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

Lecture 19: Quantum Channels and Finishing Holevo's Bound

November 11, 2015

Lecturer: Ryan O'Donnell *Scribe: Connell Donaghy*

1 Quantum Channels

As we move forward in this class, we will be transitioning more from quantum algorithms into quantum information theory. Quantum information theory is more relevant today, as there is more active physics research going on around quantum information theory than quantum computers. To begin talking about quantum information theory, we will first define and show examples of **quantum channels**. Quantum channels are of particular interest as both storing and transmitting quantum information is very non-trivial, due to issues such as noise and degradation [Gri12].

Definition 1.1. A quantum channel is any operation which takes in some density matrix ρ, and outputs a density matrix ρ'

Here's a simple sketch of a quantum channel:

$$\rho \quad \text{——}\boxed{\Phi}\text{——} \quad \rho'$$

Let's discuss a few examples of quantum channels, and we'll discorver we've been working with many of them before!

First, consider a simple unitary gate being applied to some quantum state ρ. This will output a new quantumstate ρ' We see

$$\rho' \to \mathcal{U}\rho\mathcal{U}^\dagger$$

Thus, applying a valid quantum gate is actually a valid quantum channel!

Now, we consider a more randomized quantum channel. Consider taking some input state ρ, and with probability $\frac{1}{2}$ applying a unitary gate $\mathcal{U}$, otherwise just outputting our input state ρ. In this case, we have

$$\rho' \to \frac{1}{2}\left(\mathcal{U}\rho\mathcal{U}^\dagger + \rho\right)$$

Next we consider another randomized example called a **mixed unitary channel**. Thus means we have some series of r unitary gates $\mathcal{U}_i$, and each of them get applied with probability p_i. Thus, our output state for a mixed unitary channel is

$$\rho' \to \sum_{i=1}^{r} p_i \mathcal{U}\rho\mathcal{U}^\dagger$$

Another interesting yet trivial quantum channel is to just ignore our input ρ, and output some other ρ_0 regardless of what our input is. In this case, we have

$$\rho' \to \rho_0$$

Next, we consider an interesting case in which we measure in our standard basis, but actually "forget" our outcome. Consider our simple qubit $|\psi\rangle$ that we first described in lecture 1, which has the form $\alpha|0\rangle + \beta|1\rangle$. This qubit, when considered as a density matrix takes the form

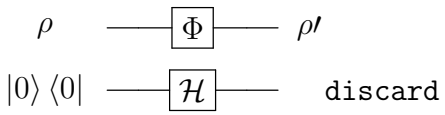
$$|\psi\rangle\langle\psi| = \rho = \begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{pmatrix}$$

We can see that this state will output $|0\rangle$ with output probability $|\alpha|^2$, and output $|1\rangle$ with probability $|\beta|^2$. Thus, if we consider dephasing this matrix, by measuring and then forgetting the outcome, our channel output will be

$$\rho' \to \begin{pmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{pmatrix}$$

This channel is commonly referred to as a **complete dephasing channel**.

Finally, as our most important example, we consider adding some register "B", which we initialize to $|0\rangle\langle 0|$. Next, we take our input state ρ along with this register, and apply some giant unitary operation $\mathcal{U}$ on the whole system $\rho \otimes |0\rangle\langle 0|$. Next, we once again measure but forget B, and let ρ' be our output. We can see this as a quantum circuit, by looking at the following:



According to physics, this type of quantum channel is actually the most general quantum channel we can describe. Examples one through five can all be described as this type of quantum channel, as well as any other valid quantum channel which can be dreamt up. Also, we must keep in mind that if we let $\dim \rho = d$ and $\dim \rho' = d'$, and we don't require at $d' = d$. Thus, our output matrix can be different dimensions than our input matrix. This is because we can either measure and forget some qubits from our input, or we can incorporate some qubits from B into ρ'. This is called a quantum channel, or a superoperator.

Definition 1.2. A quantum channel is a set of **completely positive trace preserving** operators, which from here on out we'll abbreviate as CPTP operators. This means that the operators are both linear and completely positive.

Remark 1.3. In the classical (or randomized) sense, we can actually represent a quantum channel as a $d' \times d$ stoachastic matrix. From this intuition, we can see that a quantum channel can be thought of as one step of a Markov Chain.

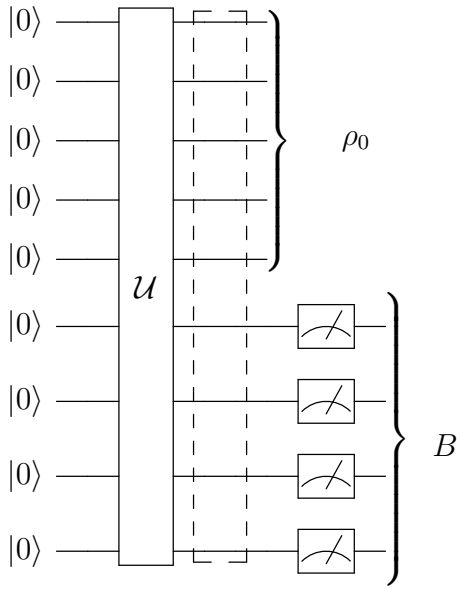
Theorem 1.4. *Every quantum channel has a "Kraus" representation of the channel This representation looks as follows*

$$\rho \longrightarrow \boxed{\Phi} \longrightarrow \quad \sum_{i=1}^{r} \mathcal{K}_i \rho \mathcal{K}_i^\dagger$$

For this representation, we require that $\mathcal{K}_i \in \mathbb{C}^{d' \times d}$*, and that* $\sum_{i=1}^{r} \mathcal{K}_i^\dagger \mathcal{K} = I$*. [NC11]*

Let's touch on some interesting things we can interpret given this definition of a quantum channel. First, consider the case where we output ρ_0, and our input state ρ is 1 dimensional. This case is called state preparation, or something that we've in general taken for granted in class. Now, we can consider any initial state we need as coming from a quantum channel which has no input. In the reverse case, we can have $d' = 1$. In this case, ρ is simply discarded, so this quantum channel would discard a state

Lets consider creating some output state ρ_0 using a quantum circuit. First, we put in some ancillary qubits, and then we discard B bits, after measuring then forgetting them, and then take our output state which is $\rho_0 = \mathrm{tr}_B \left|\psi\right\rangle_{AB}$. We can consider the circuit below to find this definition. We actually call our intermediate pure state $\left|\psi\right\rangle_{AB}$ a pure state. looking at this quantum channel and our mixed output state ρ_0, we note that *every mixed state is a purification of some pure state.*



The boxed area after the application of our giant unitary operator $\mathcal{U}$ is clearly some pure state, as it originates from a unitary transformation on our ancillas. Interestingly, we actually call this boxed state $\left|\psi\right\rangle$ the *purification* of ρ_0.

2 Finishing Holevo's Bound

Now, let's return to our story of Alice and Bob. In this story, Alice has some ensemble of mixed states, which well call $\mathcal{E}$. This ensemble consists of some density matrices σ_x, and

each density matrix has some probability $p(x)$ of being sent over to Bob. Now, we define a random variable X, which is in itself a mixed state, which sums ovar all of Alice's possible density matrices with their associated probabilities. Since X is a sum of mixed states, X itself is actually a mixed state, which has the property that the probability $P(X = x) = p(x)$. Alice prepares and sends some density matrix σ_X over to bob. The ultimate question we're looking to answer here is : how much information about X can bob determine from σ_X? To answer this question, let's look into what Bob will do with this σ_X.

Essentially, all Bob can really do with this σ_X is measure it in his favorite way, which produces a classical outcome $Y \in \Gamma$. Bob's mixed state which he measures on, is

$$\rho_B = \sum_{x \in \Sigma} p(x)\sigma_x$$

This makes the overall joint state between Alice and Bob look like the following

$$\rho_{AB} = \sum_{X \in \Sigma} \rho(x) \left|x\right\rangle \left\langle x\right| \otimes \sigma_x$$

Now, let's get back to our question : *How much classical info can bob get on X?* To look at this, we're going to go back to the previous lecture and what we learned about mutual information, or $I(X, Y)$.

Claim 2.1. *By Holevo's theorem, we have an upper limit on* $I(X, Y) \leq \chi(\mathcal{E})$

Recall our previous definition of $I(X, Y) = H(X) + H(Y) - H(X, Y)$, where each $H(X)$ is the Shannon Entropy, defined as $H(X) = \sum_{x \in \Sigma} p(x) \log(\frac{1}{p(x)})$.

Claim 2.2. $I(X, Y) = H(Y) - H(Y|X)$

Before going into the proof of this we'll first note that we mean $H(Y|X)$ to be defined as

$$H(Y|X) = \sum_{x \in \Sigma} \rho(x) H(Y_x)$$

Where $H(Y_x)$ is the Shannon entropy of Y when we condition that $X = x$.

Proof. It will suffice to show that $H(X, Y) = H(X) + H(Y|X)$, because if we prove this then we can simply show that

$$\begin{aligned} I(X, Y) &= H(X) + H(Y) - H(X, Y) && (1) \\ &= H(X) + H(Y) - H(X) - H(Y|X) && (2) \\ &= H(Y) - H(Y|X) && (3) \end{aligned}$$

To make the proof manageable, we will do this for a classical X and Y, but the result does hold for quantum random variables as well. Keep in mind that Y_x is just Y conditioned on

$X = x$

$$H(X,Y) = \sum_{x,y} p(x,y)\log(\frac{1}{p(x,y)}) \tag{4}$$

$$= \sum_{x,y} p(x)Pr[Y_x = y]\log(\frac{1}{p(x)Pr[Y_x = y]}) \tag{5}$$

$$= \sum_{x,y} p(x)Pr[Y_x = y](\log(\frac{1}{p(x)}) + \log(\frac{1}{Pr[Y_x = y]})) \tag{6}$$

$$= \sum_{x,y} p(x)Pr[Y_x = y]\log(\frac{1}{p(x)}) + \sum_{x,y} p(x)Pr[Y_x = y]\log(\frac{1}{Pr[Y_x = y]}) \tag{7}$$

$$= \sum_{x} p(x)\log(\frac{1}{p(x)}) + \sum_{x} p(x)H(Y_x) \tag{8}$$

$$= H(X) + H(Y|X) \tag{9}$$

□

Now that we've shown that $H(X,y) = H(X) + H(Y|X)$, let's go on to answer another question, what exactly is $\chi(\mathcal{E})$. (Additionally, remember that our ensemble is $\mathcal{E} = \{p(x), \sigma_x\}$ with $x \in \Sigma$, so this is the same χ as in the lecture 18 scribe notes.

Definition 2.3. Before giving this definition, we define $I(\rho_A; \rho_B)$ to be the quantum mutual information between two density matrices, which is defined whenever you have a mixed state over two registers. Also, we define $H(\rho)$ to be the quantum, or Von Nuemann entropy of some mixed state. With these definition in mind, we define our Holevo information $\chi(\mathcal{E})$ to be

$$\chi(\mathcal{E}) = I(\rho_A; \rho_B) = H(\rho_A) + H(\rho_B) - H(\rho_{AB})$$

Another common definition which is used for the Holevo Information is the following:

Definition 2.4.

$$\chi(\mathcal{E}) = H(\rho_B) - \sum_{x\in\Sigma} \rho(x)H(\sigma_x)$$

Now, lets look at a few useful pieces of information about Von Nuemann entropy and Holevo Information

Remark 2.5. The Holevo Information of an ensemble is always non-negative, because Von Nuemann entropy is concave.

Remark 2.6. $\chi(\mathcal{E}) \leq H(\rho_B)$. This is easy to see from definition 2.4 and the non-negativity of Von Nuemann entropy.

Corollary 2.7. *A quantum channel which sends* $\log_2 d$ *qubits can convey at most an equivalent* $\log_2 d$ *classical bits of information between Alice and Bob [Hol73].*

Although quantum information gains no advantage over classical information in this sense, if Alice and Bob share enough EPR pairs, they actually only need $n/2$ qubits via a technique called *superdense coding* [Aar14].

An interesting theorem which Holevo and others showed in 1997 was that the upper bound of $\chi(\mathcal{E})$ mutual information to be shared between Alice and Bob is acheivable in the case of repeated communication. [SW97]. To connect this theorem concretely back to our discussion, we consider Alice drawing n random variables X, and producing $X^n = (x_1, \cdots, x_n)$. Alice then sends $\sigma_{x_1} \otimes \sigma_{x_2} \cdots \otimes \sigma_{x_n}$ to Bob, who makes n separate measurements to produce a Y^n such that their mutual information takes the following form:

$$I(X^n, Y^n) = \chi(\mathcal{E}) - O_{n\to\infty}(1)$$

3 More comments on Holevo Information

Now, let's consider Alice sending her ensemble through some noisy channel as follows

$$\textit{Alice} \quad — \quad \{p(x), \sigma_x\} \quad \text{——}\boxed{\Phi}\text{—} \mathcal{E}' = \{p(x), \Phi(\sigma_x)\} — \quad \textit{Bob}$$

We now define the **Holevo Capacity** of the noise Φ, which is equivalent to the number of bits you can reliably communicate

$$\chi(\Phi) = \max_{\mathcal{E}}(\chi(\mathcal{E}'))$$

Now, lets delve more into the idea (which we won't prove due to time constraints and lack of relevance), that $I(X, Y) = \chi(\mathcal{E})$

Theorem 3.1. *Von Nuemann entropy has the property that it is strongly subadditive. That is to say, if we have some τ_{ABC}, a tripartite quantum state (think of a state over three registers). Now, consider τ_B to be measuring and forgetting registers A and C. Strong subadditivity is to say that*

$$H(\tau_B) + H(\tau_{ABC}) \leq H(\tau_{AB}) + H(\tau_{BC})$$

Since Von Neumann entropy is strongly subadditive, it is also subadditive. Subadditivity is actually a case of strong subadditivity where B is 0 qubits, or it has degree one. This means that $H(\tau_B) = 0$, and that $H(\tau_{ABC}) = H(\tau_{AC})$. In this case, we can define the subadditivity of Von Neumann entropy to be

$$H(\tau_{AC}) \leq H(\tau_A) + H(\tau_C)$$

Remark 3.2. Because $H(\tau_{AC}) \leq H(\tau_A) + H(\tau_C)$, we can clearly see that $I(\tau_A; \tau_C) \geq 0$, which makes sense, as a negative mutual information doesn't make sense intuitively.

Remark 3.3. Discarding a register never increases the mutual information. By combining subadditivity and strong subadditivity, we can deduce that

$$H(\tau_A) + H(\tau_{BC}) - H(\tau_{ABC}) \geq H(\tau_A) + H(\tau_B) - H(\tau_{AB}) \tag{10}$$

$$I(\tau_A; \tau_{BC}) \geq I(\tau_A; \tau_B) \tag{11}$$

From this deduction, we gain that discarding a register can never help you gain information, which also makes intuitive sense.

Corollary 3.4. *Let τ_{AB} be a 2 register state. Say we then pass the B register through some noise, Φ which praudices τ'_{AB}. We say that $I(\tau_A; \tau'_B) \leq I(\tau_A, \tau_B)$.*

By corollary 3.4, we can actually see the holevo bound, This is eessentially bob taking a register, and then measuring it and forgetting it, which produces

$$I(\tau_A, \tau'_B) \leq I(\tau_A, \tau_B) \tag{12}$$
$$I(X, Y) \leq I(\rho_A, \rho_B) \tag{13}$$

Ultimately, since an upper bound of $\chi(\mathcal{E})$ shared information can be both proven and acheived, quantum information theory doesn't get any magical improvments over classical information theory in the same way which quantum algorithms can improve classical algorithms.

References

[Aar14] Scott Aaronson. Lecture 21: Quantum communication complexity. *6.845 Quantum Complexity Theory*, page 1, 2014.

[Gri12] Robert B. Griffiths. Quantum channels, kraus operators, povms. *33-658 Spring 2012*, pages 2–4, 2012.

[Hol73] Alexander S. Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problems of Information Transmission*, 9:177–183, 1973.

[NC11] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, New York, NY, USA, 10th edition, 2011.

[SW97] Benjamin Schumacher and Michael D. Westmoreland. Sending classical information via noisy quantum channels. *Phys. Rev. A*, 56:131–138, Jul 1997.