# QI Lecture 6

Density Matrices

## The Density Matrix

If you're asked to predict the result of a coin flip, the best you can do is to say "there's a probability of $p \approx \frac{1}{2}$ of the coin being either heads or tails". Nobody knows what the result will be, and it's completely infeasible to predict the result. Therefore, we fall back on what we know about coins: heads and tails show up about equally often.

Now say you flip a coin and cover it. Alice looks at the coin while Bob does not, and you ask them both "what is the probability that this coin is heads?" Alice's answer will be either 1 or 0, and Bob's answer will still be $\frac{1}{2}$. We begin to get a sense that probability is only partly about how the coin behaves. Lack of knowledge also forces us to use probabilities; the coin is in a definite state, Bob just doesn't know what it is.

Now say you're given a string of numbers

$$11001001000011111101101010100001$$

and you're asked to predict the next digit. You might think that there's an even chance that the next digit will be either 0 or 1, because this looks random. In fact, there's no randomness whatsoever: this is $\pi$ in base 2, and the correct answer is 0. Without knowledge of the "hidden variable", that this is $\pi$, the string of bits may as well be random. Again, we find that it's a lack of knowledge that forces us to use probabilities.

But "fundamental uncertainty" is different.[1] If you try to predict which path a photon will take through the two-path interferometer, you can never do better than "there is an equal probability of finding the photon on either path".[2] We know this because we're not missing any information; we can know with certainty that the state is $|+\rangle$, but that doesn't tell us what result we'll see in the $\{|0\rangle, |1\rangle\}$ basis. The result is "irreducibly random".

Unfortunately, we have two different ways to describe these two different sources of probabilities. For regular probabilities we use a "probability distribution", which is just a list of numbers and what they correspond to. For example, if $p_k$ is the probability of result $k$, then the probability distribution for the sum of a pair of dice is

[1] It's not obvious that it's different. We'll come back to this later and explore how we can show experimentally that quantum systems really are unpredictable.

[2] That is, if you put a detector in each path, there's no way to predict which will click.

$$p_2 = \frac{1}{36}, \quad p_3 = \frac{2}{36}, \quad p_4 = \frac{3}{36}, \quad \ldots, \quad p_{11} = \frac{2}{36}, \quad p_{12} = \frac{1}{36}$$

For fundamental randomness we use probability amplitudes and quantum states. For example, if two photons are in the states

$$|\frac{\pi}{6}\rangle = \frac{\sqrt{3}}{2}|0\rangle + \frac{1}{2}|1\rangle \qquad |-\frac{\pi}{6}\rangle = \frac{\sqrt{3}}{2}|0\rangle - \frac{1}{2}|1\rangle$$

they have the same set of probabilities, $p_0 = \frac{3}{4}$ and $p_1 = \frac{1}{4}$, and yet these states are different. A polarizer aligned at $\frac{\pi}{6}$ will always pass the first state, but will only pass the second $\cos^2\left(\frac{\pi}{6} - \left(-\frac{\pi}{6}\right)\right) = \cos^2\left(\frac{\pi}{3}\right) = \left(\frac{1}{2}\right)^2 = \frac{1}{4}$ of the time. In other words, probabilities alone are insufficient for describing this situation.

To handle both sources of probability at the same time, we introduce "density matrices".

**Example** Alice is sending Bob[3] a string of states: $\frac{2}{3}$ of the time she sends $|+\rangle$ and $\frac{1}{3}$ of the time she sends $|0\rangle$. The question is, if Bob measures in the $\{|0\rangle, |1\rangle\}$ basis, what is the probability that he'll see each result?

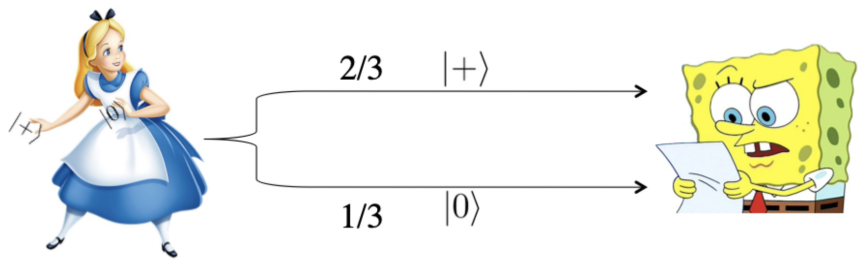


Figure 1: Alice sends one of two states to Bob, but doesn't tell him which.

First we'll consider the probability of Bob seeing the state $|1\rangle$. If the state that's sent is $|+\rangle$, the probability of seeing $|1\rangle$ is $|\langle 1|+\rangle|^2 = \left|\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2}$, and if the state that's sent is $|0\rangle$, then the probability of seeing $|1\rangle$ is $|\langle 1|0\rangle|^2 = 0$. So the total probability is the weighted sum of these two values.

$$P(1) = \frac{2}{3}\,|\langle 1|+\rangle|^2 + \frac{1}{3}\,|\langle 1|0\rangle|^2 = \frac{2}{3}\cdot\frac{1}{2} + \frac{1}{3}\cdot 0 = \frac{1}{3}$$

$P(1) = \frac{1}{3}$ is a combination of two kinds of probability: the classical probability that Alice will choose to send the state $|+\rangle$ as well as the quantum probability that Bob will look at $|+\rangle$ and see $|1\rangle$.

[3]These are the standard names for "Party A" and "Party B" in communication, information theory, cryptography, and now quantum information. In case you're wondering, a third party is usually named Carol and spies/eavesdroppers (common in cryptography literature) are invariably named Eve.

But there's another way to look at this same equation.

$$P(1) = \frac{2}{3}\left|\langle 1|+\rangle\right|^2 + \frac{1}{3}\left|\langle 1|0\rangle\right|^2 = \frac{2}{3}\langle 1|+\rangle\langle +|1\rangle + \frac{1}{3}\langle 1|0\rangle\langle 0|1\rangle = \langle 1|\left[\frac{2}{3}|+\rangle\langle +| + \frac{1}{3}|0\rangle\langle 0|\right]|1\rangle$$

This object is the **density matrix** of the quantum states being sent from Alice to Bob. It simultaneously handles the classical and non-classical probabilities of this example. Because tradition says so, we usually denote this with a $\rho$ (the Greek letter "rho", pronounced like "rowboat").

$$\rho = \frac{2}{3}|+\rangle\langle +| + \frac{1}{3}|0\rangle\langle 0|$$

Let's use this to calculate the probability of seeing $|0\rangle$. Since $P(1) = \frac{1}{3}$, we can expect that $P(0) = \frac{2}{3}$. Let's see if it is!

$$P(0) = \langle 0|\rho|0\rangle = \langle 0|\left[\frac{2}{3}|+\rangle\langle +| + \frac{1}{3}|0\rangle\langle 0|\right]|0\rangle = \frac{2}{3}\langle 0|+\rangle\langle +|0\rangle + \frac{1}{3}\langle 0|0\rangle\langle 0|0\rangle = \frac{2}{3}\cdot\frac{1}{\sqrt{2}}\cdot\frac{1}{\sqrt{2}} + \frac{1}{3}\cdot 1\cdot 1 = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$$

Finally, since this is a density <u>matrix</u>, we may as well see what that matrix is. In the $\{|0\rangle, |1\rangle\}$ basis,

$$\begin{aligned}
\rho &= \tfrac{2}{3}|+\rangle\langle +| + \tfrac{1}{3}|0\rangle\langle 0| \\
&\sim \tfrac{2}{3}\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}}, & \frac{1}{\sqrt{2}} \end{pmatrix} + \tfrac{1}{3}\begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1, & 0 \end{pmatrix} \\
&= \tfrac{2}{3}\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} + \tfrac{1}{3}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix} + \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}
\end{aligned}$$

■

An "ensemble of states" is written $\{p_k, |\psi_k\rangle\}$, signifying that each $|\psi_k\rangle$ shows up with probability $p_k$. The density matrix for such an ensemble is

$$\boxed{\rho \equiv \sum_k p_k |\psi_k\rangle\langle\psi_k|}$$

When there are multiple states in the ensemble we have a "**mixed state**", and when $\rho = |\psi\rangle\langle\psi|$ we have a "**pure state**".

From this last example, we now know that the probability of seeing the state $|\phi\rangle$ given the ensemble $\{p_k, |\psi_k\rangle\}$ with density matrix $\rho = \sum_k p_k |\psi_k\rangle\langle\psi_k|$ is

$$\boxed{p(\phi) = \langle\phi|\rho|\phi\rangle}$$

**Example** If you're using a vertical or horizontal polarizer, then a string of $|-\rangle$ states looks the same as a string of $|1\rangle$ and $|0\rangle$ states sent with equal probability. That is to say, you'll see vertical states and horizontal states with probability $\frac{1}{2}$ in both situations.

So how are the density matrices different?

First, the density matrix for the mixed state, $p_0 = p_1 = \frac{1}{2}$, in the $\{|0\rangle, |1\rangle\}$ basis.

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| \sim \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$$

And now the density matrix for the pure state, $|-\rangle$, in the $\{|0\rangle, |1\rangle\}$ basis.

$$\eta = |+\rangle\langle +| = \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right)\left(\frac{\langle 0| - \langle 1|}{\sqrt{2}}\right) = \frac{1}{2}|0\rangle\langle 0| - \frac{1}{2}|0\rangle\langle 1| - \frac{1}{2}|1\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| \sim \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Clearly there's a difference. Notice that the probability of measuring $|1\rangle$ (or $|0\rangle$) is $\frac{1}{2}$ in both cases because

$$p_\rho(1) = \langle 1|\rho|1\rangle = \frac{1}{2}\langle 1|0\rangle\langle 0|1\rangle + \frac{1}{2}\langle 1|1\rangle\langle 1|1\rangle = 0 + \frac{1}{2} = \frac{1}{2}$$

$$p_\eta(1) = \langle 1|\eta|1\rangle = \langle 1|-\rangle\langle -|1\rangle = \frac{-1}{\sqrt{2}} \cdot \frac{-1}{\sqrt{2}} = \frac{1}{2}$$

which means that if you only have access to a vertical/horizontal polarizer (only measure in the $\{|0\rangle, |1\rangle\}$ basis), you'd never know the difference. On the other hand if you measure the probability of the state $|+\rangle$:

$$p_\rho(+) = \langle +|\rho|+\rangle = \frac{1}{2}\langle +|0\rangle\langle 0|+\rangle + \frac{1}{2}\langle +|1\rangle\langle 1|+\rangle = \frac{1}{2} \cdot \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} + \frac{1}{2} \cdot \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$p_\eta(+) = \langle +|\eta|+\rangle = \langle +|-\rangle\langle -|+\rangle = 0$$

So there's a big difference if you measure in the $\{|+\rangle, |-\rangle\}$ basis.

Just for the heck of it, and so that we know how to do this another way, here's that last calculation using the matrix form in the $\{|0\rangle, |1\rangle\}$ basis:

$$\begin{aligned} & p_\eta(+) \\ = & \langle +|\eta|+\rangle \\ \sim & \begin{pmatrix} \frac{1}{\sqrt{2}}, & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \\ = & \begin{pmatrix} \frac{1}{\sqrt{2}}, & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ = & 0 \end{aligned}$$

■

## Mapping Density Matrices

If we use a unitary operator to map state kets or change the basis, then in order to map a density matrix we need to map both the kets and the bras: $\sum_k p_k U|\psi_k\rangle\langle\psi_k|U^\dagger$.

$$\boxed{|\psi\rangle \longrightarrow U|\psi\rangle \qquad\qquad \rho \longrightarrow U\rho U^\dagger}$$

## Trace

The trace of an operator is

$$\boxed{Tr[M] \equiv \sum_j \langle j|M|j\rangle}$$

which, if $M$ is written as a matrix, is the sum of the diagonal terms. Here $\{|j\rangle\}$ is any orthonormal basis. There's a cute relationship between the Identity and trace that we can use to prove a number of important properties

**Theorem** (The Trace Permutes Operators)**.**

$$Tr[ABC] = Tr[BCA] = Tr[CAB]$$

Remember that that inner products, "$\langle\cdots\rangle$", are just numbers and that $I = \sum_k |k\rangle\langle k|$.

$$\begin{aligned}
& Tr[ABC] \\
= & \sum_j \langle j|ABC|j\rangle \\
= & \sum_j \langle j|AIBC|j\rangle \\
= & \sum_j \langle j|A\left[\sum_k |k\rangle\langle k|\right]BC|j\rangle \\
= & \sum_{jk} \langle j|A|k\rangle\langle k|BC|j\rangle \\
= & \sum_{jk} \langle k|BC|j\rangle\langle j|A|k\rangle \\
= & \sum_k \langle k|BC\left[\sum_j |j\rangle\langle j|\right]A|k\rangle \\
= & \sum_k \langle k|BCIA|k\rangle \\
= & \sum_k \langle k|BCA|k\rangle \\
= & Tr[BCA]
\end{aligned}$$

QED

Notice the difference between "permute" and "commute". While it is true that $Tr[AB] = Tr[BA]$, it is <u>not</u> true that $Tr[ABC] = Tr[BAC]$. Permuting means that you can take a block of operators off of the beginning and place it at the end, but that doesn't mean you can change the order.

We can use permuting of operators to immediately see why the basis is unimportant for the trace. When we want to write an operator, $M$, in a new basis we use a "**similarity transform**", $M \to SMS^{-1}$. But the trace doesn't notice similarity transforms because

$$\boxed{Tr\left[SMS^{-1}\right] = Tr\left[S^{-1}SM\right] = Tr[M]}$$

The most important basis for an operator is its eigenbasis, in which it can be written $M = \sum_k \lambda_k |v_k\rangle\langle v_k|$, and since the trace is the same regardless of basis...

**Theorem** (The Trace is the sum of eigenvalues)**.** *If the eigenvalues of the operator $M$ are $\{\lambda_k\}$, then*

$$\boxed{Tr[M] = \sum_k \lambda_k}$$

The eigenstates of $M$, $\{|v_k\rangle\}$, form an orthonormal basis (meaning $\langle v_j|v_k\rangle = \delta_{jk}$) which we may as well use when taking the trace, and the Spectral Theorem says that $M = \sum_k \lambda_k |v_k\rangle\langle v_k|$.

$$
\begin{aligned}
& Tr[M] \\
= \ & \sum_j \langle v_j|M|v_j\rangle \\
= \ & \sum_j \langle v_j| \left[\sum_k \lambda_k |v_k\rangle\langle v_k|\right] |v_j\rangle \\
= \ & \sum_{jk} \lambda_k \langle v_j|v_k\rangle\langle v_k|v_j\rangle \\
= \ & \sum_{jk} \lambda_k \delta_{jk}\delta_{jk} \\
= \ & \sum_k \lambda_k
\end{aligned}
$$

QED

### The Trace and Projective Measurements

The trace shows up when dealing with density matrices for a number of reasons. We already know that given the state $|\psi\rangle$, the expectation value of the observable $M$ is $E[M] = \langle\psi|M|\psi\rangle$. But if we want the expectation value of $M$ over an ensemble of states, $\{p_k, |\psi_k\rangle\}$, then we have to average over all of those individual expectation values, $\langle\psi_k|M|\psi_k\rangle$. The density matrix and trace give us a good way to do this.

$$
\begin{aligned}
& E[M] \\
= \ & \sum_k p_k \langle\psi_k|M|\psi_k\rangle \\
= \ & \sum_k p_k \langle\psi_k| \left[\sum_j |j\rangle\langle j|\right] M|\psi_k\rangle \\
= \ & \sum_{jk} p_k \langle\psi_k|j\rangle\langle j|M|\psi_k\rangle \\
= \ & \sum_{jk} p_k \langle j|M|\psi_k\rangle\langle\psi_k|j\rangle \\
= \ & \sum_j \langle j|M \left[\sum_k p_k|\psi_k\rangle\langle\psi_k|\right] |j\rangle \\
= \ & \sum_j \langle j|M\rho|j\rangle \\
= \ & Tr[M\rho]
\end{aligned}
$$

If you're wondering how we can shuffle terms around, keep in mind that we're really looking at a bunch of (complex) numbers.

$$\sum_{jk} \underbrace{p_k}_{\#} \underbrace{\langle j|M|\psi_k\rangle}_{\#} \underbrace{\langle\psi_k|j\rangle}_{\#}$$

So we now have a quick way of calculating expectation values on ensembles:

$$\boxed{E[M] = Tr[M\rho]}$$

It also couldn't hurt to re-write the rules for the probability of seeing a particular result, $m$, and the effect of measurements. For a given state $|\psi\rangle$,

$$p(m) = \langle\psi|P_m|\psi\rangle \qquad\qquad |\psi\rangle \xrightarrow{m} |\psi_m\rangle = \frac{P_m|\psi\rangle}{\sqrt{p(m)}}$$

Like unitary (or any other) operators, we need to deal with both the kets and bras of density matrices. So the projection operation is $P_m\rho P_m^\dagger = P_m\rho P_m$, since projection operators are hermitian. For a given density matrix $\rho = \sum_k p_k|\psi_k\rangle\langle\psi_k|$,

$$p(m) = \sum_k p_k\langle\psi_k|P_m|\psi_k\rangle = \sum_{jk} p_k\langle\psi_k|j\rangle\langle j|P_m|\psi_k\rangle = \sum_{jk}\langle j|P_m\left(\sum_k p_k|\psi_k\rangle\langle\psi_k|\right)|j\rangle = Tr\left[P_m\rho\right]$$

With a little foresight into generalized measurements later in the class, or just to make what follows a little more aesthetic, you may want to write this as $p(m) = Tr\left[P_m\rho P_m\right]$.[4] So for a given density matrix, the probability of a given observation and the effect on the matrix are:

$$\boxed{p(m) = Tr\left[P_m\rho P_m\right] \qquad\qquad \rho \xrightarrow{m} \rho_m = \frac{P_m\rho P_m}{p(m)}}$$

## Properties of the Density Matrix

**The trace of $\rho$ is 1.**

$$\boxed{Tr[\rho] = 1}$$

We know this because

[4] Recall that the trace permute operators and that projections have the unique property that $P_m^2 = P_m$.

$$\begin{aligned}
& Tr[\rho] \\
= \; & Tr\left[\sum_k p_k|\psi_k\rangle\langle\psi_k|\right] \\
= \; & \sum_j \langle j|\left[\sum_k p_k|\psi_k\rangle\langle\psi_k|\right]|j\rangle \\
= \; & \sum_{jk} p_k\langle j|\psi_k\rangle\langle\psi_k|j\rangle \\
= \; & \sum_{jk} p_k\langle\psi_k|j\rangle\langle j|\psi_k\rangle \\
= \; & \sum_k p_k\langle\psi_k|\left[\sum_j |j\rangle\langle j|\right]|\psi_k\rangle \\
= \; & \sum_k p_k\langle\psi_k|I|\psi_k\rangle \\
= \; & \sum_k p_k\langle\psi_k|\psi_k\rangle \\
= \; & \sum_k p_k \\
= \; & 1
\end{aligned}$$

QED

Notice that $\sum_k p_k|\psi_k\rangle\langle\psi_k| \xrightarrow{Tr} \sum_k p_k\langle\psi_k|\psi_k\rangle$. This is a general rule, sometimes stated:

"The trace turns outer products into inner products."

$\rho$ **is a positive opperator.**

$$\boxed{\forall|\phi\rangle, \quad 0 \leq \langle\phi|\rho|\phi\rangle \leq 1}$$

We know this because

$$\langle\phi|\rho|\phi\rangle = \sum_k p_k\langle\phi|\psi_k\rangle\langle\psi_k|\phi\rangle = \sum_k p_k|\langle\phi|\psi_k\rangle|^2$$

Because $p_k$ are probabilities and $|\langle\phi|\psi_k\rangle|$ is a magnitude of a complex number, both of these are non-negative and therefore $\langle\phi|\rho|\phi\rangle \geq 0$. This is the definition of a "**positive operator**".

Because $p_k$ are probabilities, $\sum_k p_k = 1$, and since $|\langle\phi|\psi_k\rangle| \leq 1$, we immediately have that $\langle\phi|\rho|\phi\rangle \leq 1$.

$\rho$ **is hermitian.**

$$\rho^\dagger = \left(\sum_k p_k|\psi_k\rangle\langle\psi_k|\right)^\dagger = \sum_k p_k^*\left(\langle\psi_k|\right)^\dagger\left(|\psi_k\rangle\right)^\dagger = \sum_k p_k|\psi_k\rangle\langle\psi_k| = \rho$$

$$\boxed{\rho^\dagger = \rho}$$

**The eigenvalues of $\rho$ "look like probabilities".** Because the trace is 1 and the trace is the sum of the eigenvalues, we have that $1 = Tr[\rho] = \sum_k \lambda_k$. Because $\rho$ is hermitian we know that $\lambda_k \in \mathbb{R}$ and since $\rho$ is a positive operator we have that $\lambda_k \in \mathbb{R}^+$. It follows that

$$\boxed{0 \leq \lambda_k \leq 1 \qquad \qquad \sum_k \lambda_k = 1}$$

**Example** In the first example, where Alice is sending Bob the ensemble $p(|+\rangle) = \frac{2}{3}$ and $p(|0\rangle) = \frac{1}{3}$, we found that the density matrix is

$$\rho = \frac{2}{3}|+\rangle\langle+| + \frac{1}{3}|0\rangle\langle 0| \sim \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

where the matrix is written in the computational, $\{|0\rangle, |1\rangle\}$, basis. Because density matrices are hermitian, we can expect the spectral decomposition of $\rho$ to exist. What is that spectral decomposition?

$$\begin{aligned} C(\lambda) &= \begin{vmatrix} \frac{2}{3} - \lambda & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} - \lambda \end{vmatrix} \\ &= \left(\tfrac{2}{3} - \lambda\right)\left(\tfrac{1}{3} - \lambda\right) - \tfrac{1}{9} \\ &= \lambda^2 - \lambda + \tfrac{2}{9} - \tfrac{1}{9} \\ &= \lambda^2 - \lambda + \tfrac{1}{9} \end{aligned}$$

$$\lambda = \frac{1 \pm \sqrt{(-1)^2 - 4(1)\left(\frac{1}{9}\right)}}{2} = \frac{1}{2} \pm \frac{\sqrt{5}}{6}$$

$$\lambda_1 \approx 0.873 \qquad \qquad \lambda_2 \approx 0.127$$

These eigenvalues certainly "look like probabilities".

$$|v_1\rangle \sim \frac{1}{\sqrt{\left(\frac{1+\sqrt{5}}{2}\right)^2 + 1}} \begin{pmatrix} \frac{1+\sqrt{5}}{2} \\ 1 \end{pmatrix} \approx \begin{pmatrix} 0.851 \\ 0.526 \end{pmatrix} \qquad |v_2\rangle \sim \frac{1}{\sqrt{\left(\frac{1-\sqrt{5}}{2}\right)^2 + 1}} \begin{pmatrix} \frac{1-\sqrt{5}}{2} \\ 1 \end{pmatrix} \approx \begin{pmatrix} -0.526 \\ 0.851 \end{pmatrix}$$

Finally, we've rewritten the density matrix as

$$\rho = \frac{2}{3}|+\rangle\langle+| + \frac{1}{3}|0\rangle\langle 0| = \lambda_1|v_1\rangle\langle v_1| + \lambda_2|v_2\rangle\langle v_2|$$

What this means is that the ensembles $\left\{p(|+\rangle) = \frac{2}{3},\, p(|0\rangle) = \frac{1}{3}\right\}$ and $\{p(|v_1\rangle) = \lambda_1,\, p(|v_2\rangle) = \lambda_2\}$ have the same density matrix. In other words,

$$\boxed{\text{many different ensembles can share the same density matrix!}}$$

■

## Partial Trace

For joint systems we often do the trace over just part of the system, leaving the rest intact. This is analogous to finding the "**marginal probability**" and it's an incredibly powerful and absolutely essential technique.

Given a pair of sets $A$ and $B$ with a joint probability distribution $P(a,b)$, such that $a \in A$ and $b \in B$, the marginal probability distribution on the set $A$ is

$$P_B(b) = \sum_{a \in A} P(a,b)$$

The partial trace is a similar idea. If $A = Span\,\{|1_a\rangle, |2_a\rangle, \ldots\}$, $B = Span\,\{|1_b\rangle, |2_b\rangle, \ldots\}$ and $\rho = \sum_k p_k|\psi_k\rangle\langle\psi_k|$ is a density matrix defined on $A \otimes B$, then the "partial trace over $B$" is

$$\boxed{\rho_A = Tr_B\,[\rho] = \sum_j \langle j_b|\rho|j_b\rangle}$$

This partial trace only impacts the states in system $B$. That means being careful about how you do inner products. For example,

$$[\langle 1_a|\langle 2_b|]|2_b\rangle = \langle 1_a| \qquad [\langle 1_a|\langle 2_b|]|1_b\rangle = 0$$

We take the partial trace over system $B$ when we don't have access to $B$. What's left is the reduced density matrix of $A$ and that describes everything we can know (by gathering statistics) about the system $A$.

**Example** Every day Alice gets a Sandwich, $S = \{ham, turkey\}$, and a Beverage, $B = \{water, coffee, juice\}$. The probabilities of the combinations are

| $hw$ | $hc$ | $hj$ | $tw$ | $tc$ | $tj$ |
|---|---|---|---|---|---|
| 0.1 | 0.3 | 0 | 0.3 | 0.1 | 0.2 |

Alice always eats the sandwich before getting to the quantum lab, so Bob only ever sees the beverage. What probability distribution does Bob see?

Without access to the sandwich set, $S$, Bob's probability distribution is

$$P_B(w) = P(h,w) + P(t,w) = 0.1 + 0.3 = 0.4$$
$$P_B(c) = P(h,c) + P(t,c) = 0.3 + 0.1 = 0.4$$
$$P_B(j) = P(h,j) + P(t,j) = 0 + 0.2 = 0.2$$

This should make sense; the probability of coffee is the probability of ham-and-coffee plus the probability of turkey-and-coffee, $P_B(c) = P(h,c) + P(t,c)$.

Here's a silly way to do the same calculation. The density matrix of Alice's lunch,[5] written three times in order of increasing carefulness:

$$\begin{aligned}
\rho &= 0.1|hw\rangle\langle hw| + 0.3|hc\rangle\langle hc| + 0.3|tw\rangle\langle tw| + 0.1|tc\rangle\langle tc| + 0.2|tj\rangle\langle tj| \\
\rho &= 0.1|h\rangle|w\rangle\langle h|\langle w| + 0.3|h\rangle|c\rangle\langle h|\langle c| + 0.3|t\rangle|w\rangle\langle t|\langle w| + 0.1|t\rangle|c\rangle\langle t|\langle c| + 0.2|t\rangle|j\rangle\langle t|\langle j| \\
\rho &= 0.1|h\rangle_s|w\rangle_b\langle h|_s\langle w|_b + 0.3|h\rangle_s|c\rangle_b\langle h|_s\langle c|_b + 0.3|t\rangle_s|w\rangle_b\langle t|_s\langle w|_b + 0.1|t\rangle_s|c\rangle_b\langle t|_s\langle c|_b + 0.2|t\rangle_s|j\rangle_b\langle t|_s\langle j|_b
\end{aligned}$$

As we take the partial trace it's essential to keep the two systems straight. It often helps to write subscripts to identify the systems (bottom), but it's not necessary.

$$\begin{aligned}
& \rho_b \\
=\ & Tr_s[\rho] \\
=\ & \langle h|_s \left[0.1|h\rangle_s|w\rangle_b\langle h|_s\langle w|_b + 0.3|h\rangle_s|c\rangle_b\langle h|_s\langle c|_b + 0.3|t\rangle_s|w\rangle_b\langle t|_s\langle w|_b + 0.1|t\rangle_s|c\rangle_b\langle t|_s\langle c|_b + 0.2|t\rangle_s|j\rangle_b\langle t|_s\langle j|_b\right]|h\rangle_s \\
& +\langle t|_s \left[0.1|h\rangle_s|w\rangle_b\langle h|_s\langle w|_b + 0.3|h\rangle_s|c\rangle_b\langle h|_s\langle c|_b + 0.3|t\rangle_s|w\rangle_b\langle t|_s\langle w|_b + 0.1|t\rangle_s|c\rangle_b\langle t|_s\langle c|_b + 0.2|t\rangle_s|j\rangle_b\langle t|_s\langle j|_b\right]|t\rangle_s \\
=\ & 0.1|w\rangle_b\langle w|_b + 0.3|c\rangle_b\langle c|_b \\
& +0.3|w\rangle_b\langle w|_b + 0.1|c\rangle_b\langle c|_b + 0.2|j\rangle_b\langle j|_b \\
=\ & 0.4|w\rangle_b\langle w|_b + 0.4|c\rangle_b\langle c|_b + 0.2|j\rangle_b\langle j|_b
\end{aligned}$$

So the reduced density matrix accessible to Bob is

$$\rho_b = 0.4|w\rangle_b\langle w|_b + 0.4|c\rangle_b\langle c|_b + 0.2|j\rangle_b\langle j|_b$$

[5] The density matrix covers both kinds of probability (classical and quantum), so handling lunch isn't a problem, it's just more elaborate than you really need.

which is the same answer we got using the marginal distribution.

■

**Example** Alice and Bob each have one qubit from a pair that share the state

$$|\psi\rangle = \frac{1}{\sqrt{2}}|0\rangle_a|1\rangle_b + \frac{i}{2}|0\rangle_a|0\rangle_b - \frac{1}{2}|1\rangle_a|1\rangle_b$$

Unfortunately, Alice and Bob are on opposites sides of the universe right now, so neither of them have access to the entire state. What is Alice's reduced density matrix?

First, the full density matrix is

$$\rho = |\psi\rangle\langle\psi| = \left(\frac{1}{\sqrt{2}}|0\rangle_a|1\rangle_b + \frac{i}{2}|0\rangle_a|0\rangle_b - \frac{1}{2}|1\rangle_a|1\rangle_b\right)\left(\frac{1}{\sqrt{2}}\langle 0|_a\langle 1|_b - \frac{i}{2}\langle 0|_a\langle 0|_b - \frac{1}{2}\langle 1|_a\langle 1|_b\right)$$

We'll find that it's often easier to leave the density matrix in this form, unexpanded.

Alice's reduced density matrix, $\rho_a$, is the partial trace over Bob's qubit (because that's the part of the system she doesn't have access to).

$$\begin{aligned}
& \rho_a \\
= \; & Tr_b[\rho] \\
= \; & \begin{cases} \langle 0|_b\left(\frac{1}{\sqrt{2}}|0\rangle_a|1\rangle_b + \frac{i}{2}|0\rangle_a|0\rangle_b - \frac{1}{2}|1\rangle_a|1\rangle_b\right)\left(\frac{1}{\sqrt{2}}\langle 0|_a\langle 1|_b - \frac{i}{2}\langle 0|_a\langle 0|_b - \frac{1}{2}\langle 1|_a\langle 1|_b\right)|0\rangle_b \\ +\langle 1|_b\left(\frac{1}{\sqrt{2}}|0\rangle_a|1\rangle_b + \frac{i}{2}|0\rangle_a|0\rangle_b - \frac{1}{2}|1\rangle_a|1\rangle_b\right)\left(\frac{1}{\sqrt{2}}\langle 0|_a\langle 1|_b - \frac{i}{2}\langle 0|_a\langle 0|_b - \frac{1}{2}\langle 1|_a\langle 1|_b\right)|1\rangle_b \end{cases} \\
= \; & \left(0 + \frac{i}{2}|0\rangle_a + 0\right)\left(0 - \frac{i}{2}\langle 0|_a + 0\right) + \left(\frac{1}{\sqrt{2}}|0\rangle_a + 0 - \frac{1}{2}|1\rangle_a\right)\left(\frac{1}{\sqrt{2}}\langle 0|_a + 0 - \frac{1}{2}\langle 1|_a\right) \\
= \; & \frac{1}{4}|0\rangle_a\langle 0|_a + \frac{1}{2}|0\rangle_a\langle 0|_a - \frac{1}{2\sqrt{2}}|0\rangle_a\langle 1|_a - \frac{1}{2\sqrt{2}}|1\rangle_a\langle 0|_a + \frac{1}{4}|1\rangle_a\langle 1|_a \\
= \; & \frac{3}{4}|0\rangle_a\langle 0|_a - \frac{1}{2\sqrt{2}}|0\rangle_a\langle 1|_a - \frac{1}{2\sqrt{2}}|1\rangle_a\langle 0|_a + \frac{1}{4}|1\rangle_a\langle 1|_a
\end{aligned}$$

■

**Example** Alice and Bob each have one qubit from a pair that share the state

$$|\psi\rangle = \frac{|0\rangle_a|0\rangle_b + |1\rangle_a|1\rangle_b}{\sqrt{2}}$$

and the <u>pure</u> density matrix of their state is

$$\rho = |\psi\rangle\langle\psi| = \left(\frac{|0\rangle_a|0\rangle_b + |1\rangle_a|1\rangle_b}{\sqrt{2}}\right)\left(\frac{\langle 0|_a\langle 0|_b + \langle 1|_a\langle 1|_b}{\sqrt{2}}\right)$$

The reduced density matrix for Bob is:

$$\begin{aligned}
& \rho_b \\
= \; & Tr_a[\rho] \\
= \; & \left\{ \begin{array}{l} \langle 0|_a \left(\frac{|0\rangle_a|0\rangle_b+|1\rangle_a|1\rangle_b}{\sqrt{2}}\right)\left(\frac{\langle 0|_a\langle 0|_b+\langle 1|_a\langle 1|_b}{\sqrt{2}}\right)|0\rangle_a \\ +\langle 1|_a \left(\frac{|0\rangle_a|0\rangle_b+|1\rangle_a|1\rangle_b}{\sqrt{2}}\right)\left(\frac{\langle 0|_a\langle 0|_b+\langle 1|_a\langle 1|_b}{\sqrt{2}}\right)|1\rangle_a \end{array} \right. \\
= \; & \left\{ \begin{array}{l} \left(\frac{\langle 0|0\rangle_a|0\rangle_b+\langle 0|1\rangle_a|1\rangle_b}{\sqrt{2}}\right)\left(\frac{\langle 0|0\rangle_a\langle 0|_b+\langle 1|0\rangle_a\langle 1|_b}{\sqrt{2}}\right) \\ +\left(\frac{\langle 1|0\rangle_a|0\rangle_b+\langle 1|1\rangle_a|1\rangle_b}{\sqrt{2}}\right)\left(\frac{\langle 0|1\rangle_a\langle 0|_b+\langle 1|1\rangle_a\langle 1|_b}{\sqrt{2}}\right) \end{array} \right. \\
= \; & \left(\frac{|0\rangle_b}{\sqrt{2}}\right)\left(\frac{\langle 0|_b}{\sqrt{2}}\right) + \left(\frac{|1\rangle_b}{\sqrt{2}}\right)\left(\frac{\langle 1|_b}{\sqrt{2}}\right) \\
= \; & \tfrac{1}{2}|0\rangle_b\langle 0|_b + \tfrac{1}{2}|1\rangle_b\langle 1|_b
\end{aligned}$$

Notice that this is a <u>mixed</u> state.

■

## Exercises

### #1) What happens when Bob tilts his head?

a) Find the density matrix from the first "Alice and Bob" example near the beginning of the lecture in the $\{|+\rangle, |-\rangle\}$ basis.

b) If Bob is measuring the incoming photons in the $\{|+\rangle, |-\rangle\}$ basis, what is the probability of seeing each state?

### #2) Complimentary randomness.

a) Find the density matrix for the ensemble $p(|0\rangle) = p(|1\rangle) = \frac{1}{2}$. Write this density matrix as an actual matrix in the $\{|0\rangle, |1\rangle\}$ basis.

b) Find the density matrix for the ensemble $p(|+\rangle) = p(|-\rangle) = \frac{1}{2}$. Write this density matrix as an actual matrix in the $\{|0\rangle, |1\rangle\}$ basis. Do you expect to be able to tell the difference between these two ensembles using measurements?

c) For each of the density matrices you found in parts a and b, find the probability of measuring the state $|\theta\rangle = \cos(\theta)|0\rangle + \sin(\theta)|1\rangle$.

d) What's happening here? Why doesn't it matter what angle you do the measurement?

e) Find the density matrix for the ensemble $p(|0\rangle) = p$ and $p(|1\rangle) = q$ and find the probability of measuring the state $|\theta\rangle$ once again.

### #3) Taking my qubit and going home.

Alice and Bob each have one qubit from the shared state

$$|\psi\rangle = \frac{3i}{13}|1\rangle_a|1\rangle_b + \frac{4}{13}|2\rangle_a|1\rangle_b - \frac{12}{13}|2\rangle_a|2\rangle_b$$

but aren't on speaking terms at this moment. What are Alice and Bob's density matrices, $\rho_a$ and $\rho_b$?