PHYS 7895: Quantum Information Theory	Fall 2015
Lecture 8 — September 21, 2015	
Prof. Mark M. Wilde	*Scribe: Mark M. Wilde*

1 Overview

In the last lecture we presented the notions of an ensemble of quantum states and a density operator. We also discussed unitary evolution of the density operator and measurement of a density operator.

In this lecture we continue our development of the noisy quantum theory, discussing a more general model of measurements, the POVM formalism, product states, separable states, and the partial trace operation.

The material is coming from Chapter 4 of `http://markwilde.com/qit-notes.pdf` .

2 Measurement in the Noisy Quantum Theory

We have described measurement in the quantum theory using a set of projectors that form a resolution of the identity. For example, the set $\{\Pi_j\}_j$ of projectors that satisfy the condition $\sum_j \Pi_j = I$ form a valid von Neumann quantum measurement. A projective measurement is not the most general measurement that we can perform on a quantum system (though it is certainly one valid type of quantum measurement).

There is a more general description of quantum measurements that follows from allowing the system of interest to interact unitarily with a probe system that we measure after the interaction occurs. So suppose that the system of interest is in a state $|\psi\rangle_S$ and that the probe is in a state $|0\rangle_P$, so that the overall state before anything happens is as follows:

$$|\psi\rangle_S \otimes |0\rangle_P. \tag{1}$$

Let $\{|0\rangle_P, |1\rangle_P, \ldots, |d-1\rangle_P\}$ be an orthonormal basis for the probe system (assuming that it has dimension d). Now suppose that the system and the probe interact according to a unitary U_{SP}, and then we perform a measurement of the probe system, described by measurement operators $\{|j\rangle\langle j|_P\}$. The probability to obtain outcome j is

$$p_J(j) = \left(\langle\psi|_S \otimes \langle 0|_P U_{SP}^\dagger\right) \left(I_S \otimes |j\rangle\langle j|_P\right) \left(U_{SP}|\psi\rangle_S \otimes |0\rangle_P\right), \tag{2}$$

and the post-measurement state upon obtaining outcome j is

$$\frac{1}{\sqrt{p_J(j)}} \left(I_S \otimes |j\rangle\langle j|_P\right) \left(U_{SP}|\psi\rangle_S \otimes |0\rangle_P\right). \tag{3}$$

We can rewrite the expressions above in a simpler way. Let us expand the unitary operator U_{SP} in the orthonormal basis of the probe system P as follows:

$$U_{SP} = \sum_{j,k} M_S^{j,k} \otimes |j\rangle\langle k|_P, \tag{4}$$

where $\{M_S^{j,k}\}$ is a set of operators. Up to a permutation of the S and P systems and using the mathematics of the tensor product, this is the same as writing the unitary U_{SP} as follows:

$$\begin{bmatrix} M_S^{0,0} & M_S^{0,1} & \cdots & M_S^{0,d-1} \\ M_S^{1,0} & M_S^{1,1} & \cdots & M_S^{1,d-1} \\ \vdots & \vdots & \ddots & \vdots \\ M_S^{d-1,0} & M_S^{d-1,1} & \cdots & M_S^{d-1,d-1} \end{bmatrix}. \tag{5}$$

This set $\{M_S^{j,k}\}$ needs to satisfy some constraints corresponding to the unitarity of U_{SP}. In particular, consider the following operator:

$$\sum_j M_S^{j,0} \otimes |j\rangle\langle 0|_P, \tag{6}$$

which corresponds to the first column of operator-valued entries in U_{SP}, as illustrated in (5). In what follows, we employ the shorthand $M_S^j \equiv M_S^{j,0}$. From the fact that $U_{SP}^\dagger U_{SP} = I_{SP} = I_S \otimes I_P$, we deduce that the following equality must hold

$$I_S \otimes |0\rangle\langle 0|_P = \left(\sum_{j'} M_S^{j'\dagger} \otimes |0\rangle\langle j'|_P\right)\left(\sum_j M_S^j \otimes |j\rangle\langle 0|_P\right) \tag{7}$$

$$= \sum_{j',j} M_S^{j'\dagger} M_S^j \otimes |0\rangle \langle j'|j\rangle \langle 0|_P \tag{8}$$

$$= \sum_j M_S^{j\dagger} M_S^j \otimes |0\rangle\langle 0|_P, \tag{9}$$

where the last line follows from the fact that we chose an orthonormal basis in the representation of U_{SP} in (4). So this implies that the following condition holds

$$\sum_j M_S^{j\dagger} M_S^j = I_S. \tag{10}$$

Plugging (4) into (2) and (3), a short calculation (similar to the above one) reveals that they simplify as follows:

$$p_J(j) = \langle\psi| M_j^\dagger M_j |\psi\rangle, \tag{11}$$

$$\frac{1}{\sqrt{p_J(j)}} (I_S \otimes |j\rangle\langle j|_P)(U_{SP}|\psi\rangle_S \otimes |0\rangle_P) = \frac{M_j |\psi\rangle_S \otimes |j\rangle_P}{\sqrt{p_J(j)}}. \tag{12}$$

Since the system and the probe are in a pure product state (and thus independent of each other) after the measurement occurs, we can discard the probe system and deduce that the post-measurement state is simply $M_j |\psi\rangle_S / \sqrt{p_J(j)}$.

Motivated by the above development, we allow for a more general notion of quantum measurement, saying that it consists of a set of measurement operators $\{M_j\}_j$ that satisfy the following completeness condition:

$$\sum_j M_j^\dagger M_j = I. \tag{13}$$

Observe from the above development that this is the only constraint that the operators $\{M_j\}$ need to satisfy. This constraint is a consequence of unitarity, but can be viewed as a generalization of the completeness relation for a set of projectors that constitute a projective quantum measurement. Given a set of measurement operators of the above form, the probability for obtaining outcome j when measuring a state $|\psi\rangle$ is

$$p_J(j) \equiv \langle\psi| M_j^\dagger M_j |\psi\rangle , \tag{14}$$

and the post-measurement state when we receive outcome j is

$$\frac{M_j |\psi\rangle}{\sqrt{p_J(j)}}. \tag{15}$$

Suppose that we instead have an ensemble $\{p_X(x), |\psi_x\rangle\}$ with density operator ρ. We can carry out an analysis similar to that in the last lecture to conclude that the probability $p_J(j)$ for obtaining outcome j is

$$p_J(j) \equiv \text{Tr}\{M_j^\dagger M_j \rho\}, \tag{16}$$

and the post-measurement state when we measure result j is

$$\frac{M_j \rho M_j^\dagger}{p_J(j)}. \tag{17}$$

The expression $p_J(j) = \text{Tr}\{M_j^\dagger M_j \rho\}$ is a restatement of the Born rule.

2.1 POVM Formalism

Sometimes, we simply may not care about the post-measurement state of a quantum measurement, but instead we only care about the probability for obtaining a particular outcome. This situation arises in the transmission of classical data over a quantum channel. In this situation, we are merely concerned with minimizing the error probabilities of the classical transmission. The receiver does not care about the post-measurement state because he no longer needs it in the quantum information-processing protocol.

We can specify a measurement of this sort by some set $\{\Lambda_j\}_j$ of operators that satisfy non-negativity and completeness:

$$\Lambda_j \geq 0, \tag{18}$$

$$\sum_j \Lambda_j = I. \tag{19}$$

The set $\{\Lambda_j\}_j$ of operators is a positive operator-valued measure (POVM). The probability for obtaining outcome j is

$$\langle\psi| \Lambda_j |\psi\rangle , \tag{20}$$

if the state is some pure state $|\psi\rangle$. The probability for obtaining outcome j is

$$\text{Tr}\left\{\Lambda_j \rho\right\}, \tag{21}$$

if the state is in a mixed state described by some density operator ρ. This is another restatement of the Born rule.

3 Composite Noisy Quantum Systems

We are again interested in the behavior of two or more quantum systems when we join them together. Some of the most exotic, truly "quantum" behavior occurs in joint quantum systems, and we observe a marked departure from the classical world.

3.1 Independent Ensembles

Let us first suppose that we have two independent ensembles for quantum systems A and B. The first quantum system belongs to Alice and the second quantum system belongs to Bob, and they may or may not be spatially separated. Let $\{p_X(x), |\psi_x\rangle\}$ be the ensemble for the system A and let $\{p_Y(y), |\phi_y\rangle\}$ be the ensemble for the system B. Suppose for now that the state on system A is $|\psi_x\rangle$ for some x and the state on system B is $|\phi_y\rangle$ for some y. Then, using the composite system postulate of the noiseless quantum theory, the joint state for a given x and y is $|\psi_x\rangle \otimes |\phi_y\rangle$. The density operator for the joint quantum system is the expectation of the states $|\psi_x\rangle \otimes |\phi_y\rangle$ with respect to the random variables X and Y that describe the individual ensembles:

$$\mathbb{E}_{X,Y}\left\{\left(|\psi_X\rangle \otimes |\phi_Y\rangle\right)\left(\langle\psi_X| \otimes \langle\phi_Y|\right)\right\}. \tag{22}$$

The above expression is equal to the following one:

$$\mathbb{E}_{X,Y}\left\{|\psi_X\rangle\langle\psi_X| \otimes |\phi_Y\rangle\langle\phi_Y|\right\}, \tag{23}$$

because $\left(|\psi_x\rangle \otimes |\phi_y\rangle\right)\left(\langle\psi_x| \otimes \langle\phi_y|\right) = |\psi_x\rangle\langle\psi_x| \otimes |\phi_y\rangle\langle\phi_y|$. We then explicitly write out the expectation as a sum over probabilities:

$$\sum_{x,y} p_X(x) p_Y(y) |\psi_x\rangle\langle\psi_x| \otimes |\phi_y\rangle\langle\phi_y|. \tag{24}$$

We can distribute the probabilities and the sum because the tensor product obeys a distributive property:

$$\sum_x p_X(x)|\psi_x\rangle\langle\psi_x| \otimes \sum_y p_Y(y)|\phi_y\rangle\langle\phi_y|. \tag{25}$$

The density operator for this ensemble admits the following simple form:

$$\rho \otimes \sigma, \tag{26}$$

where $\rho = \sum_x p_X(x)|\psi_x\rangle\langle\psi_x|$ is the density operator of the X ensemble and $\sigma = \sum_y p_Y(y)|\phi_y\rangle\langle\phi_y|$ is the density operator of the Y ensemble. We can say that Alice's local density operator is ρ and Bob's local density operator is σ. The overall state is a tensor product of these two density operators.

Definition 1 (Product State)**.** *A density operator which is equal to a tensor product of two or more density operators is called a product state.*

We should expect the density operator to factor as it does above because we assumed that the ensembles are independent. There is nothing much that distinguishes this situation from the classical world, except for the fact that the states in each respective ensemble may be non-orthogonal to other states in the same ensemble. But even here, there is some equivalent description of each ensemble in terms of an orthonormal basis so that there is really not too much difference between this description and a joint probability distribution that factors as two independent distributions.

3.2 Separable States

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

$$\{p_X(x), |\psi_x\rangle \otimes |\phi_x\rangle\}. \tag{27}$$

It is straightforward to verify that the density operator of this correlated ensemble has the following form:

$$\mathbb{E}_X\{(|\psi_X\rangle \otimes |\phi_X\rangle)(\langle\psi_X| \otimes \langle\phi_X|)\} = \sum_x p_X(x)|\psi_x\rangle\langle\psi_x| \otimes |\phi_x\rangle\langle\phi_x|. \tag{28}$$

By ignoring Bob's system, Alice's local density operator is of the form

$$\mathbb{E}_X\{|\psi_X\rangle\langle\psi_X|\} = \sum_x p_X(x)|\psi_x\rangle\langle\psi_x|, \tag{29}$$

and similarly, Bob's local density operator is

$$\mathbb{E}_X\{|\phi_X\rangle\langle\phi_X|\} = \sum_x p_X(x)|\phi_x\rangle\langle\phi_x|. \tag{30}$$

States of the form in (28) can be generated by a classical procedure. A third party generates a symbol x according to the probability distribution $p_X(x)$ and sends the symbol x to both Alice and Bob. Alice prepares the state $|\psi_x\rangle$ and Bob prepares the state $|\phi_x\rangle$. If they then discard the symbol x, the state of their systems is given by (28).

We can generalize this classical preparation procedure one step further. Let us suppose that we first generate a random variable Z according to some distribution $p_Z(z)$. We then generate two other ensembles, conditional on the value of the random variable Z. Let $\{p_{X|Z}(x|z), |\psi_{x,z}\rangle\}$ be the first ensemble and let $\{p_{Y|Z}(y|z), |\phi_{y,z}\rangle\}$ be the second ensemble, where the random variables X and Y are independent when conditioned on Z. Let us label the density operators of the first and second ensembles when conditioned on a particular realization z by ρ_z and σ_z, respectively. It is then straightforward to verify that the density operator of an ensemble created from this classical preparation procedure has the following form:

$$\mathbb{E}_{X,Y,Z}\{(|\psi_{X,Z}\rangle \otimes |\phi_{Y,Z}\rangle)(\langle\psi_{X,Z}| \otimes \langle\phi_{Y,Z}|)\} = \sum_z p_Z(z)\rho_z \otimes \sigma_z. \tag{31}$$

Exercise 2. *By ignoring Bob's system, we can determine Alice's local density operator. Show that*

$$\mathbb{E}_{X,Y,Z}\left\{|\psi_{X,Z}\rangle\langle\psi_{X,Z}|\right\} = \sum_z p_Z(z)\rho_z, \tag{32}$$

so that the above expression is the density operator for Alice. It similarly follows that the local density operator for Bob is

$$\mathbb{E}_{X,Y,Z}\left\{|\phi_{Y,Z}\rangle\langle\phi_{Y,Z}|\right\} = \sum_z p_Z(z)\sigma_z. \tag{33}$$

Exercise 3. *Show that we can always write a state of the form in* (31) *as a convex combination of pure product states:*

$$\sum_z p_Z(z)\,|\phi_z\rangle\langle\phi_z| \otimes |\psi_z\rangle\langle\psi_z|, \tag{34}$$

by manipulating the general form in (31)*.*

As a consequence of Exercise 3, we see that any state of the form in (31) can be written as a convex combination of pure product states. Such states are called *separable* states, defined formally as follows:

Definition 4 (Separable State)**.** *A bipartite density operator σ_{AB} is a separable state if it can be written in the following form:*

$$\sigma_{AB} = \sum_x p_X(x)|\psi_x\rangle\langle\psi_x|_A \otimes |\phi_x\rangle\langle\phi_x|_B \tag{35}$$

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

The term "separable" implies that there is no quantum entanglement in the above state, i.e., there is a completely classical procedure that prepares the above state. In fact, this is the definition of entanglement for a general bipartite density operator:

Definition 5 (Entangled State)**.** *A bipartite density operator ρ_{AB} is entangled if it is not separable.*

3.2.1 Separable States and the CHSH Game

One motivation for Definitions 4 and 5 was already given above: for a separable state, there is a classical procedure that can be used to prepare it. Thus, for an entangled state, there is no such procedure. That is, a non-classical (quantum) interaction between the systems is necessary to prepare an entangled state.

Another related motivation is that separable states admit an explanation in terms of a classical strategy for the CHSH game. Recall from before that classical strategies $p_{AB|XY}(a,b|x,y)$ are of the following form:

$$p_{AB|XY}(a,b|x,y) = \int d\lambda\ p_\Lambda(\lambda)\ p_{A|\Lambda X}(a|\lambda,x)\ p_{B|\Lambda Y}(b|\lambda,y). \tag{36}$$

If we allow for a continuous index λ for a separable state, then we can write such a state as follows:

$$\sigma_{AB} = \int d\lambda\ p_\Lambda(\lambda)\ |\psi_\lambda\rangle\langle\psi_\lambda|_A \otimes |\phi_\lambda\rangle\langle\phi_\lambda|_B. \tag{37}$$

Recall that in a general quantum strategy, there are measurements $\{\Pi_a^{(x)}\}$ and $\{\Pi_b^{(y)}\}$, giving output bits a and b based on the input bits x and y and leading to the following strategy:

$$
\begin{aligned}
p_{AB|XY}(a,b|x,y) &= \mathrm{Tr}\{(\Pi_a^{(x)} \otimes \Pi_b^{(y)})\sigma_{AB}\} && (38)\\
&= \mathrm{Tr}\left\{(\Pi_a^{(x)} \otimes \Pi_b^{(y)})\left(\int d\lambda\ p_\Lambda(\lambda)\ |\psi_\lambda\rangle\langle\psi_\lambda|_A \otimes |\phi_\lambda\rangle\langle\phi_\lambda|_B\right)\right\} && (39)\\
&= \int d\lambda\ p_\Lambda(\lambda)\ \ \mathrm{Tr}\left\{\Pi_a^{(x)}|\psi_\lambda\rangle\langle\psi_\lambda|_A \otimes \Pi_b^{(y)}|\phi_\lambda\rangle\langle\phi_\lambda|_B\right\} && (40)\\
&= \int d\lambda\ p_\Lambda(\lambda)\ \langle\psi_\lambda|_A\Pi_a^{(x)}|\psi_\lambda\rangle_A\ \langle\phi_\lambda|_B\Pi_b^{(y)}|\phi_\lambda\rangle_B. && (41)
\end{aligned}
$$

By picking the probability distributions $p_{A|\Lambda X}(a|\lambda,x)$ and $p_{B|\Lambda Y}(b|\lambda,y)$ in (36) as follows:

$$
\begin{aligned}
p_{A|\Lambda X}(a|\lambda,x) &= \langle\psi_\lambda|_A\Pi_a^{(x)}|\psi_\lambda\rangle_A, && (42)\\
p_{B|\Lambda Y}(b|\lambda,y) &= \langle\phi_\lambda|_B\Pi_b^{(y)}|\phi_\lambda\rangle_B, && (43)
\end{aligned}
$$

we see that there is a classical strategy that can simulate any quantum strategy which uses separable states in the CHSH game. Thus, the winning probability of quantum strategies involving separable states are subject to the classical bound of 3/4 derived before. In this sense, such strategies are effectively classical.

4 Local Density Operators and Partial Trace

4.1 A First Example

Consider the entangled Bell state $|\Phi^+\rangle_{AB}$ shared on systems A and B. In the above analyses, we determined a local density operator description for both Alice and Bob. Now, we are curious if it is possible to determine such a local density operator description for Alice and Bob with respect to the state $|\Phi^+\rangle_{AB}$ or more general ones.

As a first approach to this issue, 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 merely because it is a mathematical representation that allows us to compute the probabilities resulting from a physical measurement. So, if we would like to determine a "local density operator," such a local density operator should predict the result of a local measurement.

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

$$\langle\Phi^+|_{AB}\,\Lambda_A^j \otimes I_B\,|\Phi^+\rangle_{AB} = \frac{1}{2}\sum_{k,l=0}^{1} \langle kk|_{AB}\,\Lambda_A^j \otimes I_B\,|ll\rangle_{AB} \tag{44}$$

$$= \frac{1}{2}\sum_{k,l=0}^{1} \langle k|_A\,\Lambda_A^j|l\rangle_A\,\langle k|l\rangle_B \tag{45}$$

$$= \frac{1}{2}\left(\langle 0|_A\Lambda_A^j|0\rangle_A + \langle 1|_A\Lambda_A^j\,|1\rangle_A\right) \tag{46}$$

$$= \frac{1}{2}\left(\mathrm{Tr}\left\{\Lambda_A^j|0\rangle\langle 0|_A\right\} + \mathrm{Tr}\left\{\Lambda_A^j|1\rangle\langle 1|_A\right\}\right) \tag{47}$$

$$= \mathrm{Tr}\left\{\Lambda_A^j\frac{1}{2}\left(|0\rangle\langle 0|_A + |1\rangle\langle 1|_A\right)\right\} \tag{48}$$

$$= \mathrm{Tr}\left\{\Lambda_A^j\pi_A\right\}. \tag{49}$$

The above steps follow by applying the rules of taking the inner product with respect to tensor product operators. The last line follows by recalling the definition of the maximally mixed state π, where π here is a qubit maximally mixed state.

The above calculation demonstrates that we can predict the result of any local "Alice" measurement using the density operator π. Therefore, 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.

4.2 Partial Trace

In general, we would like to determine a local density operator that predicts the outcomes of all local measurements. The general method for determining a local density operator is to employ the *partial trace operation*, which we motivate and define here, as a generalization of the example discussed at the beginning of Section 4.1.

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 rule, the probability for Alice to receive outcome j after performing the measurement is given by the following expression:

$$p_J(j) = \mathrm{Tr}\{(\Lambda_A^j \otimes I_B)\rho_{AB}\}. \tag{50}$$

In order to evaluate this trace, we can choose any orthonormal basis that we wish (recall the definition of trace and subsequent statements). Taking $\{|k\rangle_A\}$ as an orthonormal basis for Alice's Hilbert space and $\{|l\rangle_B\}$ as an orthonormal basis for Bob's Hilbert space, the set $\{|k\rangle_A \otimes |l\rangle_B\}$ constitutes an orthonormal basis for the tensor product of their Hilbert spaces. So we can evaluate

(50) as follows:

$$\mathrm{Tr}\{(\Lambda_A^j \otimes I_B)\rho_{AB}\} = \sum_{k,l} (\langle k|_A \otimes \langle l|_B) \left[(\Lambda_A^j \otimes I_B)\rho_{AB}\right] (|k\rangle_A \otimes |l\rangle_B) \tag{51}$$

$$= \sum_{k,l} \langle k|_A (I_A \otimes \langle l|_B) \left[(\Lambda_A^j \otimes I_B)\rho_{AB}\right] (I_A \otimes |l\rangle_B) |k\rangle_A \tag{52}$$

$$= \sum_{k,l} \langle k|_A \Lambda_A^j (I_A \otimes \langle l|_B) \rho_{AB} (I_A \otimes |l\rangle_B) |k\rangle_A \tag{53}$$

$$= \sum_{k} \langle k|_A \Lambda_A^j \left[\sum_l (I_A \otimes \langle l|_B) \rho_{AB} (I_A \otimes |l\rangle_B)\right] |k\rangle_A. \tag{54}$$

The first equality follows from the definition of the trace and using the orthonormal basis $\{|k\rangle_A \otimes |l\rangle_B\}$. The second equality follows because

$$|k\rangle_A \otimes |l\rangle_B = (I_A \otimes |l\rangle_B) |k\rangle_A. \tag{55}$$

The third equality follows because

$$(I_A \otimes \langle l|_B) (\Lambda_A^j \otimes I_B) = \Lambda_A^j (I_A \otimes \langle l|_B). \tag{56}$$

The fourth equality follows by bringing the sum over l inside. Using the definition of trace and the fact that $\{|k\rangle_A\}$ is an orthonormal basis for Alice's Hilbert space, we can rewrite (54) as

$$\mathrm{Tr}\left\{\Lambda_A^j \left[\sum_l (I_A \otimes \langle l|_B) \rho_{AB} (I_A \otimes |l\rangle_B)\right]\right\}. \tag{57}$$

Our final step is to define the partial trace operation as follows:

Definition 6 (Partial Trace)**.** *Let X_{AB} be a 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:*

$$\mathrm{Tr}_B\{X_{AB}\} \equiv \sum_l (I_A \otimes \langle l|_B) X_{AB} (I_A \otimes |l\rangle_B). \tag{58}$$

For simplicity, we often suppress the identity operators I_A and write this as follows:

$$\mathrm{Tr}_B\{X_{AB}\} \equiv \sum_l \langle l|_B X_{AB} |l\rangle_B. \tag{59}$$

For the same reason that the definition of the trace is invariant under the choice of an orthonormal basis, the same is true for the partial trace operation. We can also observe from the above definition that the partial trace is a linear operation. Continuing with our development above, we can define a local operator ρ_A, using the partial trace, as follows:

$$\rho_A = \mathrm{Tr}_B\{\rho_{AB}\}. \tag{60}$$

This then allows us to arrive at a rewriting of (57) as

$$\mathrm{Tr}\{\Lambda_A^j \rho_A\}, \tag{61}$$

which allows us to conclude that

$$p_J(j) = \text{Tr}\{(\Lambda_A^j \otimes I_B)\rho_{AB}\} = \text{Tr}\{\Lambda_A^j \rho_A\}. \tag{62}$$

Thus, from the operator ρ_A, we can predict the outcomes of local measurements that Alice performs on her system. Also important here is that the global picture, in which we have a density operator ρ_{AB} and a measurement of the form $\{\Lambda_A^j \otimes I_B\}$, is consistent with the local picture, in which the measurement is written as $\{\Lambda_A^j\}$ and the operator ρ_A is used to calculate the probabilities $p_J(j)$. The operator ρ_A is itself a density operator, called the *local* or *reduced density operator*, and the next exercise asks you to verify that it is indeed a density operator.

Exercise 7 (Local Density Operator)**.** *Let ρ_{AB} be a density operator acting on a bipartite Hilbert space. Prove that $\rho_A = \text{Tr}_B\{\rho_{AB}\}$ is a density operator, meaning that it is positive semi-definite and has trace equal to one.*

In conclusion, given a density operator ρ_{AB} describing the joint state held by Alice and Bob, we can always calculate a local density operator ρ_A, which describes the local state of Alice if Bob's system is inaccessible to her.

There is an alternate way of describing partial trace, of which it is helpful to be aware. For a simple state of the form

$$|x\rangle\langle x|_A \otimes |y\rangle\langle y|_B, \tag{63}$$

with $|x\rangle_A$ and $|y\rangle_B$ each unit vectors, the partial trace has the following action:

$$\text{Tr}_B\left\{|x\rangle\langle x|_A \otimes |y\rangle\langle y|_B\right\} = |x\rangle\langle x|_A \ \text{Tr}\left\{|y\rangle\langle y|_B\right\} = |x\rangle\langle x|_A, \tag{64}$$

where we "trace out" the second system to determine the local density operator for the first. If the partial trace acts on a tensor product of rank-one operators (not necessarily corresponding to a state)

$$|x_1\rangle\langle x_2|_A \otimes |y_1\rangle\langle y_2|_B\,, \tag{65}$$

its action is as follows:

$$\text{Tr}_B\left\{|x_1\rangle\langle x_2|_A \otimes |y_1\rangle\langle y_2|_B\right\} = |x_1\rangle\langle x_2|_A \ \text{Tr}\left\{|y_1\rangle\langle y_2|_B\right\} \tag{66}$$

$$= |x_1\rangle\langle x_2|_A \ \langle y_2|y_1\rangle\,. \tag{67}$$

In fact, an alternate way of defining the partial is as above and to extend it by linearity.

Exercise 8. *Show that the two notions of the partial trace operation are consistent. That is, show that*

$$\text{Tr}_B\left\{|x_1\rangle\langle x_2|_A \otimes |y_1\rangle\langle y_2|_B\right\} = \sum_i \langle i|_B\left(|x_1\rangle\langle x_2|_A \otimes |y_1\rangle\langle y_2|_B\right)|i\rangle_B \tag{68}$$

$$= |x_1\rangle\langle x_2|_A \ \langle y_2|y_1\rangle\,, \tag{69}$$

for some orthonormal basis $\{|i\rangle_B\}$ on Bob's system.

It can be helpful to see the alternate notion of partial trace worked out in detail. The most general density operator on two systems A and B is some operator ρ_{AB} that is positive semi-definite with unit trace. We can obtain the local density operator ρ_A from ρ_{AB} by tracing out the B system:

$$\rho_A = \text{Tr}_B\left\{\rho_{AB}\right\}. \tag{70}$$

In more detail, let us expand an arbitrary density operator ρ_{AB} with an orthonormal basis $\{|i\rangle_A \otimes |j\rangle_B\}_{i,j}$ for the bipartite (two-party) state:

$$\rho_{AB} = \sum_{i,j,k,l} \lambda_{i,j,k,l}(|i\rangle_A \otimes |j\rangle_B)(\langle k|_A \otimes \langle l|_B). \tag{71}$$

The coefficients $\lambda_{i,j,k,l}$ are the matrix elements of ρ_{AB} with respect to the basis $\{|i\rangle_A \otimes |j\rangle_B\}_{i,j}$, and they are subject to the constraint of non-negativity and unit trace for ρ_{AB} . We can rewrite the above operator as

$$\rho_{AB} = \sum_{i,j,k,l} \lambda_{i,j,k,l}|i\rangle\langle k|_A \otimes |j\rangle\langle l|_B. \tag{72}$$

We can now evaluate the partial trace:

$$\begin{aligned}
\rho_A &= \mathrm{Tr}_B \left\{ \sum_{i,j,k,l} \lambda_{i,j,k,l}|i\rangle\langle k|_A \otimes |j\rangle\langle l|_B \right\} && (73)\\
&= \sum_{i,j,k,l} \lambda_{i,j,k,l} \, \mathrm{Tr}_B \left\{ |i\rangle \langle k|_A \otimes |j\rangle\langle l|_B \right\} && (74)\\
&= \sum_{i,j,k,l} \lambda_{i,j,k,l}|i\rangle \langle k|_A \, \mathrm{Tr} \left\{ |j\rangle\langle l|_B \right\} && (75)\\
&= \sum_{i,j,k,l} \lambda_{i,j,k,l}|i\rangle \langle k|_A \langle j|l\rangle && (76)\\
&= \sum_{i,j,k} \lambda_{i,j,k,j}|i\rangle \langle k|_A && (77)\\
&= \sum_{i,k} \left(\sum_j \lambda_{i,j,k,j} \right) |i\rangle \langle k|_A. && (78)
\end{aligned}$$

The second equality exploits the linearity of the partial trace operation. The last equality explicitly shows how the partial trace operation earns its name—it is equivalent to performing a trace operation over the coefficients corresponding to Bob's system.