Advanced quantum information: entanglement and nonlocality

Alexander Streltsov, Jędrzej Kaniewski

April 20, 2022

Contents

Index

1 Short review of quantum theory

1.1 Quantum states

Any physical system is completely described by a state vector $|\psi\rangle$ in a Hilbert space $\mathcal{H}$. A system with a two-dimensional Hilbert space is called a qubit (quantum bit). In general, we consider a Hilbert space with an arbitrary but finite dimension.

Any system which is described by a single state vector is said to be in a pure state. If the system is in the pure state $|\psi_i\rangle$ with probability p_i, the physical state of the system is described by the density matrix

$$\rho = \sum_i p_i \, |\psi_i\rangle\langle\psi_i| \, ,$$

where $|\psi_i\rangle\langle\psi_i|$ denotes projector onto the vector $|\psi_i\rangle$. If $p_{\max} < 1$, the system is in a mixed state.

Example. For $p_0 = p_1 = 1/2$ and

$$|\psi_0\rangle = |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |\psi_1\rangle = \cos\alpha \, |0\rangle + \sin\alpha \, |1\rangle = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}$$

we have the density matrix

$$\begin{aligned} \rho &= \frac{1}{2} |0\rangle\langle 0| + \frac{1}{2} |\psi_1\rangle\langle\psi_1| = \frac{1}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} \begin{pmatrix} \cos\alpha & \sin\alpha \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \cos^2\alpha & \cos\alpha\sin\alpha \\ \cos\alpha\sin\alpha & \sin^2\alpha \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 + \cos^2\alpha & \cos\alpha\sin\alpha \\ \cos\alpha\sin\alpha & \sin^2\alpha \end{pmatrix}. \end{aligned}$$

Properties of density matrices:

- ρ has trace equal to one:

  $$\mathrm{Tr}[\rho] = 1,$$

- ρ is positive semidefinite:

  $$\langle\psi|\rho|\psi\rangle \geq 0$$

  for any vector $|\psi\rangle$.

Note that the second property also implies that ρ is Hermitian: $\rho^\dagger = \rho$.

1.2 Quantum measurements and operations

According to the measurement postulate of quantum mechanics, for a spin-$\frac{1}{2}$ particle in the state

$$|\psi\rangle = a\,|\uparrow\rangle + b\,|\downarrow\rangle = \begin{pmatrix} a \\ b \end{pmatrix}$$

the probability to measure "spin up" or "spin down" is given by

$$p(\uparrow) = |a|^2\,,$$
$$p(\downarrow) = |b|^2 = 1 - p(\uparrow).$$

The post-measurement state of the particle is either $|\uparrow\rangle$ or $|\downarrow\rangle$.

Here, we consider a more general definition. A general quantum measurement is described by a collection $\{K_i\}$ of Kraus operators that fulfill the completeness equation:

$$\sum_i K_i^\dagger K_i = \mathbb{1}_d = \begin{pmatrix} 1 & 0 & & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ & & \ddots & & \\ 0 & 0 & & 1 & 0 \\ 0 & 0 & & 0 & 1 \end{pmatrix}. \tag{1.1}$$

Given a density matrix ρ and the set of Kraus operators $\{K_i\}$, the probability that the measurement outcome i occurs is given by

$$p_i = \mathrm{Tr}[K_i \rho K_i^\dagger].$$

For $p_i \neq 0$ the post-measurement state of the system is described by the density matrix

$$\rho_i = \frac{K_i \rho K_i^\dagger}{p_i}.$$

Any set of Kraus operators corresponds to a measurement, in principle realizable in laboratory. Vice versa, for any physically realizable measurement there exists a valid set of Kraus operators.

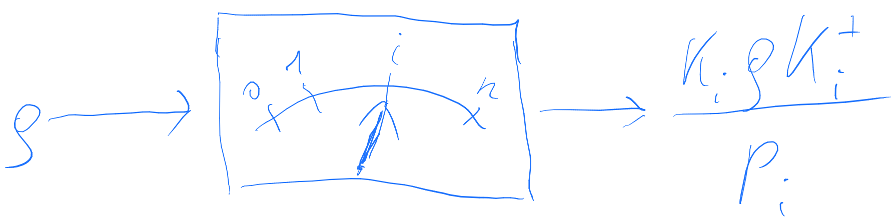

Figure 1.1: General quantum measurement

The set of operators

$$M_i = K_i^\dagger K_i$$

is called positive operator-valued measure (POVM). The completeness condition (1.1) implies $\sum_i M_i = \mathbb{1}_d$, and the probabilities of the outcome i is $p_i = \mathrm{Tr}[M_i\rho]$. For a projective measurement, the operators K_i are orthogonal projectors: $K_iK_j = \delta_{ij}K_i$. If K_i are orthogonal projectors with rank one, we have a von Neumann measurement.

Any set of Kraus operators $\{K_i\}$ also defines a quantum operation:

$$\Lambda(\rho) = \sum_i K_i \rho K_i^\dagger.$$

Quantum operations describe the most general change of a quantum state in a physical process. They correspond to a special class of linear maps, which are completely positive and trace preserving (CPTP).

1.3 Composite systems

For two parties, Alice (A) and Bob (B), with Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ the total Hilbert space is a tensor product of the subsystem spaces: $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$.

Example. Consider the states

$$|\psi\rangle^A = \cos\alpha\,|0\rangle + \sin\alpha\,|1\rangle = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}, \qquad |\psi\rangle^B = \cos\beta\,|0\rangle + \sin\beta\,|1\rangle = \begin{pmatrix} \cos\beta \\ \sin\beta \end{pmatrix}.$$

The state of the total system is

$$|\psi\rangle^{AB} = |\psi\rangle^A \otimes |\psi\rangle^B = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix} \otimes \begin{pmatrix} \cos\beta \\ \sin\beta \end{pmatrix} = \begin{pmatrix} \cos\alpha\cos\beta \\ \cos\alpha\sin\beta \\ \sin\alpha\cos\beta \\ \sin\alpha\sin\beta \end{pmatrix}.$$

If $\{|i\rangle\}$ and $\{|k\rangle\}$ are orthonormal bases of $\mathcal{H}_A$ and $\mathcal{H}_B$, then $\{|i\rangle \otimes |k\rangle\}$ is an orthonormal basis of $\mathcal{H}_{AB}$. We can expand any pure state as

$$|\psi\rangle^{AB} = \sum_{i,k} c_{ik}\,|i\rangle \otimes |k\rangle\,.$$

with $c_{ik} \in \mathbb{C}$. Any density matrix can be expanded as

$$\rho^{AB} = \sum_{i,j,k,l} c_{ijkl}\,|i\rangle\langle j| \otimes |k\rangle\langle l|$$

with $c_{ijkl} \in \mathbb{C}$. The subsystem A is described by the reduced density matrix

$$\rho^A = \mathrm{Tr}_B[\rho^{AB}] = \sum_{i,j,k,l} c_{ijkl} |i\rangle\langle j| \, \mathrm{Tr}\,[|k\rangle\langle l|] = \sum_{i,j,k,l} c_{ijkl} |i\rangle\langle j| \, \delta_{kl} = \sum_{i,j,k} c_{ijkk} |i\rangle\langle j| \,, \tag{1.2}$$

where Tr_B is the partial trace over the subsystem B.

Example. Consider the density matrix

$$\rho^{AB} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} X & Y \\ Y^\dagger & Z \end{pmatrix}$$

with matrices $X = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{pmatrix}$, $Y = \begin{pmatrix} 0 & \frac{1}{2} \\ 0 & 0 \end{pmatrix}$ and $Z = \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$.
The reduced density matrices are

$$\rho^A = \begin{pmatrix} \mathrm{Tr}\,[X] & \mathrm{Tr}\,[Y] \\ \mathrm{Tr}\left[Y^\dagger\right] & \mathrm{Tr}\,[Z] \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$
$$\rho^B = X + Z = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

For any pure state $|\psi\rangle^{AB}$ there exists a product basis $\{|i\rangle \otimes |j\rangle\}$ such that

$$|\psi\rangle^{AB} = \sum_i \sqrt{\lambda_i}\, |i\rangle \otimes |i\rangle \tag{1.3}$$

with $\lambda_i \geq 0$. This is called Schmidt decomposition of $|\psi\rangle^{AB}$. The numbers λ_i are called Schmidt coefficients of $|\psi\rangle^{AB}$. The Schmidt coefficients are equal to the eigenvalues of the reduced states $\mathrm{Tr}_A[|\psi\rangle\langle\psi|^{AB}]$ and $\mathrm{Tr}_B[|\psi\rangle\langle\psi|^{AB}]$.

For composite systems it is possible to perform local measurements on one of the subsystems. Kraus operators of local measurements on Alice's side have the form $K_i^{AB} = K_i \otimes \mathbb{1}$ with the completeness condition

$$\sum_i \left(K_i^{AB}\right)^\dagger K_i^{AB} = \sum_i K_i^\dagger K_i \otimes \mathbb{1} = \mathbb{1}_{AB}.$$

Local quantum operations on Alice's side are defined as

$$\Lambda^A(\rho^{AB}) = \sum_i (K_i \otimes \mathbb{1})\, \rho^{AB}\, (K_i \otimes \mathbb{1})^\dagger \,.$$

The state of Bob does not change upon local operations of Alice:

$$\rho^B = \mathrm{Tr}_A\left[\rho^{AB}\right] = \mathrm{Tr}_A\left[\Lambda^A(\rho^{AB})\right].$$

Purification: A pure state $|\psi\rangle^{AB}$ is called a purification of a mixed state ρ^A if

$$\rho^A = \mathrm{Tr}_B[|\psi\rangle\langle\psi|^{AB}].$$

Two states $|\psi\rangle^{AB}$ and $|\phi\rangle^{AB}$ are purifications of the same state ρ^A if and only if

$$|\psi\rangle^{AB} = (\mathbb{1} \otimes U)\,|\phi\rangle^{AB}$$

for some local unitary U.

Useful properties of square matrices

Functions of matrices: Let f be a function from $\mathbb{C}$ to $\mathbb{C}$. For a normal (diogonalizable) matrix $A = \sum_i a_i\,|\psi_i\rangle\langle\psi_i|$ with eigenvalues $a_i \in \mathbb{C}$ and eigenstates $|\psi_i\rangle$ we define

$$f(A) := \sum_i f(a_i)\,|\psi_i\rangle\langle\psi_i|\,.$$

Polar decomposition: For any square matrix A there exist unitary matrices U and V such that

$$A = U\sqrt{A^\dagger A} = \sqrt{AA^\dagger}V.$$

Every Hermitian matrix H can be decomposed into a positive and negative part $H = P_+ - P_-$ with positive matrices $P_\pm$. Moreover, P_+ and P_- are supported on orthogonal subspaces, such that $\mathrm{Tr}\,[P_+P_-] = 0$.

2 Theory of quantum entanglement

2.1 Definition

If there are states $|a\rangle \in \mathcal{H}_A$ and $|b\rangle \in \mathcal{H}_B$ such that

$$|\psi\rangle^{AB} = |a\rangle \otimes |b\rangle \,,$$

then $|\psi\rangle^{AB}$ is called separable (or product state). Otherwise the state is called entangled. $|\psi\rangle^{AB}$ is product if and only if ρ^A is pure.

Notation: For product states $|i\rangle \otimes |j\rangle$ we sometimes write $|i\rangle\,|j\rangle$ or $|ij\rangle$.

Example. $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ is entangled since $\rho^A = \frac{1}{2}\mathbb{1}_2$.

2.2 Local operations and classical communication (LOCC)

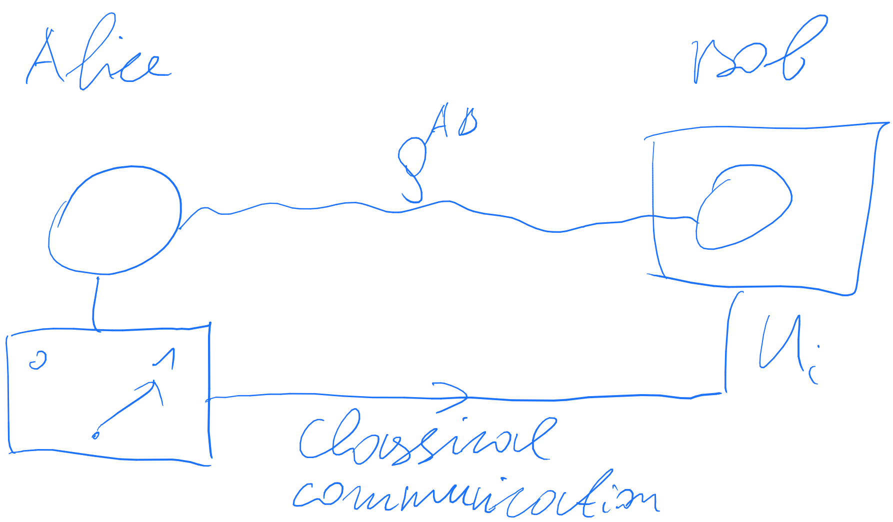

Figure 2.1: Local operations and classical communication

LOCC describes the most general procedure Alice and Bob can apply, if they can perform arbitrary quantum measurements/operations locally, and exchange classical information. Any LOCC protocol can be decomposed into the following steps:

1. Alice performs a local measurement $\{K_i\}$ on her subsystem.
2. The outcome i of Alice's measurement is communicated to Bob via a classical channel.
3. Bob performs a local measurement $\{L_j(i)\}$ on his subsystem, which depends on Alice's outcome i.
4. The outcome j of Bob's measurement is communicated classically to Alice.
5. Alice performs a local measurement on her subsystem which can depend on all outcomes of all previous measurements, and the process starts over at step 2.

2.3 Pure state conversion via LOCC

Assume that Alice and Bob share the state $|\psi\rangle^{AB}$. Which other states $|\phi\rangle^{AB}$ can be obtained via LOCC?

Proposition 2.1. *Suppose $|\psi\rangle^{AB}$ can be transformed into $|\phi\rangle^{AB}$ via LOCC. Then this transformation can be achieved by a protocol involving just the following steps: Alice performs a measurement with Kraus operators $\{K_j\}$, sends the result j to Bob, who applies a conditional unitary U_j on his system.*

Proof. Let $K_j = \sum_{k,l} K_{j,kl} |k\rangle\langle l|$ be a Kraus operator of Bob expanded in the Schmidt basis of $|\psi\rangle = \sum_i \sqrt{\lambda_i}\,|i\rangle \otimes |i\rangle$. The post-measurement state $|\mu_j\rangle$ is given as

$$|\mu_j\rangle = \frac{\mathbb{1} \otimes K_j\,|\psi\rangle}{\sqrt{p_j}} = \frac{\sum_{k,l} K_{j,kl}\sqrt{\lambda_l}\,|l\rangle \otimes |k\rangle}{\sqrt{p_j}}$$

with probability

$$p_j = \langle\psi|\mathbb{1} \otimes K_j^\dagger K_j|\psi\rangle = \sum_{k,l} \lambda_l |K_{j,kl}|^2.$$

Assume now that instead Alice performs a measurement with Kraus operator $L_j = \sum_{k,l} K_{j,kl}\,|k\rangle\langle l|$, leading to the state

$$|\nu_j\rangle = \frac{L_j \otimes \mathbb{1}\,|\psi\rangle}{\sqrt{p_j}} = \frac{\sum_{k,l} K_{j,kl}\sqrt{\lambda_l}\,|k\rangle \otimes |l\rangle}{\sqrt{p_j}}$$

with the same probability p_j. Note that $|\mu_j\rangle$ and $|\nu_j\rangle$ are the same up to interchanging A and B, which by Schmidt decomposition implies that

$$|\mu_j\rangle = \sum_i \sqrt{\alpha_{ij}} \left(U_j |i\rangle\right) \otimes \left(V_j |i\rangle\right),$$
$$|\nu_j\rangle = \sum_i \sqrt{\alpha_{ij}} \left(V_j |i\rangle\right) \otimes \left(U_j |i\rangle\right)$$

for some $\alpha_{ij} \geq 0$ and local unitaries U_j and V_j, and thus

$$|\mu_j\rangle = (U_j V_j^\dagger \otimes V_j U_j^\dagger) |\nu_j\rangle .$$

Thus, Bob performing a measurement $\{K_j\}$ on $|\psi\rangle$ is equivalent to Alice performing a measurement $\{U_j V_j^\dagger L_j\}$, followed by Bob performing the unitary $V_j U_j^\dagger$.

A measurement by Bob on a pure state can be simulated by a measurement by Alice, and a conditional unitary by Bob. If Alice and Bob perform an LOCC protocol consisting of many rounds of measurements and classical communication, we replace each round involving Bob's measurement by a corresponding measurement on Alice's side. In this way, any LOCC protocol transforming $|\psi\rangle^{AB}$ into $|\phi\rangle^{AB}$ can be simulated by a single measurement of Alice, followed by conditional unitary on Bob's side. □

Majorization: Consider two real d-dimensional vectors $\vec{x}$ and $\vec{y}$ with elements in decreasing order. Then $\vec{x} \prec \vec{y}$ if

$$\sum_{i=1}^{k} x_i \leq \sum_{i=1}^{k} y_i$$

for all $k \in [1, d-1]$, and $\sum_{i=1}^{d} x_i = \sum_{i=1}^{d} y_i$. For a Hermitian matrix H let $\vec{\lambda}_H$ be the vector of eigenvalues of H in decreasing order. For two Hermitian matrices H and K we write $H \prec K$ if $\vec{\lambda}_H \prec \vec{\lambda}_K$.

Proposition 2.2. *Let H and K be Hermitian matrices. Then $H \prec K$ if and only if there is a probability distribution p_j and unitary matrices U_j such that*

$$H = \sum_j p_j U_j K U_j^\dagger.$$

For a given state $|\psi\rangle^{AB}$, let $\vec{\lambda}_\psi$ denote the vector with eigenvalues of the reduced state $\mathrm{Tr}_B[|\psi\rangle\langle\psi|^{AB}]$ in decreasing order. Equipped with these tools, we can provide a complete characterization of LOCC transformations between bipartite pure states in the following theorem, which is also called Nielsen's theorem.

Theorem 2.1. *There exists an LOCC protocol transforming* $|\psi\rangle^{AB}$ *into* $|\phi\rangle^{AB}$ *if and only if* $\vec{\lambda}_\psi \prec \vec{\lambda}_\phi$.

Proof. Suppose $|\psi\rangle^{AB}$ can be transformed into $|\phi\rangle^{AB}$ via LOCC. By proposition 2.1, the transformation is achieved if Alice applies a measurement with local Kraus operators $\{K_j\}$ and Bob applies local unitaries $\{U_j\}$. After Alice's measurement, the total post-measurement state is equal to $|\phi\rangle^{AB}$ up to local unitaries on Bob's side:

$$K_j \otimes \mathbb{1} |\psi\rangle^{AB} = \sqrt{p_j} \mathbb{1} \otimes U_j^\dagger |\phi\rangle^{AB} .$$

Defining $\rho_\psi = \mathrm{Tr}_B[|\psi\rangle\langle\psi|^{AB}]$ and $\rho_\phi = \mathrm{Tr}_B[|\phi\rangle\langle\phi|^{AB}]$, we get

$$K_j \rho_\psi K_j^\dagger = p_j \rho_\phi$$

with $p_j = \mathrm{Tr}[K_j \rho_\psi K_j^\dagger]$. By polar decomposition there exists a unitary V_j such that

$$K_j \sqrt{\rho_\psi} = \sqrt{K_j \rho_\psi K_j^\dagger} V_j = \sqrt{p_j \rho_\phi} V_j.$$

Multiplying this equation with its adjoint from the left, we get

$$\sqrt{\rho_\psi} K_j^\dagger K_j \sqrt{\rho_\psi} = p_j V_j^\dagger \rho_\phi V_j.$$

Taking sum over j and using $\sum_j K_j^\dagger K_j = \mathbb{1}$ we obtain

$$\rho_\psi = \sum_j p_j V_j^\dagger \rho_\phi V_j,$$

and by proposition 2.2 we have $\vec{\lambda}_\psi \prec \vec{\lambda}_\phi$.

Suppose that $\vec{\lambda}_\psi \prec \vec{\lambda}_\phi$, and thus $\rho_\psi \prec \rho_\phi$. By proposition 2.2

$$\rho_\psi = \sum_j p_j U_j \rho_\phi U_j^\dagger$$

for some probabilities p_j and unitaries U_j. If ρ_ψ is invertible, we define

$$K_j := \sqrt{p_j \rho_\phi} U_j^\dagger \rho_\psi^{-1/2}.$$

It holds that

$$\sum_j K_j^\dagger K_j = \rho_\psi^{-1/2} \left(\sum_j p_j U_j \rho_\phi U_j^\dagger \right) \rho_\psi^{-1/2} = \rho_\psi^{-1/2} \rho_\psi \rho_\psi^{-1/2} = \mathbb{1},$$

thus K_j are valid Kraus operators. Suppose Alice performs the measurement $\{K_j\}$, it follows

$$K_j \rho_\psi K_j^\dagger = p_j \rho_\phi.$$

When Alice applies the measurement $\{K_j\}$ to the total state $|\psi\rangle^{AB}$, she obtains the reduced state ρ_ϕ with probability p_j. Since all purifications of ρ_ϕ are equivalent up to unitary on Bob's side (see Section 1.3), it follows that there exist unitaries U_j on Bob's side such that

$$K_j \otimes \mathbb{1} |\psi\rangle^{AB} = \sqrt{p_j} \mathbb{1} \otimes U_j |\phi\rangle^{AB}.$$

Thus, if Alice applies measurement $\{K_j\}$ to the state $|\psi\rangle^{AB}$, communicates the measurement outcome j to Bob, and he performs $U_j^\dagger$, they achieve the conversion $|\psi\rangle^{AB} \to |\phi\rangle^{AB}$. □

2.4 Probabilistic conversion and catalysis

If there is no LOCC protocol converting $|\psi\rangle^{AB}$ into $|\phi\rangle^{AB}$, there might still be a chance to perform probabilistic conversion. Here, Alice and Bob are allowed to post-select the outcomes of their local measurements, leading to a conversion $|\psi\rangle^{AB} \to |\phi\rangle^{AB}$ with probability p. For general density matrices ρ^{AB} and σ^{AB} the optimal probability can be defined as

$$P\left(\rho^{AB} \to \sigma^{AB}\right) = \max_{\{K_i\}} \left\{ \mathrm{Tr}\left[\sum_i K_i \rho^{AB} K_i^\dagger\right] : \sigma^{AB} = \frac{\sum_i K_i \rho^{AB} K_i^\dagger}{\mathrm{Tr}\left[\sum_i K_i \rho^{AB} K_i^\dagger\right]} \right\},$$

and the maximum is taken over all (incomplete) sets of Kraus operators $\{K_i\}$ which are implementable via LOCC. For bipartite pure states $|\psi\rangle^{AB}$ and $|\phi\rangle^{AB}$ the maximal conversion probability can be evaluated as

$$P\left(|\psi\rangle^{AB} \to |\phi\rangle^{AB}\right) = \min_{l \in [1,n]} \frac{\sum_{i=l}^n \alpha_i}{\sum_{j=l}^n \beta_j},$$

where α_i and β_j are the Schmidt coefficients of $|\psi\rangle^{AB}$ and $|\phi\rangle^{AB}$, respectively, sorted in decreasing order.

A catalytic conversion between the states $|\psi\rangle^{AB}$ and $|\phi\rangle^{AB}$ is possible if there exists an additional state $|c\rangle^{A'B'}$ and an LOCC protocol converting $|\psi\rangle^{AB} \otimes |c\rangle^{A'B'}$ into $|\phi\rangle^{AB} \otimes |c\rangle^{A'B'}$.

2.5 Bell states

In the Hilbert space of two qubits the following four states form an orthonormal basis:

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle),$$
$$|\Phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle),$$
$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle),$$
$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle).$$

These states are called Bell states (or EPR states). The state $|\Psi^-\rangle$ is also called singlet state. The reduced state of any Bell state is $\frac{1}{2}\mathbb{1}_2$, and for any single-qubit state ρ it holds $\frac{1}{2}\mathbb{1}_2 \prec \rho$. With Theorem 2.1 it follows that any Bell state can be converted into any two-qubit pure state via LOCC. Bell states are also called maximally entangled states (of two qubits).

For $d_A = d_B = d$, a quantum state $|\Psi_d\rangle$ is maximally entangled if and only if

$$\mathrm{Tr}_A\left[|\Psi_d\rangle\langle\Psi_d|\right] = \frac{1}{d}\mathbb{1}_d.$$

All maximally entangled states are equivalent to

$$|\Phi_d^+\rangle = \frac{1}{\sqrt{d}}\sum_{i=0}^{d-1}|ii\rangle$$

up to local unitary on one side: there exist unitaries U and V such that

$$|\Psi_d\rangle = (U \otimes \mathbb{1})\,|\Phi_d^+\rangle = (\mathbb{1} \otimes V)\,|\Phi_d^+\rangle$$

for any maximally entangled state $|\Psi_d\rangle$.

2.6 Entanglement for mixed states

A bipartite mixed state is separable if it can be written as:

$$\rho_{\mathrm{sep}}^{AB} = \sum_i p_i\,|\psi_i\rangle\langle\psi_i| \otimes |\phi_i\rangle\langle\phi_i|$$

with $p_i \geq 0$, $\sum_i p_i = 1$, $|\psi_i\rangle \in \mathcal{H}_A$ and $|\phi_i\rangle \in \mathcal{H}_B$. If the state cannot be written in this form, it is called entangled. Separable states form a convex subset in the set of all quantum states.

Any separable state can be produced by LOCC from an initial product state $|00\rangle$. No entangled state can be produced by LOCC.

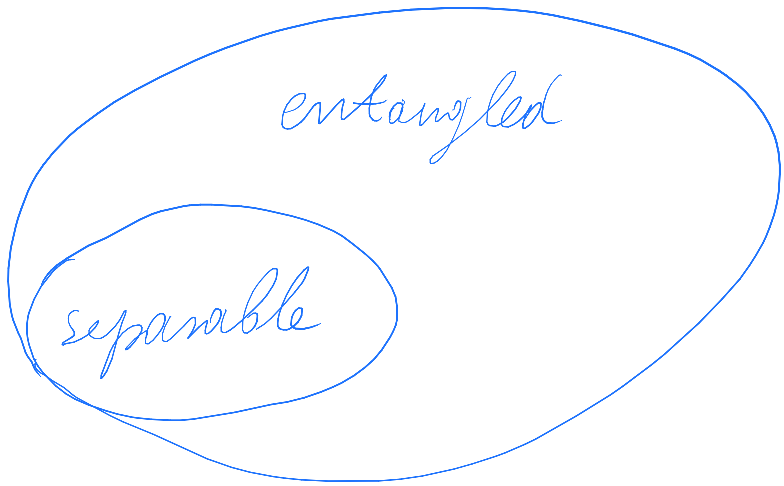

Figure 2.2: Separable states are a convex subset in the set of all states.

3 Entanglement detection

Literature: Horodecki *et al.*, Rev. Mod. Phys. **81**, 865 (2009)

3.1 Entanglement witnesses

Let W^{AB} be a Hermitian matrix such that

$$\mathrm{Tr}\left[W^{AB}\left(|\psi\rangle\langle\psi|\otimes|\phi\rangle\langle\phi|\right)\right]=\left(\langle\psi|\otimes\langle\phi|\right)W^{AB}\left(|\psi\rangle\otimes|\phi\rangle\right)\geq 0$$

for any $|\psi\rangle\in\mathcal{H}_A$ and $|\phi\rangle\in\mathcal{H}_B$. Then, for any separable state ρ_{sep}^{AB} we have

$$\mathrm{Tr}\left[W^{AB}\rho_{\mathrm{sep}}^{AB}\right]=\sum_i p_i\,\mathrm{Tr}\left[W^{AB}\left(|\psi_i\rangle\langle\psi_i|\otimes|\phi_i\rangle\langle\phi_i|\right)\right]\geq 0.$$

Thus, if

$$\mathrm{Tr}\left[W^{AB}\rho^{AB}\right]<0,$$

the state ρ^{AB} must be entangled. The matrix W^{AB} is called entanglement witness. From the Hahn-Banach theorem follows

Theorem 3.1. *For any entangled state ρ^{AB} there exists an entanglement witness such that* $\mathrm{Tr}\left[W^{AB}\rho^{AB}\right]<0$.

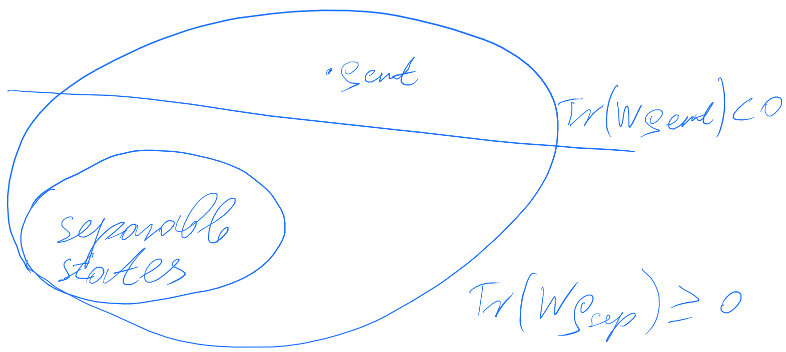

Figure 3.1: Entanglement witness.

An entanglement witness can be interpreted as an observable with expectation value $\mathrm{Tr}\left[W^{AB}\rho^{AB}\right]$.

Example. For $d_A = d_B$ the swap operation is an entanglement witness:

$$W^{AB} = \sum_{i,j=0}^{d-1} |i\rangle\langle j| \otimes |j\rangle\langle i| \,.$$

For any product state $|\psi\rangle \otimes |\phi\rangle$ we find that $W^{AB} |\psi\rangle \otimes |\phi\rangle = |\phi\rangle \otimes |\psi\rangle$, and thus

$$\big(\langle\psi| \otimes \langle\phi|\big)\, W^{AB} \big(|\psi\rangle \otimes |\phi\rangle\big) = \big(\langle\psi| \otimes \langle\phi|\big)\big(|\phi\rangle \otimes |\psi\rangle\big) = |\langle\psi|\phi\rangle|^2 \geq 0.$$

W^{AB} has negative eigenvalues:

$$W^{AB} |\Psi^-\rangle = \frac{1}{\sqrt{2}}\big(W^{AB} |01\rangle - W^{AB} |10\rangle\big) = -\,|\Psi^-\rangle\,,$$

thus W^{AB} detects entanglement in the state $|\Psi^-\rangle$.

3.2 Partial transposition

For a bipartite state $\rho = \sum_{i,j,k,l} c_{ijkl} |i\rangle\langle j| \otimes |k\rangle\langle l|$ the partial transposition on Bob's subsystem is defined as

$$\rho^{T_B} = \sum_{i,j,k,l} c_{ijkl} |i\rangle\langle j| \otimes (|k\rangle\langle l|)^T = \sum_{i,j,k,l} c_{ijkl} |i\rangle\langle j| \otimes |l\rangle\langle k| \,.$$

Note that ρ^{T_A} and ρ^{T_B} have the same eigenvalues.

Applying partial transposition to a separable state leads to another quantum state:

$$\rho_{\text{sep}}^{T_B} = \sum_i p_i |\psi_i\rangle\langle\psi_i| \otimes (|\phi_i\rangle\langle\phi_i|)^T = \sum_i p_i |\psi_i\rangle\langle\psi_i| \otimes |\phi_i^*\rangle\langle\phi_i^*| \,.$$

Thus, if ρ^{T_B} is not positive, ρ must be an entangled state.

Example. For the state $|\psi\rangle = \cos\alpha\,|00\rangle + \sin\alpha\,|11\rangle$ we have

$$\rho = |\psi\rangle\langle\psi| = \begin{pmatrix} \cos^2\alpha & 0 & 0 & \cos\alpha\sin\alpha \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \cos\alpha\sin\alpha & 0 & 0 & \sin^2\alpha \end{pmatrix} = \begin{pmatrix} X & Y \\ Y^\dagger & Z \end{pmatrix}$$

with $X = \begin{pmatrix} \cos^2\alpha & 0 \\ 0 & 0 \end{pmatrix}$, $Y = \begin{pmatrix} 0 & \cos\alpha\sin\alpha \\ 0 & 0 \end{pmatrix}$ and $Z = \begin{pmatrix} 0 & 0 \\ 0 & \sin^2\alpha \end{pmatrix}$.
We obtain

$$\rho^{T_A} = \begin{pmatrix} X^T & Y^T \\ (Y^\dagger)^T & Z^T \end{pmatrix} = \begin{pmatrix} \cos^2\alpha & 0 & 0 & 0 \\ 0 & 0 & \cos\alpha\sin\alpha & 0 \\ 0 & \cos\alpha\sin\alpha & 0 & 0 \\ 0 & 0 & 0 & \sin^2\alpha \end{pmatrix},$$

$$\rho^{T_B} = \begin{pmatrix} X & Y^\dagger \\ Y & Z \end{pmatrix} = \rho^{T_A}.$$

The eigenvalues of ρ^{T_A} are $\cos^2\alpha$, $\sin^2\alpha$, $\pm|\cos\alpha\sin\alpha|$, thus $|\psi\rangle$ is entangled for all $\alpha \neq n\frac{\pi}{2}$. In general $\rho^{T_A} \neq \rho^{T_B}$.

Choi–Jamiołkowski isomorphism, positive, and completely positive maps. A positive map is a linear map Λ acting on matrices such that $\Lambda(\rho)$ is positive semidefinite for any positive semidefinite matrix ρ. For a bipartite density matrix $\rho^{AB} = \sum_{i,j,k,l} c_{ijkl}\,|i\rangle\langle j| \otimes |k\rangle\langle l|$ we define $\mathbb{1} \otimes \Lambda(\rho^{AB})$ as

$$\mathbb{1} \otimes \Lambda(\rho^{AB}) = \sum_{i,j,k,l} c_{ijkl}\,|i\rangle\langle j| \otimes \Lambda(|k\rangle\langle l|).$$

The map Λ is completely positive (CP) if $\mathbb{1} \otimes \Lambda(\rho^{AB})$ is positive semidefinite for any positive semidefinite matrix ρ^{AB} in the extended Hilbert space of any dimension. Every quantum operation is CP (see section 1.2). Every CP map is positive, but there are positive maps which are not CP (e.g. transpose).

For a linear map Λ acting on Hilbert space of dimension d, the Choi matrix is defined as

$$M_\Lambda = (\mathbb{1} \otimes \Lambda)\,|\Phi_d^+\rangle\langle\Phi_d^+|\,.$$

A linear map Λ is positive if and only if M_Λ is an entanglement witness. Moreover, for any entanglement witness W^{AB} with $d_A = d_B$ there exists a positive map Λ such that $W^{AB} = M_\Lambda$. The map Λ is CP if and only if M_Λ is positive semidefinite.

Proposition 3.1. *For $d_A = d_B = 2$ a state ρ^{AB} is separable if and only if ρ^{T_B} is positive semidefinite.*

Proof. For any entangled state ρ^{AB} there exists an entanglement witness W^{AB} such that (see Section 3.1)

$$\mathrm{Tr}\left[W^{AB}\rho^{AB}\right] < 0.$$

With the Choi-Jamiołkowski isomorphism, there also exists a positive map Λ such that

$$\mathrm{Tr}\left[(\mathbb{1} \otimes \Lambda \,|\Phi^+\rangle\langle\Phi^+|)\, \rho^{AB}\right] < 0.$$

Every positive qubit map can be decomposed as

$$\Lambda(\rho) = \Lambda_1^{\mathrm{CP}}(\rho) + \left[\Lambda_2^{\mathrm{CP}}(\rho)\right]^T,$$

with CP maps Λ_i^{CP}, and thus

$$\begin{aligned} 0 > \mathrm{Tr}\left[(\mathbb{1} \otimes \Lambda \,|\Phi^+\rangle\langle\Phi^+|)\, \rho^{AB}\right] &= \mathrm{Tr}\left[\left(\mathbb{1} \otimes \Lambda_1^{\mathrm{CP}} \,|\Phi^+\rangle\langle\Phi^+|\right) \rho^{AB}\right] \\ &+ \mathrm{Tr}\left[\left(\mathbb{1} \otimes \Lambda_2^{\mathrm{CP}} \,|\Phi^+\rangle\langle\Phi^+|\right)^{T_B} \rho^{AB}\right] \\ &= \mathrm{Tr}\left[X_1\rho^{AB}\right] + \mathrm{Tr}\left[X_2^{T_B}\rho^{AB}\right] \end{aligned}$$

with positive matrices $X_i = \mathbb{1} \otimes \Lambda_i^{\mathrm{CP}} \,|\Phi^+\rangle\langle\Phi^+|$. Using

$$\mathrm{Tr}\left[X_2^{T_B}\rho^{AB}\right] = \mathrm{Tr}\left[X_2\rho^{T_B}\right],$$

we obtain

$$0 > \mathrm{Tr}\left[X_1\rho^{AB}\right] + \mathrm{Tr}\left[X_2\rho^{T_B}\right] \geq \mathrm{Tr}\left[X_2\rho^{T_B}\right].$$

Since X_2 is positive, ρ^{T_B} must have negative eigenvalues. □

This result is called positive partial transpose (PPT) criterion, which extends to larger dimensions as follows.

Theorem 3.2. *For $d_A d_B \leq 6$ a state ρ^{AB} is separable if and only if ρ^{T_B} is positive. For all $d_A d_B > 6$ there exist entangled states which have positive partial transpose.*

4 Applications of entanglement

4.1 Quantum teleportation

Suppose Alice and Bob share a Bell state $|\Phi^+\rangle^{AB}$. Additionally, Alice has a qubit A' in the state $|\psi\rangle^{A'} = c_0|0\rangle + c_1|1\rangle$. Alice can send the qubit A' to Bob by using quantum teleportation, see also Fig. 4.1.

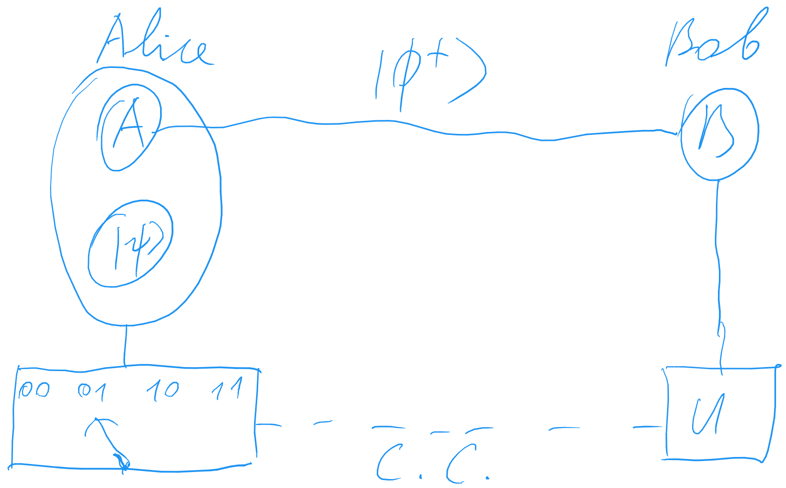

Figure 4.1: Quantum teleportation.

The total initial state of Alice and Bob has the form

$$
\begin{aligned}
|\Phi\rangle^{A'AB} &= \left(c_0|0\rangle^{A'} + c_1|1\rangle^{A'}\right) \otimes \frac{1}{\sqrt{2}}\left(|00\rangle^{AB} + |11\rangle^{AB}\right) \\
&= \frac{1}{\sqrt{2}}\left[c_0|0\rangle\left(|00\rangle + |11\rangle\right) + c_1|1\rangle\left(|00\rangle + |11\rangle\right)\right].
\end{aligned}
$$

Controlled NOT gate (CNOT): A unitary transformation acting on two qubits (control and target) as follows

Before		**After**	
Control	Target	Control	Target
$\|0\rangle$	$\|0\rangle$	$\|0\rangle$	$\|0\rangle$
$\|0\rangle$	$\|1\rangle$	$\|0\rangle$	$\|1\rangle$
$\|1\rangle$	$\|0\rangle$	$\|1\rangle$	$\|1\rangle$
$\|1\rangle$	$\|1\rangle$	$\|1\rangle$	$\|0\rangle$

Hadamard gate: A unitary transformation on one qubit acting as follows

$$|0\rangle \to \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle),$$
$$|1\rangle \to \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle).$$

In the next step, Alice performs a CNOT gate on her qubits $A'A$, where A' is the control qubit and A is the target. This leads to

$$|\Phi'\rangle = \frac{1}{\sqrt{2}}\left[c_0\,|0\rangle\,(|00\rangle + |11\rangle) + c_1\,|1\rangle\,(|10\rangle + |01\rangle)\right].$$

Finally, Alice applies a Hadamard gate to A':

$$\begin{aligned}|\Phi''\rangle &= \frac{1}{2}\left[c_0\,(|0\rangle + |1\rangle)\,(|00\rangle + |11\rangle) + c_1\,(|0\rangle - |1\rangle)\,(|10\rangle + |01\rangle)\right]\\ &= \frac{1}{2}\left[|00\rangle\,(c_0\,|0\rangle + c_1\,|1\rangle) + |01\rangle\,(c_0\,|1\rangle + c_1\,|0\rangle) + |10\rangle\,(c_0\,|0\rangle - c_1\,|1\rangle) + |11\rangle\,(c_0\,|1\rangle - c_1\,|0\rangle)\right].\end{aligned}$$

Alice measures her qubits A' and A in the computational basis $\{|0\rangle, |1\rangle\}$. Depending on the outcome of her measurement, the state of Bob's qubit B collapses to one of the following states:

Alice's outcome	State of B
00	$c_0\,\|0\rangle + c_1\,\|1\rangle$
01	$c_0\,\|1\rangle + c_1\,\|0\rangle$
10	$c_0\,\|0\rangle - c_1\,\|1\rangle$
11	$c_0\,\|1\rangle - c_1\,\|0\rangle$

Bob performs a correction on his qubit depending on Alice's measurement according to the following table (σ_i are Pauli matrices) outcome:

Alice's outcome	State of B	Correction	State of B after correction
00	$c_0 \|0\rangle + c_1 \|1\rangle$	$\mathbb{1}$	$c_0 \|0\rangle + c_1 \|1\rangle$
01	$c_0 \|1\rangle + c_1 \|0\rangle$	σ_x	$c_0 \|0\rangle + c_1 \|1\rangle$
10	$c_0 \|0\rangle - c_1 \|1\rangle$	σ_z	$c_0 \|0\rangle + c_1 \|1\rangle$
11	$c_0 \|1\rangle - c_1 \|0\rangle$	$i\sigma_y$	$c_0 \|0\rangle + c_1 \|1\rangle$

In the end of the protocol Bob's qubit B is in the state $|\psi\rangle^B = c_0 |0\rangle + c_1 |1\rangle$, which is the initial state of Alice's qubit A'.

The protocol does not depend on the state to be teleported. Also, the Bell state $|\Phi^+\rangle^{AB}$ is destroyed in this procedure, thus teleportation of one qubit consumes one Bell state.

Quantum teleportation can also be applied to teleport a part of Alice's subsystem, see also Fig. 4.2. If Alice is in possession of two qubits C and D in a quantum state $|\psi\rangle^{CD}$, she can teleport the particle D to Bob by using a Bell state. In this way, Alice and Bob will end up sharing the two-qubit state $|\psi\rangle$ which was initially in Alice's laboratory.

Systems of larger dimension $d > 2$ can be teleported as follows: if $d = 2^n$ for some integer n, then the particle A' (which is the particle Alice wants to teleport) can be regarded as an n-qubit system: $A' = A'_1 A'_2 \ldots A'_n$. Teleportation of A' can then be achieved by teleporting each of the qubits A'_i, thus consuming $n = \log_2 d$ Bell states. If $\log_2 d$ is not an integer, we define $n = \lceil \log_2 d \rceil$. The state $|\psi\rangle^{A'}$ can then be written as

$$|\psi\rangle^{A'} = \sum_{i=0}^{2^n-1} c_i |i\rangle^{A'},$$

and all coefficients c_i are zero for $i \geq d$. Thus, also in this case we can regard A' as being composed of n qubits, and teleport each of the qubits individually.

Similarly, if Alice is in possession of a bipartite state $|\psi\rangle^{CD}$, where C and D are now general quantum systems of arbitrary (but finite) dimension, Alice can teleport the particle D to Bob. The number of Bell states required for this procedure can be determined by the following proposition.

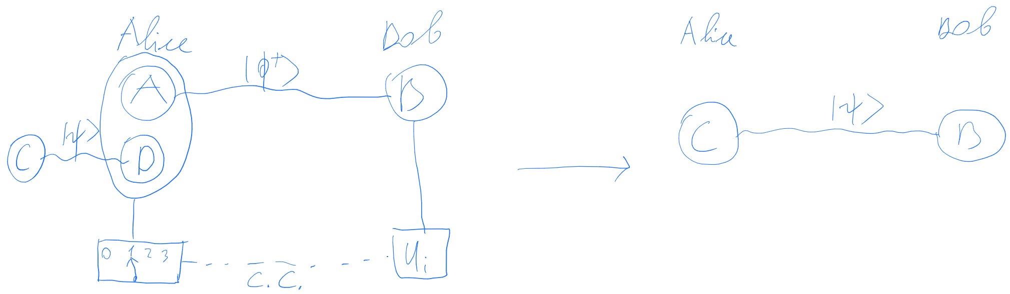

Figure 4.2: Quantum teleportation of a part of a quantum system.

Proposition 4.1. *For a state*

$$|\psi\rangle^{CD} = \sum_{i=0}^{k-1} \sqrt{\lambda_i}\, |i\rangle^C \otimes |i\rangle^D$$

with k nonzero Schmidt coefficients the teleportation of D can be done by consuming $\lceil \log_2 k \rceil$ Bell states.

4.2 Superdense coding

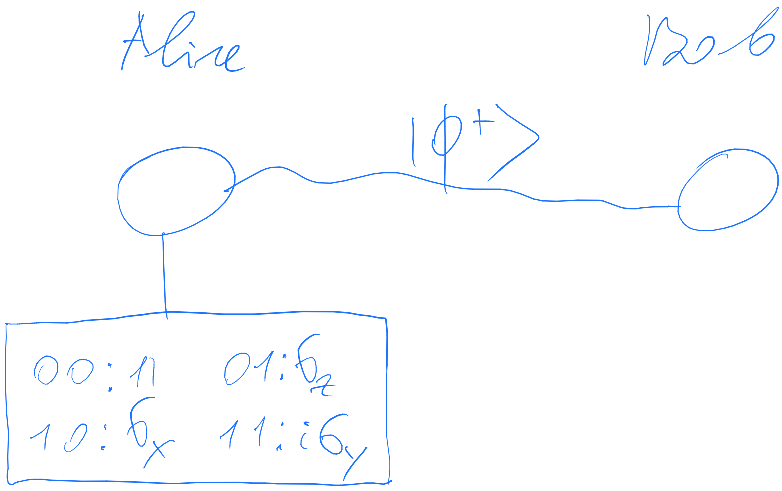

Figure 4.3: Superdense coding.

Suppose that Alice and Bob share two qubits in the state $|\Phi^+\rangle$. They can use $|\Phi^+\rangle$ to communicate two bits of information with a single qubit via the following procedure.

1. Alice applies a unitary on her qubit, depending on which two bits she wants to send to Bob. The concrete unitaries are given by

Encoded bits	00	01	10	11
Alice applies	$\mathbb{1}$	σ_z	σ_x	$i\sigma_y$

The resulting states are given as

$$
\begin{aligned}
00 &: |\Phi^+\rangle \to (\mathbb{1} \otimes \mathbb{1})\, |\Phi^+\rangle = |\Phi^+\rangle\,, \\
01 &: |\Phi^+\rangle \to (\sigma_z \otimes \mathbb{1})\, \frac{1}{\sqrt{2}}\,(|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}}\,(|00\rangle - |11\rangle) = |\Phi^-\rangle\,, \\
10 &: |\Phi^+\rangle \to (\sigma_x \otimes \mathbb{1})\, \frac{1}{\sqrt{2}}\,(|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}}\,(|10\rangle + |01\rangle) = |\Psi^+\rangle\,, \\
11 &: |\Phi^+\rangle \to \left(i\sigma_y \otimes \mathbb{1}\right) \frac{1}{\sqrt{2}}\,(|00\rangle + |11\rangle) = \frac{1}{\sqrt{2}}\,(-|10\rangle + |01\rangle) = |\Psi^-\rangle\,.
\end{aligned}
$$

2. Alice sends her qubit to Bob, who is now in possession of one of the four Bell states.

3. Bob applies a von Neumann measurement in the Bell basis. From his outcome, he can directly read off the two bits encoded by Alice.

Note that two bits is the maximal amount of classical information that one qubit can carry.

5 Entanglement distillation and dilution

5.1 Shannon and von Neumann entropy

Consider an integer random variable x with probability distribution $p(x)$. A sequence of independent and identically distributed variables x_i has probability distribution

$$p(x_1, \dots, x_m) = p(x_1)p(x_2)\dots p(x_m).$$

The Shannon entropy of the probability distribution is defined as

$$H(p(x)) = -\sum_x p(x) \log_2 p(x).$$

Correspondingly, we can define the von Neumann entropy of a quantum state ρ with eigenvalues λ_i:

$$S(\rho) = -\operatorname{Tr}[\rho \log_2 \rho] = -\sum_i \lambda_i \log_2 \lambda_i.$$

5.2 Typical sequences

Consider a sequence $x_1, x_2, \dots, x_m$ of m independent and identically distributed random variables x_i. For large m certain sequences will be suppressed, they are atypical. Typical sequences are those that are most likely to appear for large m. A sequence of independent and identically distributed random variables x_i with entropy $H(p(x))$ is called ϵ-typical if

$$2^{-m(H(p(x))+\epsilon)} \le p(x_1, \dots, x_m) \le 2^{-m(H(p(x))-\epsilon)}. \tag{5.1}$$

Example. Consider a biased coin, where we associate 0 with "heads" and 1 with "tails". We further define $p(0) = 2/3$ and $p(1) = 1/3$, see Fig. 5.1. For $\epsilon = 0.01$ and $m = 10$ we obtain

$$p(1, 1, \dots, 1, 1) = \frac{1}{3^{10}} \approx 2 \times 10^{-5},$$
$$2^{-m(H(p(x))\pm\epsilon)} \approx 2 \times 10^{-3}.$$

Thus, the sequence $1, 1, \dots, 1, 1$ is not ϵ-typical.

$p(\text{heads}) = p(0) = \frac{2}{3}$ $\qquad$ $p(\text{tails}) = p(1) = \frac{1}{3}$

Figure 5.1: Biased coin as an example of a random variable

Theorem of typical sequences.
(1) Fix $\epsilon > 0$. For any $\delta > 0$, for sufficiently large m the probability that a sequence is ϵ-typical is at least $1 - \delta$:

$$\sum_{x \ \epsilon-\text{typical}} p(x_1)p(x_2)\ldots p(x_m) > 1 - \delta.$$

(2) For any fixed $\epsilon > 0$ and $\delta > 0$, for sufficiently large m, the number $|T(m,\epsilon)|$ of ϵ-typical sequences satisfies

$$(1-\delta)2^{m(H(p(x))-\epsilon)} \leq |T(m,\epsilon)| \leq 2^{m(H(p(x))+\epsilon)}.$$

5.3 Entanglement dilution

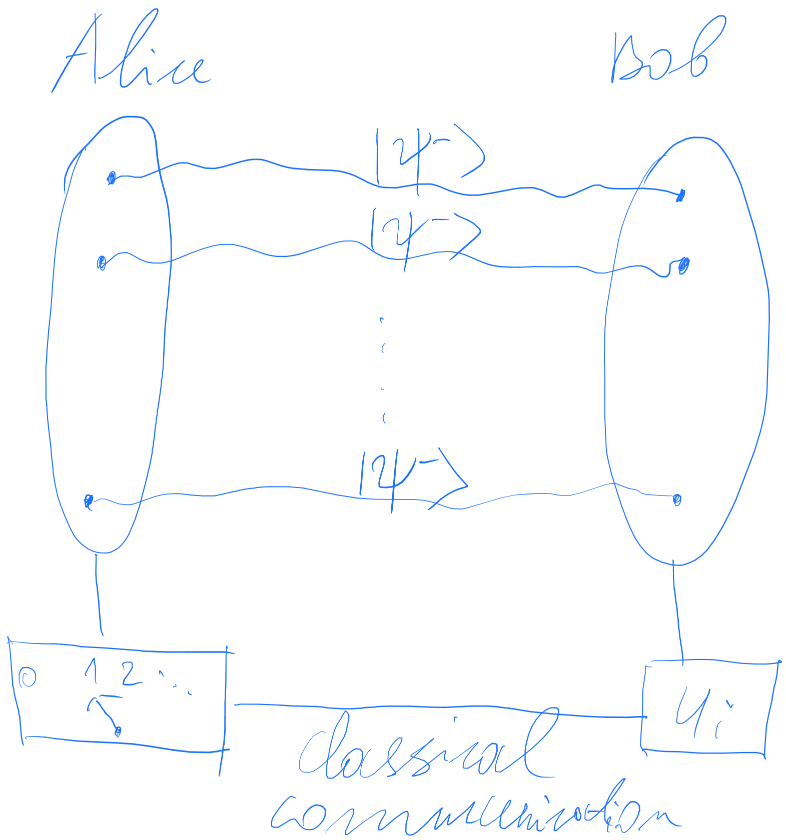

Figure 5.2: Entanglement dilution

Assume that Alice and Bob share n singlets $|\Psi^-\rangle$. Entanglement dilution is an LOCC protocol transforming $|\Psi^-\rangle^{\otimes n}$ into m copies of another state $|\psi\rangle$. The procedure can have an error which should vanish in the asymptotic limit $n \to \infty$.

The minimal fraction n/m in the limit $n \to \infty$ is called entanglement cost of $|\psi\rangle$.

Proposition 5.1. *The entanglement cost of a state $|\psi\rangle$ is at most $S(\rho_\psi)$, where $\rho_\psi = \mathrm{Tr}_B[|\psi\rangle\langle\psi|]$ is the reduced state of Alice.*

Proof. Suppose an entangled state $|\psi\rangle$ has Schmidt decomposition

$$|\psi\rangle = \sum_x \sqrt{p(x)}\,|x\rangle^A \otimes |x\rangle^B .$$

The state $|\psi_m\rangle := |\psi\rangle^{\otimes m}$ can be written as

$$|\psi_m\rangle = \sum_{x_1,x_2,\ldots,x_m} \sqrt{p(x_1)p(x_2)\ldots p(x_m)}\, |x_1x_2\ldots x_m\rangle^A \otimes |x_1x_2\ldots x_m\rangle^B \,. \tag{5.2}$$

We now define a new quantum state $|\phi_m\rangle$ by omitting terms $x_1,\ldots,x_m$ which are not ϵ-typical:

$$|\phi_m\rangle = \sum_{x\ \epsilon-\text{typical}} \sqrt{p(x_1)p(x_2)\ldots p(x_m)}\, |x_1x_2\ldots x_m\rangle^A \otimes |x_1x_2\ldots x_m\rangle^B \,. \tag{5.3}$$

Note that $|\phi_m\rangle$ is in general not normalized. We normalize it by defining

$$|\phi'_m\rangle = \frac{1}{\sqrt{\langle\phi_m|\phi_m\rangle}}\, |\phi_m\rangle \,. \tag{5.4}$$

Consider now the scalar product $\langle\psi_m|\phi'_m\rangle$:

$$\langle\psi_m|\phi'_m\rangle = \frac{1}{\sqrt{\langle\phi_m|\phi_m\rangle}} \sum_{x\ \epsilon-\text{typical}} p(x_1)p(x_2)\ldots p(x_m) = \sqrt{\sum_{x\ \epsilon-\text{typical}} p(x_1)p(x_2)\ldots p(x_m)}.$$

Part (1) of the theorem of typical sequences implies that

$$\lim_{m\to\infty}\left(\sum_{x\ \epsilon-\text{typical}} p(x_1)p(x_2)\ldots p(x_m)\right) = 1,$$

and thus

$$\lim_{m\to\infty}\langle\psi_m|\phi'_m\rangle = 1.$$

This means that for large m the state $|\phi'_m\rangle$ is a good approximation of $|\psi_m\rangle = |\psi\rangle^{\otimes m}$.

Alice now prepares the state $|\phi'_m\rangle$ locally, and then teleports what should be Bob's half of the state $|\phi'_m\rangle$ over to Bob. In this way Alice and Bob end up sharing the state $|\phi'_m\rangle$, which is a good approximation of $|\psi\rangle^{\otimes m}$. By part (2) of the theorem of typical sequences, the number of terms in the sum (5.3) is at most

$$2^{m(H(p(x))+\epsilon)} = 2^{m(S(\rho_\psi)+\epsilon)}.$$

This implies that the state $|\phi'_m\rangle$ has at most $2^{m(S(\rho_\psi)+\epsilon)}$ nonzero Schmidt coefficients. By using Proposition 4.1, Alice can teleport half of the state $|\phi'_m\rangle$ to Bob by consuming at most

$$n = \left\lceil m(S(\rho_\psi)+\epsilon)\right\rceil$$

Bell states. Since ϵ can be chosen arbitary small, we can bring the ratio n/m arbitrary close to $S(\rho_\psi)$ by choosing m large enough. Thus, the entanglement cost of $|\psi\rangle$ is at most $S(\rho_\psi)$. □

5.4 Entanglement distillation

Entanglement distillation can be seen as the reverse process of entanglement dilution. Assume that Alice and Bob share m copies of the state $|\psi\rangle$. Entanglement distillation is an LOCC protocol transforming $|\psi\rangle^{\otimes m}$ into n singlets $|\Psi^-\rangle$. The procedure can have an error which should vanish in the asymptotic limit $m \to \infty$.

The maximal fraction n/m in the limit $m \to \infty$ is called distillable entanglement of $|\psi\rangle$.

Proposition 5.2. *The distillable entanglement of a state $|\psi\rangle$ is at least $S(\rho_\psi)$.*

Proof. Suppose that Alice and Bob share m copies of the state $|\psi\rangle$, see also Eq. (5.2). Alice first performs a local projective measurement with Kraus operators

$$\Pi_0 = \sum_{x\ \epsilon-\text{typical}} |x_1 x_2 \dots x_m\rangle\langle x_1 x_2 \dots x_m|$$

and $\Pi_1 = \mathbb{1} - \Pi_0$. The probability of measurement outcome 0 is

$$p_0 = \mathrm{Tr}[(\Pi_0 \otimes \mathbb{1})\,|\psi_m\rangle\langle\psi_m|] = \sum_{x\ \epsilon-\text{typical}} p(x_1)p(x_2)\dots p(x_m),$$

and the post-measurement state of Alice and Bob is

$$\frac{1}{\sqrt{p_0}}(\Pi_0 \otimes \mathbb{1})\,|\psi_m\rangle = \frac{1}{\sqrt{p_0}} \sum_{x\ \epsilon-\text{typical}} \sqrt{p(x_1)p(x_2)\dots p(x_m)}\,|x_1 x_2 \dots x_m\rangle^A \otimes |x_1 x_2 \dots x_m\rangle^B\,, \tag{5.5}$$

which is equivalent to $|\phi_m'\rangle$ in Eq. (5.4). The probability p_0 converges to 1 in the limit $m \to \infty$ due to part (1) of the theorem of typical sequences.

The (unnormalized) vector $|\phi_m\rangle$ in Eq. (5.3) has Schmidt coefficients of the form $p(x_1)p(x_2)\dots p(x_m)$ with an ϵ-typical sequence $x_1, \dots, x_m$. By definition of ϵ-typical sequences [see Eq. (5.1)], the largest Schmidt coefficient of $|\phi_m\rangle$ is at most

$$p(x_1)p(x_2)\dots p(x_m) \le 2^{-m(H(p(x))-\epsilon)} = 2^{-m(S(\rho_\psi)-\epsilon)},$$

where $\rho_\psi = \mathrm{Tr}_B[|\psi\rangle\langle\psi|]$.

The post-measurement state of Alice and Bob in Eq. (5.5) corresponds to the state $|\phi_m'\rangle$ in Eq. (5.4), and can be written as

$$|\phi_m'\rangle = \frac{|\phi_m\rangle}{\sqrt{\langle\phi_m|\phi_m\rangle}} = \frac{|\phi_m\rangle}{\sqrt{\sum_{x\ \epsilon-\text{typical}} p(x_1)p(x_2)\dots p(x_m)}}.$$

By part (1) of the theorem of typical sequences, for any $\delta > 0$ and m large enough we have

$$\sum_{x\ \epsilon-\text{typical}} p(x_1)p(x_2)\ldots p(x_m) > 1-\delta.$$

Thus, the largest Schmidt coefficient of $|\phi'_m\rangle$ is at most $2^{-m(S(\rho_\psi)-\epsilon)}/(1-\delta)$.

Choose n such that

$$\frac{2^{-m(S(\rho_\psi)-\epsilon)}}{1-\delta} \leq 2^{-n}. \tag{5.6}$$

Since the Schmidt coefficients of $|\phi'_m\rangle$ correspond to eigenvalues of $\rho_{\phi'_m}$, all eigenvalues of $\rho_{\phi'_m}$ are at most 2^{-n}. Thus, the vector $\vec{\lambda}_{\phi'_m}$, containing eigenvalues of $\rho_{\phi'_m}$ in decreasing order, is majorized by the vector

$$\vec{v} = (\underbrace{2^{-n}, 2^{-n}, \ldots, 2^{-n}}_{2^n \text{ times}}, 0, \ldots, 0),$$

where zeros are added to make the dimension of $\vec{v}$ equal to the dimension of $\vec{\lambda}_{\phi'_m}$ (due to Eq. (5.6) the dimension of $\vec{\lambda}_{\phi'_m}$ is at least 2^n). By theorem 2.1, the state $|\phi'_m\rangle$ can then be converted into n singlets via LOCC. Since ϵ and δ can be chosen arbitrary small, analyzing Eq. (5.6) we can bring the fraction n/m arbitrary close to $S(\rho_\psi)$ in the limit of large m. □

Propositions 5.1 and 5.2 provide bounds on the distillable entanglement and entanglement cost of a pure state. We will now prove that these bounds are optimal.

Theorem 5.1. *The distillable entanglement and entanglement cost of a state $|\psi\rangle$ are equal to $S(\rho_\psi)$.*

Proof. Assume by contradiction that there exists a hypothetical LOCC protocol converting m copies of $|\psi\rangle$ into n singlets such that $n/m \approx S > S(\rho_\psi)$ with asymptotically vanishing error in the limit $m \to \infty$.

Assume now that Alice and Bob start with k singlets $|\Psi^-\rangle$. Due to proposition 5.1, for large k there exists an LOCC protocol converting the singlets into $|\psi\rangle^{\otimes m}$ such that $k/m \approx S(\rho_\psi)$. In the next step, Alice and Bob use the hypothetical protocol from the previous paragraph, converting $|\psi\rangle^{\otimes m}$ into $|\Psi^-\rangle^{\otimes n}$ with $n/m \approx S > S(\rho_\psi)$. Note that

$$n \approx mS = k\frac{S}{S(\rho_\psi)} > k.$$

In summary, Alice and Bob started with k singlets, and ended up with $n > k$ singlets. Noting that the number of singlets cannot be increased via LOCC, the contradiction follows. This proves that the distillable entanglement of $|\psi\rangle$ is equal to $S(\rho_\psi)$. The proof that the entanglement cost of $|\psi\rangle$ must be equal to $S(\rho_\psi)$ follows similar lines of reasoning. □

5.5 LOCC and separable operations

Any LOCC protocol is a separable operation:

$$\rho^{AB} \to \Lambda_{\text{LOCC}}\left(\rho^{AB}\right) = \sum_i A_i \otimes B_i \rho^{AB} A_i^\dagger \otimes B_i^\dagger,$$

where $A_i \otimes B_i$ fulfill the completeness condition for Kraus operators:

$$\sum_i A_i^\dagger A_i \otimes B_i^\dagger B_i = \mathbb{1}_{AB}.$$

Not every separable operation is an LOCC.

Any stochastic LOCC transformation has the form

$$\rho^{AB} \to \frac{1}{p} \sum_i A_i \otimes B_i \rho^{AB} A_i^\dagger \otimes B_i^\dagger, \tag{5.7}$$

where

$$\sum_i A_i^\dagger A_i \otimes B_i^\dagger B_i \leq \mathbb{1}_{AB},$$

and the probability of the transformation is given as

$$p = \text{Tr}\left[\sum_i A_i \otimes B_i \rho^{AB} A_i^\dagger \otimes B_i^\dagger\right].$$

A stochastic LOCC transformation mapping $\mathcal{H}_{AB}$ onto the space of two qubits has the form (5.7), where A_i is a $2 \times d_A$ rectangular matrix, and B_i is a $2 \times d_B$ rectangular matrix.

5.6 Entanglement distillation for mixed states

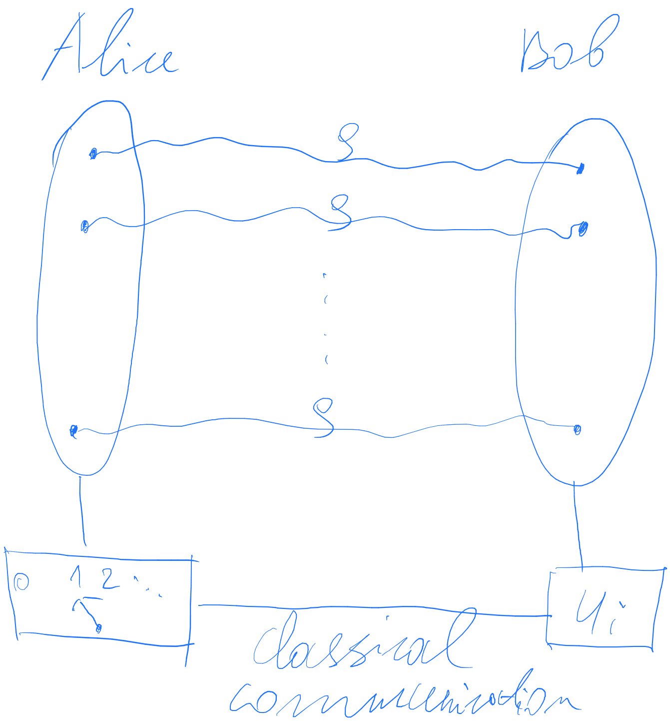

Figure 5.3: Mixed-state entanglement distillation

We now consider entanglement distillation for mixed states. Assume that Alice and Bob share many copies of a mixed state with density matrix ρ, and want to convert them into singlets via LOCC. We first note that Alice and Bob cannot distill any singlets if the shared state is separable. This is because for a separable state ρ also $\rho^{\otimes m}$ is separable. As discussed in Section 5.5, any stochastic LOCC protocol brings the state $\rho^{\otimes m}$ to the state

$$\sigma = \frac{1}{p} \sum_j A_j \otimes B_j \rho^{\otimes m} A_j^\dagger \otimes B_j^\dagger$$

with probability $p = \mathrm{Tr}[\sum_j A_j \otimes B_j \rho^{\otimes m} A_j^\dagger \otimes B_j^\dagger]$. Since $\rho^{\otimes m}$ is separable, also σ is separable. This implies that σ cannot be close to a singlet, even in the asymptotic limit $m \to \infty$.

The above arguments show that separable states cannot be distilled into singlets. The following theorem extends this observation to all states with positive partial transpose.

Theorem 5.2. *States with positive partial transpose cannot be distilled into singlets.*

Proof. If ρ can be distilled into singlets, there must exist a stochastic LOCC protocol bringing $\rho^{\otimes m}$ arbitrary close to a singlet for large m. Then, there must also exist a stochastic LOCC protocol transforming $\rho^{\otimes m}$ into an entangled two-qubit state σ_{2q}. A general stochastic LOCC transformation converting $\rho^{\otimes m}$ into a two-qubit state has the form (see Section 5.5)

$$\sigma_{2q} = \frac{1}{p} \sum_j A_j \otimes B_j \rho^{\otimes m} A_j^\dagger \otimes B_j^\dagger,$$

with $p = \mathrm{Tr}[\sum_j A_j \otimes B_j \rho^{\otimes m} A_j^\dagger \otimes B_j^\dagger]$ and rectangular matrices A_j and B_j.

Since σ_{2q} is entangled, there must be an integer i such that

$$\sigma_i = \frac{1}{p_i} A_i \otimes B_i \rho^{\otimes m} A_i^\dagger \otimes B_i^\dagger$$

is an entangled state, where $p_i = \mathrm{Tr}[A_i \otimes B_i \rho^{\otimes m} A_i^\dagger \otimes B_i^\dagger]$. Recall that A_i is a rectangular $2 \times d_A$ matrix, and B_i is a rectangular $2 \times d_B$ matrix. Thus, A_i and B_i can be written as

$$A_i = |0\rangle\langle\alpha_0| + |1\rangle\langle\alpha_1|,$$
$$B_i = |0\rangle\langle\beta_0| + |1\rangle\langle\beta_1|,$$

where $|\alpha_i\rangle \in \mathcal{H}_A$ and $|\beta_i\rangle \in \mathcal{H}_B$ are (possibly unnormalized) vectors.

Let P_A be a projector onto the subspace spanned by $|\alpha_0\rangle$ and $|\alpha_1\rangle$, and P_B be a projector onto the subspace spanned by $|\beta_0\rangle$ and $|\beta_1\rangle$. Then it holds

$$\sigma_i = \frac{1}{p_i} A_i \otimes B_i \rho^{\otimes m} A_i^\dagger \otimes B_i^\dagger = \frac{1}{p_i} A_i \otimes B_i \left(P_A \otimes P_B \rho^{\otimes m} P_A \otimes P_B \right) A_i^\dagger \otimes B_i^\dagger.$$

Since σ_i is an entangled state, also the state

$$\mu = \frac{P_A \otimes P_B \rho^{\otimes m} P_A \otimes P_B}{\mathrm{Tr}\left[P_A \otimes P_B \rho^{\otimes m} P_A \otimes P_B\right]}$$

must be entangled.

Consider now an orthonormal product basis $|f_i\rangle \otimes |g_k\rangle$ such that

$$P_A = |f_0\rangle\langle f_0| + |f_1\rangle\langle f_1|,$$
$$P_B = |g_0\rangle\langle g_0| + |g_1\rangle\langle g_1|.$$

Expanded in the basis $|f_i\rangle \otimes |g_k\rangle$, the state μ takes the form

$$\mu = \begin{pmatrix} \tau_{2q} & 0 & \cdots & 0 \\ 0 & 0 & & \\ \vdots & & \ddots & \\ 0 & & & 0 \end{pmatrix}, \tag{5.8}$$

where τ_{2q} is a 4×4 density matrix, which can be interpreted as a two-qubit state.

For evaluating the partial transpose μ^{T_A} we can focus on the partial transpose $\tau_{2q}^{T_A}$. If τ_{2q} had positive partial transpose, then by Theorem 3.2 τ_{2q} must be separable, and thus also μ is separable, which is a contradiction. Thus, $\tau_{2q}^{T_A}$ must have negative eigenvalues, i.e., there exists a vector

$$|\psi\rangle = \sum_{i,k=0}^{1} c_{ik} |f_i\rangle |g_k\rangle$$

such that

$$\langle\psi|\tau_{2q}^{T_A}|\psi\rangle < 0.$$

Due to Eq. (5.8) we have $\langle\psi|\tau_{2q}^{T_A}|\psi\rangle = \langle\psi|\mu^{T_A}|\psi\rangle$, which implies that

$$\langle\psi|\mu^{T_A}|\psi\rangle < 0.$$

Using the equalities

$$\begin{aligned}\left(P_A \otimes P_B \rho^{\otimes m} P_A \otimes P_B\right)^{T_A} &= P_A \otimes P_B \left(\rho^{\otimes m}\right)^{T_A} P_A \otimes P_B,\\ P_A \otimes P_B |\psi\rangle &= |\psi\rangle\end{aligned}$$

it follows that

$$\begin{aligned}0 > \langle\psi|\mu^{T_A}|\psi\rangle &= \frac{\langle\psi| \left(P_A \otimes P_B \rho^{\otimes m} P_A \otimes P_B\right)^{T_A} |\psi\rangle}{\mathrm{Tr}\left[P_A \otimes P_B \rho^{\otimes m} P_A \otimes P_B\right]} = \\ &= \frac{\langle\psi| P_A \otimes P_B \left(\rho^{\otimes m}\right)^{T_A} P_A \otimes P_B |\psi\rangle}{\mathrm{Tr}\left[P_A \otimes P_B \rho^{\otimes m} P_A \otimes P_B\right]} = \frac{\langle\psi| \left(\rho^{\otimes m}\right)^{T_A} |\psi\rangle}{\mathrm{Tr}\left[P_A \otimes P_B \rho^{\otimes m} P_A \otimes P_B\right]},\end{aligned}$$

implying that $\rho^{\otimes m}$ has non-positive partial transpose. This also implies that ρ^{T_A} is not positive semidefinite. □

5.7 Matrix realignment criterion and bound entanglement

Given a 2×2 matrix

$$M = \begin{pmatrix} M_{00} & M_{01} \\ M_{10} & M_{11} \end{pmatrix}$$

we can "vectorize" it by defining the vector

$$\vec{M} = (M_{00}, M_{10}, M_{01}, M_{11})^T. \tag{5.9}$$

Consider now a two-qubit density matrix

$$\rho = \begin{pmatrix} \rho_{00} & \rho_{01} & \rho_{02} & \rho_{03} \\ \rho_{10} & \rho_{11} & \rho_{12} & \rho_{13} \\ \rho_{20} & \rho_{21} & \rho_{22} & \rho_{23} \\ \rho_{30} & \rho_{31} & \rho_{32} & \rho_{33} \end{pmatrix} = \begin{pmatrix} X & Y \\ Y^\dagger & Z \end{pmatrix}$$

with 2×2 matrices X, Y and Z. We define the realigned matrix $\tilde{\rho}$ as follows:

$$\tilde{\rho} = \begin{pmatrix} \vec{X}^T \\ \overrightarrow{Y^\dagger}^T \\ \vec{Y}^T \\ \vec{Z}^T \end{pmatrix} = \begin{pmatrix} \rho_{00} & \rho_{10} & \rho_{01} & \rho_{11} \\ \rho_{20} & \rho_{30} & \rho_{21} & \rho_{31} \\ \rho_{02} & \rho_{12} & \rho_{03} & \rho_{13} \\ \rho_{22} & \rho_{32} & \rho_{23} & \rho_{33} \end{pmatrix}.$$

It is straightforward to extend these definitions to dimensions larger than qubits. In the following, we will consider the trace norm of the realigned matrix.

Trace norm. For a general matrix M with singular values s_i the trace norm is defined as

$$||M||_1 = \mathrm{Tr}\,\sqrt{M^\dagger M} = \sum_i s_i.$$

The trace norm fulfills the triangle inequality:

$$||A + B||_1 \leq ||A||_1 + ||B||_1$$

for any two matrices A and B. The trace norm is also absolutely homogeneous:

$$||aM||_1 = |a| \cdot ||M||_1$$

for any matrix M and any $a \in \mathbb{C}$. More details about the trace norm can also be found in Section 6.2.

As we will see in the following proposition, the trace norm of $\tilde{\rho}$ can be used to detect entanglement in the state ρ.

Proposition 5.3. *Any separable state ρ fulfills $||\tilde{\rho}||_1 \leq 1$.*

Proof. Let ρ be a pure product state: $\rho = |\psi\rangle\langle\psi| \otimes |\phi\rangle\langle\phi|$. We "vectorize" the matrices $|\psi\rangle\langle\psi|$ and $|\phi\rangle\langle\phi|$ in the same way as in Eq. (5.9), with the corresponding vectors $\vec{\psi}$ and $\vec{\phi}$. Note that

$$|\vec{\psi}| = |\vec{\phi}| = 1.$$

The realigned matrix $\tilde{\rho}$ can be written as

$$\tilde{\rho} = \vec{\psi} \cdot \vec{\phi}^T.$$

Note that the trace norm of $\tilde{\rho}$ is given as

$$||\tilde{\rho}||_1 = 1.$$

Consider now a separable state

$$\rho_{\text{sep}} = \sum_i p_i |\psi_i\rangle\langle\psi_i| \otimes |\phi_i\rangle\langle\phi_i| .$$

The realigned matrix $\widetilde{\rho_{\text{sep}}}$ takes the form

$$\widetilde{\rho_{\text{sep}}} = \sum_i p_i \overrightarrow{\psi_i} \cdot \overrightarrow{\phi_i}^T,$$

where $\overrightarrow{\psi_i}$ and $\overrightarrow{\phi_i}$ are "vectorized" matrices $|\psi_i\rangle\langle\psi_i|$ and $|\phi_i\rangle\langle\phi_i|$. For the trace norm of $\widetilde{\rho_{\text{sep}}}$ we obtain

$$||\widetilde{\rho_{\text{sep}}}||_1 = \left\| \sum_i p_i \overrightarrow{\psi_i} \cdot \overrightarrow{\phi_i}^T \right\|_1 \leq \sum_i p_i \left\| \overrightarrow{\psi_i} \cdot \overrightarrow{\phi_i}^T \right\|_1 = 1,$$

where we have used the fact that the trace norm is absolutely homogeneous and fulfills the triangle inequality. □

Using the above proposition, we will now show that there exist entangled states which cannot be distilled into singlets. Such states are called bound entangled, since they require singlets for their creation, but cannot be converted into singlets even asymptotically. For $d_A = d_B = 3$ consider the following state for $0 \leq a \leq 1$:

$$\rho_a = \frac{1}{8a+1} \begin{pmatrix} a & 0 & 0 & 0 & a & 0 & 0 & 0 & a \\ 0 & a & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a & 0 & 0 & 0 & 0 & 0 \\ a & 0 & 0 & 0 & a & 0 & 0 & 0 & a \\ 0 & 0 & 0 & 0 & 0 & a & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1+a}{2} & 0 & \frac{\sqrt{1-a^2}}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & a & 0 \\ a & 0 & 0 & 0 & a & 0 & \frac{\sqrt{1-a^2}}{2} & 0 & \frac{1+a}{2} \end{pmatrix}.$$

This state has positive partial transpose for $0 \leq a \leq 1$, but $||\widetilde{\rho_a}||_1 > 1$ for all $0 < a < 1$. Thus, for all $0 < a < 1$ the state ρ_a is a bound entangled state.

It is an open question whether all quantum states with non-positive partial transpose can be distilled into singlets.

6 Quantification of entanglement

Having characterized entanglement, we are now interested to quantify the amount of entanglement in a given state. For this we will consider functions of the state $E(\rho)$ which fulfill the following properties:

1. $E(\rho) \geq 0$, and equality holds if ρ is separable,

2. E does not increase under local operations and classical communication:

$$E(\Lambda_{\mathrm{LOCC}}[\rho]) \leq E(\rho) \tag{6.1}$$

   for any LOCC protocol Λ_{LOCC}.

Interestingly, the second property implies that for $d_A = d_B = d$ the state $|\Phi_d^+\rangle$ and any other maximally entangled states (see Section 2.5) has indeed the maximal amount of entanglement among all states. This is a direct consequence of Theorem 2.1, stating that $|\Phi_d^+\rangle$ can be converted into any pure state via LOCC. Note that this also implies that $|\Phi_d^+\rangle$ can be converted into any mixed state via LOCC.

Functions that fulfill the above two properties are also called entanglement measures. Many entanglement measures have additional properties, such as convexity:

$$E\left(\sum_i p_i \rho_i^{AB}\right) \leq \sum_i p_i E\left(\rho_i^{AB}\right).$$

Moreover, many entanglement measures are nonincreasing on average under LOCC:

$$\sum_i q_i E\left(\sigma_i^{AB}\right) \leq E\left(\rho^{AB}\right), \tag{6.2}$$

where the states σ_i^{AB} and probabilities q_i are obtained from ρ^{AB} by means of LOCC. Condition (6.2) is also called strong monotonicity. Note that strong monotonicity together with convexity implies Eq. (6.1).

In the following, we will study examples of entanglement measures.

6.1 Entanglement of formation

Entanglement of formation is defined for pure states as

$$E_f(|\psi\rangle^{AB}) = S(\rho^A),$$

where $\rho^A = \mathrm{Tr}_B[|\psi\rangle\langle\psi|^{AB}]$ is the reduced state of Alice. This quantity is also called entanglement entropy of $|\psi\rangle^{AB}$. For mixed states ρ^{AB} we define

$$E_f(\rho^{AB}) = \min \sum_i p_i E(|\psi_i\rangle^{AB}),$$

and the minimum is taken over all decompositions $\{p_i, |\psi_i\rangle^{AB}\}$ such that $\rho^{AB} = \sum_i p_i |\psi_i\rangle\langle\psi_i|^{AB}$. The entanglement of formation can be interpreted as the minimal average entanglement required to create the state ρ^{AB}.

We will first show that $E_f(\rho^{AB}) \geq 0$. For this, note that for any decomposition $\{p_i, |\psi_i\rangle^{AB}\}$ the average entanglement $\sum_i p_i E_f(|\psi_i\rangle^{AB})$ is nonnegative. Moreover, for a separable state σ^{AB} there exists a decomposition into product states $|\psi_i\rangle^{AB} = |\alpha_i\rangle^A \otimes |\beta_i\rangle^B$ with $E_f(|\psi_i\rangle^{AB}) = 0$, which implies that $E_f(\sigma^{AB}) = 0$ for any separable state.

Proposition 6.1. *Entanglement of formation is convex:*

$$E_f\left(\sum_i p_i \rho_i^{AB}\right) \leq \sum_i p_i E_f\left(\rho_i^{AB}\right).$$

Proof. Consider a decomposition of the state $\rho_i^{AB} = \sum_j q_{ij} |\psi_{ij}\rangle\langle\psi_{ij}|^{AB}$ with the property that

$$E_f(\rho_i^{AB}) = \sum_j q_{ij} E_f\left(|\psi_{ij}\rangle^{AB}\right).$$

We then obtain

$$\sum_i p_i E_f\left(\rho_i^{AB}\right) = \sum_{ij} p_i q_{ij} E_f\left(|\psi_{ij}\rangle^{AB}\right).$$

We now define the state $\sigma^{AB} = \sum_i p_i \rho_i^{AB}$, and note now that σ^{AB} can also be expressed as

$$\sigma^{AB} = \sum_i p_i \rho_i^{AB} = \sum_{ij} p_i q_{ij} |\psi_{ij}\rangle\langle\psi_{ij}|^{AB}.$$

Recalling that the entanglement of formation is defined as the minimal average entanglement of a state, it must be that

$$E_f\left(\sigma^{AB}\right) \leq \sum_i p_i q_{ij} E_f\left(|\psi_{ij}\rangle^{AB}\right).$$

Combining these results, we obtain

$$E_f\left(\sum_i p_i\rho_i^{AB}\right) = E_f\left(\sigma^{AB}\right) \leq \sum_i p_i E_f\left(\rho_i^{AB}\right),$$

which completes the proof. □

6.1.1 Monotonicity under LOCC

Our next aim is to show that the entanglement of formation does not increase under local operations and classical communication. For achieving this, we will first show that E_f is monotonic on average under local measurements for pure states. Consider a pure state $|\psi\rangle^{AB}$, and suppose that Alice applies a local measurement with Kraus operators $\{K_i\}$. The corresponding post-measurement states are

$$|\phi_i\rangle^{AB} = \frac{1}{\sqrt{p_i}}(K_i \otimes \mathbb{1})\,|\psi\rangle^{AB}$$

with probability

$$p_i = \mathrm{Tr}\left[K_i \otimes \mathbb{1}\,|\psi\rangle\langle\psi|^{AB}\,K_i^\dagger \otimes \mathbb{1}\right],$$

see Section 1.2. We now have the following proposition.

Proposition 6.2. *For pure states $|\psi\rangle^{AB}$ entanglement of formation does not increase on average under local measurements on Alice's side:*

$$\sum_i p_i E_f(|\phi_i\rangle^{AB}) \leq E_f(|\psi\rangle^{AB}).$$

Proof. Note that local measurements on Alice's side do not change the state of Bob, and thus

$$\rho^B = \mathrm{Tr}_A\left[|\psi\rangle\langle\psi|^{AB}\right] = \sum_i p_i\,\mathrm{Tr}_A\left[|\phi_i\rangle\langle\phi_i|^{AB}\right] = \sum_i p_i\sigma_i^B,$$

where we defined $\sigma_i^B = \mathrm{Tr}_A\left[|\phi_i\rangle\langle\phi_i|^{AB}\right]$. By definition of E_f we further have[1]

$$E_f(|\psi\rangle^{AB}) = S(\rho^B), \qquad \sum_i p_i E_f(|\phi_i\rangle^{AB}) = \sum_i p_i S(\sigma_i^B).$$

Combining these results and using the fact that the von Neumann entropy is concave we obtain

$$\sum_i p_i E_f(|\phi_i\rangle^{AB}) = \sum_i p_i S(\sigma_i^B) \leq S\left(\sum_i p_i\sigma_i^B\right) = S(\rho^B) = E_f(|\psi\rangle^{AB}).$$

□

[1] Note that for a pure state $|\psi\rangle^{AB}$ it holds $S(\rho^A) = S(\rho^B)$.

We will now show that this proposition also extends to mixed states ρ^{AB}. Now, if Alice performs a local measurement with Kraus operators $\{K_i\}$, the outcome probability and the post-measurement states are given as

$$p_i = \mathrm{Tr}\left[K_i \otimes \mathbb{1} \rho^{AB} K_i^\dagger \otimes \mathbb{1}\right],$$
$$\sigma_i^{AB} = \frac{1}{p_i} K_i \otimes \mathbb{1} \rho^{AB} K_i^\dagger \otimes \mathbb{1},$$

see Section 1.2. We will now prove the following result.

Proposition 6.3. *For all mixed states ρ^{AB} the entanglement of formation does not increase on average under local measurements on Alice's side:*

$$\sum_i p_i E_f(\sigma_i^{AB}) \leq E_f(\rho^{AB}).$$

Proof. Consider an optimal decomposition $\{q_j, |\psi_j\rangle^{AB}\}$ of the state ρ^{AB} such that $\rho^{AB} = \sum_j q_j |\psi_j\rangle\langle\psi_j|^{AB}$ and

$$E_f(\rho^{AB}) = \sum_j q_j E_f(|\psi_j\rangle^{AB}). \tag{6.3}$$

We now define

$$p_{ij} = \mathrm{Tr}\left[(K_i \otimes \mathbb{1}) |\psi_j\rangle\langle\psi_j| \left(K_i^\dagger \otimes \mathbb{1}\right)\right], \tag{6.4a}$$
$$|\phi_{ij}\rangle^{AB} = \frac{1}{\sqrt{p_{ij}}} (K_i \otimes \mathbb{1}) |\psi_j\rangle^{AB}, \tag{6.4b}$$

and note that

$$\sum_j q_j p_{ij} = p_i.$$

For the entanglement of formation of the states σ_i^{AB} we obtain

$$E_f\left(\sigma_i^{AB}\right) = E_f\left(\frac{1}{p_i} K_i \otimes \mathbb{1} \rho^{AB} K_i^\dagger \otimes \mathbb{1}\right) = E_f\left(\sum_j \frac{q_j}{p_i} K_i \otimes \mathbb{1} |\psi_j\rangle\langle\psi_j|^{AB} K_i^\dagger \otimes \mathbb{1}\right)$$
$$= E_f\left(\sum_j \frac{q_j p_{ij}}{p_i} |\phi_{ij}\rangle\langle\phi_{ij}|^{AB}\right).$$

Using convexity of E_f further gives us

$$E_f\left(\sigma_i^{AB}\right) \leq \sum_j \frac{q_j p_{ij}}{p_i} E_f\left(|\phi_{ij}\rangle^{AB}\right).$$

Multiplying this inequality with p_i on both sides and taking the sum over i gives

$$\sum_i p_i E_f\left(\sigma_i^{AB}\right) \leq \sum_{i,j} q_j p_{ij} E_f\left(|\phi_{ij}\rangle^{AB}\right).$$

Now note that the states $|\phi_{ij}\rangle^{AB}$ and probabilities p_{ij} are obtained from $|\psi_j\rangle^{AB}$ via a local measurement on Alice's side, see Eqs. (6.4). Thus, from Proposition 6.2 we have

$$\sum_i p_{ij} E_f\left(|\phi_{ij}\rangle^{AB}\right) \leq E_f\left(|\psi_j\rangle^{AB}\right),$$

which then leads to

$$\sum_i p_i E_f\left(\sigma_i^{AB}\right) \leq \sum_j q_j \sum_i p_{ij} E_f\left(|\phi_{ij}\rangle^{AB}\right) \leq \sum_j q_j E_f\left(|\psi_j\rangle^{AB}\right).$$

The proof is complete by recalling that $\{q_j, |\psi_j\rangle^{AB}\}$ is an optimal decomposition of ρ^{AB}, see Eq. (6.3). □

While the above proposition concerns only local operations on Alice's side, it is straightforward to show that it generalizes to any LOCC protocol, where Alice and Bob perform local measurements and exchange their measurement outcomes via a classical channel. In the following proposition, σ_i^{AB} denote states which can be obtained from an initial state ρ^{AB} via an arbitrary LOCC protocol, with corresponding probability p_i.

Proposition 6.4. *Entanglement of formation does not increase on average under local operations and classical communication:*

$$\sum_i p_i E_f(\sigma_i^{AB}) \leq E_f(\rho^{AB}).$$

With the above result, we can finally prove the following theorem.

Theorem 6.1. *Entanglement of formation does not increase under LOCC:*

$$E_f(\Lambda_{\mathrm{LOCC}}[\rho]) \leq E_f(\rho)$$

for any LOCC protocol Λ_{LOCC}.

Proof. Let Λ_{LOCC} be an LOCC protocol leading to states σ_i^{AB} with probability p_i when applied to a state ρ^{AB}:

$$\Lambda_{\mathrm{LOCC}}[\rho^{AB}] = \sum_i p_i \sigma_i^{AB}.$$

We use Proposition 6.4 and convexity of E_f:

$$E_f(\Lambda_{\mathrm{LOCC}}[\rho^{AB}]) = E_f\left(\sum_i p_i \sigma_i^{AB}\right) \leq \sum_i p_i E_f(\sigma_i^{AB}) \leq E_f(\rho^{AB}).$$

□

6.1.2 Evaluating entanglement of formation for two qubits

Given a general state of two qubits, we will now give a formula for calculating the entanglement of formation. For this, we first define the concurrence of a state ρ^{AB}:

$$C(\rho^{AB}) = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\},$$

where λ_i are the square roots (in decreasing order) of the eigenvalues of $\rho\tilde{\rho}$, with

$$\tilde{\rho} = (\sigma_y \otimes \sigma_y)\rho^*(\sigma_y \otimes \sigma_y),$$

the Pauli matrix $\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, and ρ^* denotes entry-wise complex conjugation. Concurrence can be seen as a measure of entanglement on its own right, as it is nonnegative, and zero for any separable state.

Having defined the concurrence, the entanglement of formation of ρ^{AB} can be given as

$$E_f(\rho^{AB}) = h\left(\frac{1 + \sqrt{1 - C^2(\rho^{AB})}}{2}\right)$$

with the binary entropy $h(x) = -x \log_2 x - (1 - x)\log_2(1 - x)$.

6.2 Trace distance and fidelity

For two quantum states ρ and σ the trace distance is defined as

$$D_t(\rho, \sigma) = \frac{1}{2}\left\|\rho - \sigma\right\|_1 \tag{6.5}$$

with the trace norm $||M||_1 = \mathrm{Tr}\sqrt{M^\dagger M}$, see also page 36. It holds that $D_t(\rho, \sigma) = 0$ if and only if $\rho = \sigma$, and $1 \geq D(\rho, \sigma) > 0$ otherwise. Moreover, the trace distance does not increase under quantum operations, i.e.,

$$D_t(\Lambda[\rho], \Lambda[\sigma]) \leq D_t(\rho, \sigma) \tag{6.6}$$

for any quantum operation Λ. Eq. (6.6) is also called data-processing inequality, and is a consequence of the following theorem.

Theorem 6.2. *For any Hermitian $d \times d$ matrix H and any trace preserving positive linear map Λ acting on the Hilbert space of dimension d it holds that*

$$||\Lambda(H)||_1 \leq ||H||_1\,. \tag{6.7}$$

Proof. Let $\Lambda(H) = Q_+ - Q_-$ and $H = P_+ - P_-$ be decompositions into orthogonal parts $Q_\pm \geq 0$ and $P_\pm \geq 0$. It follows that

$$\mathrm{Tr}\,(Q_+) \leq \mathrm{Tr}\,(\Lambda\,[P_+])\,,$$
$$\mathrm{Tr}\,(Q_-) \leq \mathrm{Tr}\,(\Lambda\,[P_-])\,.$$

Recalling that Λ is trace preserving, we further have

$$\mathrm{Tr}\,(Q_+ + Q_-) \leq \mathrm{Tr}\,(P_+ + P_-)\,.$$

The proof is complete using the fact that $\|H\|_1 = \mathrm{Tr}(P_+ + P_-)$ and $\|\Lambda(H)\|_1 = \mathrm{Tr}(Q_+ + Q_-)$. □

In the following, we will make use of the following proposition.

Proposition 6.5. *For any unitary U it holds that*

$$|\mathrm{Tr}\,(AU)| \leq \|A\|_1\,.$$

Proof. By the polar decomposition we have

$$|\mathrm{Tr}\,(AU)| = \left|\mathrm{Tr}\left(V\sqrt{A^\dagger A}U\right)\right| = \left|\mathrm{Tr}\left(\left[A^\dagger A\right]^{1/4}\left[A^\dagger A\right]^{1/4}UV\right)\right|.$$

Using the Cauchy-Schwarz inequality $\left|\mathrm{Tr}\left(X^\dagger Y\right)\right|^2 \leq \mathrm{Tr}\left(X^\dagger X\right)\mathrm{Tr}\left(Y^\dagger Y\right)$ and setting

$$X = \left[A^\dagger A\right]^{1/4}, \quad Y = \left[A^\dagger A\right]^{1/4}UV$$

we obtain

$$|\mathrm{Tr}\,(AU)| \leq \sqrt{\mathrm{Tr}\,\sqrt{A^\dagger A}\,\mathrm{Tr}\left(V^\dagger U^\dagger\sqrt{A^\dagger A}UV\right)} = \mathrm{Tr}\,\sqrt{A^\dagger A} = \|A\|_1\,.$$

□

A quantity which is closely related to the trace distance is the fidelity. For two quantum states ρ and σ the fidelity is defined as

$$F(\rho,\sigma) = \mathrm{Tr}\,\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}.$$

The fidelity is related to the trace distance as follows:

$$1 - F(\rho,\sigma) \leq D_t(\rho,\sigma) \leq \sqrt{1 - F(\rho,\sigma)^2}. \qquad (6.8)$$

From Eq. (6.8) we see that $0 \leq F(\rho,\sigma) \leq 1$, and $F(\rho,\sigma) = 1$ if and only if $\rho = \sigma$.

Let now $|\psi\rangle = |\psi\rangle^{AB}$ and $|\phi\rangle = |\phi\rangle^{AB}$ be purifications of the states $\rho = \rho^B$ and $\sigma = \sigma^B$. The following theorem provides a connection between the fidelity and the purifications of the states.

Theorem 6.3. *For any two states ρ and σ it holds that*

$$F(\rho,\sigma) = \max_{|\psi\rangle,|\phi\rangle} \left|\langle\psi|\phi\rangle\right|,$$

where the maximum is taken over all purifications $|\psi\rangle$ of ρ and $|\phi\rangle$ of σ.

Proof. We can write any purification of ρ and σ as follows:

$$|\psi\rangle = \left(U_A \otimes \sqrt{\rho}U_B\right)|m\rangle,$$
$$|\phi\rangle = \left(V_A \otimes \sqrt{\sigma}V_B\right)|m\rangle$$

with $d_A = d_B$, $|m\rangle = \sum_i |i\rangle|i\rangle$ and some unitaries U_A, U_B, V_A, and V_B. We obtain

$$\left|\langle\psi|\phi\rangle\right| = \left|\langle m|U_A^\dagger V_A \otimes U_B^\dagger \sqrt{\rho}\sqrt{\sigma}V_B|m\rangle\right|.$$

Using the equality

$$|\langle m|A \otimes B|m\rangle| = \mathrm{Tr}\left(A^\dagger B\right)$$

we further obtain

$$\left|\langle\psi|\phi\rangle\right| = \left|\mathrm{Tr}\left(V_A^\dagger U_A U_B^\dagger \sqrt{\rho}\sqrt{\sigma}V_B\right)\right|.$$

Defining the unitary $U = V_B V_A^\dagger U_A U_B^\dagger$ we arrive at

$$\left|\langle\psi|\phi\rangle\right| = \left|\mathrm{Tr}\left(\sqrt{\rho}\sqrt{\sigma}U\right)\right|.$$

Using Proposition 6.5, we see that

$$\left|\langle\psi|\phi\rangle\right| \leq \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1 = \mathrm{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}} = F(\rho,\sigma).$$

The equality can be attained by choosing a unitary V such that $M = \sqrt{MM^\dagger}V$, where $M = \sqrt{\rho}\sqrt{\sigma}$. Setting $V_B = V^\dagger$ and $U_B = U_A = V_A = \mathbb{1}$ the equality is attained. □

Using Theorem 6.3, we will now prove that the fidelity is monotonic under quantum operations.

Theorem 6.4. *For any two quantum states ρ and σ and any quantum operation Λ it holds that*

$$F(\Lambda[\rho], \Lambda[\sigma]) \geq F(\rho,\sigma).$$

Proof. Every quantum operation Λ on the system B can be written as

$$\Lambda\left[\rho^B\right] = \mathrm{Tr}_E\left[U_{BE}\left(\rho^B \otimes |0\rangle\langle 0|^E\right)U_{BE}^\dagger\right].$$

Let $|\psi\rangle^{AB}$ and $|\phi\rangle^{AB}$ be purifications of ρ^B and σ^B, such that $F(\rho,\sigma) = \left|\langle\psi|\phi\rangle\right|$ (see Theorem 6.3). Then $\mathbb{1} \otimes U_{BE} |\psi\rangle^{AB} |0\rangle^E$ is a purification of $\Lambda[\rho^B]$ and $\mathbb{1} \otimes U_{BE} |\phi\rangle^{AB} |0\rangle^E$ is a purification of $\Lambda[\sigma^B]$. Using Theorem 6.3 we obtain

$$F\left(\Lambda\left[\rho\right], \Lambda\left[\sigma\right]\right) \geq \left|\langle\psi| \langle 0| \mathbb{1} \otimes U_{BE}^\dagger U_{BE} |\phi\rangle |0\rangle\right| = \left|\langle\psi|\phi\rangle\right| = F(\rho,\sigma).$$

□

Using fidelity, it is possible to define the Bures distance

$$D_b(\rho,\sigma) = \sqrt{2 - 2F(\rho,\sigma)} \tag{6.9}$$

which has similar properties as the trace distance. In particular, $D_b(\rho,\sigma) \geq 0$ with equality if and only if $\rho = \sigma$. Moreover, D_b fulfills the data-processing inequality:

$$D_b(\Lambda[\rho], \Lambda[\sigma]) \leq D_b(\rho,\sigma)$$

for any quantum operation Λ.

6.3 Distance-based entanglement measures

Given a distance function $D(\rho,\sigma)$ for any pair of density matrices, it is possible to construct an entanglement measure as

$$E(\rho) = \inf_{\sigma\in\mathcal{S}} D(\rho,\sigma), \tag{6.10}$$

where the infimum is taken over the set of separable states $\mathcal{S}$, see also Fig. 6.1.

For any distance which fulfills $D(\rho,\sigma) \geq 0$ with equality if $\rho = \sigma$, the corresponding entanglement measure is nonnegative, and zero for separable states. Moreover, if the distance D fulfills the data-processing inequality, i.e.,

$$D(\Lambda[\rho], \Lambda[\sigma]) \leq D(\rho,\sigma) \tag{6.11}$$

for any quantum operation Λ, the corresponding entanglement quantifier does not increase under LOCC. To see this, let σ be a separable state realizing the minimum in Eq. (6.10), such that $E(\rho) = D(\rho,\sigma)$. Noting that $\Lambda_{\mathrm{LOCC}}[\sigma]$ is a separable state, we have

$$E\left(\Lambda_{\mathrm{LOCC}}[\rho]\right) = \min_{\mu\in\mathcal{S}} D\left(\Lambda_{\mathrm{LOCC}}[\rho], \mu\right) \leq D\left(\Lambda_{\mathrm{LOCC}}[\rho], \Lambda_{\mathrm{LOCC}}[\sigma]\right) \leq D(\rho,\sigma) = E(\rho).$$

Examples for distances fulfilling Eq. (6.11):

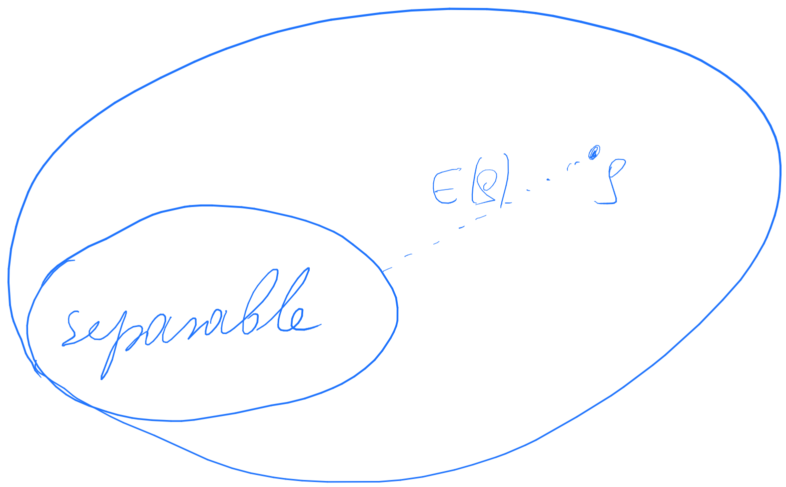

Figure 6.1: Quantifying entanglement via a distance from the set of separable states.

- Quantum relative entropy[2]

$$S(\rho\|\sigma) = \mathrm{Tr}[\rho \log_2 \rho] - \mathrm{Tr}[\rho \log_2 \sigma],$$

and the corresponding entanglement measure is called relative entropy of entanglement: $E_r(\rho) = \min_{\sigma\in\mathcal{S}} S(\rho\|\sigma)$. It is an upper bound on distillable entanglement. For any pure state $|\psi\rangle^{AB}$ it holds $E_r(|\psi\rangle^{AB}) = S(\rho^A)$.

- Bures distance $D_b(\rho,\sigma) = \sqrt{2 - 2F(\rho,\sigma)}$, see also Eq. (6.9).

- Trace distance $D_t(\rho,\sigma) = \frac{1}{2}\|\rho - \sigma\|_1$, see also Eq. (6.5).

6.4 Negativity

Given a bipartite state ρ^{AB}, the negativity is defined as

$$E_n(\rho^{AB}) = \frac{\|\rho^{T_B}\|_1 - 1}{2}.$$

The negativity is nonnegative and $E_n(\rho^{AB}) = 0$ when ρ^{AB} is separable or (more generally) when ρ^{AB} has a positive partial transpose.

Theorem 6.5. *Negativity does not increase under LOCC:*

$$E_n\left(\Lambda_{\mathrm{LOCC}}\left[\rho^{AB}\right]\right) \le E_n\left(\rho^{AB}\right).$$

[2]Note that the quantum relative entropy is in general not symmetric and does not fulfill the triangle inequality.

Proof. Recall that any LOCC protocol can be written as (see Section 5.5)

$$\Lambda_{\mathrm{LOCC}}[\rho^{AB}] = \sum_i A_i \otimes B_i \rho^{AB} A_i^\dagger \otimes B_i^\dagger$$

with Kraus operators $A_i \otimes B_i$ fulfilling the completeness relation

$$\sum_i A_i^\dagger A_i \otimes B_i^\dagger B_i = \mathbb{1}_{AB}.$$

Taking partial transpose with respect to Bob's system on both sides of this equality we get

$$\sum_i A_i^\dagger A_i \otimes B_i^T B_i^* = \mathbb{1}_{AB},$$

where we used the fact that the identity matrix $\mathbb{1}_{AB}$ is invariant under partial transpose. This implies that $A_i \otimes B_i^*$ are also valid Kraus operators.

In the next step, we apply partial transpose onto $\Lambda_{\mathrm{LOCC}}[\rho^{AB}]$:

$$\left(\Lambda_{\mathrm{LOCC}}\left[\rho^{AB}\right]\right)^{T_B} = \left(\sum_i A_i \otimes B_i \rho^{AB} A_i^\dagger \otimes B_i^\dagger\right)^{T_B} = \sum_i A_i \otimes B_i^* \rho^{T_B} A_i^\dagger \otimes B_i^T.$$

Taking the trace norm of this expression gives

$$\left\|\left(\Lambda_{\mathrm{LOCC}}\left[\rho^{AB}\right]\right)^{T_B}\right\|_1 = \left\|\sum_i A_i \otimes B_i^* \rho^{T_B} A_i^\dagger \otimes B_i^T\right\|_1 = \left\|\tilde{\Lambda}\left[\rho^{T_B}\right]\right\|_1,$$

where $\tilde{\Lambda}$ is the quantum operation corresponding to the Kraus operators $\{A_i \otimes B_i^*\}$. Using the fact that the trace norm does not increase under quantum operations, we obtain

$$\left\|\tilde{\Lambda}\left[\rho^{T_B}\right]\right\|_1 \leq \left\|\rho^{T_B}\right\|_1,$$

which in summary gives us

$$\left\|\left(\Lambda_{\mathrm{LOCC}}\left[\rho^{AB}\right]\right)^{T_B}\right\|_1 \leq \left\|\rho^{T_B}\right\|_1.$$

Using this in the definition of negativity completes the proof. □

Negativity is also convex, and fulfills strong monotonicity, see Eq. (6.2).

6.5 Distillable entanglement and entanglement cost

Distillable entanglement has been defined in Section 5.4 as the singlet rate obtainable from a quantum state ρ via LOCC in the asymptotic limit. Correspondingly, entanglement cost has been defined in Section 5.3 as the singlet rate required to create a state ρ via LOCC in the asymptotic limit. An explicit formula for distillable entanglement can be given as

$$E_d(\rho) = \sup\left\{r : \lim_{n\to\infty}\left(\inf_{\Lambda}\left\|\Lambda\left[\rho^{\otimes n}\right] - |\Phi^+\rangle\langle\Phi^+|^{\otimes\lfloor rn\rfloor}\right\|_1\right) = 0\right\},$$

where the infimum is taken over all LOCC protocols Λ. Correspondingly, entanglement cost can be given as

$$E_c(\rho) = \inf\left\{r : \lim_{n\to\infty}\left(\inf_{\Lambda}\left\|\rho^{\otimes n} - \Lambda\left[|\Phi^+\rangle\langle\Phi^+|^{\otimes\lfloor rn\rfloor}\right]\right\|_1\right) = 0\right\}.$$

Distillable entanglement and entanglement cost are special cases of asymptotic state-conversion rates, which can in general be given as

$$R(\rho\to\sigma) = \sup\left\{r : \lim_{n\to\infty}\left(\inf_{\Lambda}\left\|\Lambda\left[\rho^{\otimes n}\right] - \sigma^{\otimes\lfloor rn\rfloor}\right\|_1\right) = 0\right\}.$$

It holds that

$$E_d(\rho) = R(\rho\to|\Phi^+\rangle\langle\Phi^+|), \qquad E_c(\rho) = \left[R(|\Phi^+\rangle\langle\Phi^+|\to\rho)\right]^{-1}.$$

Moreover, for pure states $|\psi\rangle$ and $|\phi\rangle$ we obtain

$$R(|\psi\rangle\to|\phi\rangle) = \frac{S(\rho_\psi)}{S(\rho_\phi)},$$

where ρ_ψ is the reduced state of $|\psi\rangle$.

Distillable entanglement and entanglement cost are bounded as

$$E_d(\rho^{AB}) \le E_c(\rho^{AB}) \le E_f(\rho^{AB}), \tag{6.12a}$$
$$E_r(\rho^{AB}) \ge E_d(\rho^{AB}) \ge S(\rho^A) - S(\rho^{AB}), \tag{6.12b}$$

where E_f is the entanglement of formation and E_r is the relative entropy of entanglement.

As an application, consider a state of the form

$$\rho_{\mathrm{mc}}^{AB} = \sum_{i,j}\alpha_{ij}\,|ii\rangle\langle jj|$$

with $\alpha_{ij} \in \mathbb{C}$. States of this form are also called maximally correlated. Note that every pure state is maximally correlated. For the separable state $\sigma_{\text{sep}}^{AB} = \sum_i \alpha_{ii} |ii\rangle\langle ii|$ it is straightforward to verify the equality

$$S(\rho_{\text{mc}}^{AB} \| \sigma_{\text{sep}}^{AB}) = S(\rho_{\text{mc}}^{A}) - S(\rho_{\text{mc}}^{AB}). \tag{6.13}$$

From the definition of E_r and Eqs. (6.12) we further see that

$$S(\rho_{\text{mc}}^{AB} \| \sigma_{\text{sep}}^{AB}) \geq E_r(\rho^{AB}) \geq E_d(\rho^{AB}) \geq S(\rho_{\text{mc}}^{A}) - S(\rho_{\text{mc}}^{AB}).$$

Together with Eq. (6.13) we arrive at the final expression for the distillable entanglement of any maximally correlated state:

$$E_d(\rho_{\text{mc}}^{AB}) = S(\rho_{\text{mc}}^{A}) - S(\rho_{\text{mc}}^{AB}).$$

7 Monogamy of entanglement

Consider two qubits A and B in the maximally entangled state $|\Phi^+\rangle$. Then, neither A nor B can be entangled (or even correlated) with another qubit C, see Fig. 7.1. This phenomenon is called entanglement monogamy. Note that this is a purely quantum phenomenon, since a classical random variable can be maximally correlated with arbitrary many classical systems at the same time.

Quantitatively, there is a tradeoff between the amount of entanglement between the qubits A and B and the qubits A and C. For a pure three-qubit state $|\psi\rangle^{ABC}$ it can be formulated in terms of of concurrence C (see Section 6.1.2):

$$C^2_{A:B} + C^2_{A:C} \leq C^2_{A:BC}.$$

Here, $C_{A:B}$ and $C_{A:C}$ is the concurrence of the reduced state ρ^{AB} and ρ^{AC}, respectively, and

$$C_{A:BC} = \sqrt{2\left(1 - \mathrm{Tr}\left[(\rho^A)^2\right]\right)}$$

is the concurrence of the total state $|\psi\rangle^{ABC}$.

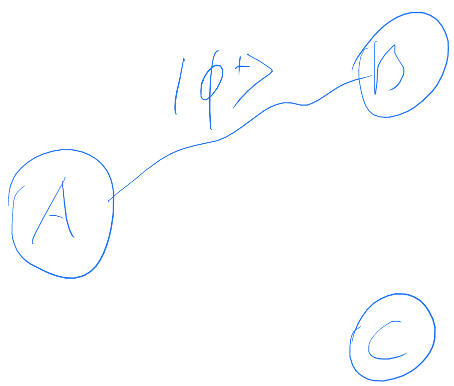

Figure 7.1: Monogamy of entanglement