Frontiers of Quantum Complexity and Cryptography Spring 2022

Lecture 2 - Mixed States and Density Matrices

Lecturer: Henry Yuen *Scribes: Melody Hsu*

1 Overview

In the third lecture, we wrapped up a review of the foundational concepts we will need to tackle later topics, and then started our first quantum complexity topic, state tomography. We briefly discussed several ways to analyze probabilistic mixtures of quantum states; after the conclusion of the review, we explored lower bounds on complexity of a classic tomography protocol.

2 Probabilistic Mixtures of Pure Quantum States

Previously, we had only encountered pure states $|\psi\rangle$, which are described by unit vectors in $\mathbb{C}^d$. Consider the following pure state, called $|+\rangle$:

$$H|0\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = |+\rangle$$

Measuring $|+\rangle$ in the standard basis yields a random outcome (i.e. $|0\rangle$ and $|1\rangle$ with equal probability), but measuring $|+\rangle$ in the diagonal basis (i.e. the basis $\{|+\rangle, |-\rangle\}$) yields a deterministic outcome (we will get $|+\rangle$ with probability 1).

Not all quantum systems can be adequately described by deterministic pure states; what if we want to describe a probabilistic state? In other words, a quantum state where, no matter what measurement basis you take, the outcome is always random? We will need a new formalism to describe such quantum states.

2.1 Mixed states

Mixed states states are probabilistic mixtures of pure quantum states. They are useful in a number of applications; for example, if we want to consider a quantum system affected by noise, we can use mixed states to describe the state of the system while incorporating the inherent uncertainty due to randomness. Furthermore, we will need mixed states are to describe subsystems of quantum states, in particular entangled states. Probabilistic mixtures will give us a formalized way of describing individual qubits which, using only the pure state formalism, cannot be described by themselves.

2.2 Probabilistic Mixtures

Consider the following scenario. A qubit's state is assigned to be either $|0\rangle$ or $|+\rangle$ depending on a coin toss, the outcome of which is unknown to us. Without observing the qubit, how can we

describe our prior belief of this system? To do this, we can describe a probabilistic distribution over pure states:

$$(\frac{1}{2}, |0\rangle), (\frac{1}{2}, |+\rangle)$$

This description fully captures our beliefs of a *complete system* prior to observing an outcome. But for larger, more complex systems that might take values over a very large number of pure states, such a description can become unwieldy very quickly. Thus, we will need a tool that will enable us to completely and succinctly capture all relevant information about a mixture of states.

3 Density Matrices

3.1 Background

Before we begin discussion of density matrices, we need to recall some facts from linear algebra.

Spectral theorem. Let $A \in \mathbb{C}^{d \times d}$ be a $d \times d$ Hermitian matrix. Then, there exists some orthonormal basis $\{|b_1\rangle, \ldots, |b_d\rangle\}$ of $\mathbb{C}^d$ and real eigenvalues $\lambda_1, \ldots, \lambda_d$ such that $A = \sum_j \lambda_j |b_j\rangle\langle b_j|$. In other words, the spectral theorem describes a method of diagonalization on a $d \times d$ Hermitian matrix.

Positive semi-definite. A Hermitian matrix A is positive semi-definite (we will use the shorthand PSD for brevity) if and only if all of its eigenvalues are nonnegative.

Armed with the above two concepts, we can now formally describe a density matrix.

3.2 Overview

A **density matrix** is a d-dimensional matrix $\rho \in \mathbb{C}^{d \times d}$ such that ρ is PSD, and has $\mathrm{Tr}(\rho) = 1$.

Given a pure state $|\psi\rangle \in \mathbb{C}^d$, the density matrix corresponding to it is the outer product $|\psi\rangle\langle\psi|$, which is also the projection onto the vector $|\psi\rangle$. This is a valid density matrix because its trace is one ($\mathrm{Tr}(|\psi\rangle\langle\psi|) = \langle\psi|\psi\rangle = 1$) and it is PSD (rank-1 outer products of a vector are always PSD).

More generally, a mixture over quantum states $\{(p_1, |\psi_1\rangle), \ldots, (p_k, |\psi_k\rangle)\}$ corresponds to the linear combination $\sum_i p_i |\psi_i\rangle\langle\psi_i|$, where p_i corresponds to the probability of $|\psi_i$, and $\sum_i p_i = 1$. This is a $d \times d$ matrix and it is a valid density matrix because it has trace one:

$$\mathrm{Tr}\left(\sum_i p_i |\psi_i\rangle\langle\psi_i|\right) = \sum_i p_i \mathrm{Tr}(|\psi_i\rangle\langle\psi_i|) = \sum_i p_i = 1$$

and it is PSD because it is a nonnegative linear combination of PSD matrices.

Given some density matrix, we can informally interpret it as follows: Along the diagonal are our beliefs of the system being in that corresponding pure state. (Thus, in a 2×2 density matrix, the upper left element would correspond to the probability of being in the $|0\rangle$ state, and the lower right element would correspond to the probability of being in the $|1\rangle$ state.) The off-diagonal elements are a measure of how "quantum" a state is - a measure of the superposition, or phase, of the state.

Example 1. What are the density matrices for $|0\rangle$ and $|1\rangle$?

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

Both of the above matrices are PSD and have trace one.

Example 2. What is the density matrix for $\left(\frac{1}{2}, |0\rangle\right), \left(\frac{1}{2}, |+\rangle\right)$?

$$\frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|+\rangle\langle +| = \begin{bmatrix}\frac{1}{2} & 0\\0 & 0\end{bmatrix} + \begin{bmatrix}\frac{1}{4} & \frac{1}{4}\\\frac{1}{4} & \frac{1}{4}\end{bmatrix} = \begin{bmatrix}\frac{3}{4} & \frac{1}{4}\\\frac{1}{4} & \frac{1}{4}\end{bmatrix}$$

Again, this matrix also is PSD and has $\mathrm{Tr} = 1$.

Example 3. In this example we will see that a density matrix does *not* necessarily identify a unique mixture.

$$\left(\frac{1}{2}, |0\rangle\right), \left(\frac{1}{2}|1\rangle\right) \to \frac{1}{2}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} + \frac{1}{2}\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} = \begin{bmatrix}\frac{1}{2} & 0\\0 & \frac{1}{2}\end{bmatrix} = \frac{I}{2}$$
$$\left(\frac{1}{2}, |+\rangle\right), \left(\frac{1}{2}|-\rangle\right) \to \frac{1}{4}\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix} + \frac{1}{4}\begin{bmatrix}1 & -1\\-1 & 1\end{bmatrix} = \begin{bmatrix}\frac{1}{2} & 0\\0 & \frac{1}{2}\end{bmatrix} = \frac{I}{2}$$

Because these two mixtures both have density matrices that are the maximally mixed state $\frac{I}{2}$, there is no way to distinguish between them.

3.3 Mixtures of Density Matrices

We can also define mixtures of density matrices, $(p, \rho), (q, \sigma)$, for two probabilities p and q and two density matrices ρ and σ. The mixture of density matrices would result in the density matrix $\tau = p\rho + q\sigma$.

We can check that τ is indeed a density matrix by confirming that its trace is equal to 1. We have $\mathrm{Tr}(\tau) = p\mathrm{Tr}(\rho) + q\mathrm{Tr}(\sigma) = p + q = 1$. Thus, τ is a density matrix.

3.4 Projective Measurements

We define a projective measurement to be a measurement over some set of Hermitian projection matrices of a state which assesses the probability of obtaining some outcome. Formally, the **projective measurement** $M = \{M_1, M_2, \ldots, M_k\}$ is a $k-$outcome projective measurement if each M_i is a Hermitian projection matrix, i.e., $M_i^\dagger = M_i$ and $M_i^2 = M_i$ and $M_1 + M_2 + \ldots + M_k = I$ (all M_i are orthogonal to each other).

Given a pure state $|\psi\rangle$, using M, we would obtain outcome i with probability $||M_i|\psi\rangle||^2$. After measurement, that state would collapse to $\frac{M_i|\psi\rangle}{||M_i|\psi\rangle||^2}$.

3.5 Unitary Evolution on Density Matrices

Consider a unitary evolution on a density matrix $\rho \to U\rho U^\dagger$ where $U\rho U^\dagger$ is the new state of the system after evolution. How do we compute this evolution?

In the pure state case, we start with our state vector and apply unitary evolution $|\psi\rangle \to U|\psi\rangle$. The corresponding mixed state representation would be $|\psi\rangle\langle\psi| \to U|\psi\rangle\langle\psi|U^\dagger$. This is the rank-1 outer product of the vector $U|\psi\rangle$.

Now, consider calculating the probability of calculating a certain outcome i from measuring ρ with the projective measurement M. Given $M = \{M_1, \ldots, M_k\}$, we should obtain outcome i with probability $\mathrm{Tr}(M_i\rho)$. Taking $\rho = |\psi\rangle\langle\psi|$ as an example,

$$\begin{aligned}\mathrm{Tr}(M_i\rho) &= \mathrm{Tr}(M_i|\psi\rangle\langle\psi|) \\ &= \langle\psi|M_i|\psi\rangle = \langle\psi|M_i^\dagger M_i|\psi\rangle = ||M_i|\psi\rangle||^2\end{aligned}$$

which is what we were supposed to get if we were just working with pure states.

3.6 Observables

An **observable** in quantum mechanics is a way of specifying a measurement as well as a *weighting* of the outcomes of the measurement. Mathematically, observables are simply Hermitian matrices.

Given some Hermitian matrix $A \in \mathbb{C}^{d\times d}$, we can diagonalize A such that $A = \sum_j \lambda_j |b_j\rangle\langle b_j|$ (by the spectral theorem). We can take the orthonormal basis and the measurement corresponding to the basis $\{|b_1\rangle, |b_2\rangle, \ldots, |b_d\rangle\}$; then, we can take the eigenvalues λ_j as weights which we assign to outcome $|b_j\rangle$.

As an example, consider the measurement of a quantum state ρ.

$$\begin{aligned}\mathrm{Tr}(A\rho) &= \mathrm{Tr}(\sum_j \lambda_i |b_j\rangle\langle b_j|\rho) \\ &= \sum_j \lambda_j \mathrm{Tr}(|b_j\rangle\langle b_j|\rho) \\ &= \sum_j \lambda_j \Pr[\text{obtain outcome } |b_j\rangle \text{ when measuring } \rho] \ .\end{aligned}$$

3.7 Density Matrices of Multiple Systems

If we have two quantum states, their joint density matrix is given by their tensor product; i.e., given ρ, σ are density matrices, their joint density matrix is $\rho \otimes \sigma$. We can also abbreviate n copies of ρ as $\rho^{\otimes n}$.

Note that not all composite quantum systems can be written as a tensor product. However, this does **not** necessarily mean that such systems are entangled!

Consider the state $\rho = \frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11|$. These two qubits are *correlated*, but *not entangled*.

4 Traces and Partial Traces

Traces can be used to analyze a subsystem of a quantum system by discarding the irrelevant subsystems. In other words, if we had a density matrix ρ_{AB} on systems AB, but we only wanted to analyze system A, we would apply a linear map called a **partial trace** to isolate A to obtain the **reduced density matrix on system** A. The resulting state $\rho_A = \mathrm{Tr}_B(\rho_{AB})$ would be a matrix with dimensionality corresponding to A only. For example, if we have a two qubit state ρ_{AB} where the first qubit is labelled A and the second qubit is labelled B, then tracing out the B system yields a density matrix on qubit A.

Formally, we define the partial trace map $\mathrm{Tr}_B(\cdot)$ (which one can think of as "ignoring the B system") as follows. Let $|a_1\rangle, |a_2\rangle$ be orthogonal basis states for system A, and let $|b_1\rangle, |b_2\rangle$ be orthogonal basis states for system B. Consider the states $|a_1, b_1\rangle$ and $|a_2, b_2\rangle$ for the composite system AB. Then we define

$$\mathrm{Tr}_B(|a_1, b_1\rangle\langle a_2, b_2|) = \mathrm{Tr}_B(|a_1\rangle \otimes |b_1\rangle\langle a_2| \otimes \langle b_2|) = |a_1\rangle\langle a_2| \cdot \langle b_2|b_1\rangle \ .$$

This gives us a recipe to calculate the partial trace of *any* density matrix ρ_{AB}; this is because we can express ρ_{AB} as a linear combination of outer products $|a_1, b_1\rangle\langle a_2, b_2|$, and computing $\mathrm{Tr}_B(\rho_{AB})$ can be done by computing $\mathrm{Tr}_B(\cdot)$ on the linear combination.

Every mixed state ρ_A is also the result of taking a partial trace of *some* pure state $|\psi\rangle_{AB}$ on a *larger* system AB, called a *purification* of ρ_A. This implies that for every mixed state, there exists a pure state from which it is derived – in fact, there are infinitely many.

4.1 Distinguishability of Density Matrices

There are several ways we can measure the degree of similarity between two density matrices. We describe one such metric below.

Trace distance. We can calculate the trace distance between two density matrices ρ, σ as follows:

$$D(\rho, \sigma) = \frac{1}{2}||\rho - \sigma||_1 = \frac{1}{2}\mathrm{Tr}|\rho - \sigma|)$$

The trace distance gives the maximum probability that an observer can distinguish any two density matrices using any operation; i.e., if $D(\rho, \sigma) = \frac{1}{2}$, then the distance is at least $\frac{1}{2}$. Thus, if there does not exist any method of distinguishing these matrices, the distance D would necessarily be 0, and the matrices must be the same.

Note: $|\rho - \sigma|$ is the matrix absolute value, which is not the same as taking the elementwise absolute value, but rather the sum of the absolute values of the eigenvalues of $\rho - \sigma$.

D as a distance metric has a number of nice properties. Namely, it is nonzero; it is symmetric; it satisfies the triangle inequality; it is convex; it does not increase when tracing out systems; and it is unitarily equivalent ($D(U\rho U^\dagger, U\sigma U^\dagger) = D(\rho, \sigma)$).