathcal{Z}^{\otimes m}\big)$ of order m are defined as follows:

$$V_0 = \sigma_z^{\otimes m} \tag{6.385}$$

and

$$\begin{aligned} V_{2k-1} &= \sigma_z^{\otimes(k-1)} \otimes \sigma_x \otimes \mathbb{1}^{\otimes(m-k)}, \\ V_{2k} &= \sigma_z^{\otimes(k-1)} \otimes \sigma_y \otimes \mathbb{1}^{\otimes(m-k)}, \end{aligned} \tag{6.386}$$

for $k = 1, \ldots, m$, where $\mathbb{1}$ denotes the identity operator on $\mathcal{Z}$ and σ_x, σ_y, and σ_z are given by the Pauli operators. In matrix form, these operators are as follows:

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{6.387}$$

Example 6.70. In the case $m = 3$, the Weyl–Brauer operators $V_0, \ldots, V_6$ are

$$\begin{aligned} V_0 &= \sigma_z \otimes \sigma_z \otimes \sigma_z \\ V_1 &= \sigma_x \otimes \mathbb{1} \otimes \mathbb{1} \\ V_2 &= \sigma_y \otimes \mathbb{1} \otimes \mathbb{1} \\ V_3 &= \sigma_z \otimes \sigma_x \otimes \mathbb{1} \\ V_4 &= \sigma_z \otimes \sigma_y \otimes \mathbb{1} \\ V_5 &= \sigma_z \otimes \sigma_z \otimes \sigma_x \\ V_6 &= \sigma_z \otimes \sigma_z \otimes \sigma_y. \end{aligned} \tag{6.388}$$

A proposition summarizing the properties of the Weyl–Brauer operators that are relevant to the proof of Tsirelson's theorem follows.

Proposition 6.71. *Let m be a positive integer, let $V_0, \ldots, V_{2m}$ denote the Weyl–Brauer operators of order m, and let $(\alpha_0, \ldots, \alpha_{2m}), (\beta_0, \ldots, \beta_{2m}) \in \mathbb{R}^{2m+1}$ be vectors of real numbers. It holds that*

$$\left(\sum_{k=0}^{2m} \alpha_k V_k\right)^2 = \left(\sum_{k=0}^{2m} \alpha_k^2\right) \mathbb{1}^{\otimes m} \tag{6.389}$$

and

$$\frac{1}{2^m}\left\langle \sum_{j=0}^{2m} \alpha_j V_j, \sum_{k=0}^{2m} \beta_k V_k \right\rangle = \sum_{k=0}^{2m} \alpha_k \beta_k. \tag{6.390}$$

Proof. The Pauli operators anti-commute in pairs:

$$\sigma_x \sigma_y = -\sigma_y \sigma_x, \quad \sigma_x \sigma_z = -\sigma_z \sigma_x, \quad \text{and} \quad \sigma_y \sigma_z = -\sigma_z \sigma_y. \tag{6.391}$$

By an inspection of the definition of the Weyl–Brauer operators, it follows that $V_0, \ldots, V_{2m}$ also anti-commute in pairs:

$$V_j V_k = -V_k V_j \tag{6.392}$$

for distinct choices of $j, k \in \{0, \ldots, 2m\}$. Moreover, each V_k is both unitary and Hermitian, and therefore $V_k^2 = \mathbb{1}^{\otimes m}$. It follows that

$$\begin{aligned} \left(\sum_{k=0}^{2m} \alpha_k V_k\right)^2 &= \sum_{k=0}^{2m} \alpha_k^2 V_k^2 + \sum_{0 \le j < k \le 2m} \alpha_j \alpha_k \big(V_j V_k + V_k V_j\big) \\ &= \left(\sum_{k=0}^{2m} \alpha_k^2\right) \mathbb{1}^{\otimes m}. \end{aligned} \tag{6.393}$$

Moreover,

$$\langle V_j, V_k \rangle = \begin{cases} 2^m & \text{if } j = k \\ 0 & \text{if } j \neq k, \end{cases} \tag{6.394}$$

and therefore

$$\frac{1}{2^m}\left\langle \sum_{j=0}^{2m} \alpha_j V_j, \sum_{k=0}^{2m} \beta_k V_k \right\rangle = \frac{1}{2^m} \sum_{j=0}^{2m} \sum_{k=0}^{2m} \alpha_j \beta_k \langle V_j, V_k \rangle = \sum_{k=0}^{2m} \alpha_k \beta_k, \tag{6.395}$$

as required. □

Proof of Theorem 6.68. The implications to be proved among the statements, which suffice to prove the theorem, may be summarized as follows:

$$(2) \Rightarrow (1) \Rightarrow (3) \Rightarrow (4) \Rightarrow (2). \tag{6.396}$$

One has that statement 2 trivially implies statement 1.

Assume statement 1 holds, define an operator

$$K = \sum_{a \in \Sigma_{\mathrm{A}}} e_a \operatorname{vec}\bigl((A_a \otimes \mathbb{1})\sqrt{\rho}\bigr)^* + \sum_{b \in \Sigma_{\mathrm{B}}} e_b \operatorname{vec}\bigl((\mathbb{1} \otimes B_b)\sqrt{\rho}\bigr)^*, \tag{6.397}$$

and consider the operator $KK^* \in \operatorname{Pos}\bigl(\mathbb{C}^{\Sigma_{\mathrm{A}} \sqcup \Sigma_{\mathrm{B}}}\bigr)$, which may be written in a block form as

$$KK^* = \begin{pmatrix} P & Y \\ Y^* & Q \end{pmatrix} \tag{6.398}$$

for $P \in \operatorname{Pos}\bigl(\mathbb{C}^{\Sigma_{\mathrm{A}}}\bigr)$, $Q \in \operatorname{Pos}\bigl(\mathbb{C}^{\Sigma_{\mathrm{B}}}\bigr)$, and $Y \in \operatorname{Pos}\bigl(\mathbb{C}^{\Sigma_{\mathrm{B}}}, \mathbb{C}^{\Sigma_{\mathrm{A}}}\bigr)$. It holds that

$$Y(a,b) = \bigl\langle (A_a \otimes \mathbb{1})\sqrt{\rho}, (\mathbb{1} \otimes B_b)\sqrt{\rho} \bigr\rangle = \bigl\langle A_a \otimes B_b, \rho \bigr\rangle = X(a,b) \tag{6.399}$$

for every $a \in \Sigma_{\mathrm{A}}$ and $b \in \Sigma_{\mathrm{B}}$, and therefore $Y = X$. Moreover, for each $a \in \Sigma_{\mathrm{A}}$ one has

$$P(a,a) = \bigl\langle (A_a \otimes \mathbb{1})\sqrt{\rho}, (A_a \otimes \mathbb{1})\sqrt{\rho} \bigr\rangle = \bigl\langle A_a^2 \otimes \mathbb{1}, \rho \bigr\rangle, \tag{6.400}$$

which is necessarily a nonnegative real number in the interval $[0,1]$; and through a similar calculation, one finds that $Q(b,b)$ is also a nonnegative integer in the interval $[0,1]$ for each $b \in \Sigma_{\mathrm{B}}$. A nonnegative real number may be added to each diagonal entry of this operator to yield another positive semidefinite operator, so one has that statement 3 holds. It has therefore been proved that statement 1 implies statement 3.

Next, assume statement 3 holds, and observe that

$$\frac{1}{2}\begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} + \frac{1}{2}\begin{pmatrix} P & X \\ X^* & Q \end{pmatrix}^{\mathsf{T}} = \begin{pmatrix} \frac{P+\overline{P}}{2} & X \\ X^* & \frac{Q+\overline{Q}}{2} \end{pmatrix} \tag{6.401}$$

is a positive semidefinite operator having real number entries, and all of its diagonal entries are equal to 1. Define

$$u_a = \begin{pmatrix} \frac{P+\overline{P}}{2} & X \\ X^* & \frac{Q+\overline{Q}}{2} \end{pmatrix}^{\frac{1}{2}} \begin{pmatrix} e_a \\ 0 \end{pmatrix} \quad \text{and} \quad v_b = \begin{pmatrix} \frac{P+\overline{P}}{2} & X \\ X^* & \frac{Q+\overline{Q}}{2} \end{pmatrix}^{\frac{1}{2}} \begin{pmatrix} 0 \\ e_b \end{pmatrix} \tag{6.402}$$

for each $a \in \Sigma_{\mathrm{A}}$ and $b \in \Sigma_{\mathrm{B}}$. As the square root of a positive semidefinite operator having real number entries also has real number entries, one has that u_a and v_b are unit vectors with real number entries, and moreover it holds that

$$\langle u_a, v_b \rangle = X(a,b) \tag{6.403}$$

for all $a \in \Sigma_{\mathrm{A}}$ and $b \in \Sigma_{\mathrm{B}}$. It has therefore been proved that statement 3 implies statement 4.

Finally, assume statement 4 holds. Let

$$m = \left\lceil \frac{|\Sigma_{\mathrm{A}}| + |\Sigma_{\mathrm{B}}| - 1}{2} \right\rceil, \tag{6.404}$$

so that $2m + 1 \geq |\Sigma_{\mathrm{A}}| + |\Sigma_{\mathrm{B}}|$, and let $f : \Sigma_{\mathrm{A}} \sqcup \Sigma_{\mathrm{B}} \to \{0, \ldots, 2m\}$ be a fixed but otherwise arbitrarily chosen one-to-one function. Let $\Gamma = \{0, 1\}$, let $\mathcal{Z} = \mathbb{C}^{\Gamma}$, and define

$$A_a = \sum_{c \in \Sigma_{\mathrm{A}} \sqcup \Sigma_{\mathrm{B}}} u_a(c) V_{f(c)} \quad \text{and} \quad B_b = \sum_{c \in \Sigma_{\mathrm{A}} \sqcup \Sigma_{\mathrm{B}}} v_b(c) V_{f(c)}^{\mathsf{T}} \tag{6.405}$$

for each $a \in \Sigma_{\mathrm{A}}$ and $b \in \Sigma_{\mathrm{B}}$, for $V_0, \ldots, V_{2m}$ being the Weyl–Brauer operators of order m, regarded as operators acting on $\mathcal{Z}^{\otimes m}$. As the vectors

$$\{u_a \,:\, a \in \Sigma_{\mathrm{A}}\} \quad \text{and} \quad \{v_b \,:\, b \in \Sigma_{\mathrm{B}}\} \tag{6.406}$$

are unit vectors having real number entries, it follows from Proposition 6.71 that the operators

$$\{A_a \,:\, a \in \Sigma_{\mathrm{A}}\} \quad \text{and} \quad \{B_b \,:\, b \in \Sigma_{\mathrm{B}}\} \tag{6.407}$$

are unitary, and it is evident that they are Hermitian as well. Define

$$\tau = \frac{1}{2^m} \operatorname{vec}\bigl(\mathbb{1}_{\mathcal{Z}}^{\otimes m}\bigr) \operatorname{vec}\bigl(\mathbb{1}_{\mathcal{Z}}^{\otimes m}\bigr)^{*}. \tag{6.408}$$

For each choice of $a \in \Sigma_{\mathrm{A}}$ and $b \in \Sigma_{\mathrm{B}}$ it holds that

$$\begin{aligned} \bigl\langle A_a \otimes B_b, \tau \bigr\rangle &= \frac{1}{2^m} \operatorname{Tr}\bigl(A_a B_b^{\mathsf{T}}\bigr) \\ &\frac{1}{2^m} \sum_{c,d \in \Sigma_{\mathrm{A}} \sqcup \Sigma_{\mathrm{B}}} \bigl\langle u_a(c) V_{f(c)}, v_b(d) V_{f(d)} \bigr\rangle = \langle u_a, v_b \rangle, \end{aligned} \tag{6.409}$$

again by Proposition 6.71. This is equivalent to statement 2 (taking Γ^m in place of Γ). It has therefore been proved that statement 4 implies statement 2, which completes the proof. □

As a consequence of Tsirelson's theorem (Theorem 6.68), there exists a semidefinite program for the supremum value of the inner product $\langle K, C \rangle$, for K taking the form (6.377) and for C ranging over all quantum correlation operators of the form

$$C \in \mathrm{L}\big(\mathbb{R}^{\Sigma_{\mathrm{A}} \times \Gamma_{\mathrm{A}}}, \mathbb{R}^{\Sigma_{\mathrm{B}} \times \Gamma_{\mathrm{B}}}\big), \tag{6.410}$$

for Σ_{A} and Σ_{B} being arbitrary alphabets and Γ_{A} and Γ_{B} both being equal to the binary alphabet $\Gamma = \{0, 1\}$.

To understand why this is so, consider an arbitrary quantum correlation operator C, which must be given by

$$C\big((a, c), (b, d)\big) = \big\langle \mu_a(c) \otimes \nu_b(d), \rho \big\rangle \tag{6.411}$$

for every $a \in \Sigma_{\mathrm{A}}$, $b \in \Sigma_{\mathrm{B}}$, and $c, d \in \Gamma$, for some choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$, a state $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$, and two collections of measurements $\{\mu_a : a \in \Sigma_{\mathrm{A}}\}$ and $\{\nu_b : b \in \Sigma_{\mathrm{B}}\}$ whose elements take the form

$$\mu_a : \Gamma \to \mathrm{Pos}(\mathcal{X}) \quad \text{and} \quad \nu_b : \Gamma \to \mathrm{Pos}(\mathcal{Y}). \tag{6.412}$$

For an operator K of the form (6.377) for some choice of $M \in \mathrm{L}\big(\mathbb{R}^{\Sigma_{\mathrm{B}}}, \mathbb{R}^{\Sigma_{\mathrm{A}}}\big)$, one has that the value of the inner product $\langle K, C \rangle$ is given by

$$\sum_{(a,b) \in \Sigma_{\mathrm{A}} \times \Sigma_{\mathrm{B}}} M(a, b) \big\langle (\mu_a(0) - \mu_a(1)) \otimes (\nu_b(0) - \nu_b(1)), \rho \big\rangle. \tag{6.413}$$

Now, an operator H, acting on an arbitrary complex Euclidean space, may be written as

$$H = \mu(0) - \mu(1) \tag{6.414}$$

for some binary-valued measurement μ if and only if H is Hermitian and satisfies $\|H\| \leq 1$. Thus, an optimization of the expression (6.413) over all choices of the measurements $\{\mu_a : a \in \Sigma_{\mathrm{A}}\}$ and $\{\nu_b : b \in \Sigma_{\mathrm{B}}\}$ is equivalent to an optimization of the expression

$$\sum_{(a,b) \in \Sigma_{\mathrm{A}} \times \Sigma_{\mathrm{B}}} M(a, b) \big\langle A_a \otimes B_b, \rho \big\rangle \tag{6.415}$$

over all collections

$$\{A_a : a \in \Sigma_{\mathrm{A}}\} \subset \mathrm{Herm}(\mathcal{X}) \quad \text{and} \quad \{B_b : b \in \Sigma_{\mathrm{B}}\} \subset \mathrm{Herm}(\mathcal{Y}) \tag{6.416}$$

of Hermitian operators satisfying $\|A_a\| \leq 1$ and $\|B_b\| \leq 1$, for every $a \in \Sigma_{\mathrm{A}}$ and $b \in \Sigma_{\mathrm{B}}$, respectively.

By optimizing over all complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$ and density operators $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$, one finds (by Theorem 6.68) that the supremum value of $\langle K, C \rangle$ over all quantum correlation operators C is equal to the supremum value of the inner product $\langle M, X \rangle$ over all choices of operators $X \in \mathrm{L}\big(\mathbb{R}^{\Sigma_{\mathrm{B}}}, \mathbb{R}^{\Sigma_{\mathrm{A}}}\big)$ for which it holds that

$$\begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \in \mathrm{Pos}\big(\mathbb{C}^{\Sigma_{\mathrm{A}}} \oplus \mathbb{C}^{\Sigma_{\mathrm{B}}}\big), \tag{6.417}$$

for $P \in \mathrm{Pos}\big(\mathbb{C}^{\Sigma_{\mathrm{A}}}\big)$ and $Q \in \mathrm{Pos}\big(\mathbb{C}^{\Sigma_{\mathrm{B}}}\big)$ satisfying $P(a, a) = 1$ and $Q(b, b) = 1$ for every $a \in \Sigma_{\mathrm{A}}$ and $b \in \Sigma_{\mathrm{B}}$. Such an optimization corresponds directly to the following primal problem of a semidefinite program:

$$\underline{\text{Primal problem}}$$

$$\begin{aligned} \text{maximize:} \quad & \frac{1}{2}\langle M, X \rangle + \frac{1}{2}\langle M^*, X^* \rangle \\ \text{subject to:} \quad & \begin{pmatrix} P & X \\ X^* & Q \end{pmatrix} \geq 0, \\ & \Delta(P) = \mathbb{1},\ \Delta(Q) = \mathbb{1}, \\ & P \in \mathrm{Pos}\big(\mathbb{C}^{\Sigma_{\mathrm{A}}}\big),\ Q \in \mathrm{Pos}\big(\mathbb{C}^{\Sigma_{\mathrm{B}}}\big), \\ & X \in \mathrm{L}\big(\mathbb{C}^{\Sigma_{\mathrm{B}}}, \mathbb{C}^{\Sigma_{\mathrm{A}}}\big). \end{aligned}$$

In this problem, Δ refers to the completely dephasing channel, defined with respect to either $\mathbb{C}^{\Sigma_{\mathrm{A}}}$ or $\mathbb{C}^{\Sigma_{\mathrm{B}}}$, and $\mathbb{1}$ denotes the identity operator on either of these spaces, as the context dictates without ambiguity.

The dual problem of this semidefinite program is as follows:

$$\underline{\text{Dual problem}}$$

$$\begin{aligned} \text{minimize:} \quad & \frac{1}{2}\,\mathrm{Tr}(Y) + \frac{1}{2}\,\mathrm{Tr}(Z) \\ \text{subject to:} \quad & \begin{pmatrix} \Delta(Y) & -M \\ -M^* & \Delta(Z) \end{pmatrix} \geq 0, \\ & Y \in \mathrm{Herm}\big(\mathbb{C}^{\Sigma_{\mathrm{A}}}\big), \\ & Z \in \mathrm{Herm}\big(\mathbb{C}^{\Sigma_{\mathrm{B}}}\big). \end{aligned}$$

It follows from Slater's theorem (Theorem 1.18) that strong duality holds for this semidefinite program—strict feasibility holds for both the primal and dual problems.

Example 6.72 (Tsirelson's bound)**.** Consider the operator

$$K = \begin{pmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{pmatrix} = M \otimes \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \tag{6.418}$$

for

$$M = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{6.419}$$

which was examined in Example 6.67. One has $\|M\| = \sqrt{2}$, so that

$$\begin{pmatrix} \sqrt{2}\mathbb{1} & -M \\ -M^* & \sqrt{2}\mathbb{1} \end{pmatrix} \geq 0. \tag{6.420}$$

By taking $Y = \sqrt{2}\mathbb{1}$ and $Z = \sqrt{2}\mathbb{1}$ in the dual problem above, a feasible dual solution achieving the objective value $2\sqrt{2}$ is obtained. Therefore,

$$\langle K, C \rangle \leq 2\sqrt{2} \tag{6.421}$$

for every quantum correlation operator C. The Bell inequality violation exhibited in Example 6.67 is therefore optimal for this choice of K.

6.4 Exercises

6.1. Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces and let $\Phi \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$ be a channel. Prove that the following three statements are equivalent:

1. For every complex Euclidean space $\mathcal{Z}$ and every state $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Z})$, it holds that
$$\big(\Phi \otimes \mathbb{1}_{\mathrm{L}(\mathcal{Z})}\big)(\rho) \in \mathrm{SepD}(\mathcal{Y} : \mathcal{Z}). \tag{6.422}$$
2. $J(\Phi) \in \mathrm{Sep}(\mathcal{Y} : \mathcal{X})$.

3. There exists an alphabet Σ, a measurement $\mu : \Sigma \to \mathrm{Pos}(\mathcal{X})$, and a collection of states $\{\sigma_a : a \in \Sigma\} \subseteq \mathrm{D}(\mathcal{Y})$ such that

$$\Phi(X) = \sum_{a \in \Sigma} \langle \mu(a), X \rangle \sigma_a \tag{6.423}$$

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

Channels for which these statements hold are called *entanglement-breaking* channels.

6.2. Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces, both being of dimension n, let $\{U_1, \ldots, U_m\} \in \mathrm{U}(\mathcal{Y}, \mathcal{X})$ be a set of pairwise orthogonal isometries, and let $u_k \in \mathcal{X} \otimes \mathcal{Y}$ be the vector defined as

$$u_k = \frac{1}{\sqrt{n}} \mathrm{vec}(U_k) \tag{6.424}$$

for each $k \in \{1, \ldots, m\}$. Let $\mu : \{1, \ldots, m\} \to \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$ be a measurement such that $\mu(k) \in \mathrm{PPT}(\mathcal{X} : \mathcal{Y})$ for every $k \in \{1, \ldots, m\}$. Prove that

$$\sum_{k=1}^{m} \langle \mu(k), u_k u_k^* \rangle \leq n. \tag{6.425}$$

(Observe that a correct solution to this exercise generalizes Theorem 6.33.)

6.3. Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces having dimension at least 2. Prove that there exist entanglement-breaking channels $\Phi_0, \Phi_1 \in \mathrm{C}(\mathcal{X}, \mathcal{Y})$, as defined in Exercise 6.1, such that

$$\bigl|\!\bigl|\!\bigl| \Phi_0 - \Phi_1 \bigr|\!\bigr|\!\bigr|_1 > \bigl\| \Phi_0(\rho) - \Phi_1(\rho) \bigr\|_1 \tag{6.426}$$

for every $\rho \in \mathrm{D}(\mathcal{X})$. Such channels have the seemingly strange property that they destroy entanglement, and yet evaluating them on an entangled state helps to discriminate between them.

6.4. Let X and Y be registers and let $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ be a state of the pair (X, Y). With respect to ρ, one defines the *entanglement of formation* between the pair (X, Y) as

$$\mathrm{E_F}(\mathsf{X} : \mathsf{Y}) = \inf \left\{ \sum_{a \in \Sigma} p(a)\, \mathrm{H}\bigl(\mathrm{Tr}_{\mathcal{Y}}(u_a u_a^*)\bigr) \;:\; \sum_{a \in \Sigma} p(a) u_a u_a^* = \rho \right\}, \tag{6.427}$$

where the infimum is over all choices of an alphabet Σ, a probability vector $p \in \mathcal{P}(\Sigma)$, and a collection of unit vectors $\{u_a : a \in \Sigma\} \subset \mathcal{X} \otimes \mathcal{Y}$ for which it holds that

$$\sum_{a \in \Sigma} p(a) u_a u_a^* = \rho. \tag{6.428}$$

(a) Prove that the infimum in (6.427) is achieved for some choice of Σ, p, and $\{u_a : a \in \Sigma\}$ for which $|\Sigma| \leq \dim(\mathcal{X} \otimes \mathcal{Y})^2$.

(b) Prove that $\mathrm{E}_{\mathrm{D}}(\mathsf{X} : \mathsf{Y}) \leq \mathrm{E}_{\mathrm{F}}(\mathsf{X} : \mathsf{Y}) \leq \mathrm{E}_{\mathrm{C}}(\mathsf{X} : \mathsf{Y})$.

(c) Suppose that Z and W are registers and $\Phi \in \mathrm{LOCC}(\mathcal{X}, \mathcal{Z} : \mathcal{Y}, \mathcal{W})$ is an LOCC channel. Prove that

$$\mathrm{E}_{\mathrm{F}}(\mathsf{Z} : \mathsf{W})_\sigma \leq \mathrm{E}_{\mathrm{F}}(\mathsf{X} : \mathsf{Y})_\rho \tag{6.429}$$

where $\sigma = \Phi(\rho)$ and $\mathrm{E}_{\mathrm{F}}(\mathsf{X} : \mathsf{Y})_\rho$ and $\mathrm{E}_{\mathrm{F}}(\mathsf{Z} : \mathsf{W})_\sigma$ denote the entanglement of formation of the pairs (X, Y) and (Z, W) with respect to the states ρ and σ, respectively.

(d) Prove a more general statement than the one required of a solution to part (c), holding not only for all LOCC channels, but for all *separable* channels of the form $\Phi \in \mathrm{SepC}(\mathcal{X}, \mathcal{Z} : \mathcal{Y}, \mathcal{W})$.

6.5. Let Σ be an alphabet, let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces of the form $\mathcal{X} = \mathbb{C}^\Sigma$ and $\mathcal{Y} = \mathbb{C}^\Sigma$, let $n = |\Sigma|$, and consider the projection operators Δ_0, Δ_1, Π_0, and Π_1 defined in Example 6.11. The states

$$\rho_0 = \frac{\Pi_0}{\binom{n+1}{2}} \quad \text{and} \quad \rho_1 = \frac{\Pi_1}{\binom{n}{2}} \tag{6.430}$$

are therefore Werner states, while

$$\sigma_0 = \Delta_0 \quad \text{and} \quad \sigma_1 = \frac{\Delta_1}{n^2 - 1} \tag{6.431}$$

are isotropic states.

(a) Prove that if $\mu : \{0, 1\} \to \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$ is a measurement satisfying $\mu(0), \mu(1) \in \mathrm{PPT}(\mathcal{X} : \mathcal{Y})$, then

$$\frac{1}{2}\langle \mu(0), \rho_0 \rangle + \frac{1}{2}\langle \mu(1), \rho_1 \rangle \leq \frac{1}{2} + \frac{1}{n+1}. \tag{6.432}$$

Prove that there exists an LOCC measurement μ for which (6.432) holds with equality.

(b) Prove that if $\nu : \{0,1\} \to \mathrm{Pos}(\mathcal{X} \otimes \mathcal{Y})$ is a measurement satisfying $\nu(0), \nu(1) \in \mathrm{PPT}(\mathcal{X} : \mathcal{Y})$, then

$$\frac{1}{2}\langle \nu(0), \sigma_0 \rangle + \frac{1}{2}\langle \nu(1), \sigma_1 \rangle \leq 1 - \frac{1}{2n+2}. \tag{6.433}$$

Prove that there exists an LOCC measurement ν for which (6.433) holds with equality.

6.6. Let N and m be positive integers, and assume that there exist unitary and Hermitian operators $U_0, \ldots, U_{2m} \in \mathrm{L}(\mathbb{C}^N)$ that anti-commute in pairs: $U_j U_k = -U_k U_j$ for distinct choices of $j, k \in \{0, \ldots, 2m\}$. Prove that the collection

$$\left\{ U_0^{a_0} \cdots U_{2m}^{a_{2m}} \,:\, a_0, \ldots, a_{2m} \in \{0,1\},\ a_0 + \cdots + a_{2m} \text{ is even} \right\} \tag{6.434}$$

is an orthogonal collection, and conclude that $N \geq 2^m$.

Observe that a correct solution to this exercise implies that the Weyl–Brauer operators have the minimum possible dimension required to possess the properties mentioned above.

6.5 Bibliographic remarks

The phenomenon of entanglement was first recognized in a 1935 paper of Einstein, Podolsky, and Rosen [68], although it was not formally defined or called entanglement therein. Einstein, Podolsky, and Rosen's work inspired Schrödinger to investigate the phenomenon of entanglement, and to give it its name; he published a three-part paper in German [180, 181, 182], as well as two related English-language papers [183, 184] discussing entanglement and other issues, as they pertained to the nature of quantum physics at that time. (An English translation of Schrödinger's three-part paper in German was published later [204].) The identification of entanglement with a lack of separability is due to Werner [225], who used the terms *classically correlated* and *EPR correlated* rather than *separable* and *entangled*.

The equivalence of the first two statements in Theorem 6.10 was proved by M. Horodecki, P. Horodecki, and R. Horodecki [116], and Proposition 6.7 was proved by P. Horodecki [119]. Several elementary analytic facts about the set of separable states that have been discussed in Section 6.1.1 were also observed in the papers proving these facts. The equivalence of the third

statement in Theorem 6.10 to the first two was proved a few years later by P. Horodecki [120]. In general, it is likely to be a computationally difficult task to test a bipartite density operator for separability, as suggested by the hardness result proved by Gurvits [82].

The fact that any operator sufficiently close to the identity operator in a bipartite tensor product space is separable was first proved by Życzkowski, P. Horodecki, Sanpera, and Lewenstein [237]. Theorem 6.14 is due to Gurvits and Barnum [83].

The local operations and classical communication paradigm, also called the *distant labs* paradigm, arose naturally in quantum information theory as various quantum information processing tasks were considered. Among the first researchers to consider this paradigm were Peres and Wootters [171], who compared the capabilities of LOCC measurements to general measurements in a setting in which information is encoded into bipartite product states. The teleportation procedure of Bennett, Brassard, Crepéau, Josza, Peres, and Wootters [31]—certainly one of the most important LOCC procedures, both historically and theoretically speaking—followed shortly after.

There are natural extensions of the definition of LOCC channels that have not been discussed in this chapter. In particular, the definition of LOCC channels in the present chapter requires an LOCC channel to be a finite composition of one-way LOCC channels, corresponding to a fixed number of classical message transmissions between two individuals implementing the channel, but one may also consider channels implemented by a potentially unbounded number of message transmissions. It is known that the set of LOCC channels, as they have been defined in this chapter, is generally not closed for a fixed choice of spaces; this was proved (for bipartite channels) by Chitambar, Leung, Mančinska, Ozols, and Winter [49]. The definition of LOCC channels presented in this chapter is based on one of the definitions considered by these authors.

The class of separable channels was identified by Vedral, Plenio, Rippin, and Knight [212], although they did not raise the possibility (first suggested by Rains [174]) that some separable channels might not be LOCC channels. The existence of separable measurements that are not LOCC measurements (and, in fact, not even approached by a sequence of LOCC measurements in the limit) was proved by Bennett, DiVincenzo, Fuchs, Mor, Rains, Shor, Smolin, and Wootters [33]. Childs, Leung, Mančinska, and Ozols [45] give a

simplified proof of this fact, along with some generalizations of it.

Bennett, Bernstein, Popescu, and Schumacher [30] defined the distillable entanglement and entanglement cost, and proved Theorem 6.42 through the design and analysis of LOCC channels for entanglement distillation and its reverse for pure states. (Bennett, Bernstein, Popescu, and Schumacher used the term *entanglement of formation* rather than entanglement cost—but that terminology has since come to refer to a related but different measure of entanglement.)

Entanglement distillation for general quantum states was considered by Bennett, Brassard, Popescu, Schumacher, Smolin, and Wootters [32] and Bennett, DiVincenzo, Smolin, and Wootters [36] around the same time. It is known that the entanglement cost of every bipartite entangled state is nonzero [233].

The entanglement rank was first defined by Terhal and P. Horodecki [200], who referred to it as the *Schmidt number* of a density operator (as it generalizes the number of nonzero terms in a Schmidt decomposition of the vector representation of a given pure state). They also proved that the entanglement rank of a state cannot increase under the action of an LOCC channel, based on related observations by Lo and Popescu [151] regarding pure states, and that it is generally not multiplicative with respect to tensor products.

Theorem 6.33 was proved by Nathanson [160], and Theorem 6.35 was proved by Walgate, Short, Hardy, and Vedral [218].

The equivalence of statements 1, 2, and 3 in Theorem 6.37, as well as statement 4 for LOCC channels rather than separable channels, was proved by Nielsen [163]. Nielsen's proof made use of the fact that every bipartite pure state transformation induced by an LOCC channel is also induced by a one-way LOCC channel, which was proved earlier by Lo and Popescu [151]. The equivalent of statement 4 of Nielsen's theorem with the first three was proved by Gheorghiu and Griffiths [77]. The proof of Theorem 6.42 concerning entanglement distillation and cost for pure states also appears in the same paper of Nielsen.

Peres [170] proposed the computationally efficient partial transpose test for separability of bipartite density operators; he observed that separable states are necessarily PPT, and that interesting families of entangled states were revealed to be entangled through this test. By the Horodecki criterion (Theorem 6.10) proved shortly after, it follows that the partial transpose test correctly identifies all entangled state in a tensor product of two complex Euclidean spaces, both of dimension 2 or one of dimension 2 and one of dimension 3, based on work of Størmer [198] and Woronowicz [232], but that entangled PPT states in higher dimensions must exist [116]. The first explicit examples of entangled PPT states were given by P. Horodecki [119]; the unextendable product set construction of such states is due to Bennett, DiVincenzo, Mor, Shor, Smolin, and Terhal [34], who introduced the notion of an unextendable product set as well as the specific example given in this

chapter. Proposition 6.49 and Theorem 6.50 were proved by M. Horodecki, P. Horodecki, and R. Horodecki [117].

The teleportation procedure described in Example 6.56 is, as mentioned above, due to Bennett, Brassard, Crepéau, Josza, Peres, and Wootters [31]. The dense coding procedure described in Example 6.61 is due to Bennett and Wiesner [40]. Various generalizations of these procedures have been discovered—the general presentation of teleportation and dense coding in this chapter is based on work of Werner [227].

The fact that entangled states may induce non-classical correlations was discovered by Bell in a highly influential 1964 paper [27]. The Bell inequality described in Example 6.67 is due to Clauser, Horn, Shimony, and Holt [52]. Some entangled states fail to induce non-classical correlations—this was demonstrated for the special case in which only projective measurements are made on the two parts of a bipartite state by Werner [225], and for general measurements by Barrett [23]. The entangled states constructed by Werner that have this property are among those described in Example 6.11. Theorem 6.68 is due to Tsirelson [205].

This chapter has presented just a small part of an extensive body of work on entanglement. Readers interested in learning more about this topic are referred to the survey of R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki [121].