PHYS 7895: Quantum Information Theory	Fall 2015

Lecture 7 — Sept. 16, 2015

Prof. Mark M. Wilde *Scribe: Mark M. Wilde*

1 Overview

In the last lecture we discussed Bell's theorem and the CHSH game, and we showed how a quantum strategy can outperform a classical strategy at winning the CHSH game (in particular, we showed that a classical strategy can win the CHSH game with probability 3/4 and no larger, whereas a quantum strategy can win with probability $\cos^2(\pi/8) \approx 0.85$). We also proved Tsirelson's bound, which demonstrates that the quantum winning probability cannot exceed $\cos^2(\pi/8) \approx 0.85$.

In this lecture we will discuss the Schmidt decomposition, which is one of the most important theorems for understanding pure bipartite states. We will also venture into the noisy quantum theory, discussing density operators, evolution of density operators, and measurement in the noisy quantum theory. The material is coming from Sections 3.7.4, 4.1, and 4.2 of `http://markwilde.com/qit-notes.pdf` .

2 Schmidt decomposition

The Schmidt decomposition is one of the most important tools for analyzing bipartite pure states in quantum information theory. The Schmidt decomposition shows that it is possible to decompose any pure bipartite state as a superposition of corresponding states. We state this result formally as the following theorem:

Theorem 1 (Schmidt decomposition)**.** *Suppose that we have a bipartite pure state,*

$$|\psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B. \tag{1}$$

Then it is possible to express this state as follows:

$$|\psi\rangle_{AB} \equiv \sum_{i=0}^{d-1} \lambda_i \, |i\rangle_A \, |i\rangle_B \,, \tag{2}$$

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

$$d \leq \min \{\dim(\mathcal{H}_A), \dim(\mathcal{H}_B)\} \,. \tag{3}$$

Proof. This is essentially a restatement of the singular value decomposition of a matrix. Consider an arbitrary bipartite pure state $|\psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$. Let $d_A \equiv \dim(\mathcal{H}_A)$ and $d_B \equiv \dim(\mathcal{H}_B)$. We can express $|\psi\rangle_{AB}$ as follows:

$$|\psi\rangle_{AB} = \sum_{j=0}^{d_A-1} \sum_{k=0}^{d_B-1} \alpha_{j,k} |j\rangle_A |k\rangle_B , \tag{4}$$

for some amplitudes $\alpha_{j,k}$ and some orthonormal bases $\{|j\rangle_A\}$ and $\{|k\rangle_B\}$ on the respective systems A and B. Let us write the matrix formed by the coefficients $\alpha_{j,k}$ as some $d_A \times d_B$ matrix A where

$$[A]_{j,k} = \alpha_{j,k}. \tag{5}$$

Since every matrix has a singular value decomposition, we can write A as

$$A = U\Lambda V, \tag{6}$$

where U is a $d_A \times d_A$ unitary matrix, V is a $d_B \times d_B$ unitary matrix, and Λ is a $d_A \times d_B$ matrix with d real, strictly positive numbers λ_i along the diagonal and zeros elsewhere. Let us write the matrix elements of U as $u_{j,i}$ and those of V as $v_{i,k}$. The above matrix equation is then equivalent to the following set of equations:

$$\alpha_{j,k} = \sum_{i=0}^{d-1} u_{j,i} \lambda_i v_{i,k}. \tag{7}$$

Let us make this substitution into the expression for the state in (4):

$$|\psi\rangle_{AB} = \sum_{j=0}^{d_A-1} \sum_{k=0}^{d_B-1} \left(\sum_{i=0}^{d-1} u_{j,i} \lambda_i v_{i,k} \right) |j\rangle_A |k\rangle_B . \tag{8}$$

Readjusting some terms by exploiting the properties of the tensor product, we find that

$$|\psi\rangle_{AB} = \sum_{i=0}^{d-1} \lambda_i \left(\sum_{j=0}^{d_A-1} u_{j,i} |j\rangle_A \right) \otimes \left(\sum_{k=0}^{d_B-1} v_{i,k} |k\rangle_B \right) \tag{9}$$

$$= \sum_{i=0}^{d-1} \lambda_i |i\rangle_A |i\rangle_B , \tag{10}$$

where we define the orthonormal basis on the A system as $|i\rangle_A \equiv \sum_j u_{j,i} |j\rangle_A$ and we define the orthonormal basis on the B system as $|i\rangle_B \equiv \sum_k v_{i,k} |k\rangle_B$. This final step completes the proof of the theorem. You can verify that the set of states $\{|i\rangle_A\}$ form an orthonormal basis (the proof for the set of states $\{|i\rangle_B\}$ is similar). □

Remark 2. *The Schmidt decomposition applies not only to bipartite systems but to any number of systems where we can make a bipartite cut of the systems. For example, suppose that there is a state $|\phi\rangle_{ABCDE}$ on systems $ABCDE$. We could say that AB are part of one system and CDE are part of another system and write a Schmidt decomposition for this state as follows:*

$$|\phi\rangle_{ABCDE} = \sum_y \sqrt{p_Y(y)} |y\rangle_{AB} |y\rangle_{CDE} , \tag{11}$$

where $\{|y\rangle_{AB}\}$ is an orthonormal basis for the joint system AB and $\{|y\rangle_{CDE}\}$ is an orthonormal basis for the joint system CDE.

3 Noisy Quantum States

We generally may not have perfect knowledge of a prepared quantum state. Suppose a third party, Bob, prepares a state for us and only gives us a probabilistic description of it. We may only know that Bob selects the state $|\psi_x\rangle$ with a certain probability $p_X(x)$. Our description of the state is then as an ensemble $\mathcal{E}$ of quantum states where

$$\mathcal{E} \equiv \{p_X(x), |\psi_x\rangle\}_{x\in\mathcal{X}} . \tag{12}$$

In the above, X is a random variable with distribution $p_X(x)$. Each realization x of random variable X belongs to an alphabet $\mathcal{X}$. For our purposes, it is sufficient for us to say that $\mathcal{X} \equiv \{1, \ldots, |\mathcal{X}|\}$. Thus, the realization x merely acts as an index, meaning that the quantum state is $|\psi_x\rangle$ with probability $p_X(x)$. We also assume that each state $|\psi_x\rangle$ is a d-dimensional qudit state.

A simple example is the following ensemble:

$$\left\{\left\{\frac{1}{3}, |1\rangle\right\}, \left\{\frac{2}{3}, |3\rangle\right\}\right\}. \tag{13}$$

The states $|1\rangle$ and $|3\rangle$ are in a four-dimensional Hilbert space with basis states

$$\{|0\rangle, |1\rangle, |2\rangle, |3\rangle\}. \tag{14}$$

The interpretation of this ensemble is that the state is $|1\rangle$ with probability $1/3$ and the state is $|3\rangle$ with probability $2/3$.

3.1 The Density Operator

Suppose now that we have the ability to perform a perfect measurement of a system with ensemble description $\mathcal{E}$ in (12). Let Π_j be the elements of this projective measurement so that $\sum_j \Pi_j = I$, and let J be the random variable that denotes the index j of the measurement outcome. Let us suppose at first, without loss of generality, that the state in the ensemble is $|\psi_x\rangle$ for some $x \in \mathcal{X}$. Then the Born rule of the noiseless quantum theory states that the conditional probability $p_{J|X}(j|x)$ of obtaining measurement result j (given that the state is $|\psi_x\rangle$) is

$$p_{J|X}(j|x) = \langle\psi_x| \Pi_j |\psi_x\rangle, \tag{15}$$

and the post-measurement state is

$$\frac{\Pi_j |\psi_x\rangle}{\sqrt{p_{J|X}(j|x)}}. \tag{16}$$

But, we would also like to know the actual probability $p_J(j)$ of obtaining measurement result j for the ensemble description $\mathcal{E}$. By the *law of total probability*, the unconditional probability $p_J(j)$ is

$$p_J(j) = \sum_{x\in\mathcal{X}} p_{J|X}(j|x) p_X(x) \tag{17}$$

$$= \sum_{x\in\mathcal{X}} \langle\psi_x| \Pi_j |\psi_x\rangle p_X(x). \tag{18}$$

Definition 3 (Trace)**.** *The trace* $\text{Tr}\{A\}$ *of an operator* A *is*

$$\text{Tr}\{A\} \equiv \sum_i \langle i| A |i\rangle , \tag{19}$$

where $\{|i\rangle\}$ *is some complete, orthonormal basis.*

(Observe that the trace operation is *linear*. It is also independent of which orthonormal basis we choose.) We can then show the following useful property with the above definition:

$$\begin{aligned} \text{Tr}\{\Pi_j |\psi_x\rangle \langle\psi_x|\} &= \sum_i \langle i| \Pi_j |\psi_x\rangle \langle\psi_x|i\rangle \quad (20)\\ &= \sum_i \langle\psi_x|i\rangle \langle i| \Pi_j |\psi_x\rangle \quad (21)\\ &= \langle\psi_x| \left(\sum_i |i\rangle \langle i|\right) \Pi_j |\psi_x\rangle \quad (22)\\ &= \langle\psi_x| \Pi_j |\psi_x\rangle . \quad (23) \end{aligned}$$

The last equality uses the completeness relation $\sum_i |i\rangle \langle i| = I$. Thus, we continue with the development in (18) and show that

$$\begin{aligned} p_J(j) &= \sum_{x\in\mathcal{X}} \text{Tr}\{\Pi_j |\psi_x\rangle \langle\psi_x|\} \, p_X(x) \quad (24)\\ &= \text{Tr}\left\{\Pi_j \sum_{x\in\mathcal{X}} p_X(x) |\psi_x\rangle \langle\psi_x|\right\}. \quad (25) \end{aligned}$$

We can rewrite the last equation as follows:

$$p_J(j) = \text{Tr}\{\Pi_j \rho\} , \tag{26}$$

where we define the *density operator* ρ as

$$\rho \equiv \sum_{x\in\mathcal{X}} p_X(x) |\psi_x\rangle \langle\psi_x| . \tag{27}$$

The above operator is known as the density operator because it is the quantum generalization of a probability density function.

We sometimes refer to the density operator as the *expected density operator* because there is a sense in which we are taking the expectation over all of the states in the ensemble in order to obtain the density operator. We can equivalently write the density operator as follows:

$$\rho = \mathbb{E}_X \{|\psi_X\rangle \langle\psi_X|\} , \tag{28}$$

where the expectation is with respect to the random variable X. Note that we are careful to use the notation $|\psi_X\rangle$ instead of the notation $|\psi_x\rangle$ for the state inside of the expectation because the state $|\psi_X\rangle$ is a random quantum state, random with respect to a classical random variable X.

3.1.1 Properties of the Density Operator

What are the properties that a given density operator must satisfy? Let us consider taking the trace of ρ:

$$\begin{aligned}
\operatorname{Tr}\{\rho\} &= \operatorname{Tr}\left\{\sum_{x\in\mathcal{X}} p_X(x)\left|\psi_x\right\rangle\left\langle\psi_x\right|\right\} && (29)\\
&= \sum_{x\in\mathcal{X}} p_X(x) \operatorname{Tr}\{\left|\psi_x\right\rangle\left\langle\psi_x\right|\} && (30)\\
&= \sum_{x\in\mathcal{X}} p_X(x) \left\langle\psi_x|\psi_x\right\rangle && (31)\\
&= \sum_{x\in\mathcal{X}} p_X(x) && (32)\\
&= 1. && (33)
\end{aligned}$$

The above development shows that every density operator should have *unit trace* because it arises from an ensemble of quantum states. Every density operator is also *positive semi-definite*, meaning that

$$\forall \left|\varphi\right\rangle : \qquad \left\langle\varphi\right|\rho\left|\varphi\right\rangle \geq 0. \tag{34}$$

We write $\rho \geq 0$ to indicate that an operator is positive semi-definite. The proof of non-negativity of any density operator ρ is as follows:

$$\begin{aligned}
\left\langle\varphi\right|\rho\left|\varphi\right\rangle &= \left\langle\varphi\right|\left(\sum_{x\in\mathcal{X}} p_X(x)\left|\psi_x\right\rangle\left\langle\psi_x\right|\right)\left|\varphi\right\rangle && (35)\\
&= \sum_{x\in\mathcal{X}} p_X(x) \left\langle\varphi|\psi_x\right\rangle\left\langle\psi_x|\varphi\right\rangle && (36)\\
&= \sum_{x\in\mathcal{X}} p_X(x) \left|\left\langle\varphi|\psi_x\right\rangle\right|^2 \geq 0. && (37)
\end{aligned}$$

The inequality follows because each $p_X(x)$ is a probability and is therefore non-negative.

Let us consider taking the conjugate transpose of the density operator ρ:

$$\begin{aligned}
\rho^\dagger &= \left(\sum_{x\in\mathcal{X}} p_X(x)\left|\psi_x\right\rangle\left\langle\psi_x\right|\right)^\dagger && (38)\\
&= \sum_{x\in\mathcal{X}} p_X(x) \left(\left|\psi_x\right\rangle\left\langle\psi_x\right|\right)^\dagger && (39)\\
&= \sum_{x\in\mathcal{X}} p_X(x) \left|\psi_x\right\rangle\left\langle\psi_x\right| && (40)\\
&= \rho. && (41)
\end{aligned}$$

Every density operator is thus a *Hermitian* operator as well because the conjugate transpose of ρ is ρ.

3.1.2 Ensembles and the Density Operator

Every ensemble has a unique density operator, but the opposite does not necessarily hold: every density operator does not correspond to a unique ensemble and could correspond to many ensembles.

This last result has profound implications for the predictions of the quantum theory because it is possible for two or more completely different ensembles to have the same probabilities for measurement results. It also has important implications for quantum Shannon theory as well.

By the spectral theorem, it follows that every density operator ρ has a spectral decomposition in terms of its eigenstates $\{|\phi_x\rangle\}_{x\in\{0,\ldots,d-1\}}$ because every ρ is Hermitian:

$$\rho = \sum_{x=0}^{d-1} \lambda_x \left|\phi_x\right\rangle \left\langle\phi_x\right| , \tag{42}$$

where the coefficients λ_x are the eigenvalues.

Thus, given any density operator ρ, we can define a "canonical" ensemble $\{\lambda_x, |\phi_x\rangle\}$ corresponding to it. This observation is so important for quantum Shannon theory that we see this idea arise again and again throughout this book.

3.1.3 Density Operator as the State

We can also refer to the density operator as the *state* of a given quantum system because it is possible to use it to calculate all of the predictions of the quantum theory. We can make these calculations without having an ensemble description—all we need is the density operator. The noisy quantum theory also subsumes the noiseless quantum theory because any state $|\psi\rangle$ has a corresponding density operator $|\psi\rangle\langle\psi|$ in the noisy quantum theory, and all calculations with this density operator in the noisy quantum theory give the same results as using the state $|\psi\rangle$ in the noiseless quantum theory. For these reasons, we will say that the *state* of a given quantum system is a density operator.

One of the most important states in the noisy quantum theory is the maximally mixed state π. The maximally mixed state π arises as the density operator of a uniform ensemble of orthogonal states $\left\{\frac{1}{d}, |x\rangle\right\}$, where d is the dimensionality of the Hilbert space. The maximally mixed state π is then equal to

$$\pi = \frac{1}{d} \sum_{x\in\mathcal{X}} |x\rangle \langle x| = \frac{I}{d}. \tag{43}$$

3.2 Noiseless Evolution of an Ensemble

Quantum states can evolve in a noiseless fashion either according to a unitary operator or a measurement. In this section, we determine the noiseless evolution of an ensemble and its corresponding density operator.

3.2.1 Noiseless Unitary Evolution of a Noisy State

We first consider noiseless evolution according to some unitary U. Suppose we have the ensemble $\mathcal{E}$ in (12) with density operator ρ. Suppose without loss of generality that the state is $|\psi_x\rangle$. Then the evolution postulate of the noiseless quantum theory gives that the state after the unitary evolution is as follows:

$$U|\psi_x\rangle . \tag{44}$$

This result implies that the evolution leads to a new ensemble

$$\mathcal{E}_U \equiv \{p_X(x), U|\psi_x\rangle\}_{x\in\mathcal{X}} . \tag{45}$$

The density operator of the evolved ensemble is

$$\sum_{x\in\mathcal{X}} p_X(x) U|\psi_x\rangle\langle\psi_x| U^\dagger = U\left(\sum_{x\in\mathcal{X}} p_X(x)|\psi_x\rangle\langle\psi_x|\right)U^\dagger \tag{46}$$

$$= U\rho U^\dagger. \tag{47}$$

Thus, the above relation shows that we can keep track of the evolution of the density operator ρ, rather than worrying about keeping track of the evolution of every state in the ensemble $\mathcal{E}$. It suffices to keep track of only the density operator evolution because this operator is sufficient to determine the predictions of the quantum theory.

3.2.2 Noiseless Measurement of a Noisy State

In a similar fashion, we can analyze the result of a measurement on a system with ensemble description $\mathcal{E}$ in (12). Suppose that we perform a projective measurement with projection operators $\{\Pi_j\}_j$ where $\sum_j \Pi_j = I$. Suppose further without loss of generality that the state in the ensemble is $|\psi_x\rangle$. Then the noiseless quantum theory predicts that the probability of obtaining outcome j conditioned on the index x is

$$p_{J|X}(j|x) = \langle\psi_x|\Pi_j|\psi_x\rangle , \tag{48}$$

and the resulting state is

$$\frac{\Pi_j|\psi_x\rangle}{\sqrt{p_{J|X}(j|x)}}. \tag{49}$$

Supposing that we receive outcome j, then we have a new ensemble:

$$\mathcal{E}_j \equiv \left\{p_{X|J}(x|j), \Pi_j|\psi_x\rangle/\sqrt{p_{J|X}(j|x)}\right\}_{x\in\mathcal{X}} . \tag{50}$$

The density operator for this ensemble is

$$\begin{aligned}
&\sum_{x\in\mathcal{X}} p_{X|J}\left(x|j\right) \frac{\Pi_j \left|\psi_x\right\rangle \left\langle\psi_x\right| \Pi_j}{p_{J|X}\left(j|x\right)} \\
&= \Pi_j \left(\sum_{x\in\mathcal{X}} \frac{p_{X|J}\left(x|j\right)}{p_{J|X}(j|x)} \left|\psi_x\right\rangle \left\langle\psi_x\right|\right) \Pi_j && (51)\\
&= \Pi_j \left(\sum_{x\in\mathcal{X}} \frac{p_{J|X}(j|x)p_X(x)}{p_{J|X}(j|x)p_J(j)} \left|\psi_x\right\rangle \left\langle\psi_x\right|\right) \Pi_j && (52)\\
&= \frac{\Pi_j \left(\sum_{x\in\mathcal{X}} p_X(x) \left|\psi_x\right\rangle \left\langle\psi_x\right|\right) \Pi_j}{p_J(j)} && (53)\\
&= \frac{\Pi_j \rho \Pi_j}{p_J(j)}. && (54)
\end{aligned}$$

The second equality follows from applying the Bayes rule:

$$p_{X|J}\left(x|j\right) = p_{J|X}(j|x)p_X\left(x\right)/p_J(j). \tag{55}$$

The above expression gives the evolution of the density operator under a measurement. We can again employ the law of total probability to compute that $p_J(j)$ is

$$\begin{aligned}
p_J(j) &= \sum_{x\in\mathcal{X}} p_{J|X}(j|x)p_X(x) && (56)\\
&= \sum_{x\in\mathcal{X}} p_X(x) \left\langle\psi_x\right| \Pi_j \left|\psi_x\right\rangle && (57)\\
&= \sum_{x\in\mathcal{X}} p_X(x) \operatorname{Tr}\left\{\left|\psi_x\right\rangle \left\langle\psi_x\right| \Pi_j\right\} && (58)\\
&= \operatorname{Tr}\left\{\sum_{x\in\mathcal{X}} p_X(x) \left|\psi_x\right\rangle \left\langle\psi_x\right| \Pi_j\right\} && (59)\\
&= \operatorname{Tr}\left\{\rho \Pi_j\right\}. && (60)
\end{aligned}$$

We can think of $\operatorname{Tr}\{\rho\Pi_j\}$ intuitively as the area of the shadow of ρ onto the space that the projector Π_j projects.