Introduction to Quantum Information and Computation

Agrawal, Vaibhav
vaibhav.agrawal@students.iiit.ac.in

Bishnoi, Pranjali
pranjali.bishnoi@students.iiit.ac.in

February 6, 2023

Contents

1 Linear Algebra: The Hilbert Space

- In quantum mechanics the state of a physical system is represented by a vector in a *Hilbert Space*: a complex vector space with an inner product.
- Every d dimensional Hilbert space is isomorphic to the complex vector space $\mathcal{C}^d$, whose elements are d-tuples of complex numbers.
- Here we will be dealing only with finite dimensional vector spaces.
- The *Dirac* notation will be used, in which the vectors are denoted by $|v\rangle$, called a *ket*, where v is some symbol which identifies the vector. The complex conjugate (also called dual) of this vector is denoted by $\langle v|$, called a *bra*.
- The inner product of the two vectors $|v\rangle$ and $|w\rangle$ is represented as $\langle v|w\rangle$. This is similar to the ordinary dot product $\vec{v}.\vec{w}$ except that here we take the complex conjugate of the vector on the left, and hence can be thought as $\vec{v}^*\vec{w}$.
- The inner product is conjugate in the first argument and linear in the second argument, i.e.,

$$\langle\psi_1|a\psi_2 + b\psi_3\rangle = a\langle\psi_1|\psi_2\rangle + b\langle\psi_1|\psi_3\rangle$$
$$\langle a\psi_1 + b\psi_2|\psi_3\rangle = a^*\langle\psi_1|\psi_3\rangle + b^*\langle\psi_2|\psi_3\rangle$$

- The inner product is conjugate symmetric:

$$(\langle\psi_1|\psi_2\rangle)^* = \langle\psi_2|\psi_1\rangle$$

- The inner product of a vector with itself is always greater than or equal to zero, i.e.

$$\langle\psi|\psi\rangle = |\psi|^2 \geq 0$$

1.1 The Hilbert space of dimension d

- A collection of linearly independent vectors $\{|a_i\rangle\}$ form a basis of $\mathcal{H}$ given any $|\psi\rangle$ can be written as a linear combination.

$$|\psi\rangle = \sum_j c_j|a_j\rangle$$

- The standard basis for the Hilbert space of d dimensions is given by:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ ... \\ 0 \end{pmatrix}, |1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ ... \\ 0 \end{pmatrix}, |2\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ ... \\ 0 \end{pmatrix}, ..., |d\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ ... \\ 1 \end{pmatrix}$$

2 Operators

- *Operators* are (usually) linear maps of the Hilbert space onto itself. If O is an operator, then for any $|\phi\rangle$, $A|\phi\rangle$ is another element in $\mathcal{H}$, and the linearity means that

$$A(b|\psi\rangle + c|\phi\rangle) = bA|\psi\rangle + cA|\phi\rangle$$

- Operators can be represented as *matrices*.

2.1 Dagger or adjoint

- The dagger or adjoint operation $\dagger$ when applied on a vector $|v\rangle$, transforms it into it's complex conjugate, denoted by $\langle v|$ or $|v\rangle^\dagger$:

$$|v\rangle = \begin{pmatrix} v_1 \\ v_2 \\ ... \\ v_d \end{pmatrix} \implies \langle v| = \begin{pmatrix} v_1^* & v_2^* & ... & v_d^* \end{pmatrix}$$

- The operation when applied on a matrix, results in the adjoint of the matrix. Hence $A^\dagger$ is called the adjoint of the operator A.

- Note that the dagger operation is antilinear in that the scalars are changed to their complex conjugates

2.2 Hermitian Operators

- A Hermitian or self-adjoint operator is defined by the property that $A = A^\dagger$. The eigenvalues of a Hermitian operator are real numbers.

3 Introduction to Quantum Mechanics

Quantum mechanics is a mathematical framework for the development of physical theories. On it's own quantum mechanics does not tell you what laws a physical system must obey, but it does provide a mathematical and conceptual framework for the development of such laws. These postulates provide a connection between the physical world and the mathematical formalism of quantum mechanics. The postulates have been derived after a long process of trial and (mostly) error. This involved a considerable amount of guessing and fumbling by the originators of the theory. Let us look at these postulates:

3.1 Postulate 1: State

Associated to any isolated physical system is a complex vector space with inner product (that is, a Hilbert space) known as the *state space* of the system. The system is completely described by its *state vector*, which is a **unit vector** in the system's state space.

A state can be understood as a snapshot of the system at any given instant. This snapshot would contain details about all the physical properties of the system, for example, temperature, pressure, etc. This information is contained in the *state vector*. Remember that this vector must be a unit vector.
The *state space* of any system is the vector space of all the state vectors that could correspond to that system. This is a complex vector space, i.e., a Hilbert space.

Let us begin by considering the simplest quantum mechanical system, a *qubit*. A qubit is any system whose state space is of dimension 2. We know from linear algebra, that the vectors

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

form an orthonormal basis for the two dimensional Hilbert space. This is to say that any given state vector $|\psi\rangle$ in the two dimensional Hilbert space can be written in the following form:

$$|\psi\rangle = a|0\rangle + b|1\rangle \tag{2}$$

where a and b are complex numbers.

Notice that **the unit-norm constraint on $|\psi\rangle$ implies that $|a|^2+|b|^2 = 1$.** This property is used extensively in our further study, as you will soon see. We say that $|\psi\rangle$ is a *superposition* of states $|0\rangle$ and $|1\rangle$, with amplitudes a and b respectively. This is somewhat similar to the superposition principle in wave mechanics.

3.2 Postulate 2: Evolution

How does the state, $|\psi\rangle$, of a quantum mechanical system change with time? The following postulate describes exactly this:

The evolution of a *closed* system is described by a unitary transformation. That is, the state $|\psi\rangle$ of the system at time t_1 is related to the state $|\psi'\rangle$ of the system at time t_2 by a unitary operator U which depends only on the times t_1 and t_2.

$$|\psi'\rangle = U|\psi\rangle$$

Note that unitary transformations have this beautiful property that they are *norm-preserving*, i.e., the norm of the transformed vector is same as the norm of the input vector (prove this!). We know that the state vector must be a unit vector. Unitary operations are consistent with this fact because the output state vector also has norm 1.

3.3 Postulate 3: Measurement

Suppose that we would like to learn something about the quantum state of a system, given by the state vector $|\psi\rangle$, where

$$|\psi\rangle = a|0\rangle + b|1\rangle \tag{3}$$

In other words, we wish to *measure* this system. Nature prevents us from learning anything about the amplitudes a and b if we have only one quantum measurement that we can perform on a copy of the state. Nature only allows us to measure *observables*. Observables are physical variables such as position or momentum of a particle. In quantum theory, we represent observables as **Hermitian operators**. This is partly because Hermitian operators have real eigenvalues, and the outcomes of any measurement are real values only. **We always measure an observable**.

According to *Dirac*, one of the founding fathers of quantum theory, *A measurement always causes the system to jump into an eigenstate of the dynamical variable (i.e., observable) that is being measured.*

Let us consider a physical variable (i.e., an observable), A with eigenvectors $|0\rangle$ and $|1\rangle$. Let us try to measure the state of a system represented by the following state vector:

$$|\psi\rangle = a|0\rangle + b|1\rangle \tag{4}$$

We say that we are performing a "measurement on the basis $\{|0\rangle, |1\rangle\}$" or a "measurement of the observable A ". The measurement postulate of quantum theory, also known as the *Born rule*, states that the system reduces to the state $|0\rangle$ with probability $|a|^2$ and to the state $|1\rangle$ with probability $|b|^2$. The outcome of the measurement is 0 for the first case and 1 for the second. Here 0 and 1 are not representing their numerical value. Instead, they are the eigenvalues corresponding to the eigenvectors $|0\rangle$ and $|1\rangle$ respectively.

To summarise, when a measurement is performed, the system is *thrown* into one of the eigenstates, say $|a\rangle$ of the observable A, and the outcome of the measurement is a, the eigenvalue corresponding to the eigenvector $|a\rangle$. If the amplitude of $|a\rangle$ in the initial state $|\psi\rangle$ was α, then this event happens with probability $|\alpha|^2$.
Notice that the amplitude α of $|a\rangle$ in the initial state $|\psi\rangle$ is nothing but the magnitude of projection of $|\psi\rangle$ onto the axis $|a\rangle$. Also the probability that the state collapses to $|a\rangle$ is equal to the square of the amplitude α^2. This means that the *closer* a vector is to one of the eigenstates, more likely it is to collapse to that state upon the measurement of the corresponding observable.
This goes to say that if the initial state *coincides* with one of the eigenstates of the observable, then the state after measurement of that observable should not change. This is indeed mathematically true, because we know that the eigenvectors of the observable A must be perpendicular to each other (because A is a Hermitian operator), and therefore if the initial state vector lies completely in the direction of one of the eigenstates, then it's components along the other eigenstates is zero, and hence there is zero probability of the state getting changed after the measurement.

3.3.1 Measurement Operators and Projective Measurements

There is a very neat mathematical representation of the fact that the initial state $|\psi\rangle$ *projects* itself in one of the eigenstates after the measurement of an observable.

Suppose A is an observable with eigenvectors $|0\rangle$, $|1\rangle$, ... $|d\rangle$, where d is the dimension of the Hilbert space associated with the system. We define a *measurement operator* corresponding to the eigenstate $|i\rangle$ as follows:

$$\Lambda^i = |i\rangle\langle i| \tag{5}$$

(The superscript i does not represent exponentiation!) We see that Λ^i is a projection operator, and it returns the projection of the input vector along $|i\rangle$(read about projection operators in linear algebra!).

We will soon come to the motivation for definition of this operator. But first let us consider some important things. The probability that the initial state $|\psi\rangle$ collapses to the eigenstate $|i\rangle$ is given by the square of the magnitude of the projection of $|\psi\rangle$ along $|i\rangle$. Mathematically, we can denote it as follows:

$$\begin{aligned}\text{Prob(i)} &= |(|i\rangle^\dagger|\psi\rangle)|^2 \\ &= |\langle i|\psi\rangle|^2\end{aligned}$$

Notice that we can rewrite this as:

$$\begin{aligned}\text{Prob(i)} &= \langle i|\psi\rangle^\dagger \cdot \langle i|\psi\rangle \\ &= \langle\psi|i\rangle\langle i|\psi\rangle \\ &= \langle\psi| \cdot (|i\rangle\langle i|) \cdot |\psi\rangle \\ &= \langle\psi|\Lambda^i|\psi\rangle\end{aligned}$$

Suppose that the state of the system after measurement is $|i\rangle$. We can represent this final state as follows:

$$\frac{\Lambda_i|\psi\rangle}{\sqrt{\langle\psi|\Lambda^i|\psi\rangle}}$$

This comes from the fact that the measurement operator is nothing but a projection operator, and hence the numerator of the expression represents the projection of $|\psi\rangle$ onto the eigenvector $|i\rangle$. The denominator has been chosen to normalise the final state vector, and is equal to the magnitude of the projection of the state vector along $|i\rangle$.

We can measure any orthonormal basis in this way. Such a measurement is known as *Von Neumann measurement.*

4 Composite Quantum States

A single physical qubit is an interesting physical system that exhibits unique quantum phenomena, but it is not particularly useful on it's own. We can perform interesting quantum information-processing tasks when we combine multiple qubits together. For understanding the mathematics of composite systems, we need to have knowledge of the tensor product operation.

Tensor Product: The tensor product of vectors, operators and Hilbert spaces plays an important role in quantum theory. For example, it is used to describe the state of multiple quantum systems. For two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ with dimensions d_A and d_B, we can define the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, with dimension $d_A \times d_B$. The calculation of tensor product is best understood through an example. For any two vectors $|\psi\rangle_A$ and $|\psi\rangle_B$ belonging to the Hilbert spaces $\mathcal{H}_A \otimes \mathcal{H}_B$, the tensor product $|\psi_A\rangle \otimes |\psi_B\rangle$ is defined as follows:

$$|\psi_A\rangle \otimes |\psi_B\rangle = \begin{pmatrix} \alpha_0 \\ \alpha_1 \end{pmatrix} \otimes \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \alpha_0 \cdot \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} \\ \alpha_1 \cdot \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} \alpha_0\beta_0 \\ \alpha_0\beta_1 \\ \alpha_0\beta_2 \\ \alpha_1\beta_0 \\ \alpha_1\beta_1 \\ \alpha_1\beta_2 \end{pmatrix}.$$

The tensor product Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ with the corresponding bases $\{|i\rangle_A\}$ and $\{|j\rangle_B\}$ is defined to be the Hilbert space spanned by the vectors $|i\rangle_A \otimes |j\rangle_B$:

$$\mathcal{H}_A \otimes \mathcal{H}_B := \text{span}\{|i\rangle_A \otimes |j\rangle_B : 0 \leq i \leq d_A - 1, 0 \leq j \leq d_B - 1\}$$

We will explore the other properties of tensor product as we study about composite systems in more detail.

For a two qubit system, consisting of qubits A and B with bases $\{|0\rangle_A, |1\rangle_A\}$

and $\{|0\rangle_B, |1\rangle_B\}$, we can define the basis for the composite system $A \otimes B$ to be $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, where $|00\rangle$ is the shorthand notation for $|0\rangle_A \otimes |0\rangle_B$, and similarly for the rest of the terms.
This goes to say that any given state vector $|\psi\rangle_{AB}$ belonging to $\mathcal{H}_A \otimes \mathcal{H}_B$ can be written as

$$\alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle \tag{6}$$

where α, β, γ and δ are amplitudes.

Here we have only provided a brief introduction to the theory of composite systems. We will come back to it after developing some more mathematical tools, which will be useful to study these systems in greater detail.

5 Noisy Quantum States

Until now we have only considered the case where we had full information about the prepared quantum state. Or at least, there had been no meaningful *probabilistic* value of the current state of the system (of course, we are talking about the *initial* state of the system, i.e., no measurement has been performed yet). But in some cases, we might have a probabilistic description of the state of the system. Our description of the state is then an *ensemble* $\mathcal{E}$ of quantum states where

$$\mathcal{E} \equiv \{p_X(x), |\psi_x\rangle\}_{x \in X}. \tag{7}$$

In the above, X is a random variable, and each realisation x occurs with a probability $p_X(x)$.
For example, $\{\{1/3, |0\rangle\}, \{2/3, |2\rangle\}\}$ is a valid ensemble, where the states $|0\rangle$ and $|2\rangle$ belong to a three dimensional Hilbert space with basis states $\{|0\rangle, |1\rangle, |2\rangle\}$.
Let us measure this noisy state (also called *mixed* state) along some basis A. We wish to find out the probability of obtaining the outcome j.

$$\begin{aligned} p_J(j) &= \sum_{x \in X} p_{J|X}(j|x) p_X(x) \\ &= \sum_{x \in X} \langle\psi_x|\Lambda^j|\psi_x\rangle p_X(x) \end{aligned}$$

The first equality comes from the fact that if we are given the state is $|\psi\rangle_x$, then we can easily determine the probability of obtaining outcome j.

At this point we define a very useful mathematical tool, called the *density operator*

Density operator: The density operator ρ corresponding to an ensemble $\mathcal{E} \equiv \{p_X(x), |\psi_x\rangle\}_{x \in X}$ is defined as follows:

$$\rho \equiv \sum_{x \in X} p_X(x)|\psi_x\rangle\langle\psi_x| \tag{8}$$

Density operator as the state: We have describe the density operator for a ensemble. We can view the density operator as representing the state of a quantum system. Consider the density operator of the pure state $|\psi\rangle$:

$$\rho = |\psi\rangle\langle\psi| \tag{9}$$

This is because the probability of the state being $|\psi\rangle$ is 1, and for all other states, the probability is 0.
For a mixed state, density operator can be viewed as the *expected value of the state*:

$$\rho = \mathbb{E}\{|\psi_X\rangle\langle\psi_X|\} \tag{10}$$

Notice that the subscript to $|\psi\rangle$ is X rather than x. This has been done intentionally because we always calculate the expected value of a random variable.

We would also like to define the trace operator.

Trace operator: The trace of a square operator A acting on Hilbert space $\mathcal{H}$ is defined as follows:

$$\mathrm{Tr}A \equiv \sum_i \langle i|A|i\rangle \tag{11}$$

where $\{|i\rangle\}$ is some complete, orthonormal basis for $\mathcal{H}$.

Trace has some interesting properties:

- Trace operation is cyclic, i.e., $\mathrm{Tr}\{ABC\} = \mathrm{Tr}\{BCA\} = \mathrm{Tr}\{CAB\}$
- Trace operation is invariant under the choice of the basis $\{|i\rangle\}$.
- For any vector $|\psi\rangle$ and operator M,

$$\langle\psi|M|\psi\rangle = \mathrm{Tr}\{M|\psi\rangle\langle\psi|\} \tag{12}$$

The trace of the density operator is 1. It is a positive semidefinite operator. You can try proving these properties.

Let us come back to the probability of obtaining outcome j on measurement A:

$$
\begin{aligned}
p_J(j) &= \sum_{x \in X} \langle \psi_x | \Lambda^j | \psi_x \rangle p_X(x) \\
&= \sum_{x \in X} \mathrm{Tr}\{\Lambda^i |\psi\rangle\langle\psi|\} p_X(x) \\
&= \mathrm{Tr}\{\Lambda^i \sum_{x \in X} p_X(x) |\psi\rangle\langle\psi|\} \\
&= \mathrm{Tr}\{\Lambda^i \rho\}
\end{aligned}
$$

Now suppose that we have performed measurement of A on this ensemble, and obtained the outcome j. We have a new ensemble:

$$
\mathcal{E}_j \equiv \left\{ p_{X|J}(x|j), \frac{\Lambda^j |\psi_x\rangle}{\sqrt{p_{J|X}(j|x)}} \right\}
$$

We would like to write the density operator for this new ensemble:

$$
\begin{aligned}
\rho_j &= \sum_{x \in X} p_{X|J}(x|j) \frac{\Lambda^j |\psi_x\rangle\langle\psi_x| \Lambda^j}{p_{J|X}(j|x)} \\
&= \Lambda^j \left(\sum_{x \in X} \frac{p_{X|J} x|j}{p_{J|X} j|x} |\psi_x\rangle\langle\psi_x| \right) \Lambda^j \\
&= \Lambda^j \left(\frac{\sum_{x \in X} p_X(x) |\psi_x\rangle\langle\psi_x|}{p_J(j)} \right) \Lambda^j \\
&= \Lambda^j \left(\frac{\rho}{p_J(j)} \right) \Lambda^j \\
&= \frac{\Lambda^j \rho \Lambda^j}{p_J(j)}
\end{aligned}
$$

The second equality follows from the Baye's rule in probability.

6 The General Measurement Postulate

When we described the measurement postulate earlier, we kind of described a *special case* of the measurement postulate. Here we describe the general measurement postulate. Do not worry if you cannot completely understand it, the author also did not understand this completely at the time of writing this!

Quantum measurements are described by a collection $\{M_m\}$ of *measurement operators*. These are the operators acting on the state space of the system being measured. The index m refers to the measurement outcomes that may occur in the experiment. If the state of the quantum system is $|\psi\rangle$ immediately before the measurement, then the probability that result m occurs is given by

$$p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle$$

and the state of the state of the system after measurement is

$$\frac{M_m|\psi\rangle}{\sqrt{\langle\psi|M_m^\dagger M_m|.\psi\rangle}}$$

The measurement operators satisfy the completeness equation:

$$\sum_m M_m^\dagger M_m = I.$$

The completeness equation expresses the fact that the probabilities sum to one:

$$1 = \sum_m p(m) = \sum_m \langle\psi|M_m^\dagger M|\psi\rangle.$$

You might be surprised to know that the projective measurements (the one described earlier when discussing the postulates of quantum mechanics), when augmented with the unitary evolution postulate (also described earlier) give us the general measurement postulate. (You might want to read about noisy evolution of a quantum system). The general measurement postulate relates to the way we perform quantum measurements in the real world.

6.1 POVM Formalism

For some applications, the final state of a system after measurement is of little interest to us, with the main item of interest being the probability of obtaining a particular outcome. This formalism is a simple consequence of the general measurement postulate described earlier.
Suppose a measurement described by the measurement operators M_m is performed upon a quantum system in the state $|\psi\rangle$. Then the probability of outcome m is given by $p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle$. Suppose we define

$$E_m = M_m^\dagger M_m \tag{13}$$

Then we have that $p(m) = \langle\psi|E_m|\psi\rangle$ and $\sum_m E_m = I$. It turns out that the latter expression is the only condition that needs to be satisfied for $\{E_m\}_m$ to be a valid POVM.

To understand the power of POVMs, we need to understand a concept called the distinguishability of states.

6.2 Distinguishability of States

Suppose we have two parties, Alice and Bob. Alice chooses a state $|\psi_i\rangle$ $(1 \leq i \leq n)$ from some fixed set of states known to both the parties. She gives the state $|\psi_i\rangle$ to Bob, whose task it is to identify the index i of the state that Alice has given him.
Suppose the states are $|\psi_i\rangle$ are orthonormal. Then Bob can do a quantum measurement to distinguish these states, using the following procedure. Define the measurement $M_i \equiv |\psi_i\rangle\langle\psi_i|$, one for each possible index i, and an additional measurement M_0 defined as the positive square root of the positive operator $I - \sum_{i\neq 0} |\psi\rangle\langle\psi|$. This comes from the general measurement postulate, which states that $\sum_m M_m^\dagger M_m = I$. For projective measurements,

we have $M_m^\dagger M_m = M_m$. Hence we have

$$
\begin{aligned}
&\sum_{i=0}^{n} M_i^\dagger M_i = I \\
&= \sum_{i=1}^{n} M_i + M_0^\dagger M_0 = I \\
&= M_0^\dagger M_0 = I - \sum_{i=1}^{n} M_i \\
&= M_0 = \sqrt{I - \sum_{i=1}^{n} M_i}
\end{aligned}
$$

Since these operators satisfy the completeness relation, and if the state $|\psi_i\rangle$ is prepared, then $p(i) = \langle\psi_i|M_i|\psi_i\rangle = 1$, so the result i occurs with certainity. Thus it is possible to distinguish these orthonormal states.

On the other hand, we can prove that *there exists no reliable measurement that can distinguish between non orthogonal states.*

Proof: Let us assume that Alice prepares one of the two non orthonormal states $|\psi_1\rangle$ and $|\psi_2\rangle$. Bob uses the set of measurement operators $\{M_j\}$ and depending on the outcome of his measurement he tries to guess the index i of Alice's state using some rule $i = f(j)$, e.g. $f(1) = 1, f(2) = 2$ and $f(3) = 2$.
Now, let us assume that Bob can distinguish Alice's non-orthonormal states.

- If Alice prepares the state $|\psi\rangle_1$ then the probability of Bob obtaining the outcome j such that $f(j) = 1$ is 1.
- If Alice prepares the state $|\psi\rangle_2$ then the probability of Bob obtaining the outcome j such that $f(j) = 2$ is 1.

Hence we can say that:

$$
\begin{aligned}
\langle\psi_1|E_1|\psi_1\rangle = 1 \\
\langle\psi_2|E_2|\psi_2\rangle = 1
\end{aligned}
$$

We also have that $\langle\psi_2|E_1|\psi_2\rangle = 0$.
Now, since $|\psi_1\rangle$ and $|\psi_2\rangle$ are not orthonormal, therefore we can write:

$$|\psi_1\rangle = \alpha|\psi_2\rangle + \beta|\phi\rangle \tag{14}$$

where $|\phi\rangle$ is a non zero vector such that $|\phi\rangle$ is orthonormal to $|\psi_2\rangle$, and $|\alpha|^2 + |\beta|^2 = 1$.

Now, let us consider:

$$\begin{aligned}\langle\psi_1|E_1|\psi_1\rangle &= (\alpha_2^*\langle\psi_2| + \beta^*\langle\phi|)E_1(\alpha|\psi_2\rangle + \beta|\phi\rangle)\\ &= |\beta|^2\langle\phi|E_1|\phi\rangle\\ &\leq |\beta|^2\\ &\leq 1\end{aligned}$$

where the second last inequality comes from the fact that $\langle\phi|E_1|\phi\rangle \leq \sum_i\langle\phi|E_i|\phi\rangle = \langle\phi|\phi\rangle = 1$

This contradicts our basic assumption.
Hence there exists no reliable measurement that can perfectly distinguish between non orthonormal states.

7 Usefulness of POVM

While it is not perfectly possible to distinguish between non-orthonormal states, we can do it *some* of the times. This is where POVMs come into play. Let us look at an example:
Suppose Alice gives Bob a qubit prepared in one of the two states, $|\psi_1\rangle = |0\rangle$ or $|\psi_2\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. As explained earlier it is impossible for Bob to determine whether he has been given $|\psi_1\rangle$ or $|\psi_2\rangle$. *However it is possible for him to make a measurement that enables him to measure the state some of the time, but never make an error of mis-identification.* Consider a POVM of three elements:

$$\begin{aligned}E_1 &= \frac{\sqrt{2}}{1+\sqrt{2}}|1\rangle\langle 1|\\ E_2 &= \frac{\sqrt{2}}{1+\sqrt{2}}\frac{(|0\rangle - |1\rangle)(\langle 0| - \langle 1|)}{2}\\ E_3 &= I - E_1 - E_2\end{aligned}$$

Since $\sum_i E_i = I$, therefore this is a valid POVM. Suppose that Bob is given the state ψ_1. Then there is zero probability that he would observe the

result E_1. Thus if he observes the result E_1 for some qubit given to him, he can safely conclude that it is not $|\psi_1\rangle$, and therefore must be $|\psi_2\rangle$. Similarly if he observes the result E_2 then he can safely conclude that the state given to him was $|\psi_1\rangle$ (because it could not have been $|\psi_2\rangle$).
But suppose now that the outcome of his measurement is E_3. In such a case he can never find out which state was given to him. The point here is that **he never makes an error of mis-identification**.

8 Back to Composite Systems

Let us suppose that we have two independent ensembles for quantum states A and B. The first quantum state belongs to Alice and the second quantum state belongs to Bob, they may or may not be spatially separated. Let $\{p_X(x), |\psi_x\rangle\}$ be the ensemble for system A and $\{p_Y(y), |\phi_y\rangle\}$ be the ensemble for system B. If state of system A is $|\psi_x\rangle$ and that of system B is $|\phi_y\rangle$, then the state of the joint quantum system is given by $|\psi_x\rangle \otimes |\phi_y\rangle$. The density operator for the joint quantum system would be given by:

$$\mathbb{E}\{(|\psi_X\rangle \otimes |\phi_Y\rangle)(\langle\psi_X| \otimes \langle\phi_Y|)\} \tag{15}$$

The above expression is equal to this one:

$$\mathbb{E}\{|\psi_X\rangle\langle\psi_X| \otimes |\phi_Y\rangle\langle\phi_Y|\} \tag{16}$$

We can explicitly write the above expression as a sum over probabilities:

$$\sum_{x,y} p_X(x) p_Y(y) |\psi_X\rangle\langle\psi_X| \otimes |\phi_Y\rangle\langle\phi_Y| \tag{17}$$

Due to the distributive property of the cross product we can write:

$$\sum_{x} p_X(x) |\psi_X\rangle\langle\psi_X| \otimes \sum_{y} p_Y(y) |\psi_Y\rangle\langle\psi_Y| \tag{18}$$

and thus we can write the density operator for the joint system in the following form:

$$\rho \otimes \sigma \tag{19}$$

where ρ is the density operator for A and σ is the density operator for B.

We see the individual density operators *factored out* in the expression for the join density operator, and this was kind of expected, because we assumed that A and B were independent systems.

8.1 Separable States

Let us now consider two systems A and B whose corresponding ensembles are correlated in a certain way. We describe this correlated joint ensemble:

$$\{p_X(x), |\psi_x\rangle \otimes |\psi_y\rangle\} \tag{20}$$

By ignoring Bob's system, Alice's local density operator is given by:

$$\mathbb{E}_X\{|\psi_X\rangle\langle\psi_X|\} = \sum_x p_X(x)|\psi_x\rangle\langle\psi_x| \tag{21}$$

Similarly, ignoring Alice's system, Bob's local density operator is given by:

$$\mathbb{E}_Y = \{|\phi_Y\rangle\langle\phi_Y| = \sum_y p_Y(y)|\phi_y\rangle\langle\phi_y| \tag{22}$$

Separable State : *A bipartite density operator σ_{AB} is a separable state if it can be written in the following form:*

$$\sigma_{AB} = \sum_x p_X(x)|\psi_x\rangle\langle\psi_x|_A \otimes |\phi_y\rangle\langle\phi_y|_B \tag{23}$$

for some probability distribution $p_X(x)$ and sets $\{|\psi_x\rangle_A\}$ and $\{|\phi_y\rangle_B\}$ of pure states.

The term "separable" means that there is no quantum entanglement in the above state. This leads us to the definition of entanglement for density operators:

Entangled State *A bipartite density operator ρ_{AB} is entangled if it is not separable.*

There is another definition for entanglement that makes use of state vector instead of the density operator:

Entangled State (2) *A pure bipartite state $|\psi_{AB}\rangle$ is entangled if it cannot be written in as a product state $|\phi\rangle_A \otimes |\phi\rangle_B$ for any choices of states $|\phi\rangle_A$ and $|\phi\rangle_B$.*

8.2 Local Density Operators and Partial Trace

Consider the Bell state $|\Phi^+\rangle_{AB}$.

$$|\Phi^+\rangle_{AB} = \frac{1}{\sqrt{2}}(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B) \tag{24}$$

Bell states are a set of four states that form a maximally entangled basis for the joint space of two qubits.

Recall that the density operator description arises from its usefulness in determining the probabilities of the outcomes of a particular measurement. We say that the density operator is the "state" of the system because it is a mathematical representation that allows us to compute the probabilities resulting from a physical measurement.

Let us consider a local POVM $\{\Lambda^j\}_j$ that Alice can perform on her system. The global measurement operators for this local measurement are $\{\Lambda^j_A \otimes I_B\}_j$ because nothing (the identity) happens to Bob's system. The probability of obtaining the outcome j when performing this measurement on the state $|\Phi^+\rangle_{AB}$ is:

$$\begin{aligned}
\langle\Phi^+|_{AB}\Lambda^j_A \otimes I_B|\Phi^+\rangle_{AB} &= \frac{1}{2}((\langle 00|_{AB} + \langle 11|_{AB})\Lambda^j_A \otimes I_B(|00\rangle_{AB} + |11\rangle_{AB}) \\
&= \frac{1}{2}\sum_{k,l=0}^{1} \langle kk|_{AB}\Lambda^j_A \otimes I_B|ll\rangle_{AB} \\
&= \frac{1}{2}(\langle 0|_A\Lambda^j_A|0\rangle_A + \langle 1|_A\Lambda^j_A|1\rangle_A) \\
&= \frac{1}{2}(\mathrm{Tr}\{\Lambda^j_A|0\rangle\langle 0|_A\} + \mathrm{Tr}\{\Lambda^j_B|0\rangle\langle 0|_A\}) \\
&= \mathrm{Tr}\left\{\Lambda^j_A\frac{1}{2}(|0\rangle\langle 0|_A + |1\rangle\langle 1|_A)\right\} \\
&= \mathrm{Tr}\{\Lambda^j_A\pi_A\}
\end{aligned}$$

where π is the qubit maximally mixed state as defined below:

Maximally Mixed State: *The maximally mixed state π is the density operator corresponding to a uniform ensemble of orthogonal states $\left\{\frac{1}{d}, |x\rangle\right\}$, where d is the dimensionality of the Hilbert space. The maximally mixed* *state π is then equal to:*

$$\pi = \frac{1}{d} \sum_{x \in X} |x\rangle\langle x| \tag{25}$$

Thus we can predict the result of any local "Alice" measurement using the density operator π. **Thus, it is reasonable to say that Alice's local density operator is π, and we even go as far to say that her *local state* is π. A symmetric calculation shows that Bob's local state is also π.**

8.2.1 Partial Trace

In general we would like to determine a local density operator that predicts the outcomes of all local measurements (the above result was only for the Bell state $|\Phi^+\rangle_{AB}$!). The general method for doing this is the *partial trace operation.*

Suppose that Alice and Bob share a bipartite state ρ_{AB} and that Alice performs a local measurement on her system, described by a POVM $\{\Lambda_A^j\}$. Then the overall POVM on the joint system is $\{\Lambda_A^j \otimes I_B\}$ because we are assuming that Bob is not doing anything to his system. According to the Born's rule, the probability for Alice to receive outcome j after performing the measurement is given by the following expression:

$$p_J(j) = \mathrm{Tr}\{(\Lambda_A^j \otimes I_B)\rho_{AB}\} \tag{26}$$

In order to evaluate the trace, we can choose any orthonormal basis we wish (property of the trace operation!). Let $\{|k\rangle_A\}$ be an orthonormal basis for Alice's Hilbert space and $\{|l\rangle_B\}$ form an orthonormal basis for Bob's Hilbert space. We have that $\{|k\rangle_A \otimes |l\rangle_B\}$ forms an orthonormal basis for the tensor product of their Hilbert spaces. So the probability of obtaining outcome j

can be written as:

$$
\begin{aligned}
\mathrm{Tr}\{(\Lambda_A^j \otimes I_B)\rho_{AB}\} &= \sum_{k,l=0}^{1} \langle kl|_{AB}(\Lambda_A^j \otimes I_B)\rho_{AB}|kl\rangle_{AB} \\
&= \sum_{k,l=0}^{1} \langle k|_A(I_A \otimes \langle l|_B)(\Lambda_A^j \otimes I_B)\rho_{AB}(I_A \otimes |l_B\rangle)|k\rangle_A \\
&= \sum_{k,l=0}^{1} \langle k|_A\Lambda_A^j(I_A \otimes \langle l|_B)\rho_{AB}(I_A \otimes |l\rangle_B)|k\rangle_A \\
&= \sum_{k=0}^{1} \langle k|_A\Lambda_A^j\left[\sum_{l=0}^{1}(I_A \otimes \langle l|_B)\rho_{AB}(I_A \otimes |l\rangle_B)\right]|k\rangle_A
\end{aligned}
$$

Using the fact that $\{|k\rangle_A\}$ is an orthonormal basis for A, we can write:

$$
\mathrm{Tr}\{(\Lambda_A^j \otimes I_B)\rho_{AB}\} = \mathrm{Tr}\left\{\Lambda_A^j\left[\sum_{l=0}^{1}(I_A \otimes \langle l|_B)\rho_{AB}(I_A \otimes |l\rangle_B)\right]\right\} \tag{27}
$$

Do you see something interesting? In the square bracket, ρ_{AB} is getting transformed to a different operator, which is exactly the local density operator of A. If we supress the identity operators, then with a slight abuse of notation, we can write

$$
\begin{aligned}
\rho_A &= \left[\sum_{l=0}^{1}(I_A \otimes \langle l|_B)\rho_{AB}(I_A \otimes |l\rangle_B)\right] \\
&= \sum_{l=0}^{1}\langle l|_B\rho_{AB}|l\rangle_B
\end{aligned}
$$

This looks familiar. This is almost like the trace operator that we have studied. We call this the partial trace:

Partial Trace: *Let ρ_{AB} denote the square operator acting on a tensor product Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, and let $\{|l\rangle_B\}$ be an orthonormal basis for $\mathcal{H}_B$. Then the partial trace over the Hilbert space $\mathcal{H}_B$ is defined as follows:*

$$
\mathrm{Tr}_B\{\rho_{AB}\} = \sum_l (I_A \otimes \langle l|_B)\rho_{AB}(I_A \otimes |l\rangle_B). \tag{28}
$$

We can see that the partial trace operation is invariant under the choice of an orthonormal basis. Also partial trace is a linear operation.

Continuing with the development above, the local density operator for Alice is given by:

$$\rho_A = \mathrm{Tr}_B \rho_{AB} \tag{29}$$

and hence we can write:

$$p_J(j) = \mathrm{Tr}\{(\Lambda_A^j \otimes I_B)\rho_{AB}\} = \mathrm{Tr}\{\Lambda_A^j \rho_A\} \tag{30}$$

Thus from the local density operator ρ_A, we can predict the outcomes of local measurements that Alice performs on her system. It is important to note that the global picture, where we have the density operator ρ_{AB} and measurement of the form $\{\Lambda_A^j \otimes I_B\}$ is consistent with the local picture.

9 Quantum Channels

We will end this section of the notes with Quantum channels.
In simple terms a quantum channel transforms one density operator to another.

Definition : Let $\mathcal{B}(\mathcal{H})$ denote the space of trace class linear operators acting on a Hilbert space $\mathcal{H}$. Let $\mathcal{N}$ denote a map which takes linear operators in $\mathcal{B}(\mathcal{H}_A)$ to $\mathcal{B}(\mathcal{H}_B)$. For $\mathcal{N}$ to be a quantum channel, $\mathcal{N}$ has to be linear, completely positive and trace preserving.

Linearity simply implies that

$$\mathcal{N}(\alpha X_A + \beta Y_A) = \alpha\mathcal{N}(X_A) + \beta\mathcal{N}(Y_A) \tag{31}$$

for $X_A, Y_A \in \mathcal{B}(A)$

Completely Positive Map: *A linear map $\mathcal{M} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ is completely positive if $id_R \otimes \mathcal{M}$ is a positive map for a reference system R of arbitrary size.*

Trace Preservation: A quantum channel is trace preserving, in the sense that $\mathrm{Tr}\{X_A\} = \mathrm{Tr}\{\mathcal{N}(X_A)\}$ for all $X_A \in \mathcal{B}(\mathcal{H}_A)$.

Choi-Kraus Decomposition: *A map $\mathcal{N} : \mathcal{B}(\mathcal{H}_A) \to \mathcal{B}(\mathcal{H}_B)$ (denoted also by $\mathcal{N}_{A\to B}$) is a quantum channel if and only if it has a Choi-Kraus decomposition as follows:*

$$\mathcal{N}_{A\to B}(X_A) = \sum_{i=0}^{d-1} V_l X_A V_l^\dagger \tag{32}$$

At this point we would like to define the maximally entangled state for d dimensions.

Maximally Entangled State: *Let $\{|i\rangle_A\}$ form an orthonormal basis for $\mathcal{H}_A$ and $\{|i\rangle_B\}$ form an orthonormal basis for $\mathcal{H}_B$, where $\mathcal{H}_A$ and $\mathcal{H}_B$ are both Hilbert spaces of d dimensions. Then the maximally entangled state of d dimensions is given by:*

$$|\psi\rangle_{entangled} = \frac{1}{\sqrt{d}} \sum_{i=0}^{d} |i\rangle_A \otimes |i\rangle_B \tag{33}$$

Choi State: *Choi state is the result of applying a quantum channel to the perfectly entangled state.*

If a linear, trace class operator is completely positive for a maximally entangled state (i.e., if the Choi State is positive), then it is a completely positive operator.

10 Schmidt Decomposition

Suppose we have a bipartite pure state:

$$|\psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B, \tag{34}$$

where $\mathcal{H}_A$ and $\mathcal{H}_B$ are finite dimensional Hilbert spaces, not necessarily of the same dimension, and $|||\psi\rangle_{AB}||_2 = 1$. Then it is possible to express this state as follows:

$$|\psi\rangle_{AB} = \sum_{i=0}^{d-1} \lambda_i |i\rangle_A |i\rangle_B, \tag{35}$$

where the amplitudes λ_i are real, strictly positive and normalized so that $\sum_i \lambda_i^2 = 1$, the states $\{|i\rangle_A\}$ form an orthonormal basis for system A, and the states $\{|i\rangle_B\}$ form an orthonormal basis for system B. The vector $[\lambda_i]_{i\in\{0,\ldots,d-1\}}$ is called the vector of Schmidt coefficients. The Schmidt rank d of a bipartite state is equal to the number of Schmidt coefficients λ_i in its Schmidt decompositon and satisfies

$$d \leq \min\{\dim(\mathcal{H}_A), \dim(\mathcal{H}_B)\} \tag{36}$$

Proof: A bipartite pure state can be expressed in a standard form (*the Schmidt decomposition*) that is often very useful.
To arrive at this form, note that an arbitrary vector in $\mathcal{H}_A \otimes \mathcal{H}_B$ can be written in the following form:

$$|\psi\rangle_{AB} = \sum_{i,\mu} \psi_{i\mu} |i\rangle_A \otimes |\mu\rangle_B = \sum_i |i\rangle_A \otimes |i'\rangle_B. \tag{37}$$

where $\{|i\rangle_A\}$ form an orthonormal basis for $\mathcal{H}_A$ and $\{|\mu\rangle_B\}$ form an orthonormal basis for $\mathcal{H}_B$. For writing the second equality, we have defined the following:

$$|i'\rangle_B = \sum_\mu \psi_{i\mu} |\mu\rangle_B \tag{38}$$

Note that (till now) $\{|i'\rangle_B\}$ are not orthogonal to each other.
Now, let us choose $\{i\rangle_A\}$ basis so that ρ_A is is a diagonal in this basis. Hence we can computer the local density operator for A as follows:

$$\rho_A = \sum_i p_i |i\rangle\langle i| \tag{39}$$

We can also compute the local density operator for A by tracing out B:

$$\begin{aligned}
\rho_A &= \mathrm{Tr}_B\{|\psi_{AB}\rangle\langle\psi_{AB}|\} \\
&= \mathrm{Tr}_B\{\sum_{i,j}(|i\rangle_A \otimes |i'\rangle_B)(\langle j|_A \otimes \langle j'|_B)\} \\
&= \mathrm{Tr}_B\{\sum_{i,j} |i\rangle_A\langle j|_A \otimes |i'\rangle_B\langle j'|_B\} \\
&= \sum_k (I_A \otimes \langle k|_B)\{\sum_{i,j} |i\rangle_A\langle j|_A \otimes |i'\rangle_B\langle j'|_B\}(I_A \otimes |k\rangle_B) \\
&= \sum_{i,j,k}(|i\rangle_A\langle j|_A) \otimes (\langle k|_B|i'\rangle_B\langle j'|_B|k\rangle_B) \\
&= \sum_{i,j,k}(|i\rangle_A\langle j|_A) \otimes (\langle j'|_B|k\rangle_B\langle k|_B|i'\rangle_B) \\
&= \sum_{i,j}(|i\rangle_A\langle j|_A) \otimes \sum_k(\langle j'|_B(|k\rangle\langle k|)|i'\rangle_B) \\
&= \sum_{i,j}(|i\rangle_A\langle j|_A) \otimes \langle j'|_B|i'\rangle_B
\end{aligned}$$

Comparing the two expressions for the density operator, we have the following:

$$\langle j'|_B|i'\rangle_B = \delta_{ij}p_i \tag{40}$$

This is precisely the condition for orthogonality of $\{|i'\rangle_B\}$. Thus it turns out that $\{|i'\rangle_B\}$ are indeed orthogonal to each other. To normalize them, we can divide $|i'\rangle_B$ by $\sqrt{p_i}$. Let us define

$$|i''\rangle_B = |i'\rangle_B/\sqrt{p_i} \tag{41}$$

so that $\{i''\rangle_B\}$ forms a set of orthonormal vectors in $\mathcal{H}_B$. Using our initial definition of $|\psi\rangle_{AB}$, we can write:

$$|\psi\rangle_{AB} = \sum_i \sqrt{p_i}|i\rangle_A \otimes |i''\rangle_B \tag{42}$$

in terms of a *particular* orthonormal basis of $\mathcal{H}_A$ and $\mathcal{H}_B$.