Lecture 13: October 15, 2020

Lecturer: Aram Harrow *Scribe: Thiago Bergamaschi, Yuan Lee*

We begin by exploring the intuition for the connections between the Holevo χ quantity, the minimum entropy and the entanglement of formation in Shor's 2003 paper quant-ph/0305035, following the discussion in the previous lecture.

13.1 Sub-additivity of S_{min} ⇒ Super-additivity of χ

Given an ensemble of states $\{p_i, \rho_i\}$ with average state $\bar{\rho} = \sum p_i \rho_i$, the Holevo χ quantity is

$$\chi(N) = S(N(\bar{\rho})) - \sum p_i S(N(\rho_i)) \tag{13.1}$$

by definition.

We can bound χ from above using

$$\chi(N) \leq S(N(\bar{\rho})) - S_{min}(N) \leq S_{max}(N) - S_{min}(N) \leq \log d_B - S_{min}(N) \tag{13.2}$$

Shor showed that for every channel N, one can construct a channel N' that makes the inequalities above tight. In particular, if N has dimension d_B,

$$\chi(N') = \log d_B - S_{min}(N) \tag{13.3}$$

The construction is quite straightforward. After the application of the quantum channel N, apply a classically-controlled random Pauli operator σ_x, so that $N'(\rho) = \sigma_x N(\rho) \sigma_x^\dagger$. In this manner, the first term in equation (13.1) is $S(N(\bar{\rho})) = \log d_B$, because

$$N'\left(\sum_x \frac{1}{d_B^2} |x\rangle\langle x| \otimes \rho\right) = \sum_x \frac{1}{d_B^2} \sigma_x N(\rho) \sigma_x^\dagger = \frac{\mathbb{I}}{d_B} \tag{13.4}$$

Moreover, the second term is $S_{min}(N') = S_{min}(N)$. It follows that if S_{min} is subadditive, then χ is superadditive.

The other direction is non-trivial.

13.2 Renyi Entropies

S_{min} is generally computationally more tractable than χ, as we can view them as the limit of Renyi entropies. Recall

$$S_\alpha(\rho) = \frac{1}{1-\alpha}\mathrm{Tr}[\rho^\alpha] \tag{13.1}$$

There are a few particular cases of α to highlight: $S_0 = \log \mathrm{rank}\, \rho, S_\infty = -\log ||\rho||_\infty$ and $S_1(\rho) = S(\rho)$, the standard Von Neumann entropy. Analogously, we can define the min Renyi Entropy via

$$S_{\alpha,min}(N) = \min_\psi S_\alpha(N(\psi)) = \frac{\alpha}{1-\alpha}\log ||N||_{1\to\alpha} \tag{13.2}$$

where $||N||_{\beta\to\alpha}$ is the "beta to alpha norm" defined as follows

$$||N||_{\beta\to\alpha} = \sup\frac{||N(X)||_\alpha}{||X||_\beta} \tag{13.3}$$

Finding $S_{\alpha,min}$ is still a hard optimization problem, but $S_{\alpha,min}$ is more helpful to us because the norms it is related to obey useful inequalities.

13.3 The Connection to the Entanglement of Formation

We can analogously extend the Holevo information of a state ρ, by decomposing over an ensemble of pure states $\{p, \phi\}$ that averages ρ

$$\chi(N,\rho) = \max_{\{p,\phi\} \text{ s.t. } \sum_x p_x\phi_x=\rho} S(N(\rho)) - \sum p_x S(N(\phi_x)) \tag{13.1}$$

where to conclude $\chi(N) = \max_\rho \chi(N,\rho)$. If we consider applying the Stinespring dilation theorem to N s.t. $N(\omega) = \mathrm{Tr}_E(V\omega V^\dagger)$, then $S(N(\phi_x))$ is simply the entanglement of $V|\phi_x\rangle$, and it follows

$$\chi(N,\rho) = S(N(\rho)) - E_F(V\rho V^\dagger) \tag{13.2}$$

We might be concerned that not all entanglements of formation $E_F(V\rho V^\dagger)$ correspond to the *minimum* average entropy of the corresponding channel N. However, the MSW correspondence states that

$$E_F\left(\rho^{BE}\right) = \min_{\{p,\phi\} \text{ s.t. } \sum_x p_x\phi_x=\rho} \sum_x p_x S\left(\mathrm{tr}_E \phi_x^{BE}\right)$$

for any bipartite state ρ.

Applying it to the Stinespring dilation of $N_{A\to B}$, we have

$$E_F(V\rho V^\dagger) = \min_{\{p,\phi\}\text{ s.t. }\sum_x p_x\phi_x=\rho} \sum_x p_x S\left(\mathrm{tr}_E V\phi_x V^\dagger\right)$$

which guarantees that equation (13.2) holds.

13.4 Entanglement-Assisted Capacity

The discussion of S_{min} and χ above only delays the pain of performing the optimization problem over $\{p, \phi\}$ such that $\sum_x p_x\phi_x = \rho$. Now we turn to the more well-understood problem of entanglement-assisted capacities.

We want to analyze the additivity of the entanglement-assisted capacity of two independent channels $C_E(N_1 \otimes N_2)$. Define systems A_1', A_2' upon which N_1, N_2 act, respectively, and consider an environment system A

$$C_E(N_1\otimes N_2) = \max I(A:B_1B_2)_\tau, \tau = (\mathbb{I}_A\otimes N_1^{A_1'\to B_1}\otimes N_2^{A_2'\to B_2})(\phi^{AA_1'A_2'}) \tag{13.1}$$

$$= \max I(A:B_1B_2)_\psi, |\psi\rangle = (\mathbb{I}_A\otimes V_1^{A_1'\to B_1E_1}\otimes V_2^{A_2'\to B_2E_2})|\phi\rangle \tag{13.2}$$

note the distinction between the two definitions, where in the second we purify the two systems independently s.t. they are separable in ψ but not in τ. We will use this independence later. It follows now that we can apply the chain rule sequentially

$$I(A:B_1B_2)_\psi = I(A:B_1) + I(A:B_2|B_1) = \tag{13.3}$$

$$= I(A:B_1) + I(AB_1:B_2) - I(B_1:B_2) \le I(A:B_1) + I(AB_1:B_2) \tag{13.4}$$

as the mutual information is non-negative. Let us consider the terms above independently, starting by $I(AB_1 : B_2)$. Intuitively, if it was helpful to include B_1 in addition to A, we could have included it into the definition of the 'environment system' WLOG. In this manner, $I(AB_1 : B_2) \le I(AB_1E_1 : B_2)$, and symmetrically for $1 \leftrightarrow 2$. We conclude

$$C_E(N_1\otimes N_2) \le I(AB_1E_1:B_2) + I(AB_2E_2:B_1) \le C_E(N_1) + C_E(N_2) \tag{13.5}$$

Finally, note that this upper bound is always achievable as we can run the channels independently. We conclude

$$C_E(N_1\otimes N_2) = C_E(N_1) + C_E(N_2) \tag{13.6}$$

and therefore we conclude C_E is additive and has a single-letter formula that is concave in ρ. Moreover, through superdense coding and quantum teleportation (problem set 5, problem 1) it determines the quantum entanglement-assisted capacity $Q_E = C_E/2$ as well.

We can contrast these nice properties of the entanglement-assisted capacities (C_E additive and $C_E = 2Q_E$) with the difficulty of the unassisted capacities (C, Q). We only know that $Q \leq C$, but this bound can have large gaps. For instance, the completely dephasing channel has $C = 1$ but $Q = 0$. On the other hand, the noiseless channel has $Q = C$. This makes it difficult to think of channels as equivalent resources in the absence of free entanglement.

13.5 Quantum Reverse Shannon Theorem and Embezzling States

The additivity of C_E and the reversibility of the quantum Shannon theorem allows us to think of channels as equivalent resources, associated with a resource theory. In particular, reversibility (defined formally below) allows us to convert between channels at a common "exchange rate".

The quantum reverse Shannon theorem states that any quantum channel can be simulated by an 'unlimited amount' of shared entanglement and C_E classical bits, where C_E is the entanglement-assisted classical capacity of the channel. In informal resource notation,

$$\text{unlimited entanglement} + C_E[c \to c] \geq \langle N \rangle \tag{13.1}$$

The lecturer traces a key distinction here between 'unlimited entanglement' and $\infty[qq]$, an arbitrary amount of EPR pairs. The key intuition is that the channel simulation may consume a different amount of EPR pairs for different inputs, and therefore it doesn't suffice to feed some amount of EPR pairs to the protocol. Instead, *embezzling states* are bipartite states that allow the removal of a small amount of entanglement under local operations into an additional set of registers, while the original state remains approximately the same. That is, heuristically,

$$|\Gamma\rangle_{AB} \to\approx |\Gamma\rangle_{AB} \otimes |\psi\rangle_{A'B'} \tag{13.2}$$

where the A'B' registers are much smaller than AB. A motivating example is the following state.

$$|\Gamma\rangle = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} |\Phi_2\rangle^{\otimes i} \otimes |00\rangle^{\otimes n-i} |ii\rangle \tag{13.3}$$

Note that if we define Γ' based on the removal of the first Bell pair Φ_2, i.e.

$$|\Gamma'\rangle = \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} |\Phi_2\rangle^{\otimes i} \otimes |00\rangle^{\otimes n-i} |ii\rangle \tag{13.4}$$

then the fidelity $F(\Gamma, \Gamma') = 1 - \frac{1}{n}$ and we have 'stolen an EPR pair'.

Another example of an embezzling state is

$$|\Psi\rangle \propto \sum_{i=1}^{N} \frac{1}{\sqrt{i}} |ii\rangle$$

for some finite N. Embezzling entanglement then looks like

$$(U^A \otimes V^B) |\Psi\rangle^{AB} |00\rangle^{AB} \approx |\Psi\rangle^{AB} |\Phi_2\rangle^{AB}$$

for some local unitaries U^A and V^B.

We can show that there exist local unitaries U, V such that $F((U \otimes V) |\Psi\rangle |00\rangle, |\Psi\rangle |\Phi_2\rangle) \geq 1 - 1/\log n$. Let $\sum_{i=1}^{N} 1/i = C_N$. The Schmidt coefficients of $|\Psi\rangle |00\rangle$ are

$$\frac{1}{\sqrt{C_N}}, \frac{1}{\sqrt{2C_N}}, \frac{1}{\sqrt{3C_N}}, \frac{1}{\sqrt{4C_N}}, \dots \frac{1}{\sqrt{NC_N}}, 0, \dots 0$$

whereas the Schmidt coefficients of $|\Psi\rangle |\Phi_2\rangle$ are

$$\frac{1}{\sqrt{2C_N}}, \frac{1}{\sqrt{2C_N}}, \frac{1}{\sqrt{4C_N}}, \frac{1}{\sqrt{4C_N}}, \frac{1}{\sqrt{6C_N}}, \frac{1}{\sqrt{6C_N}}, \dots, \frac{1}{\sqrt{2NC_N}}, \frac{1}{\sqrt{2NC_N}}$$

Therefore, the maximum fidelity is

$$\begin{aligned}
&\left(\frac{1}{\sqrt{C_N}}, \frac{1}{\sqrt{2C_N}}, \dots \frac{1}{\sqrt{NC_N}}, 0, \dots 0\right) \cdot \left(\frac{1}{\sqrt{2C_N}}, \frac{1}{\sqrt{2C_N}}, \frac{1}{\sqrt{4C_N}}, \frac{1}{\sqrt{4C_N}}, \dots \frac{1}{\sqrt{2NC_N}}\right) \\
&\geq \frac{1}{C_N} \left(\frac{1}{2} + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{6} + \frac{1}{6} + \dots + \frac{1}{N}\right) \\
&\geq \frac{1}{C_N} \left(\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{\lfloor N/2 \rfloor}\right) \\
&\geq \frac{\ln N/2}{\ln N} = 1 - \frac{1}{\log N}
\end{aligned}$$

Note that entanglement embezzlement preserves the original superposition across the bipartite state $|\Psi\rangle$, which is crucial for the quantum reverse Shannon theorem.

13.6 Quantum Capacity

In resource notation, we define the quantum capacity by

$$\langle N \rangle \geq Q(N)[q \to q] \tag{13.1}$$

In general, $Q \leq Q_F \leq Q_2$, that is feedback and two-way channels increase the capacity of quantum information over the channel, however, sending additional classical communication $[c \to c]$ does not help. Mathematically, the quantum capacity is defined by the maximum amount distallable entanglement that can be generated with the channel

$$Q(N) = \max_{\psi_{AA'}} E_D((\mathbb{I}_A \otimes N_{A' \to B})\psi_{AA'}) \tag{13.2}$$

The Choi-Jamiolkowski state $\omega(N)$ is

$$\omega(N) = (\mathbb{I}_A \otimes N_{A' \to B})(\Phi_{AA'}) = \frac{1}{d_A} \sum_{ij} |i\rangle\langle j| \otimes N(|i\rangle\langle j|) \tag{13.3}$$

where $\Phi_{AA'} = \frac{1}{\sqrt{d_A}} \sum_i |ii\rangle$ is the maximally mixed state. This state presents an interesting interpretation of the channel, as the mapping $N \to \omega(N)$ is an isomorphism (known as the *Choi-Jamiolkowski isomorphism*). We can show that this mapping is isomorphic by identifying the inverse map: conditioning on the first subsystem of $\omega(N)$, we obtain $N(|i\rangle \langle j|)$ for every basis element $|i\rangle \langle j|$, which suffices to define the channel N.

We can use ω to simulate N as follows. Consider three registers E, A, A', where E holds a state ρ and A, A' share the maximally mixed state $\Phi_{AA'}$. Consider the quantum circuit defined by feeding A' through the quantum channel N, and a Bell state measurement is jointly made on the registers E, A. If the bell state measurement returns a string j, then the state resulting on the register $A' \to B$ is $N(\sigma_j \rho \sigma_j^\dagger)$. In this manner, $j = 0$ with probability d_A^{-2}, and then $N(\sigma_j \rho \sigma_j^\dagger) = N(\rho)$. It follows $\omega(N)$ can simulate N with probability d_A^{-2} and in this manner,

$$E_D(\omega(N)) > 0 \iff Q(N) > 0 \tag{13.4}$$

Unfortunately, it is still largely unknown when $Q(N) = 0$. A case that $Q(N) = 0$ is when N is entanglement-breaking, or equivalently when $\omega(N) \in \text{Sep}$ is separable.

We can generalize entanglement-breaking channels in two different ways. One way to generalize the entanglement-breaking property is to consider antidegradable channels.

We say N is *antidegradable* if there exists some map ε such that $N = \varepsilon \circ N^c$ (i.e. Bob gets less information than Eve). Conversely, we say that N is *degradable* if there exists some map ε such that $N^c = \varepsilon \circ N$ (i.e. Bob gets more information than Eve). For example, the erasure channel with erasure probability p is degradable for $p \leq 1/2$ and antidegradable for $p \geq 1/2$. In general, however, not all channels are either degradable or antidegradable.

We can show using the no-cloning theorem that antidegradable channels also have zero quantum capacity (without classical feedback). Interestingly, degradable channels have additive capacity.