Lecture 14: Quantum information revisited

March 14, 2006

So far, this course has focused almost entirely on quantum algorithms. The next topics to be discussed will be of a somewhat different flavor, starting with quantum error correction and moving on to quantum cryptography. Before discussing these topics, however, it will be helpful to say more about the mathematics of quantum information in general.

In the beginning of the course I described the model of quantum information that we have used up to this point. Specifically, the model describes states of qubits as unit vectors, and the allowed operations come from a fairly restricted set (unitary operations and measurements of a simple type). This description was sufficient for discussing quantum algorithms, so it has not been necessary to extend it until now.

To describe the more general model of quantum information, some notation will be helpful. First, for any finite, nonempty set Σ, let $\mathbb{C}(\Sigma)$ denote the vector space of all column vectors indexed by Σ. Just as before, elements of such spaces are denoted by kets, e.g., $|\psi\rangle \in \mathbb{C}(\Sigma)$. It will be typical and often helpful to assign specific names to spaces of the form $\mathbb{C}(\Sigma)$, and scripted letters such as $\mathcal{A}$, $\mathcal{B}$, $\mathcal{X}$, $\mathcal{Y}$, etc., are generally used. For example, we might write $\mathcal{X} = \mathbb{C}(\{0,1\}^n)$ to indicate that the space $\mathcal{X}$ is indexed by the set $\{0,1\}^n$, and simply use the symbol $\mathcal{X}$ to refer to that space from that point on. Although it will seldom be necessary, we may also write $\mathbb{C}(\Sigma)^\dagger$ or $\mathcal{X}^\dagger$ (for example) to refer to the space of all row vectors (or bra vectors) $\langle\psi|$, for $|\psi\rangle \in \mathbb{C}(\Sigma)$ or $|\psi\rangle \in \mathcal{X}$. You will also see these things written with a star $*$ instead of a dagger $\dagger$ sometimes, as in $\mathcal{X}^*$.

Also similar to before, when we consider a particular quantum system we assume that it has some finite set Σ of associated *classical states*. We will typically use the term *register* from now on to refer to abstract physical devices (such as qubits or collections of qubits). Associated with any register having classical state set Σ is the vector space $\mathbb{C}(\Sigma)$. It is sometimes helpful, but certainly not necessary, to use the same letter (in different fonts) to refer to a register and its associated space. For example, register X may have associated space $\mathcal{X}$.

Density matrices

So far nothing is new except for a little bit of notation. In order to describe what really is new, it is helpful to start with a special case. Suppose that in the "old" representation, a register X having classical state set Σ is in a quantum state $|\psi\rangle \in \mathcal{X}$ for $\mathcal{X} = \mathbb{C}(\Sigma)$. The "new" way of representing this state will be:

$$|\psi\rangle\langle\psi|.$$

We have seen objects like this before (in the analysis of Grover's algorithm. It is effectively a matrix. For example, suppose $\Sigma = \{0, 1\}$ and $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$. Then

$$|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \text{and} \quad \langle\psi| = \begin{pmatrix} \overline{\alpha} & \overline{\beta} \end{pmatrix}$$

so

$$|\psi\rangle\langle\psi| = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \begin{pmatrix} \overline{\alpha} & \overline{\beta} \end{pmatrix} = \begin{pmatrix} \alpha\overline{\alpha} & \alpha\overline{\beta} \\ \overline{\alpha}\beta & \beta\overline{\beta} \end{pmatrix} = \begin{pmatrix} |\alpha|^2 & \alpha\overline{\beta} \\ \overline{\alpha}\beta & |\beta|^2 \end{pmatrix}.$$

This is called a *density matrix* or *density operator*. (The term *operator* usually just means a linear map from a space to itself.) Not all density operators have the special form $|\psi\rangle\langle\psi|$ for some unit vector $|\psi\rangle$, as we will see. In this special case, the density matrix describes what is called a *pure state*. All of the quantum states we have considered so far in the course have actually been of this special kind of state.

Suppose we have a probability distribution $(p_1, \ldots, p_k)$, for some positive integer k, as well as unit vectors $|\psi_1\rangle, \ldots, |\psi_k\rangle \in \mathcal{X}$, where $\mathcal{X} = \mathbb{C}(\Sigma)$ is the space corresponding to some register X. Somebody randomly chooses $j \in \{1, \ldots, k\}$ according to the probability distribution $(p_1, \ldots, p_k)$, prepares the register X in the state $|\psi_j\rangle$ for the chosen j, and then hands you X without telling you the value of j. The collection $\{(p_1, |\psi_1\rangle), \ldots, (p_k, |\psi_k\rangle)\}$ that describes the different possible states $|\psi_j\rangle$ along with their associated probabilities is called a *mixture*. How do you represent your knowledge of the state of X? In the "old" representation, there is no convenient way to do this beyond specifying the mixture. In the "new" representation, the density matrix corresponding to the above mixture is:

$$\sum_{j=1}^{k} p_k\,|\psi_j\rangle\langle\psi_j|.$$

In other words, it is meaningful to take a weighted average of the pure states $|\psi_j\rangle\langle\psi_j|$.

Example 1. Suppose Alice has a qubit A. She flips a fair coin: if the result is HEADS she prepares A in the state $|0\rangle$, and if the result is TAILS she prepares A in the state $|1\rangle$. She gives Bob the qubit without revealing the result of the coin-flip. Bob's knowledge of the qubit is described by the density matrix

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

Example 2. Suppose Alice has a qubit A as in the previous example. As before she flips a fair coin, but now if the result is HEADS she prepares A in the state $|+\rangle = \frac{1}{\sqrt{2}}\,|0\rangle + \frac{1}{\sqrt{2}}\,|1\rangle$, and if the result is TAILS she prepares A in the state $|-\rangle = \frac{1}{\sqrt{2}}\,|0\rangle - \frac{1}{\sqrt{2}}\,|1\rangle$. Bob's knowledge of the qubit is now described by the density matrix

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

Same as before!

The previous two examples demonstrate that different mixtures can yield precisely the same density matrix. This is not an accident, but rather is one of the strengths of the density matrix formalism. A given density matrix in essence represents a perfect description of the state of a quantum system—two different mixtures can be distinguished (in a statistical sense) if and only if they yield different density matrices.

Typically, lower case Greek letters such as ρ, ξ, σ, and τ are used to denote density matrices. The state of a system that corresponds to a given density matrix is called a ***mixed state***.

Some important facts about density matrices:

1. **The trace of a density matrix is 1.**

   (The trace of a matrix is the sum of its diagonal entries.)

   In fact, the diagonal entries describe precisely the probability distribution that would result if the system in question were measured (with respect to the restricted type of measurement we have so far considered).

   The fact that the trace is a linear function together with the useful formula $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$ (which holds for any choice of matrices A and B for which the product AB is square) makes the above fact easy to verify for any given mixture:

$$\mathrm{Tr}\left(\sum_{j=1}^{k} p_k \left|\psi_j\right\rangle \left\langle\psi_j\right|\right) = \sum_{j=1}^{k} p_k \,\mathrm{Tr}\left(\left|\psi_j\right\rangle \left\langle\psi_j\right|\right) = \sum_{j=1}^{k} p_k \,\mathrm{Tr}\left(\left\langle\psi_j\right| \left|\psi_j\right\rangle\right) = \sum_{j=1}^{k} p_k \left\langle\psi_j|\psi_j\right\rangle = 1.$$

2. **Every density matrix is *positive semidefinite*.**

   In general, a square matrix A is positive semidefinite if $\langle\psi|A|\psi\rangle$ is a nonnegative real number for every vector $|\psi\rangle$. An equivalent definition is that A is positive semidefinite if (i) $A = A^\dagger$ (i.e., A is Hermitian), and (ii) all eigenvalues of A are nonnegative real numbers.

   Again this condition is easy to check for mixtures: if

$$\rho = \sum_{j=1}^{k} p_k \left|\psi_j\right\rangle \left\langle\psi_j\right|$$

   and $|\phi\rangle$ is any vector, we have

$$\langle\phi|\rho|\phi\rangle = \sum_{j=1}^{k} p_k \left\langle\phi|\psi_j\right\rangle \left\langle\psi_j|\phi\right\rangle = \sum_{j=1}^{k} p_k \left|\left\langle\phi|\psi_j\right\rangle\right|^2 \geq 0.$$

You can interpret the above facts as the *defining* properties of density matrices: by definition, a density matrix is any matrix that is positive semidefinite and has trace equal to 1. Given such a matrix, it is possible to find a mixture that yields the given density matrix by letting $p_1, \ldots, p_k$ be the nonzero eigenvalues and $|\psi_1\rangle, \ldots, |\psi_k\rangle$ a collection of corresponding eigenvectors.

Operations on density matrices

In our "old" description of quantum information, a unitary operation U applied to a state $|\psi\rangle$ resulted in the state $U|\psi\rangle$. In the density matrix description, a unitary operation U applied to a pure state $|\psi\rangle\langle\psi|$ results in the density matrix

$$U|\psi\rangle\langle\psi|U^\dagger.$$

This is consistent with the observation that $(U|\psi\rangle)^\dagger = \langle\psi|U^\dagger$. More generally, applying U to the mixed state ρ results in state $U\rho U^\dagger$. You can easily check that the two required conditions of density matrices are necessarily met by this new matrix.

We can, however, consider a much more general set of possible operations than just unitary operations. Any operation Φ that can be written as

$$\Phi(\rho) = \sum_{j=1}^{k} A_j \rho A_j^\dagger$$

for some collection of matrices $A_1, \ldots, A_k$ satisfying

$$\sum_{j=1}^{k} A_j^\dagger A_j = I$$

represents an operation that can (in an idealized sense) be physically implemented. Such operations are called *admissible operations*. (There are several other names that are used as well, such as *completely positive trace preserving* operations and other variations on these words.)

If ρ is a matrix and Φ is admissible, then $\Phi(\rho)$ is also a density matrix. In fact a somewhat stronger property holds, which is that if Φ is applied to just part of a larger system whose state is described by some density matrix, then the resulting state will also be described by a density matrix.

Example 3. Suppose we have a single qubit X. Consider the operation that corresponds to measuring the qubit and forgetting the result. An admissible operation that describes this process is given by

$$\begin{aligned} A_0 &= |0\rangle\langle 0|, \\ A_1 &= |1\rangle\langle 1|, \\ \Phi(\rho) &= \sum_{j=0}^{1} A_j \rho A_j^\dagger = |0\rangle\langle 0|\rho|0\rangle\langle 0| + |1\rangle\langle 1|\rho|1\rangle\langle 1|. \end{aligned}$$

First let us check that this is a valid admissible operation:

$$\sum_{j=0}^{1} A_j^\dagger A_j = |0\rangle\langle 0|0\rangle\langle 0| + |1\rangle\langle 1|1\rangle\langle 1| = |0\rangle\langle 0| + |1\rangle\langle 1| = I.$$

It satisfies the required property, so indeed it is admissible.

What does it do to the state $\alpha|0\rangle + \beta|1\rangle$, for example? First we need to represent $|\psi\rangle$ as a density matrix: $\rho = |\psi\rangle\langle\psi|$. Now

$$\begin{aligned}\Phi(\rho) &= \Phi(|\psi\rangle\langle\psi|) \\ &= |0\rangle\langle 0|\psi\rangle\langle\psi|0\rangle\langle 0| + |1\rangle\langle 1|\psi\rangle\langle\psi|1\rangle\langle 1| \\ &= |\langle 0|\psi\rangle|^2 |0\rangle\langle 0| + |\langle 1|\psi\rangle|^2 |1\rangle\langle 1| \\ &= |\alpha|^2 |0\rangle\langle 0| + |\beta|^2 |1\rangle\langle 1|.\end{aligned}$$

In terms of matrices:

$$\begin{pmatrix} |\alpha|^2 & \alpha\overline{\beta} \\ \overline{\alpha}\beta & |\beta|^2 \end{pmatrix} \xrightarrow{\Phi} \begin{pmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{pmatrix}.$$

In general, off diagonal entries get zeroed out.

In general, admissible operations do not need to preserve the sizes of quantum systems. For example, one might consider the situation in which one of a collection of qubits is lost or somehow destroyed. More notation will help to discuss this issue in greater specificity.

Given two spaces $\mathcal{X}$ and $\mathcal{Y}$, we let

$$\mathrm{L}(\mathcal{X}, \mathcal{Y})$$

denote the set of all linear mappings from $\mathcal{X}$ to $\mathcal{Y}$. The shorthand $\mathrm{L}(\mathcal{X})$ is used to mean $\mathrm{L}(\mathcal{X}, \mathcal{X})$. Also let

$$\mathrm{D}(\mathcal{X})$$

denote the set of all density matrices on $\mathcal{X}$ (so that $\mathrm{D}(\mathcal{X}) \subset \mathrm{L}(\mathcal{X})$). For example, if X is a quantum register with classical state set Σ and $\mathcal{X} = \mathbb{C}(\Sigma)$, then any mixed state of the register X is represented by some element of $\mathrm{D}(\mathcal{X})$.

Now, if we have a collection of mappings (or matrices) $A_1, \ldots, A_k \in \mathrm{L}(\mathcal{X}, \mathcal{Y})$ that satisfy

$$\sum_{j=1}^{k} A_j^\dagger A_j = I$$

(the identity matrix in $\mathrm{L}(\mathcal{X})$) then the admissible operation Φ defined by

$$\Phi(\rho) = \sum_{j=1}^{k} A_j \rho A_j^\dagger$$

maps elements of $\mathrm{D}(\mathcal{X})$ to elements of $\mathrm{D}(\mathcal{Y})$.

Example 4. Let us suppose that we have two qubits: X and Y. We will consider the admissible operation that corresponds to discarding the second qubit. Thus, once the operation has been performed, we will be left with a single qubit X. The vector space corresponding to X will be $\mathcal{X}$ and the space corresponding to Y will be $\mathcal{Y}$.

Now, if Φ is to describe the operation of discarding the second qubit, then we must have that $\Phi(\rho) \in \mathrm{D}(\mathcal{X})$ whenever $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$. This means that if

$$\Phi(\rho) = \sum_{j=1}^{k} A_j \rho A_j^\dagger$$

for some choice of matrices $A_1, \ldots, A_k$, then these matrices must come from the set $\mathrm{L}(\mathcal{X} \otimes \mathcal{Y}, \mathcal{X})$. This means that they must be 2×4 matrices.

Let

$$A_0 = I \otimes \langle 0| \,, \quad A_1 = I \otimes \langle 1| \,,$$

where I denotes the identity operator on the first qubit, and define Φ by

$$\Phi(\rho) = \sum_{j=0}^{1} A_j \rho A_j^\dagger = A_0 \rho A_0^\dagger + A_1 \rho A_1^\dagger$$

for all ρ. Writing A_0 and A_1 in ordinary matrix notation gives

$$A_0 = I \otimes \langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

$$A_1 = I \otimes \langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

To see that Φ is admissible, we compute:

$$\begin{aligned} A_0^\dagger A + A_1^\dagger A_1 &= (I \otimes |0\rangle)(I \otimes \langle 0|) + (I \otimes |1\rangle)(I \otimes \langle 1|) \\ &= I \otimes |0\rangle\langle 0| + I \otimes |1\rangle\langle 1| \\ &= I \otimes (|0\rangle\langle 0| + |1\rangle\langle 1|) \\ &= I \otimes I \\ &= I_{\mathcal{X} \otimes \mathcal{Y}}. \end{aligned}$$

(The subscript on the identity operator in the last line is just representing what space the identity is acting on.)

Let us now consider the effect of this operation on a couple of states. First, suppose that (X, Y) is in the state $\xi \otimes \sigma$ for two 2×2 density matrices ξ and σ. Just as for vectors in the "old" description of quantum information, we view states of the form $\xi \otimes \sigma$ as representing completely uncorrelated qubits. What do you expect the state to be if you discard the second qubit? Now let us check:

$$\begin{aligned} \Phi(\xi \otimes \sigma) &= A_0(\xi \otimes \sigma)A_0^\dagger + A_1(\xi \otimes \sigma)A_1^\dagger \\ &= (I \otimes \langle 0|)(\xi \otimes \sigma)(I \otimes |0\rangle) + (I \otimes \langle 1|)(\xi \otimes \sigma)(I \otimes |1\rangle) \\ &= \xi \otimes \langle 0|\sigma|0\rangle + \xi \otimes \langle 1|\sigma|1\rangle \\ &= \left(\langle 0|\sigma|0\rangle + \langle 1|\sigma|1\rangle\right)\, \xi \\ &= \mathrm{Tr}(\sigma)\, \xi \\ &= \xi. \end{aligned}$$

Indeed this is what you presumably expected.

In the next lecture we'll see what happens when this operation is applied to entangled qubits.

The previous example can be generalized to describe the operation that corresponds to discarding part of a system. Suppose X and Y are registers with corresponding spaces $\mathcal{X}$ and $\mathcal{Y}$. A mixed state of these two registers is represented by some element of $\mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$. If, say, the register Y is discarded, we will be left with some mixed state of X that is represented by some element of $\mathrm{D}(\mathcal{X})$. Specifically, if $\rho \in \mathrm{D}(\mathcal{X} \otimes \mathcal{Y})$ is a density matrix representing the state of (X, Y) and Y is discarded, the resulting state of X is denoted $\mathrm{Tr}_{\mathcal{Y}}(\rho)$. We say that this state is the *reduced state* of X, and we call the admissible operation $\mathrm{Tr}_{\mathcal{Y}}$ the *partial trace*. (It is called the partial trace because when extended by linearity to arbitrary matrices it satisfies $\mathrm{Tr}_{\mathcal{Y}}(X \otimes Y) = \mathrm{Tr}(Y)X$ for all $X \in \mathrm{L}(\mathcal{X})$ and $Y \in \mathrm{L}(\mathcal{Y})$.) We also refer to the action corresponding to this operation as *tracing out* the space $\mathcal{Y}$. To express $\mathrm{Tr}_{\mathcal{Y}}$ in the form of an admissible operation, let Σ denote the set of classical states of Y. Then

$$\mathrm{Tr}_{\mathcal{Y}}(\rho) = \sum_{a \in \Sigma} (I \otimes \langle a|)\rho(I \otimes |a\rangle).$$

The partial trace is a particularly important admissible operation that we will continue discussing in the next lecture.