Quantum Computation **(CMU 18-859BB, Fall 2015)**

Lecture 16: Mixed States and Measurement

November 2, 2015

Lecturer: John Wright *Scribe: Zilin Jiang*

1 Mixed States

Today, we will change the topic from quantum algorithm to quantum information theory. To lay the foundation for the qunatum information theory, we first need to generalize the definition of quantum states.

1.1 Motivation

Recall that, in the second lecture, we defined quantum states as vectors. Specifically, a quantum state in the d-dimensional qudit system is a superposition of d basis states. We can write:

$$|\psi\rangle = \alpha_1 |1\rangle + \alpha_2 |2\rangle + \cdots + \alpha_d |d\rangle,$$

where $|\alpha_1|^2 + |\alpha_2|^2 + \cdots + |\alpha_d|^2 = 1$.

However, using vectors to describe the state of a quantum system sometimes is not enough. One motivation to generalize the definition of a quantum state is to model quantum noise. When you implement a quantum system, the qunatum processes involved are naturally "noisy", whatever that means, and they are modeled as devices producing quantum states $|\psi_i\rangle$ with probability p_i. A schematic diagram of such a device is as follows.

$$\text{The devise outputs} \begin{cases} |\psi_1\rangle & \text{with probability } p_1, \\ |\psi_2\rangle & \text{with probability } p_2, \\ & \vdots \end{cases}$$

Roughly speaking, its quantum state is sometimes $|\psi_1\rangle$, sometimes $|\psi_2\rangle$ and so on.

One might be attempted to use a vector, for example $\sum_i p_i |\psi_i\rangle$, to represent the state of such a quantum device. But vectors are just not the correct notions to capture the quantum state of such a device.

1.2 Mixed state represented by matrix

In order to come up with the right notions to describe the physical system of such a devise, let's measure the devise in the basis $|v_1\rangle, |v_2\rangle, \ldots, |v_d\rangle$. We compute

$$\mathbf{Pr}[\text{observe } |v_i\rangle] = \sum_j p_j |\langle v_i|\psi_j\rangle|^2 = \sum_j p_j \langle v_i|\psi_j\rangle\langle\psi_j|v_i\rangle = \langle v_i| \left(\sum_j p_j |\psi_j\rangle \langle\psi_j|\right) |v_i\rangle \quad (1)$$

Definition 1.1. The *mixed state* $\{p_i, |\psi_i\rangle\}$ is represented by the matrix $\rho = \sum_j p_j |\psi_j\rangle \langle\psi_j|$. The state is called a pure state if $p_i = 1$ for some i.

The matrix ρ embodies everything related to the mixed state. In (1), we have seen that the outcome of the measurement in the basis can be expressed in terms of ρ, that is

$$\mathbf{Pr}\left[\text{observe } |v_i\rangle\right] = \langle v_i| \rho |v_i\rangle .$$

In the following example, we compute the matrices representing various mixed states.

Example 1.2. *The mixed state*

$$S_1 = \begin{cases} |0\rangle & \textit{with probability } 1 \\ |1\rangle & \textit{with probability } 0 \end{cases}$$

is represented by

$$|0\rangle \langle 0| = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

The mixed state

$$S_2 = \begin{cases} -|0\rangle & \textit{with probability } 1 \\ |1\rangle & \textit{with probability } 0 \end{cases}$$

is represented by the same matrix

$$(-|0\rangle)(-\langle 0|) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

The mixed state

$$S_3 = \begin{cases} |0\rangle & \textit{with probability } 1/2 \\ |1\rangle & \textit{with probability } 1/2 \end{cases}$$

is represented by

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

Though S_1 and S_2 in Example 1.2 look different, they are not distinguishable from an observer, and so they share the same representing matrix. This reflects the fact that the representing matrix contains only the observable information and does not contain redundant information.

Analogous to the vector formulation of quantum state, we can axiomatize the matrix formulation of mixed state as follows.

1. (Measurement) If ρ represents a mixed state and you measure in the basis $|v_1\rangle, \ldots, |v_d\rangle$, then $\mathbf{Pr}\left[\text{observe } |v_i\rangle\right] = \langle v_i| \rho |v_i\rangle$. We have computed this probability in (1).
2. (Evolution) Suppose a mixed state $S_1 = \{p_i, |\psi_i\rangle\}$ goes though a unitary gate U and transforms to $S_2 = \{p_i, U|\psi_i\rangle\}$. If the matrix representing S_1 is ρ, then the matrix representing S_2 is

$$\sum_i p_i(U|\psi_i\rangle)(U|\psi_i\rangle)^\dagger = \sum_i p_i U|\psi_i\rangle \langle\psi_i| U^\dagger = U\left(\sum_i p_i |\psi_i\rangle \langle\psi_i|\right) U^\dagger = U\rho U^\dagger.$$

1.3 Density matrix

The matrix presenting a mixed state has the following property.

Proposition 1.3. *If ρ represents a mixed state, then* $\mathrm{tr}(\rho) = 1$ *and ρ is positive semidefinite (PSD).*

Proof. Note that for a pure state $|\psi\rangle = a_1|1\rangle + \cdots + a_d|d\rangle$, $\mathrm{tr}(|\psi\rangle\langle\psi|) = |a_1|^2 + \cdots + |a_d|^2 = 1$. Suppose the mixed state is $\{p_i, |\psi_i\rangle\}$. On one hand, we have that $\mathrm{tr}(\rho) = \mathrm{tr}\left(\sum_i p_i |\psi_i\rangle\langle\psi_i|\right) = \sum_i p_i \,\mathrm{tr}(|\psi_i\rangle\langle\psi_i|) = \sum_i p_i = 1$. On the other hand, for any $|v\rangle$, we have that $\langle v|\rho|v\rangle = \mathbf{Pr}[\text{observe } |v\rangle] \geq 0$, hence that ρ is PSD. □

Definition 1.4. We say that a matrix ρ is a *density matrix* if and only if $\mathrm{tr}(\rho) = 1$ and ρ is PSD.

Proposition 1.3 simply says that a matrix representing a mixed state is a density matrix. In fact, the converse is also true.

Proposition 1.5. *If ρ is a density matrix, then it represents a mixed state.*

Proof. By the spectral theorem for Hermitian matrices, we know that $\rho = \sum_{i=1}^{d} \lambda_i |v_i\rangle\langle v_i|$, where λ_i's are the (real) eigenvalues and $|v_i\rangle$'s from an orthonormal basis. Since ρ is PSD, we know that $\lambda_i \geq 0$ for all i. The assumption $\mathrm{tr}(\rho) = 1$ implies that $\sum_{i=1}^{d} \lambda_i = 1$. So ρ represents the mixed state $\{p_i, |v_i\rangle : i \in [d]\}$. □

Remark 1.6. Even though a mixed state, say $\{p_i, |\psi_i\rangle\}$, can be rather complicated, for example, there are infinitely many nonzero p_i's, from the proof of Proposition 1.5 we see that $\{p_i, |\psi_i\rangle\}$ is indistinguishable from $\{\lambda_i, |v_i\rangle\}$ as they represent the same density matrix.

We have already seen in Example 1.2 that density matrix is a succinct way to represent a mixed state. One can actually use the density matrices to check whether two mixed states are distinguishable. Here is an example.

Example 1.7. *Suppose mixed state*

$$S_1 = \begin{cases} |0\rangle & \text{with probability } 3/4 \\ |1\rangle & \text{with probability } 1/4 \end{cases}$$

and mixed state

$$S_2 = \begin{cases} \frac{\sqrt{3}}{2}|0\rangle + \frac{1}{2}|1\rangle & \text{with probability } 1/2 \\ \frac{\sqrt{3}}{2}|0\rangle - \frac{1}{2}|1\rangle & \text{with probability } 1/2 \end{cases}.$$

Both mixed states are represented by

$$\frac{3}{4}|0\rangle\langle 0| + \frac{1}{4}|0\rangle\langle 0| = \begin{bmatrix} 3/4 & 0 \\ 0 & 1/4 \end{bmatrix}$$
$$= \frac{1}{2}\left(\frac{\sqrt{3}}{2}|0\rangle + \frac{1}{2}|1\rangle\right)\left(\frac{\sqrt{3}}{2}\langle 0| + \frac{1}{2}\langle 1|\right) + \frac{1}{2}\left(\frac{\sqrt{3}}{2}|0\rangle - \frac{1}{2}|1\rangle\right)\left(\frac{\sqrt{3}}{2}\langle 0| - \frac{1}{2}\langle 1|\right),$$

and hence indistinguishable.

1.4 Density matrix in quantum algorithms

In the hidden subgroup problem (HSP) for group G, if f "hides" a subgroup $H \leq G$, then the "standard method" outputs uniformly random coset

$$|gH\rangle = \frac{1}{\sqrt{|H|}} \sum_{h \in H} |gh\rangle .$$

Each coset $|gH\rangle$ is output with probability $\frac{1}{|G|}$. So the density matrix representing the mixed state is

$$\rho_H = \sum_{g \in G} \frac{1}{|G|} |gH\rangle \langle gH| = \frac{1}{|G|} \sum_{g \in G} |gH\rangle \langle gH| .$$

In general, partial measurements would generate mixed states.

2 Measurements

The next step is to generalize the definition of measurements.

2.1 Simple measurements

We first review the "simple measurements" (not standard terminology). Given a pure state $|\psi\rangle$, a "simple measurement" is as follows.

1. Pick orthonormal basis $|v_1\rangle, \ldots, |v_d\rangle$.
2. Receive outcome "i" with probability $|\langle v_i|\psi\rangle|^2$.
3. $|\psi\rangle$ "collapses" to $|v_i\rangle$.

2.2 Most general quantum measurement

The most general quantum measurement can be described using matrices. We first focus on the measurement rules for pure states.

1. Pick $d \times d$ matrices $M_1, \ldots, M_m$ satisfying the *completeness* condition that

$$M_1^\dagger M_1 + \cdots + M_m^\dagger M_m = I. \tag{2}$$

2. Receive outcome "i" with probability

$$|M_i |\psi\rangle|^2 = \langle\psi| M_i^\dagger M_i |\psi\rangle . \tag{3}$$

3. ψ collapses to

$$\frac{M_i |\psi\rangle}{|M_i |\psi\rangle|} = \frac{M_i |\psi\rangle}{\sqrt{\langle\psi| M_i^\dagger M_i |\psi\rangle}} \tag{4}$$

2.3 Sanity check

We first check that the $|M_i |\psi\rangle|^2$'s should sum to 1. Using the completeness condition (2), we have that

$$\sum_i |M_i |\psi\rangle|^2 = \sum_i \langle\psi| M_i^\dagger M_i |\psi\rangle = \langle\psi| \left(\sum_i M_i^\dagger M_i\right) |\psi\rangle = \langle\psi|\psi\rangle = 1.$$

Secondly, we check that the general measurement extends the "simple measurement". In the case of simple measurement, we measure with respect to orthonormal basis $|v_1\rangle, \dots, |v_d\rangle$. Take $M_i = |v_i\rangle\langle v_i|$ for all $i \in [d]$. Note that $M_i^\dagger M_i = |v_i\rangle\langle v_i| |v_i\rangle\langle v_i| = |v_i\rangle\langle v_i|$. According to the definition of the general quantum measurement, we check that

1. Completeness condition:
$$M_1^\dagger M_1 + \cdots + M_d^\dagger M_d = |v_1\rangle\langle v_1| + \cdots + |v_d\rangle\langle v_d| = I.$$
2. Receive "i" with probability
$$\langle\psi| M_i^\dagger M_i |\psi\rangle = \langle\psi|v_i\rangle\langle\psi|v_i\rangle = |\langle v_i|\psi_i\rangle|^2 .$$
3. ψ collapses to
$$\frac{M_i |\psi\rangle}{|M_i |\psi\rangle|} = \frac{|v_i\rangle\langle v_i|\psi\rangle}{|\langle v_i|\psi\rangle|} = e^{i\theta} |v_i\rangle .$$

We further define a special kind of general measurement, called projective measurement.

2.4 Projective measurement

Recall that a projection Π is a PSD matrix such that $\Pi^2 = \Pi$ (cf. Problem 6 in Homework 2). Equivalently, Π is a projection provided that $\Pi = \sum_{i=1}^k |v_i\rangle\langle v_i|$, where $|v_i\rangle$'s are orthonormal (but they do not necessarily form a basis).

Definition 2.1. A projective measurement is the case when $M_1 = \Pi_1, \dots, M_m = \Pi_m$, where Π's are projections such that

$$\Pi_1 + \cdots + \Pi_m = I. \tag{5}$$

Equation (5) implies the completeness condition (2). This is simply because $M_i^\dagger M_i = \Pi_i^\dagger \Pi_i = \Pi_i^2 = \Pi_i$. Moreover, we receive outcome "i" with probability $|\Pi |\psi\rangle|^2 = \langle\psi| \Pi_i^\dagger \Pi_i |\psi\rangle = \langle\psi| \Pi_i |\psi\rangle$ and the state collapses to $\Pi_i |\psi\rangle / |\Pi_i |\psi\rangle|$.

One way to think about a projective measurement is to imagine that you pick an orthonormal basis $|v_1\rangle, \dots, |v_d\rangle$ and "bundle" those vectors into projectors: $\Pi_i = \sum_{j\in S_i} |v_j\rangle\langle v_j|$ for all $i \in [m]$, where $S_1, \dots, S_m$ form a partition of $[d]$. Moreover, it is easy to see that the "simple measurement" is a projective measurement.

Example 2.2. *Given state* $|\psi\rangle = \frac{1}{\sqrt{3}}(|0\rangle + |1\rangle + |2\rangle)$. *If we pick* $\Pi_1 = |0\rangle\langle 0|, \Pi_2 = |1\rangle\langle 1| + |2\rangle\langle 2|$ *and carry out a projective measurement on* $|\psi\rangle$. *Since* $\Pi_1 |\psi\rangle = \frac{1}{\sqrt{3}} |0\rangle$ *and* $\Pi_2 |\psi\rangle = \frac{1}{\sqrt{3}}(|1\rangle + |2\rangle)$, *we will observe "1" and get* $|0\rangle$ *with probability* $\frac{1}{3}$, *and we will observe "2" and get* $\frac{1}{\sqrt{2}}(|1\rangle + |2\rangle)$ *with probability* $\frac{2}{3}$.

2.5 Measurement rules for mixed states

Now, we derive the measurement rules for mixed states.

Proposition 2.3. *Given mixed state $\{p_i, |\psi_i\rangle\}$ represented by density matrix ρ. If we measure the mixed state with respect to $M_1, \ldots, M_m$ satisfying the completeness condition, then*

1. *Receive outcome "i" with probability*
$$\operatorname{tr}\left(M_i^\dagger M_i \rho\right).$$
2. *ρ collapses to*
$$\frac{M_i \rho M_i^\dagger}{\operatorname{tr}\left(M_i^\dagger M_i \rho\right)}.$$

Proof. By the definition of density matrix, we have that $\rho = \sum_i p_i |\psi_i\rangle \langle\psi_i|$. Note that

$$\operatorname{tr}\left(M_i^\dagger M_i \rho\right) = \operatorname{tr}\left(M_i^\dagger M_i \sum_j p_j |\psi_j\rangle \langle\psi_j|\right) = \sum_j p_j \operatorname{tr}\left(M_i^\dagger M_i |\psi_j\rangle \langle\psi_j|\right) = \sum_j \langle\psi_j| M_i^\dagger M_i |\psi_j\rangle .$$

In the last equality, we used the simple fact that $\operatorname{tr}(|a\rangle \langle b|) = \langle b|a\rangle$ Using conditional probability, it is easy to see

$$p_i' := \mathbf{Pr}[\text{receive } i] = \sum_j p_j \mathbf{Pr}[\text{receive } i \mid \text{measure } |\psi_j\rangle] = \sum_j p_j \langle\psi_j| M_i^\dagger M_i |\psi_j\rangle = \operatorname{tr}\left(M_i^\dagger M_i \rho\right).$$

Because pure state $|\psi_j\rangle$ collapses to pure state

$$|\psi_j'\rangle = \frac{M_i |\psi_j\rangle}{\sqrt{\langle\psi_j| M_i^\dagger M_i |\psi_j\rangle}} =: \frac{M_i |\psi_j\rangle}{\sqrt{p_{ji}}},$$

with probability p_{ji}, the mixed state $\{p_j, |\psi_j\rangle\}$ collapses to $\left\{p_j p_{ji}/p_i', |\psi_j'\rangle\right\}$ represented by

$$\sum_j \frac{p_j p_{ji}}{p_i'} \left(\frac{M_i |\psi_j\rangle}{\sqrt{p_{ji}}}\right)\left(\frac{M_i |\psi_j\rangle}{\sqrt{p_{ji}}}\right)^\dagger = \frac{1}{p_i'} \sum_j p_j M_i |\psi_j\rangle \langle\psi_j| M_i^\dagger = \frac{M_i \rho M_i^\dagger}{\operatorname{tr}\left(M_i^\dagger M_i \rho\right)}.$$

□

3 Mixed state in entanglement

Suppose Alice and Bob share a pair of qudits with joint state $|\psi\rangle = \sum_{i,j\in[d]} \alpha_{i,j} |i\rangle \otimes |j\rangle$. Sadly, Bob is light-years away from Alice and got forgotten by everyone. The question is

"How do you describe Alice's qudit?"[1].

We claim that from Alice's perspective, her qudit is represented as a mixed state. In other words, if she performs a measurement on her qudit, then her measurement outcomes are consistent with her qudit being a mixed state. Which mixed state?

Suppose prior to Alice's measurement, Bob measures his qudit (in standard basis). He sees $|j\rangle$ with probability $p_j := \sum_i |\alpha_{i,j}|^2$ and Alice's state becomes $|\psi_j\rangle := \frac{1}{\sqrt{p_j}}\sum_i \alpha_{i,j} |i\rangle$. Had Bob measured before Alice, Alice's state would become a mixed state represented by

$$\sum_j p_j |\psi_j\rangle \langle\psi_j| = \sum_j p_j \left(\frac{1}{\sqrt{p_j}} \sum_{i_1} \alpha_{i_1,j} |i_1\rangle\right) \left(\frac{1}{\sqrt{p_j}} \sum_{i_2} \alpha^\dagger_{i_2,j} \langle i_2|\right) = \sum_{i_1,i_2} |i_1\rangle \langle i_2| \sum_j \alpha_{i_1,j}\alpha_{i_2,j}.$$

However, by relativity, Bob's measurement information hasn't propagated to Alice's world since Bob is light-years away. This means that, without loss of generality, we may always assume Bob had measured before Alice.

Well, if you believe in relativity, then you're forced to conclude that Alice's state is given by the mixed state

$$\rho_A = \sum_{i_1,i_2\in[d]} |i_1\rangle \langle i_2| \sum_{j\in[d]} \alpha_{i_1,j}\alpha^\dagger_{i_2,j}.$$

Of course, you might not believe in relativity (your loss), in which case this argument isn't too convincing. So let's see a quantum mechanical proof of this fact.

Before this, let me describe the standard way that ρ_A is defined. Suppose R and S are $d \times d$ matrices, where we think of R as Alice's matrix and S as Bob's matrix. Then the partial trace tr_B is defined as $\mathrm{tr}_B(R \otimes S) := R \cdot \mathrm{tr}(S)$.

The notation is suggestive: you take the trace of Bob's matrix S and multiply it by R. This shows how to define tr_B for $d^2 \times d^2$ matrices of the form $R \otimes S$. If we further specify that tr_B is linear, then $\mathrm{tr}_B(M)$ is defined for any $d^2 \times d^2$ matrix M. This is because any such M can be expanded as $M = \sum_i R_i \otimes S_i$, where the R_i's and S_i's are $d \times d$.

Now, I claim that ρ_A (as defined above) is equal to $\mathrm{tr}_B(|\psi\rangle\langle\psi|)$. To see this, first note that

$$|\psi\rangle\langle\psi| = \sum_{i_1,j_1,i_2,j_2} \alpha_{i_1,j_1}\alpha^\dagger_{i_2,j_2} |i_1\rangle \langle i_2| \otimes |j_1\rangle \langle j_2|.$$

Thus, by linearity of tr_B and the fact that $\mathrm{tr}(|a\rangle\langle b|) = 1$ if $a = b$ and 0 otherwise,

$$\begin{aligned}\mathrm{tr}_B(|\psi\rangle\langle\psi|) &= \sum_{i_1,j_1,i_2,j_2} \alpha_{i_1,j_1}\alpha^\dagger_{i_2,j_2} \mathrm{tr}_B(|i_1\rangle \langle i_2| \otimes |j_1\rangle \langle j_2|)\\ &= \sum_{i_1,j_1,i_2,j_2} \alpha_{i_1,j_1}\alpha^\dagger_{i_2,j_2} |i_1\rangle \langle i_2| \cdot \mathrm{tr}(|j_1\rangle \langle j_2|)\\ &= \sum_{i_1,i_2}\sum_j \alpha_{i_1,j}\alpha^\dagger_{i_2,j} |i_1\rangle \langle i_2|,\end{aligned}$$

[1] The notes below are mostly based on a Piazza post by John Wright.

which, after some trivial manipulations, is exactly ρ_A from above. So any time you see the notation tr_B applied to a multi-qudit state, it just means that you're supposed to compute the mixed state that Alice sees. And the above argument shows how to carry out this process in a simple manner. By the way, if we wanted to figure out which mixed state Bob's qudit is in, then it would be in the state $\mathrm{tr}_A(|\psi\rangle\langle\psi|)$, where tr_A is defined so that $\mathrm{tr}_A(R \otimes S) = \mathrm{tr}(R) \cdot S$.

Okay, so now let's prove that $\rho_A = \mathrm{tr}_B(|\psi\rangle\langle\psi|)$ is the state of Alice's qudit. This means that if she performs a measurement on her qudit, then her measurement outcome is consistent with her qudit being in the mixed state ρ_A. If Alice measures her qudit in the standard basis $|1\rangle, \ldots, |d\rangle$, then she sees "i" with with probability $\sum_j |\alpha_{i,j}|^2$. This is identical to performing the projective measurement $|1\rangle\langle 1| \otimes I, \ldots, |d\rangle\langle d| \otimes I$ (where I is the $d \times d$ identity matrix) on the whole state $|\psi\rangle$:

$$\mathbf{Pr}\left[\text{Alice observe } |i\rangle\right] = |(|i\rangle\langle i| \otimes I)|\psi\rangle|^2 = \left|\sum_{i',j} \alpha_{i',j}(|i\rangle\langle i|i'\rangle) \otimes |j\rangle\right|^2 = \sum_j |\alpha_{i,j}|^2.$$

Then we claimed that (without proof) if Alice measures her qudit in the basis $|v_1\rangle, \cdots, |v_d\rangle$, then this is identical to performing the projective measurement $|v_1\rangle\langle v_1| \otimes I, \cdots, |v_d\rangle\langle v_d| \otimes I$. Let's now prove this fact. If Alice measures in the basis $|v_1\rangle, \cdots, |v_d\rangle$, this means first applying the unitary $U = \sum_i |i\rangle\langle v_i|$ to her qudit and then measuring in the standard basis. This is equivalent to applying the unitary $U \otimes I$ to the total state $|\psi\rangle$ and then performing the projective measurement $|1\rangle\langle 1| \otimes I, \cdots, |d\rangle\langle d| \otimes I$. So the probability Alice observes outcome $|v_i\rangle$ is the probability she observes outcome $|i\rangle$ when measuring the state $(U \otimes I)|\psi\rangle$. And this probability is equal to

$$\langle\psi|(U^\dagger \otimes I)(|i\rangle\langle i| \otimes I)(U \otimes I)|\psi\rangle = \langle\psi|(U^\dagger|i\rangle)(\langle i|U) \otimes I|\psi\rangle = \langle\psi|(|v_i\rangle\langle v_i| \otimes I)|\psi\rangle,$$

which is identical to the measurement distribution of the projective measurement $|v_1\rangle\langle v_1| \otimes I, \cdots, |v_d\rangle\langle v_d| \otimes I$, as claimed.

With this fact in hand, we can now prove the main result, which is that Alice's state is given by the mixed state ρ_A. In particular, if she measures in a basis $|v_1\rangle, \cdots, |v_d\rangle$, then she should see outcome $|v_i\rangle$ with probability $\langle v_i|\rho|v_i\rangle$. Let's verify this. By the above fact, the probability that she gets outcome i when measuring is equal to

$$\begin{aligned}
\langle\psi|\left(|v_i\rangle\langle v_i|\otimes I\right)|\psi\rangle &= \left(\sum_{i_1,j_1}\alpha^\dagger_{i_1,j_1}\langle i_1|\otimes\langle j_1|\right)\left(|v_i\rangle\langle v_i|\otimes I\right)\left(\sum_{i_2,j_2}\alpha_{i_2,j_2}|i_2\rangle\otimes|j_2\rangle\right)\\
&= \sum_{i_1,j_1,i_2,j_2}\alpha^\dagger_{i_1,j_1}\alpha_{i_2,j_2}\langle i_1|\otimes\langle j_1|\left(|v_i\rangle\langle v_i|\otimes I\right)|i_2\rangle\otimes|j_2\rangle\\
&= \sum_{i_1,j_1,i_2,j_2}\alpha^\dagger_{i_1,j_1}\alpha_{i_2,j_2}\langle i_1|v_i\rangle\langle v_i|i_2\rangle\langle j_1|j_2\rangle\\
&= \sum_{i_1,i_2,j}\alpha^\dagger_{i_1,j}\alpha_{i_2,j}\langle v_i|i_2\rangle\langle i_1|v_i\rangle\\
&= \langle v_i|\left(\sum_{i_1,i_2,j}\alpha^\dagger_{i_1,j}\alpha_{i_2,j}|i_2\rangle\langle i_1|\right)|v_i\rangle\\
&= \langle v_i|\rho_A|v_i\rangle,
\end{aligned}$$

exactly as promised. So ρ_A represents Alice's mixed state, and we're done.

One nice thing about the partial trace definition is it allows us to prove the following fact: suppose that Bob applies the unitary U to his qudit. Then Alice's state should remain unchanged. In particular, her mixed state should be the same before and after Bob applies his unitary. To see this, her mixed state after Bob applies his unitary is given by

$$\begin{aligned}
\mathrm{tr}_B\left((I\otimes U)|\psi\rangle\langle\psi|(I\otimes U^\dagger)\right) &= \mathrm{tr}_B\left(\sum_{i_1,j_1,i_2,j_2}\alpha_{i_1,j_1}\alpha^\dagger_{i_2,j_2}\left(|i_1\rangle\langle i_2|\right)\otimes\left(U|j_1\rangle\langle j_2|U^\dagger\right)\right)\\
&= \sum_{i_1,j_1,i_2,j_2}\alpha_{i_1,j_1}\alpha^\dagger_{i_2,j_2}|i_1\rangle\langle i_2|\otimes\mathrm{tr}\left(U|j_1\rangle\langle j_2|U^\dagger\right)\\
&= \sum_{i_1,j_1,i_2,j_2}\alpha_{i_1,j_1}\alpha^\dagger_{i_2,j_2}|i_1\rangle\langle i_2|\otimes\mathrm{tr}\left(|j_1\rangle\langle j_2|\right)\\
&= \mathrm{tr}_B\left(|\psi\rangle\langle\psi|\right)\\
&= \rho_A.
\end{aligned}$$

Here we used the fact that $\mathrm{tr}(ABC) = \mathrm{tr}(BCA)$. So Bob can apply any unitary rotation to his qudit that he wants, but from Alice's perspective her qudit's state never changes. We saw this previously in lecture 3 in the special case of the quantum teleportation circuit.