Quantum Information

Andras Molnar

February 4, 2025

Remark 1. *In this note all Hilbert spaces are complex and finite dimensional unless otherwise stated.*

1 Introduction

1.1 Bra-ket notation

1.1.1 Vectors ("kets")

Let V be an d-dimensional vector space. For any vector $v \in V$ we will (usually) write $|v\rangle$ instead of v, and read *ket-v*. When one performs calculations, they have to fix a basis $v_0, v_1, \ldots, v_{d-1}$, or in ket notation, $|v_0\rangle, \ldots, |v_{d-1}\rangle$. This specific basis is then referred to as the *computational basis*. Since these vectors are used a lot, we shorten the notation and simply write $|0\rangle, \ldots, |d-1\rangle$ instead of $|v_0\rangle, \ldots, |v_{d-1}\rangle$. That is, we can write every vector $|v\rangle \in V$ as

$$|v\rangle = v_0|0\rangle + v_1|1\rangle + \ldots v_{d-1}|d-1\rangle, \tag{1}$$

with $v_0, \ldots, v_{d-1} \in \mathbb{C}$ numbers that are uniquely determined by $|v\rangle$ and our choice of the basis. As usual, we can write this equation in matrix notation; we will think of vectors as column vectors. The computational basis in this notation is given by

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad |d-1\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

and Eq. (1) becomes

$$|v\rangle = v_0|0\rangle + v_1|1\rangle + \ldots v_{d-1}|d-1\rangle = v_0 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + v_1 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + v_{d-1} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} v_0 \\ v_1 \\ \vdots \\ v_{d-1} \end{pmatrix}.$$

Given two different vector spaces V and W, of dimension n and m, respectively, both of their computational bases are denoted the same way: $|0\rangle, \ldots, |n-1\rangle$, and $|0\rangle, \ldots, |m-1\rangle$. Notice that $|0\rangle$ denotes two different vectors in two different vector spaces: once $|0\rangle \in V$, and once $|0\rangle \in W$. It usually is clear from the context which vector we are referring to; if confusion arises, we can specify the vector space by attaching an extra subscript to the ket notation; for example, we could write $|0\rangle_V \in V$ and $|0\rangle_W \in W$.

Similarly, at certain calculations we do not even specify the vector space V. Instead we give only the (basis) vectors $|0\rangle, \ldots, |d-1\rangle$. In these cases one can assume that the highest index appearing is one less than the dimension of V, that is, if the highest index is $|d-1\rangle$, then $V \simeq \mathbb{C}^d$.

Example 1. Let us consider $\mathbb{C}^2$. The computational basis is given by

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

and any vector $|v\rangle \in \mathbb{C}^2$ can be written as

$$|v\rangle = \alpha|0\rangle + \beta|1\rangle = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix},$$

for two numbers $\alpha, \beta \in \mathbb{C}$.

1.1.2 Linear functionals ("bras")

MAYBE BETTER NOT HAVING THIS? IT OVERLOADS THE BRAKET. On the other hand, $\mathcal{B}(\mathcal{H}) \simeq V \otimes V^*$ is a nice example.

We denote the set of $V \to \mathbb{C}$ linear functionals as V^*. For any linear functional $f \in V^*$ we write $\langle f|$ instead of f, and read *bra-f*. In the bra-ket notation the action of $f \in V^*$ on the vector $v \in V$, $f(v)$, is denoted by joining bra-f with ket-v: $\langle f|v\rangle = f(v)$. Recall that if V is d-dimensional, then so is V^*, and if we fix a basis $v_0, \ldots, v_{d-1}$ on V, then there is a basis $f_0, \ldots, f_{d-1}$ of V^*, called the *dual basis*, such that $f_i(v_j) = \delta_{ij}$. Here, and in the rest of the notes, δ denotes the Kroenecker delta,

$$\delta_{ij} = \begin{cases} 1 & \text{if i=j,} \\ 0 & \text{else.} \end{cases}$$

The vectors in this basis, dual to the computational basis, are denoted by $\langle 0|, \langle 1|, \ldots, \langle d-1|$ instead of $\langle f_0|, \ldots, \langle f_{d-1}|$. Using this basis, any linear functional $\langle f| \in V^*$ can be written as

$$\langle f| = f_0\langle 0| + f_1\langle 1| + \cdots + f_{d-1}\langle d-1|, \tag{2}$$

with $f_0, \ldots, f_{d-1} \in \mathbb{C}$ numbers that are uniquely determined by $\langle f|$ and our choice of the basis. Again, we can write this equation in matrix notation; we will think of linear functionals as row vectors. The dual of the computational basis is given by

$$\langle 0| = \begin{pmatrix} 1 & 0 & \ldots & 0 \end{pmatrix}, \quad \ldots, \quad \langle d-1| = \begin{pmatrix} 0 & 0 & \ldots & 1 \end{pmatrix},$$

and Eq. (2) becomes

$$\langle f| = f_0 \begin{pmatrix} 1 & 0 & \ldots & 0 \end{pmatrix} + \cdots + f_{d-1} \begin{pmatrix} 0 & 0 & \ldots & 1 \end{pmatrix} = \begin{pmatrix} f_0 & f_1 & \ldots & f_{d-1} \end{pmatrix}.$$

Given a vector $|v\rangle \in V$ and a linear functional $\langle f| \in V^*$, we can write

$$\langle f|v\rangle = \left(\sum_i f_i\langle i|\right)\left(\sum_j v_j|j\rangle\right) = \sum_{ij} f_i v_j \langle i|j\rangle = \sum_i f_i v_i.$$

This can also be expressed as the following matrix multiplication:

$$\langle f|v\rangle = \begin{pmatrix} f_0 & \ldots & f_{d-1} \end{pmatrix} \cdot \begin{pmatrix} v_0 \\ \vdots \\ v_{d-1} \end{pmatrix} = \sum_i f_i v_i.$$

1.1.3 Linear operators (matrices)

Linear operators act from the left: $A : V \to W$ acts as $|v\rangle \mapsto A|v\rangle \equiv |Av\rangle$. Due to linearity A is uniquely determined by its action on the basis vectors. We write $A|i\rangle_V = \sum_j A_{ji}|j\rangle_W$; as usual, we can think of A as a matrix, where the k^{th} column is the vector $A|k-1\rangle$:

$$A = \begin{pmatrix} A|0\rangle & A|1\rangle & \cdots & A|n-1\rangle \end{pmatrix} = \begin{pmatrix} A_{00} & A_{01} & \cdots & A_{0,n-1} \\ A_{10} & A_{11} & \cdots & A_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m-1,0} & A_{m-1,1} & \cdots & A_{m-1,n-1} \end{pmatrix}$$

Then the action of A on the vector v is described by the usual matrix-vector multiplication: $A|v\rangle = \sum_i Av_i|i\rangle = \sum_{ji} A_{ji}v_i \cdot |j\rangle$, i.e.,

$$A|v\rangle = \begin{pmatrix} A_{00} & A_{01} & \cdots & A_{0,n-1} \\ A_{10} & A_{11} & \cdots & A_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m-1,0} & A_{m-1,1} & \cdots & A_{m-1,n-1} \end{pmatrix} \cdot \begin{pmatrix} v_0 \\ v_1 \\ \vdots \\ v_{n-1} \end{pmatrix} = \begin{pmatrix} \sum_i A_{0i}v_i \\ \sum_i A_{1i}v_i \\ \vdots \\ \sum_i A_{m-1,i}v_i \end{pmatrix}.$$

1.1.4 Scalar product

1.1.5 Adjoint

1.1.6 Eigenvalues

1.2 Positivity

Definition 1 (Positivity)**.** *A matrix $A \in \mathcal{B}(\mathcal{H})$ is* positive semidefinite *if $\langle\Psi|A|\Psi\rangle \geq 0$ for all $|\Psi\rangle \in \mathcal{H}$. It is* positive definite *if $\langle\Psi|A|\Psi\rangle > 0$ for all $|\Psi\rangle \in \mathcal{H}$.*

Theorem 1. *Let $\mathcal{H}$ be a Hilbert space, $A \in \mathcal{B}(\mathcal{H})$ be a matrix. The following are equivalent.*

- *A is positive semidefinite (positive definite)*
- *A is Hermitian and it has only non-negative (positive) eigenvalues*
- *There is a(n invertible) matrix $X \in \mathcal{B}(\mathcal{H})$ such that $A = X^\dagger X$.*

Proof. Exercise. □

1.3 Tensor product

1.3.1 The maximally entangled state

Let $\mathcal{H} = \mathbb{C}^d$ and consider the state $|\Omega\rangle = \frac{1}{\sqrt{d}}\sum_{i=0}^{d-1}|ii\rangle \in \mathcal{H}\otimes\mathcal{H}$. This state is called the maximally entangled state and it has the following properties.

Proposition 1.

$$(\mathrm{Id}\otimes \mathrm{Tr})(|\Omega\rangle\langle\Omega|) = (\mathrm{Tr}\otimes\mathrm{Id})(|\Omega\rangle\langle\Omega|) = \mathrm{Id}.$$

Proof. Explicit calculation; exercise. □

Proposition 2. *Let $A \in \mathcal{M}_d$ be any square matrix, then*

$$(A\otimes\mathrm{Id})|\Omega\rangle = (\mathrm{Id}\otimes A^T)|\Omega\rangle.$$

Proof. Exercise: explicit calculation. □

More generally, let d and D be natural numbers, $|\Omega_d\rangle$ and $|\Omega_D\rangle$ be the d- and the D dimensional maximally entagled states. Then

Proposition 3. *Let $A \in \mathcal{M}_{d\times D}$ be a $d \times D$ matrix, then*

$$(A \otimes \mathrm{Id})|\Omega_D\rangle = (\mathrm{Id} \otimes A^T)|\Omega_d\rangle.$$

Proof. Exercise: explicit calculation. □

A special case of this equation is if $d = 1$. Then $|\Omega_d\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d \simeq \mathbb{C}$ is just 1, and the $d \times D$ matrix is just a linear functional. Writing out this case explicitly, we obtain

Proposition 4. *Let $|\Psi\rangle \in \mathbb{C}^d$ and $|\Omega\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ be the maximally entangled state. Then*

$$(\langle\Psi| \otimes \mathrm{Id})|\Omega\rangle = |\bar{\Psi}\rangle,$$

the complex conjugate of $|\Psi\rangle$.

Proof. Exercise: explicit calculation. □

2 Quantum mechanics

2.1 The rules of quantum mechanics

- States
- Time evolution
- Measurements: destructive, then projective

2.2 The rules of quantum mechanics: mixed states

- Mixed states
- POVMs
- CP maps

Definition 2 (Ensemble)**.** *Let $\mathcal{H}$ be a Hilbert space. En* ensemble *is a collection of normalized vectors with probabilities: it is given by $n \in \mathbb{N}$ and $\big((p_i, |\Psi_i\rangle)\big)_{i=1}^n \subseteq (\mathbb{R}_{\geq 0} \times \mathcal{H})^n$ such that $\|\Psi_i\| = 1$ for all $i = 1, \ldots, n$ and $\sum_{i=1}^n p_i = 1$.*

2.2.1 Mixed states

Definition 3 (Density matrix)**.** *Let $\mathcal{H}$ be a Hilbert space. A* state *or a* density matrix *on $\mathcal{H}$ is a matrix $\rho \in \mathcal{B}(\mathcal{H})$ such that $\operatorname{Tr}\rho = 1$, and $\rho \geq 0$, i.e., it is positive semidefinite. The state ρ is called* pure *if it is rank-one.*

Proposition 5. *Let $\mathcal{H}$ be a Hilbert space and $\rho \in \mathcal{B}(\mathcal{H})$ be a pure state. Then there is $|\Psi\rangle \in \mathcal{H}$, $\|\Psi\| = 1$, such that $\rho = |\Psi\rangle\langle\Psi|$.*

Proof. ρ is rank-one and Hermitian, thus it can be written as $\rho = |\Psi\rangle\langle\Psi|$. $\operatorname{Tr}\rho = \operatorname{Tr}|\Psi\rangle\langle\Psi| = \langle\Psi|\Psi\rangle = \|\Psi\|^2 = 1$. □

2.2.2 CP maps

Definition 4. *Let $\mathcal{H}$ and $\mathcal{K}$ be two Hilbert spaces. Let $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ be linear. We say that T is* trace preserving *if $\operatorname{Tr} T(\rho) = \operatorname{Tr} \rho$, for all $\rho \in \mathcal{B}(\mathcal{H})$.*

Definition 5. *Let $\mathcal{H}$ and $\mathcal{K}$ be two Hilbert spaces. Let $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ be linear. We say that T is a Kraus map if there are $A_i \in \mathcal{B}(\mathcal{H} \to \mathcal{K})$ such that*

$$T(\rho) = \sum_{i=1}^{n} A_i \rho A_i^\dagger. \tag{3}$$

The operators A_i are called the Kraus operators, and Eq. (3) the Kraus representation of T.

Proposition 6. *Let $\mathcal{H}$ and $\mathcal{K}$ be two Hilbert spaces. Let $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ be a Kraus map with Kraus operators $\{A_i\}_{i=1}^n$. Then T is trace preserving iff $\sum_i A_i^\dagger A_i = \mathrm{Id}_{\mathcal{H}}$.*

Proof. Exercise. □

Definition 6. *Let $\mathcal{H}$ and $\mathcal{K}$ be two Hilbert spaces. Let $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ be linear. We say that*

- *T is* positivity preserving, *or simply* positive, *if $T(\rho) \geq 0$ for all $\rho \in \mathcal{B}(\mathcal{H})$.*
- *T is n-positive if $T \otimes \mathrm{Id} : \mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathbb{C}^n) \to \mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathbb{C}^n)$ is positive.*
- *T is* completely positive (CP) *if it is n-positive for all $n \in \mathbb{N}$.*

Theorem 2. *Let $\mathcal{H}$ and $\mathcal{K}$ be two Hilbert spaces, and $|\Omega\rangle \in \mathcal{H} \otimes \mathcal{H}$ be a maximally entangled state. Let $T : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$ be linear. The following are equivalent:*

1. *T is completely positive.*
2. *T is $\dim(H)$-positive.*
3. *$(\mathrm{Id} \otimes T)(|\Omega\rangle\langle\Omega|) \geq 0$.*
4. *T is Kraus.*

Proof. The implication 1 ⇒ 2 ⇒ 3 is by definition. To check 3 ⇒ 4, note that

$$T(\rho) = (\operatorname{Tr} \otimes T)((\mathrm{Id} \otimes \rho)|\Omega\rangle\langle\Omega|) = (\operatorname{Tr} \otimes T)((\rho^T \otimes \mathrm{Id})|\Omega\rangle\langle\Omega|) = (\operatorname{Tr} \otimes \mathrm{Id})((\rho^T \otimes \mathrm{Id})X),$$

where $X = (\mathrm{Id} \otimes T)(|\Omega\rangle\langle\Omega|) \geq 0$. As $X \geq 0$, we can write it as $\sum_i |\Psi_i\rangle\langle\Psi_i|$ for some (unnormalized) vectors $|\Psi_i\rangle \in \mathcal{H} \otimes \mathcal{K}$. But $|\Psi_i\rangle$ then in turn can be written as $(\mathrm{Id} \otimes A_i)|\Omega\rangle$, and thus X can be written as

$$X = \sum_i (\mathrm{Id} \otimes A_i)|\Omega\rangle\langle\Omega|(\mathrm{Id} \otimes A_i^\dagger).$$

Therefore

$$T(\rho) = (\operatorname{Tr} \otimes \mathrm{Id})((\rho^T \otimes \mathrm{Id})X) = \sum_i (\operatorname{Tr} \otimes \mathrm{Id})((\rho^T \otimes A_i)|\Omega\rangle\langle\Omega|(\mathrm{Id} \otimes A_i^\dagger)).$$

Using now Proposition 2 in the form $(\rho^T \otimes \mathrm{Id})|\Omega\rangle = (\mathrm{Id} \otimes \rho)|\Omega\rangle$, and Proposition 1, we obtain

$$T(\rho) = \sum_i (\operatorname{Tr} \otimes \mathrm{Id})((\mathrm{Id} \otimes A_i \rho)|\Omega\rangle\langle\Omega|(\mathrm{Id} \otimes A_i^\dagger)) = \sum_i A_i \rho A_i^\dagger.$$

Finally to check 4 ⇒ 1, let the Kraus operators of T be A_i, $n \in \mathbb{N}$ arbitrary, and $\rho \in \mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathbb{C}^n)$ be positive. Then

$$(T \otimes \mathrm{Id}_{\mathcal{B}(\mathbb{C}^n)})(\rho) = \sum_i (A_i \otimes \mathrm{Id}_{\mathbb{C}^n})\rho(A_i^\dagger \otimes \mathrm{Id}_{\mathbb{C}^n}) \geq 0,$$

i.e., T is completely positive. □

Proposition 7. *The transposition map is positive but not completely positive.*

Proof. Exercise. □

3 Entanglement

Definition 7 (Entanglement of pure states)**.** *A pure state $|\Psi\rangle \in \mathcal{H} \otimes \mathcal{K}$ is called entangled if it is not a product.*

Definition 8 (Separability and entanglement)**.** *A state $\rho \in \mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{K})$ is called* separable *if there are density matrices $\mu_i \in \mathcal{B}(\mathcal{H})$, $\nu_i \in \mathcal{B}(\mathcal{H})$, $i = 1 \dots n$, and a probability distribution $p \in \mathbb{R}^n$ such that*

$$\rho = \sum_{i=1}^{n} p_i \cdot \mu_i \otimes \nu_i.$$

The state ρ is called entangled *if it is not separable.*

Proposition 8. *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces, $\rho \in \mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{K})$ be a* separable *state. Let T be a positive but not necessarily CP map. Then $(\mathrm{Id} \otimes T)(\rho) \geq 0$.*

Proof. As ρ is separable, there are density matrices $\mu_i \in \mathcal{B}(\mathcal{H})$, $\nu_i \in \mathcal{B}(\mathcal{H})$, $i = 1 \dots n$, and a probability distribution $p \in \mathbb{R}^n$ such that

$$\rho = \sum_{i=1}^{n} p_i \cdot \mu_i \otimes \nu_i.$$

Then

$$(\mathrm{Id} \otimes T)(\rho) = \sum_{i=1}^{n} p_i \cdot \mu_i \otimes T(\nu_i) \geq 0,$$

as p_i, μ_i and $T(\nu_i)$ are all positive. □

This can be used to detect entanglement: if T is a positive but not necessarily CP map, and ρ is a density matrix in a bipartite space, then if $(\mathrm{Id} \otimes T)(\rho) \ngeq 0$, then we can conclude, apart from the fact that T is not CP, that ρ is entangled. A standard example is the transposition map.

4 Teleportation and dense coding

Let us assume that Alice and Bob share a maximally entangled state $|\Omega\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$. Assume Alice has an additional, unknown, state $|\Psi\rangle \in \mathbb{C}^2$. Quantum state teleportation is a protocol where Alice transfers the state $|\Psi\rangle$ to Bob using only two bits of classical communication. Even though it is required that at some point in time there is a qubit communication between Alice and Bob, one half of the maximally entangled state, this communication can happen Alice even gets the state $|\Psi\rangle$[1], thus it carries no information about $|\Psi\rangle$.

To describe the protocol, let us define the Bell basis. Let $|\Omega\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$ be the maximally entangled state, $|\Omega\rangle = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$. Let us define $|\Omega_{\alpha\beta}\rangle$ as $|\Omega_{\alpha\beta}\rangle = (\mathrm{Id} \otimes X^\alpha Z^\beta)|\Omega\rangle$, $\alpha, \beta \in \{0, 1\}$. Explicitly,

$$\begin{aligned}
|\Omega_{00}\rangle &= (\mathrm{Id} \otimes \mathrm{Id})|\Omega\rangle = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle, \\
|\Omega_{01}\rangle &= (\mathrm{Id} \otimes Z)|\Omega\rangle = \frac{1}{\sqrt{2}}|00\rangle - \frac{1}{\sqrt{2}}|11\rangle, \\
|\Omega_{10}\rangle &= (\mathrm{Id} \otimes X)|\Omega\rangle = \frac{1}{\sqrt{2}}|01\rangle + \frac{1}{\sqrt{2}}|10\rangle, \\
|\Omega_{11}\rangle &= (\mathrm{Id} \otimes XZ)|\Omega\rangle = \frac{1}{\sqrt{2}}|01\rangle - \frac{1}{\sqrt{2}}|10\rangle.
\end{aligned}$$

At the beginning of the protocol, Alice and Bob have the state $|\Psi\rangle \otimes |\Omega\rangle \in \mathcal{H}_{A'} \otimes \mathcal{H}_A \otimes \mathcal{H}_B$; in Alice's lab the qubits A' and A, in Bob's lab the qubit B. The protocol consists of the following steps:

[1] and actually it might be Bob who sends a qubit to Alice

1. Alice measures the qubits A and A' in the Bell basis, and obtains outcome $(\alpha, \beta) \in \{0,1\}^2$.
2. Alice sends the classical bits α and β to Bob
3. Bob applies on the particle B the operator $x^\alpha Z^\beta$.

We will show that at the end of the protocol, Bob obtains the state $|\Psi\rangle$. For that, note the post-measurement state after Alice's measurement, up to normalization, is

$$|\Phi_{\alpha\beta}\rangle = (|\Omega_{\alpha\beta}\rangle\langle\Omega_{\alpha\beta}| \otimes \mathrm{Id})(|\Psi\rangle \otimes |\Omega\rangle),$$

where the order of the tensor components is $A'AB$. As the measurement is rank-1, the outcome is of the form $|\Phi_{\alpha\beta}\rangle = |\Omega_{\alpha\beta}\rangle \otimes |\Psi_{\alpha\beta}\rangle$, where

$$|\Psi_{\alpha\beta}\rangle = (\langle\Omega_{\alpha\beta}| \otimes \mathrm{Id})(|\Psi\rangle \otimes |\Omega\rangle).$$

Let us write out now the definition of $|\Omega_{\alpha\beta}\rangle$ in this formula:

$$|\Psi_{\alpha\beta}\rangle = (\langle\Omega_{\alpha\beta}| \otimes \mathrm{Id})(|\Psi\rangle \otimes |\Omega\rangle) = (\langle\Omega| \otimes \mathrm{Id})(\mathrm{Id} \otimes Z^\beta X^\alpha \otimes \mathrm{Id})(|\Psi\rangle \otimes |\Omega\rangle).$$

Using now Proposition 2, i.e., that the maximally entangled state satisfies $(A \otimes \mathrm{Id}) \cdot |\Omega\rangle = (\mathrm{Id} \otimes A^T) \cdot |\Omega\rangle$, for any square matrix A, we obtain that

$$|\Psi_{\alpha\beta}\rangle = (\langle\Omega| \otimes \mathrm{Id})(\mathrm{Id} \otimes Z^\beta X^\alpha \otimes \mathrm{Id})(|\Psi\rangle \otimes |\Omega\rangle) = (\langle\Omega| \otimes \mathrm{Id})(\mathrm{Id} \otimes \mathrm{Id} \otimes X^\alpha Z^\beta)(|\Psi\rangle \otimes |\Omega\rangle) = X^\alpha Z^\beta |\Psi_{00}\rangle,$$

where

$$|\Psi_{00}\rangle = (\langle\Omega| \otimes \mathrm{Id})(|\Psi\rangle \otimes |\Omega\rangle).$$

Using now Proposition 4 twice, we obtain that

$$|\Psi_{00}\rangle = (\langle\Omega| \otimes \mathrm{Id})(|\Psi\rangle \otimes |\Omega\rangle) = (\langle\bar{\Psi}| \otimes \mathrm{Id})|\Omega\rangle = |\Psi\rangle.$$

We thus obtain that the state after Alice's measurement, if she obtains ouctcome $\alpha\beta$, is

$$|\Phi_{\alpha\beta}\rangle = |\Omega_{\alpha\beta}\rangle \otimes X^\alpha Z^\beta |\Psi\rangle.$$

Therefore, if Bob applies $Z^\beta X^\alpha$ on his part of the state, he will obtain the state $|\Psi\rangle$ independent of Alice's side.

Break down the measurement into rotation and two single-qubit measurements

5 Schmidt decomposition

Definition 9 (Schmidt decomposition)**.** *Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces, $|\Psi\rangle \in \mathcal{H} \otimes \mathcal{K}$. A Schmidt decomposition of $|\Psi\rangle$ is a finite set $\{(\lambda_i, |l_i\rangle, |r_i\rangle)\}_{i=1}^n \subseteq \mathbb{C} \times \mathcal{H} \times \mathcal{K}$ such that*

$$|\Psi\rangle = \sum_{i=1}^n \lambda_i \cdot |l_i\rangle \otimes |r_i\rangle,$$

and such that

- $\lambda_i > 0\ \forall i = 1, \ldots, n$,
- $\langle l_i | l_j \rangle = \delta_{ij}$ *for all* $i, j = 1, \ldots, n$, *and*
- $\langle r_i | r_j \rangle = \delta_{ij}$ *for all* $i, j = 1, \ldots, n$.

The vectors $\{|l_i\rangle\}_{i=1}^n$ and $\{(|r_i\rangle\}_{i=1}^n$ are linearly independent, and thus a Schmidt decomposition is a minimal rank decomposition. Not every minimal rank decomposition is a Schmidt decomposition, as there the tensor components are not required to be orthonormal.

Example 2. Consider the following states.

- $|\Omega\rangle = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$. This is a Schmidt decomposition of $|\Omega\rangle$.
- The same state admits another Schmidt decomposition: $|\Omega\rangle = \frac{1}{\sqrt{2}}|++\rangle + \frac{1}{\sqrt{2}}|--\rangle$. So the Schmidt decomposition is not unique. Note: the lack of uniqueness is due to the degeneracy of its Schmidt values.
- $|\Psi\rangle = \frac{1}{\sqrt{2}}|+0\rangle + \frac{1}{\sqrt{2}}|11\rangle$ is not a Schmidt decomposition, as $\langle +|1\rangle \neq 0$.
- $|\Phi\rangle = \frac{1}{\sqrt{2}}|+0\rangle + \frac{1}{\sqrt{2}}|-1\rangle$ is a Schmidt decomposition. The basis on the left and on the right is not necessarily the same (they, in general, are even bases of different Hilbert spaces).

Theorem 3. *Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces, and $|\Psi\rangle \in \mathcal{H} \otimes \mathcal{K}$, $|\Psi\rangle \neq 0$. Then $|\Psi\rangle$ has a Schmidt decomposition.*

Proof. Let $\rho := \mathrm{Tr}_{\mathcal{K}} |\Psi\rangle\langle\Psi|$. This operator is positive semidefinite, and thus we can consider an eigen decomposition

$$\rho = \sum_i \lambda_i^2 \cdot |l_i\rangle\langle l_i|, \tag{4}$$

with $\lambda_i \in \mathbb{R}$, $\lambda_i \geq 0$ for all $i = 1 \dots \dim \mathcal{H}$. The vectors $\{|l_i\rangle\}_{i=1}^{\dim \mathcal{H}}$ form a basis of $\mathcal{H}$, we can thus write

$$|\Psi\rangle = \sum_i |l_i\rangle \otimes |\hat{r}_i\rangle,$$

for some $\{|\hat{r}_i\rangle\}_{i=1}^{\dim \mathcal{H}} \subseteq \mathcal{K}$. Expressing ρ with the help of this form of $|\Psi\rangle$, we obtain

$$\rho = \mathrm{Tr}_{\mathcal{K}} |\Psi\rangle\langle\Psi| = \sum_{i,j=1}^{\dim \mathcal{H}} |l_i\rangle\langle l_j| \cdot \mathrm{Tr}\{|\hat{r}_i\rangle\langle\hat{r}_j|\} = \sum_{i,j=1}^{\dim \mathcal{H}} |l_i\rangle\langle l_j| \cdot \langle\hat{r}_j|\hat{r}_i\rangle.$$

Comparing this to Eq. (4), we obtain $\langle\hat{r}_j|\hat{r}_i\rangle = \delta_{ij}\lambda_i^2$. For $\lambda_i \neq 0$ we can thus define $|r_i\rangle = \lambda_i^{-1}|\hat{r}_i\rangle$. These states are, by definition, orthonormal, and

$$|\Psi\rangle = \sum_{i:\lambda_i \neq 0} \lambda_i \cdot |l_i\rangle \otimes |r_i\rangle$$

is a Schmidt decomposition of $|\Psi\rangle$. □

Consider $|\Psi\rangle \in \mathcal{H} \otimes \mathcal{K}$, and set $\rho_{\mathcal{H}} := \mathrm{Tr}_{\mathcal{K}} |\Psi\rangle\langle\Psi|$ and $\rho_{\mathcal{K}} := \mathrm{Tr}_{\mathcal{H}} |\Psi\rangle\langle\Psi|$. These two matrices – the reduced densities of the state $|\Psi\rangle$ – have the same spectrum, except for the eigenvalue 0: they are both the squared Schmidt values of $|\Psi\rangle$, plus maybe additional zeros.

6 LOCC

The laws of physics are inherently local (gravity might be an interesting exception): that is, only physical systems that are close to each other tend to interact. In our description, this means that if the two components of a bipartite quantum system AB are "far from each other", as for example, a pair of atoms located in two different labs, then there are operations that are "easy" to do and operations that are "hard" to do. While the exact boundary between easy and hard operations depend on the actual physical system, we can assume that *local* operations are easy to do: unitaries that are of the form $U_A \otimes \mathrm{Id}_B$ or $\mathrm{Id}_A \otimes U_B$, and measurements of the form $\{M_i^A \otimes \mathrm{Id}_B\}_{i=1}^n$ and $\{\mathrm{Id} \otimes M_i^B\}_{i=1}^n$. Applying a sequence of "easy" operations is considered to be "easy" again.

LOCC operations (local operations assisted by classical communication) consist of these local operations, but where we assume in addition that the parties can communicate with each other (by classical means): they can share their measurement outcomes, and thus the subsequent unitaries and measurements of all parties might depend on the measurement outcomes. We have already seen such a protocol: teleportation.

In this section, we will try to understand when a state $|\Phi\rangle$ can be transformed into another state $|\Psi\rangle$ via LOCC operations; if it is possible, we write $|\Phi\rangle \xrightarrow{LOCC} |\Psi\rangle$. In this case $|\Phi\rangle$ is a better resource for carrying out any task where the parties are restricted by locality, but where they are allowed to communicate: for example, when they want to carry out an experiment on a collection of entangled particles where the particles are located in different labs.

Theorem 4 (Ky-Fan)**.** *Let $\mathcal{H}$ be a Hilbert space, $k \in \mathbb{N}$, and*

$$\mathcal{P}_k(\mathcal{H}) := \left\{P \in \mathcal{B}(\mathcal{H}) \middle| P = P^\dagger = P^2,\ \mathrm{Tr}(P) = k\right\}.$$

Let $\rho \in \mathcal{B}(\mathcal{H})$ Hermitian, and

$$\rho = \sum_{i=1}^{\dim \mathcal{H}} \lambda_i |\Phi_i\rangle\langle\Phi_i|$$

be its eigen decomposition, with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{\dim \mathcal{H}}$. Then

$$\sup_{P \in \mathcal{P}_k} \mathrm{Tr}\{\rho P\} = \sum_{i=1}^{k} \lambda_i.$$

In other words, given a Hermitian operator ρ, the maximum of $\mathrm{Tr}\{\rho P\}$, where P is a rank-k Hermitian projector, is the sum of the k largest eigenvalues of ρ.

Proof. Let $Q = \sum_i |\Phi_i\rangle\langle\Phi_i|$. Then $Q \in \mathcal{P}_k$ and thus

$$\sup_{P \in \mathcal{P}_k} \mathrm{Tr}\{\rho P\} \geq \mathrm{Tr}\{Q\rho\} = \sum_{i=1}^{k} \langle\Phi_i|\rho|\Phi_i\rangle = \sum_{i=1}^{k} \lambda_i. \tag{5}$$

To obtain the opposite bound, use the spectral decomposition of ρ to write

$$\mathrm{Tr}\{\rho P\} = \sum_{i=1}^{\dim \mathcal{H}} \lambda_i \,\mathrm{Tr}\left\{P|\Phi_i\rangle\langle\Phi_i|\right\} = \sum_{i=1}^{\dim \mathcal{H}} \lambda_i \cdot \langle\Phi_i|P|\Phi_i\rangle = \sum_{i=1}^{\dim \mathcal{H}} \lambda_i \cdot \omega_i,$$

where $\omega_i = \langle\Phi_i|P|\Phi_i\rangle$ satisfies $0 \leq \omega_i \leq 1$ and $\sum_{i=1}^k \omega_i = k$. We thus obtain that

$$\sup_{P \in \mathcal{P}_k} \mathrm{Tr}\{\rho P\} \leq \sup_{\substack{\omega \in [0,1]^{\dim \mathcal{H}} \\ \sum_i \omega_i = k}} \sum_{i=1}^{\dim \mathcal{H}} \lambda_i \cdot \omega_i.$$

The maximum of the r.h.s. is reached at $\omega_1 = \cdots = \omega_k = 1$, and $\omega_{k+1} = \cdots = \omega_{\dim \mathcal{H}} = 0$, and its value is $\sum_{i=1}^k \lambda_i$. Therefore

$$\sup_{P \in \mathcal{P}_k} \mathrm{Tr}\{\rho P\} \leq \sup_{\substack{\omega \in [0,1]^{\dim \mathcal{H}} \\ \sum_i \omega_i = k}} \sum_{i=1}^{\dim \mathcal{H}} \lambda_i \cdot \omega_i = \sum_{i=1}^{k} \lambda_i.$$

This, together with Eq. (5) is the desired statement. □

Definition 10 (Majorization)**.** *Let $p, q \in \mathbb{R}^n$ be probability distributions. Let $p^{\downarrow}$ ($q^{\downarrow}$) be the vector obtained by listing the entries of p (q) in descending order. We say that p is majorized by q, and write $p \preccurlyeq q$, if for all $k = 1, \dots, n$,*

$$\sum_{i=1}^{k} p_i^{\downarrow} \leq \sum_{i=1}^{k} q_i^{\downarrow}.$$

TODO: Here it is better to define Schmidt coefficient with zero in it.

Proposition 9. *Let $|\Psi\rangle, |\Phi\rangle \in \mathcal{H} \otimes \mathcal{K}$. Let the squared Schmidt coefficients of $|\Psi\rangle$ be p, and the squared Schmidt coefficients of $|\Phi\rangle$ be q. If $|\Psi\rangle \xrightarrow{LOCC} |\Psi\rangle$, then $q \preccurlyeq p$.*

Proof. Let $\rho = \mathrm{Tr}_{\mathcal{H}} |\Psi\rangle\langle\Psi|$ and $\eta = \mathrm{Tr}_{\mathcal{H}} |\Psi\rangle\langle\Psi|$. Then the eigenvalues of ρ are precisely the entries of p and the eigenvalues of η are the entries of q. Using MISSING, $|\Psi\rangle \xrightarrow{LOCC} |\Psi\rangle$ if and only if

$$\sqrt{r_i}|\Phi\rangle = M_i \otimes U_i |\Psi\rangle,$$

for a measurement $\{M_i\}_{i=1}^n$, a probability distribution $r \in \mathbb{R}^n$ and some unitaries $\{U_i\}_{i=1}^n$. Therefore

$$r_i \eta = U_i \rho_i U_i^{\dagger},$$

where $\rho_i = \mathrm{Tr}_{\mathcal{H}}\{(M_i^{\dagger} M_i \otimes \mathrm{Id})|\Psi\rangle\langle\Psi|\}$, and thus

$$\sup_{P \in \mathcal{P}_k} \mathrm{Tr}\{P\eta\} = \sup_{P \in \mathcal{P}_k} \sum_i \mathrm{Tr}\{r_i P \eta\} = \sup_{P \in \mathcal{P}_k} \sum_i \mathrm{Tr}\{P U_i \rho_i U_i^{\dagger}\} \leq \sum_i \sup_{P \in \mathcal{P}_k} \mathrm{Tr}\{U_i^{\dagger} P U_i \rho\} = \sum_i \sup_{P \in \mathcal{P}_k} \mathrm{Tr}\{P \rho_i\} = \sup_{P \in \mathcal{P}_k} \mathrm{Tr}\{P\rho\}.$$

Using now Theorem 4, we obtain that $p \preccurlyeq q$. □

Let us try to understand now a *sufficient* criterion for $|\Psi\rangle \xrightarrow{LOCC} |\Phi\rangle$.

Example 3. Let $|\Psi\rangle = \sqrt{1/2}|00\rangle + \sqrt{1/2}|11\rangle$ and $|\Phi\rangle = \sqrt{2/3}|00\rangle + \sqrt{1/3}|11\rangle$. Notice that

$$M_0 = \begin{pmatrix} \sqrt{2/3} & 0 \\ 0 & \sqrt{1/3} \end{pmatrix},$$

is a matrix such that $(M_0 \otimes \mathrm{Id})|\Psi\rangle = |\Phi\rangle$. Moreover, we can complete it to a measurement, for example, with the matrix

$$M_1 = \begin{pmatrix} \sqrt{1/3} & 0 \\ 0 & \sqrt{2/3} \end{pmatrix},$$

i.e., $M_0^{\dagger} M_0 + M_1^{\dagger} M_1 = \mathrm{Id}$. Notice that $(M_1 \otimes \mathrm{Id})|\Psi\rangle = X \otimes X|\Phi\rangle$. Therefore the measurement $N_i = X^i M_i$ ($i = 1, 2$) is such that

$$M_i \otimes X^i |\Psi\rangle = \sqrt{p_i}|\Phi\rangle,$$

for some probability distribution p_i, i.e., $|\Psi\rangle \xrightarrow{LOCC} |\Phi\rangle$.

The essence of this example is that the squared Schmidt coefficients of the states are related to each other as

$$\begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} = \frac{1}{2} \cdot (\mathrm{Id} + X) \cdot \begin{pmatrix} 1/3 \\ 2/3 \end{pmatrix}.$$

This construction can be easily generalized for the case when the squared Schmidt coefficients of $|\Psi\rangle$ can be obtained by mixing permutations of the squared Schmidt coefficients of $|\Phi\rangle$. Before we show how to obtain the LOCC operations for this generalized case, notice that this constraint is equivalent with majorization:

Lemma 1. *Let $p, q \in \mathbb{R}^n$ be probability distributions. If $p \preccurlyeq q$, then there is a probability distribution $r \in \mathbb{R}^{n!}$ such that*

$$p = \sum_{\pi \in S_n} r_\pi \pi q,$$

where $S_n \subseteq \mathcal{B}(\mathbb{R}^n)$ is the set of $n \times n$ permutation matrices.

Proof. Exercise. □

Again, number of non-zero Schmidt coeffs might be different!

Proposition 10. *Let $|\Psi\rangle, |\Phi\rangle \in \mathcal{H} \otimes \mathcal{K}$. Let the squared Schmidt coefficients of $|\Psi\rangle$ be p, and the squared Schmidt coefficients of $|\Phi\rangle$ be q. If $p \preccurlyeq q$, then $|\Psi\rangle \xrightarrow{LOCC} |\Phi\rangle$.*

Proof. Using Lemma 1, we obtain that there is a probability distribution r such that $p = \sum_{\pi \in S_n} r_\pi \pi q$, where S_n is the group of $n \times n$ permutation matrices. Let the Schmidt decomposition of the states be

$$|\Psi\rangle = \sum_{i=1}^n p_i \cdot |l_i\rangle \otimes |r_i\rangle$$
$$|\Phi\rangle = \sum_{i=1}^n q_i \cdot |\hat{l}_i\rangle \otimes |\hat{r}_i\rangle,$$

and let us define operators M_π as

$$M_\pi = \sum_i \left(r_\pi \frac{q_{\pi(i)}}{p_i} \right)^{1/2} |\hat{l}_{\pi(i)}\rangle\langle l_i|.$$

These operators form a POVM as

$$\sum_\pi M_\pi^\dagger M_\pi = \sum_{ij} \sum_\pi r_\pi \frac{q_{\pi(i)}}{p_i} |l_i\rangle\langle \hat{l}_{\pi(i)}|\hat{l}_{\pi(j)}\rangle\langle l_j| = \sum_i \sum_\pi r_\pi \frac{q_{\pi(i)}}{p_i} |l_i\rangle\langle l_i| = \sum_i |l_i\rangle\langle l_i| = \mathrm{Id}.$$

Note as well that

$$(M_\pi \otimes \mathrm{Id})|\Psi\rangle = \sum_i \sqrt{r_\pi q_{\pi(i)}} |\hat{l}_{\pi(i)}\rangle \otimes |r_i\rangle,$$

and thus, setting $V_\pi = \sum_i |r_{\pi(i)}\rangle\langle r_i|$, V_π is a unitary and

$$(M_\pi \otimes V_\pi)|\Psi\rangle = \sum_i \sqrt{r_\pi q_{\pi(i)}} |\hat{l}_{\pi(i)}\rangle \otimes |r_{\pi(i)}\rangle = \sqrt{r_\pi}|\Phi\rangle,$$

i.e., we have shown that $|\Psi\rangle \xrightarrow{LOCC} |\Phi\rangle$. □

We have thus seen in Proposition 9 and Proposition 10 that

Theorem 5. *Let $|\Psi\rangle, |\Phi\rangle \in \mathcal{H} \otimes \mathcal{K}$. Let the squared Schmidt coefficients of $|\Psi\rangle$ be p, and the squared Schmidt coefficients of $|\Phi\rangle$ be q. Then $|\Psi\rangle \xrightarrow{LOCC} |\Phi\rangle$ iff $p \preccurlyeq q$.*