PHYS 7895: Quantum Information Theory Fall 2015

Lecture 25 — November 18, 2015

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

1 Overview

In the last lecture, we proved the HSW classical capacity theorem.

In this lecture, we discuss entanglement-assisted classical communication.

2 Introduction

We have learned that shared entanglement is often helpful in quantum communication. This is certainly true for the case of a noiseless qubit channel. Without shared entanglement, the most classical information that a sender can reliably transmit over a noiseless qubit channel is just one classical bit. With shared entanglement, they can achieve the super-dense coding resource inequality:

$$[q \to q] + [qq] \geq 2\,[c \to c]\,. \tag{1}$$

That is, with one noiseless qubit channel and one shared noiseless ebit, the sender can reliably transmit two classical bits.

A natural question then for us to consider is whether shared entanglement could be helpful in transmitting classical information over a noisy quantum channel $\mathcal{N}$. As a first simplifying assumption, we let Alice and Bob have access to an infinite supply of entanglement, in whatever form they wish, and we would like to know how much classical information Alice can reliably transmit to Bob over such an entanglement-assisted quantum channel. That is, we would like to determine the highest achievable rate C of classical communication in the following resource inequality:

$$\langle \mathcal{N} \rangle + \infty\,[qq] \geq C\,[c \to c]\,. \tag{2}$$

The answer to this question is one of the strongest known results in quantum Shannon theory, and it is given by the entanglement-assisted classical capacity theorem. This theorem states that the mutual information $I(\mathcal{N})$ of a quantum channel $\mathcal{N}$ is equal to its entanglement-assisted classical capacity, where

$$I(\mathcal{N}) \equiv \max_{\phi_{AA'}} I(A;B)_\rho, \tag{3}$$

$\rho_{AB} \equiv \mathcal{N}_{A' \to B}(\phi_{AA'})$, and the maximization is over all pure bipartite states of the form $\phi_{AA'}$. We should stress that there is no need to regularize this formula in order to characterize the capacity (as done in the previous chapter and as is so often needed in quantum Shannon theory). The value of this formula *is* the capacity. Also, the optimization task that the formula in (3) sets

out is a straightforward convex optimization program. Any local maximum is a global maximum because the quantum mutual information is concave in the input state $\phi_{A'}$ (recall Theorem **??** from Chapter **??**) and the set of density operators is convex.

From the perspective of an information theorist, we should only say that a capacity theorem has been solved if there is a tractable formula equal to the optimal rate for achieving a particular operational task. The formula should apply to an arbitrary quantum channel, and it should be a function of that channel. Otherwise, the capacity theorem is still unsolved. There are several operative words in the above sentences that we should explain in more detail. The formula should be tractable, meaning that it sets out an optimization task which is efficient to solve in the dimension of the channel's input system. The formula should give the optimal achievable rate for the given information-processing task, meaning that if a rate exceeds the capacity of the channel, then the probability of error for any such protocol should be bounded away from zero as the number of channel uses grows large.[1] Finally, perhaps the most stringent (though related) criterion is that the formula itself (and *not* its regularization) should give the capacity of an arbitrary quantum channel. Despite the success of the HSW coding theorem in demonstrating that the Holevo information of a channel is an achievable rate for classical communication, the classical capacity of a quantum channel is still unsolved because there is an example of a channel for which the Holevo information is not equal to that channel's capacity (see Section **??**). Thus, it is rather impressive that the formula in (3) is equal to the entanglement-assisted classical capacity of an arbitrary channel, given the stringent requirements that we have established for a formula to give the capacity. In this sense, shared entanglement simplifies quantum Shannon theory.

This chapter presents a comprehensive study of the entanglement-assisted classical capacity theorem. We begin by defining the information-processing task, consisting of all the steps in a general protocol for classical communication over an entanglement-assisted quantum channel. We then present a simple example of a strategy for entanglement-assisted classical coding that is inspired by dense coding, and in turn, that inspires a strategy for the general case. Section 5 states the entanglement-assisted classical capacity theorem. Section 6 gives a proof of the direct coding theorem, making use of quantum typicality from Chapter **??**, the packing lemma from Chapter **??**, and ideas in the entanglement concentration protocol from Chapter **??**. It demonstrates that the rate in (3) is an achievable rate for entanglement-assisted classical communication. After taking a step back from the protocol, we can realize that it is merely a glorified super-dense coding applied to noisy quantum channels. Section 7 gives a proof of the converse of the entanglement-assisted classical capacity theorem. It exploits familiar tools such as the Alicki–Fannes inequality, the quantum data-processing inequality, and the chain rule for quantum mutual information (all from Chapter **??**), and the last part of it exploits additivity of the mutual information of a quantum channel (from Chapter **??**). The converse theorem establishes that the rate in (3) is optimal. With the proof of the capacity theorem complete, we then show the interesting result that the classical capacity of a quantum channel assisted by a quantum feedback channel is equal to the entanglement-assisted classical capacity of that channel. We close the chapter by computing the entanglement-assisted classical capacity of both a quantum erasure channel and an amplitude damping channel, and we leave the computation of the entanglement-assisted capacities of two other channels as exercises.

[1] We could strengthen this requirement even more by demanding that the probability of error increases exponentially to one in the asymptotic limit. Fulfilling such a demand would constitute a proof of a *strong converse theorem*.

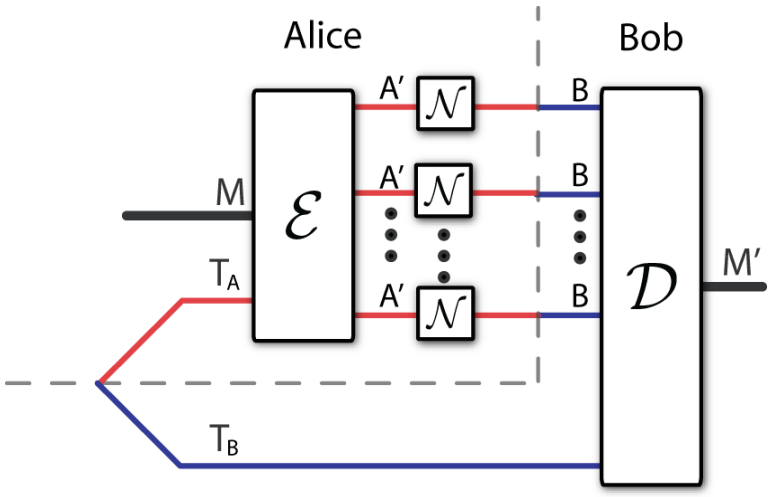

Figure 1: The most general protocol for entanglement-assisted classical communication. Alice applies an encoder to her classical message M and her share T_A of the entanglement, and she inputs the encoded systems A'^n to many uses of the channel. Bob receives the outputs of the channel, combines them with his share of the entanglement, and performs some decoding operation to estimate Alice's transmitted message.

3 The Information-Processing Task

We begin by explicitly defining the information-processing task of entanglement-assisted classical communication, i.e., we define an (n, C, ε) entanglement-assisted classical code and what it means for a rate C to be achievable. Prior to the start of the protocol, we assume that Alice and Bob share pure-state entanglement in whatever form they wish. For simplicity, we can just assume that they share a maximally entangled state of the following form:

$$|\Phi\rangle_{T_A T_B} \equiv \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle_{T_A} |i\rangle_{T_B}, \tag{4}$$

where the dimension d is as large as they would like it to be. Alice selects some message m uniformly at random from a set $\mathcal{M}$ of messages. Let M denote the random variable corresponding to Alice's random choice of message, and let $|\mathcal{M}|$ be the cardinality of the set $\mathcal{M}$. She applies some encoding channel $\mathcal{E}^m_{T_A \to A'^n}$ to her share of the entangled state $\Phi_{T_A T_B}$ depending on her choice of message m. The global state then becomes

$$\mathcal{E}^m_{T_A \to A'^n}(\Phi_{T_A T_B}). \tag{5}$$

Alice transmits the systems A'^n over n independent uses of a noisy channel $\mathcal{N}_{A' \to B}$, leading to the following state:

$$\mathcal{N}_{A'^n \to B^n}(\mathcal{E}^m_{T_A \to A'^n}(\Phi_{T_A T_B})), \tag{6}$$

where $\mathcal{N}_{A'^n \to B^n} \equiv (\mathcal{N}_{A' \to B})^{\otimes n}$. Bob receives the systems B^n, combines them with his share T_B of the entanglement, and performs a POVM $\{\Lambda^m_{B^n T_B}\}$ on the channel outputs B^n and his share T_B of the entanglement in order to detect the message m that Alice transmits. Figure 1 depicts such a general protocol for entanglement-assisted classical communication.

Let M' denote the random variable for the output of Bob's decoding POVM (this represents Bob's estimate of the message). The probability of Bob correctly decoding Alice's message is

$$\Pr\left\{M' = m | M = m\right\} = \mathrm{Tr}\{\Lambda^m_{B^n T_B} \mathcal{N}_{A'^n \to B^n}(\mathcal{E}^m_{T_A \to A'^n}(\Phi_{T_A T_B}))\}, \tag{7}$$

and thus the probability of error $p_e(m)$ for message m is

$$p_e(m) \equiv \mathrm{Tr}\{\left(I - \Lambda^m_{B^n T_B}\right) \mathcal{N}_{A'^n \to B^n}(\mathcal{E}^m_{T_A \to A'^n}(\Phi_{T_A T_B}))\}. \tag{8}$$

The maximal probability of error p^*_e for the coding scheme is

$$p^*_e \equiv \max_{m \in \mathcal{M}} p_e(m). \tag{9}$$

The rate C of communication is

$$C \equiv \frac{1}{n} \log_2 |\mathcal{M}|, \tag{10}$$

and the code has ε error if $p^*_e \leq \varepsilon$.

A rate C of entanglement-assisted classical communication is *achievable* if there exists an $(n, C - \delta, \varepsilon)$ entanglement-assisted classical code for all $\varepsilon \in (0, 1)$, $\delta > 0$, and sufficiently large n. The entanglement-assisted classical capacity $C_{\mathrm{EA}}(\mathcal{N})$ of a quantum channel $\mathcal{N}$ is equal to the supremum of all achievable rates of entanglement-assisted classical communication.

4 A Preliminary Example

Let us first recall a few items about qudits. The maximally entangled qudit state is

$$|\Phi\rangle_{AB} \equiv \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i\rangle_A |i\rangle_B. \tag{11}$$

Recall from Section **??** that the Heisenberg–Weyl operators $X(x)$ and $Z(z)$ are an extension of the Pauli matrices to d dimensions:

$$X(x) \equiv \sum_{x'=0}^{d-1} \left|x + x'\right\rangle \left\langle x'\right|, \qquad Z(z) \equiv \sum_{z'=0}^{d-1} e^{2\pi i z z'/d} \left|z'\right\rangle \left\langle z'\right|. \tag{12}$$

Let $|\Phi^{x,z}\rangle_{AB}$ denote the state that results when Alice applies the operator $X(x)Z(z)$ to her share of the maximally entangled state $|\Phi\rangle_{AB}$:

$$|\Phi^{x,z}\rangle_{AB} \equiv (X_A(x) Z_A(z) \otimes I_B) |\Phi\rangle_{AB}. \tag{13}$$

Recall from Exercise **??** that the set of states $\{|\Phi^{x,z}\rangle_{AB}\}_{x,z=0}^{d-1}$ forms a complete orthonormal basis:

$$\langle \Phi^{x_1,z_1} | \Phi^{x_2,z_2} \rangle = \delta_{x_1,x_2} \delta_{z_1,z_2}, \qquad \sum_{x,z=0}^{d-1} |\Phi^{x,z}\rangle \langle \Phi^{x,z}| = I_{AB}. \tag{14}$$

Let π_{AB} denote the maximally mixed state on Alice and Bob's system: $\pi_{AB} \equiv I_{AB}/d^2$, and let π_A and π_B denote the respective maximally mixed states on Alice and Bob's systems: $\pi_A \equiv I_A/d$ and $\pi_B \equiv I_B/d$. Observe that $\pi_{AB} = \pi_A \otimes \pi_B$.

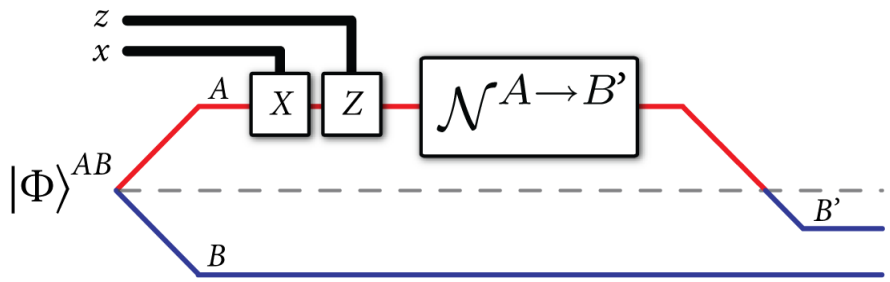

Figure 2: A simple scheme, inspired by super-dense coding, for Alice and Bob to exploit shared entanglement and a noisy channel in order to establish an ensemble at Bob's receiving end.

We now consider a simple strategy, inspired by super-dense coding and the HSW coding scheme from Theorem **??**, that Alice and Bob can employ for entanglement-assisted classical communication. That is, we show how a strategy similar to super-dense coding induces a particular ensemble at Bob's receiving end, to which we can then apply the HSW coding theorem in order to establish the existence of a good code for entanglement-assisted classical communication. Suppose that Alice and Bob possess a maximally entangled qudit state $|\Phi\rangle_{AB}$. Alice chooses two symbols x and z uniformly at random, each in $\{0,\ldots,d-1\}$. She applies the operators $X(x)Z(z)$ to her side of the maximally entangled state $|\Phi\rangle_{AB}$, and the resulting state is $|\Phi^{x,z}\rangle_{AB}$. She then sends her system A over the noisy channel $\mathcal{N}_{A\to B'}$, and Bob receives the output B' from the channel. The noisy channel on the whole system is $\mathcal{N}_{A\to B'}\otimes \mathrm{id}_B$, and the ensemble that Bob receives is as follows:

$$\left\{\frac{1}{d^2},\quad (\mathcal{N}_{A\to B'}\otimes \mathrm{id}_B)\left(\Phi^{x,z}_{AB}\right)\right\}. \tag{15}$$

This constitutes an ensemble that they can prepare with one use of the channel and one shared entangled state (Figure 2 depicts all of these steps). But, in general, we allow them to exploit many uses of the channel and however much entanglement that they need. Bob can then perform a collective measurement on both his half of the entanglement and the channel outputs in order to determine a message that Alice is transmitting.

Consider that the above scenario is similar to HSW coding. Theorem **??** from the previous chapter proves that the Holevo information of the above ensemble is an achievable rate for classical communication over this entanglement-assisted quantum channel. Thus, we can already state and prove the following corollary of Theorem **??**, simply by calculating the Holevo information of the ensemble in (15).

Corollary 1. *The quantum mutual information $I(A;B)_\sigma$ of the state $\sigma_{AB}\equiv\mathcal{N}_{A'\to B}(\Phi_{AA'})$ is an achievable rate for entanglement-assisted classical communication over a quantum channel $\mathcal{N}_{A'\to B}$.*

Proof. Observe that we can map the ensemble in (15) to the following classical–quantum state:

$$\rho_{XZB'B}\equiv\sum_{x,z=0}^{d-1}\frac{1}{d^2}|x\rangle\langle x|_X\otimes|z\rangle\langle z|_Z\otimes(\mathcal{N}_{A\to B'}\otimes\mathrm{id}_B)\left(\Phi^{x,z}_{AB}\right). \tag{16}$$

The Holevo information of this classical–quantum state is

$$I(XZ;B'B)_\rho = H(B'B)_\rho - H(B'B|XZ)_\rho, \tag{17}$$

and it is an achievable rate for entanglement-assisted classical communication over the channel $\mathcal{N}_{A'\to B}$ by Theorem **??**. We now proceed to calculate it. First, we determine the entropy $H(B'B)_\rho$ by tracing over the classical registers XZ:

$$\begin{aligned}
\operatorname{Tr}_{XZ}\{\rho_{XZB'B}\} &= \sum_{x,z=0}^{d-1} \frac{1}{d^2} \left(\mathcal{N}_{A\to B'} \otimes \mathrm{id}_B\right)\left(\Phi_{AB}^{x,z}\right) && (18)\\
&= \left(\mathcal{N}_{A\to B'} \otimes \mathrm{id}_B\right)\left(\sum_{x,z=0}^{d-1} \frac{1}{d^2} \Phi_{AB}^{x,z}\right) && (19)\\
&= \left(\mathcal{N}_{A\to B'} \otimes \mathrm{id}_B\right)\left(\pi_{AB}\right) && (20)\\
&= \mathcal{N}_{A\to B'}\left(\pi_A\right) \otimes \pi_B, && (21)
\end{aligned}$$

where the third equality follows from (14). Thus, the entropy $H(B'B)$ is as follows:

$$H(B'B) = H(\mathcal{N}_{A\to B'}(\pi_A)) + H(\pi_B). \tag{22}$$

We now determine the conditional quantum entropy $H(B'B|XZ)_\rho$:

$$\begin{aligned}
&H(B'B|XZ)_\rho \\
&= \sum_{x,z=0}^{d-1} \frac{1}{d^2} H\left(\left(\mathcal{N}_{A\to B'} \otimes \mathrm{id}_B\right)\left(\Phi_{AB}^{x,z}\right)\right) && (23)\\
&= \frac{1}{d^2} \sum_{x,z=0}^{d-1} H\left(\mathcal{N}_{A\to B'}\left[\left(X_A(x)Z_A(z)\right)\left(\Phi_{AB}\right)\left(Z_A^\dagger(z)X_A^\dagger(x)\right)\right]\right) && (24)\\
&= \frac{1}{d^2} \sum_{x,z=0}^{d-1} H\left(\mathcal{N}_{A\to B'}\left[Z_B^T(z)X_B^T(x)(\Phi_{AB})X_B^*(x)Z_B^*(z)\right]\right) && (25)\\
&= \frac{1}{d^2} \sum_{x,z=0}^{d-1} H\left(Z_B^T(z)X_B^T(x)\left[\left(\mathcal{N}_{A\to B'}\right)\left(\Phi_{AB}\right)\right]\left(X_B^*(x)Z_B^*(z)\right)\right) && (26)\\
&= H(\mathcal{N}_{A\to B'}(\Phi_{AB})). && (27)
\end{aligned}$$

The first equality follows because the system XZ is classical (recall the result in Section **??**). The second equality follows from the definition of the state $\Phi_{AB}^{x,z}$. The third equality follows by exploiting the Bell-state matrix identity in Exercise **??**. The fourth equality follows because the unitaries that Alice applies commute with the action of the channel. Finally, the entropy of a state is invariant under any unitaries applied to that state. So the Holevo information $I(XZ;B'B)_\rho$ of the state $\rho_{XZB'B}$ in (16) is

$$I(XZ;B'B)_\rho = H(\mathcal{N}(\pi_A)) + H(\pi_B) - H(\left(\mathcal{N}_{A\to B'} \otimes \mathrm{id}_B\right)\left(\Phi_{AB}\right)). \tag{28}$$

Equivalently, we can write it as the following quantum mutual information:

$$I(A;B)_\sigma, \tag{29}$$

with respect to the state $\sigma_{AB} \equiv \mathcal{N}_{A'\to B}(\Phi_{AA'})$. □

For some channels, the quantum mutual information in Corollary 1 is equal to that channel's entanglement-assisted classical capacity. This occurs for the depolarizing channel, a dephasing channel, and an erasure channel to name a few. But there are examples of channels, such as the amplitude damping channel, where the quantum mutual information in Corollary 1 is not equal to the entanglement-assisted capacity. In the general case, it might perhaps be intuitive that the quantum mutual information of the channel in (3) is equal to the entanglement-assisted capacity of the channel, and it is the goal of the next sections to prove this result.

5 Entanglement-Assisted Capacity Theorem

We now state the entanglement-assisted classical capacity theorem. Section 6 proves the direct part of this theorem, and Section 7 proves its converse part.

Theorem 2 (Bennett–Shor–Smolin–Thapliyal)**.** *The entanglement-assisted classical capacity of a quantum channel is equal to the channel's mutual information:*

$$C_{\text{EA}}(\mathcal{N}) = I(\mathcal{N}), \tag{30}$$

where the mutual information $I(\mathcal{N})$ of a channel $\mathcal{N}$ is defined as $I(\mathcal{N}) \equiv \max_{\varphi_{AA'}} I(A;B)_\rho$, $\rho_{AB} \equiv \mathcal{N}_{A'\to B}(\varphi_{AA'})$, and $\varphi_{AA'}$ is a pure bipartite state.

6 The Direct Coding Theorem

The direct coding theorem is a statement of achievability:

Theorem 3 (Direct Coding)**.** *The following resource inequality corresponds to an achievable protocol for entanglement-assisted classical communication over a noisy quantum channel:*

$$\langle \mathcal{N} \rangle + H(A)_\rho \left[qq \right] \geq I(A;B)_\rho \left[c \to c \right], \tag{31}$$

where $\rho_{AB} \equiv \mathcal{N}_{A'\to B}(\varphi_{AA'})$.

We will not prove this here, but instead point to the book for a detailed proof.

7 The Converse Theorem

This section contains a proof of the converse part of the entanglement-assisted classical capacity theorem. Let us begin by supposing that Alice and Bob are trying to use the entanglement-assisted channel many times to accomplish the task of randomness distribution (recall that we took this approach for the converse of the classical capacity theorem in Section **??**). An upper bound on the rate at which Alice can distribute randomness to Bob also serves as an upper bound on the rate at which they can communicate because a noiseless classical channel can distribute randomness. In such a task, Alice and Bob share entanglement in some pure state $|\Phi\rangle_{T_A T_B}$ (note however that our proof below applies to any shared state). Alice first prepares the maximally correlated state $\overline{\Phi}_{MM'}$, and the rate of randomness in this state is $C - \delta \equiv \frac{1}{n} \log |M|$. Alice then applies some

encoding map $\mathcal{E}_{M'T_A\to A^n}$ to the classical system M' and her share T_A of the shared entanglement. The resulting state is

$$\mathcal{E}_{M'T_A\to A^n}(\overline{\Phi}_{MM'}\otimes\Phi_{T_AT_B}). \tag{32}$$

She sends her A^n systems through many uses $\mathcal{N}_{A^n\to B^n}$ of the channel $\mathcal{N}_{A\to B}$, and Bob receives the systems B^n, producing the state

$$\omega_{MT_BB^n}\equiv\mathcal{N}_{A^n\to B^n}(\mathcal{E}_{M'T_A\to A^n}(\overline{\Phi}_{MM'}\otimes\Phi_{T_AT_B})). \tag{33}$$

Finally, Bob performs some decoding map $\mathcal{D}_{B^nT_B\to\hat{M}}$ on the above state to give

$$\omega'_{M\hat{M}}\equiv\mathcal{D}_{B^nT_B\to\hat{M}}(\omega_{MT_BB^n}). \tag{34}$$

If the protocol is ε-good for randomness distribution, then the actual state $\omega'_{M\hat{M}}$ resulting from the protocol should be ε-close in trace distance to the ideal shared randomness state:

$$\left\|\omega'_{M\hat{M}}-\overline{\Phi}_{M\hat{M}}\right\|_1\leq\varepsilon. \tag{35}$$

We now show that the quantum mutual information of the channel serves as an upper bound on the rate C of any reliable protocol for entanglement-assisted randomness distribution (a protocol meeting the error criterion in (35))). Consider the following chain of inequalities:

$$\begin{aligned}
\log|M| &= I(M;\hat{M})_{\overline{\Phi}} && (36)\\
&\leq I(M;\hat{M})_{\omega'}+f(n,\varepsilon) && (37)\\
&\leq I(M;B^nT_B)_\omega+f(n,\varepsilon) && (38)\\
&= I(T_BM;B^n)_\omega+I(M;T_B)_\omega-I(B^n;T_B)_\omega+f(n,\varepsilon) && (39)\\
&= I(T_BM;B^n)_\omega-I(B^n;T_B)_\omega+f(n,\varepsilon) && (40)\\
&\leq I(T_BM;B^n)_\omega+f(n,\varepsilon) && (41)\\
&\leq \max_{\rho_{XAA'^n}} I(AX;B^n)_\rho+f(n,\varepsilon). && (42)
\end{aligned}$$

The first equality follows by evaluating the quantum mutual information of the shared randomness state $\overline{\Phi}_{M\hat{M}}$. The first inequality follows from the assumption that the protocol satisfies the error criterion in (35) and by applying the Alicki–Fannes inequality from Exercise **??** with $f(n,\varepsilon)\equiv 6\varepsilon\log|M|+4h_2(\varepsilon)$. This function has the property that $\lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}f(n,\varepsilon)=0$. The second inequality follows from quantum data processing (Theorem **??**)—Bob processes the state ω with the decoder $\mathcal{D}$ to get the state ω'. The second equality follows from the chain rule for quantum mutual information (see Exercise **??**). The third equality follows because the systems M and T_B are in a product state, so $I(M;T_B)_\omega=0$. The third inequality follows because $I(B^n;T_B)_\omega\geq 0$. Observe that the state $\omega_{MT_BB^n}$ is a classical–quantum state of the form

$$\rho_{XAB^n}\equiv\sum_x p_X(x)|x\rangle\langle x|_X\otimes\mathcal{N}_{A'^n\to B^n}(\rho^x_{AA'^n}), \tag{43}$$

where the classical system X in ρ_{XAB^n} plays the role of M in $\omega_{MT_BB^n}$ and the quantum system A in ρ_{XAB^n} plays the role of T_B in $\omega_{MT_BB^n}$. Then the final inequality follows because the quantum mutual information $I(T_BM;B^n)_\omega$ can never be greater than the maximum of $I(AX;B^n)_\rho$ over all input states of the form in (43).

We can strengthen this converse proof considerably. First, observe that the most general form of an encoding is an arbitrary CPTP map $\mathcal{E}_{M'T_A\to A^n}$ that acts on a classical register M' and a quantum register T_A. From Section **??**, we know that this map takes the following form:

$$\mathcal{E}_{M'T_A\to A^n}(\overline{\Phi}_{MM'}\otimes\Phi_{T_AT_B}) = \frac{1}{M}\sum_m |m\rangle\langle m|_M \otimes \mathcal{E}^m_{T_A\to A^n}(\Phi_{T_AT_B}), \tag{44}$$

where each $\mathcal{E}^m_{T_A\to A^n}$ is a CPTP map. This particular form follows because the first register M' on which the map $\mathcal{E}_{M'T_A\to A^n}$ acts is a classical register. Now, it would seem strange if performing a conditional noisy encoding $\mathcal{E}^m_{T_A\to A^n}$ for each message m could somehow improve performance. So, we would like to prove that conditional noisy encodings can never outperform conditional isometric (noiseless) encodings. In this vein, since Alice is in control of the encoder, we allow her to simulate the noisy encodings $\mathcal{E}^m_{T_A\to A^n}$ by acting with their isometric extensions $U^{\mathcal{E}_m}_{T_A\to A^nE'}$ and tracing out the environments E' (to which she has access). Then the value of the quantum mutual information $I(T_BM;B^n)_\omega$ is unchanged by this simulation. Now suppose instead that Alice performs a complete projective measurement of the environment of the encoding and she places the outcome of the measurement in some classical register L. Then the quantum mutual information can only increase, a result that follows from the quantum data-processing inequality:

$$I(T_BLM;B^n)_\omega \geq I(T_BM;B^n)_\omega. \tag{45}$$

Thus, isometric encodings are sufficient for achieving the entanglement-assisted classical capacity.

We can view this result in a less operational (and more purely mathematical) way as well. Consider a state of the form in (43). Suppose that each $\rho^x_{AA'^n}$ has a spectral decomposition

$$\rho^x_{AA'^n} = \sum_y p_{Y|X}(y|x)\psi^{x,y}_{AA'^n}, \tag{46}$$

where the states $\psi^{x,y}_{AA'^n}$ are pure. We can define the following augmented state:

$$\rho_{XYAB^n} \equiv \sum_{x,y} p_X(x)p_{Y|X}(y|x)|x\rangle\langle x|_X\otimes|y\rangle\langle y|_Y\otimes\mathcal{N}_{A'^n\to B^n}(\psi^{x,y}_{AA'^n}), \tag{47}$$

such that $\rho_{XAB^n} = \mathrm{Tr}_Y\{\rho_{XYAB^n}\}$. Then the quantum data-processing inequality implies that

$$I(AX;B^n)_\rho \leq I(AXY;B^n)_\rho. \tag{48}$$

By joining the classical Y register with the classical X register, the following equality holds:

$$\max_{\rho_{XAA'^n}} I(AX;B^n)_\rho = \max_{\sigma_{XAA'^n}} I(AX;B^n)_\sigma, \tag{49}$$

where

$$\sigma_{XAB^n} \equiv \sum_x p_X(x)|x\rangle\langle x|_X\otimes\mathcal{N}_{A'^n\to B^n}(\psi^x_{AA'^n}), \tag{50}$$

so that the maximization is over only pure states $\psi^x_{AA'^n}$. Then we know from the result of Exercise **??** that

$$\max_{\sigma_{XAA'^n}} I(AX;B^n)_\omega = \max_{\phi_{AA'^n}} I(A;B^n)_\omega, \tag{51}$$

where the maximization on the right-hand side is with respect to pure states $\phi_{AA'^n}$. Finally, from additivity of the quantum mutual information of a quantum channel (Theorem **??**) and an inductive argument similar to that in Corollary **??**, the following equality holds

$$\max_{\phi_{AA'^n}} I(A;B^n)_\omega = nI(\mathcal{N}). \tag{52}$$

Thus, the bound on the classical rate C of a reliable protocol for entanglement-assisted randomness distribution is

$$C - \delta = \frac{1}{n}\log|M| \leq I(\mathcal{N}) + \frac{1}{n}f(n,\varepsilon), \tag{53}$$

and it also serves as an upper bound for entanglement-assisted classical communication. Taking the limit as $n \to \infty$ and as $\varepsilon, \delta \to 0$ then establishes that an achievable rate C necessarily satisfies $C \leq I(\mathcal{N})$. This demonstrates a single-letter upper bound on the entanglement-assisted classical capacity of a quantum channel and completes the proof of Theorem 2.

7.1 Feedback Does Not Increase Capacity

The entanglement-assisted classical capacity formula is the closest formal analogy to Shannon's capacity formula for a classical channel. The mutual information $I(\mathcal{N})$ of a quantum channel $\mathcal{N}$ is the optimum of the quantum mutual information over all bipartite input states:

$$I(\mathcal{N}) = \max_{\phi_{AA'}} I(A;B), \tag{54}$$

and it is equal to the channel's entanglement-assisted classical capacity by Theorem 3. The mutual information $I(p_{Y|X})$ of a classical channel $p_{Y|X}$ is the optimum of the classical mutual information over all correlated inputs to the channel:

$$I(p_{Y|X}) = \max_{XX'} I(X;Y), \tag{55}$$

where XX' are correlated random variables with the distribution $p_{X,X'}(x,x') = p_X(x)\delta_{x,x'}$. The formula is equal to the classical capacity of a classical channel by Shannon's noisy coding theorem. Both formulas not only appear similar in form, but they also have the important property of being "single-letter," meaning that the above formulas are equal to the capacity (this was not the case for the Holevo information from the previous chapter).

We now consider another way in which the entanglement-assisted classical capacity is a good candidate for being the fully quantum generalization of Shannon's formula to the quantum world. Though it might be surprising, it is well known that free access to a classical feedback channel from receiver to sender does not increase the capacity of a classical channel. We state this result as the following theorem (without proof).

Theorem 4 (Feedback Does Not Increase Classical Capacity)**.** *The feedback capacity of a classical channel $p_{Y|X}(y|x)$ is equal to the mutual information of that channel:*

$$\sup\{C : C \text{ is achievable with feedback }\} = I(p_{Y|X}), \tag{56}$$

where $I(p_{Y|X})$ is defined in (55)*.*

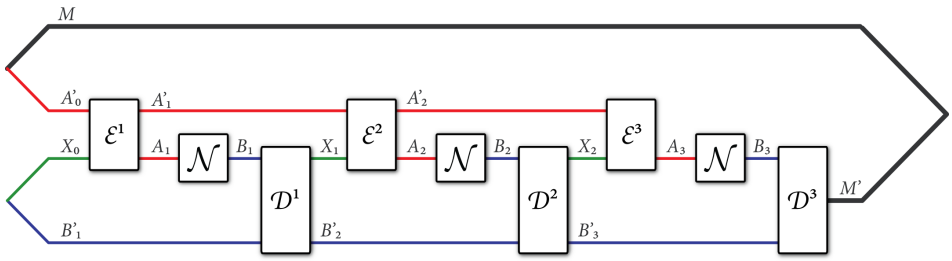

Figure 3: Three rounds of the most general protocol for classical communication with a quantum feedback channel.

Given the above result, we might wonder if a similar result could hold for the entanglement-assisted classical capacity. Such a result would more firmly place the entanglement-assisted classical capacity as a good generalization of Shannon's coding theorem. Indeed, the following theorem states that this result holds.

Theorem 5 (Quantum Feedback Does Not Increase the EAC Capacity)**.** *The classical capacity of a quantum channel assisted by a quantum feedback channel is equal to that channel's entanglement-assisted classical capacity:*

$$\sup\{C \mid C \text{ is achievable with quantum feedback}\} = I(\mathcal{N}), \tag{57}$$

where $I(\mathcal{N})$ *is defined in* (54)*.*

Proof. We define free access to a quantum feedback channel to mean that there is a noiseless quantum channel of arbitrarily large dimension going from the receiver Bob to the sender Alice. The bound LHS $\geq$ RHS follows because Bob can use the quantum feedback channel to establish an arbitrarily large amount of entanglement with Alice. They then just execute the protocol from Section 6 to achieve a rate equal to the entanglement-assisted classical capacity.

The bound LHS $\leq$ RHS is much less obvious, and it requires a proof that is different from the proof of Theorem 4. We first need to determine the most general protocol for classical communication with the assistance of a quantum feedback channel. Figure 3 depicts such a protocol. The protocol begins with Alice preparing a classical register M with a uniformly random message to be sent, which is correlated with some system A'_0. Bob uses the quantum feedback channel to send a quantum system X_0 to Alice, which is correlated with some quantum system B_0. Alice performs an encoding $\mathcal{E}^1_{A'_0 X_0 \to A'_1 A_1}$. Alice sends system A_1 through the first use of the channel $\mathcal{N}$. Bob now applies the decoding map $\mathcal{D}^1_{B_1 B'_1 \to X_1 B'_2}$. The next encoder of Alice occurs, and the procedure repeats. The last decoding map of Bob outputs a classical system M' which contains Bob's estimate of the message that Alice transmitted. The state of registers $MB_nB'_n$ after the nth channel $\mathcal{N}_{A_n \to B_n}$ has been applied has the following form:

$$\omega^{(n)}_{MB_nB'_n} \equiv \mathcal{N}_{A_n \to B_n}(\rho^{(n)}_{MB'_nA_n}), \tag{58}$$

where $\rho^{(n)}_{MB'_nA_n}$ is the state of registers MB'_nA_n after the nth encoding map has been applied. Let $\psi^{(n)}_{R^{(n)}MB'_nA_n}$ be a purification of $\rho^{(n)}_{MB'_nA_n}$, and let

$$\omega^{(n)}_{R^{(n)}MB_nB'_n} \equiv \mathcal{N}_{A_n \to B_n}(\psi^{(n)}_{R^{(n)}MB'_nA_n}). \tag{59}$$

This protocol is the most general for classical communication with quantum feedback. We can now proceed with proving the upper bound LHS $\leq$ RHS. To do so, we assume that the random variable M modeling Alice's message selection is a uniform random variable, and Bob obtains a random variable M' by measuring all of his systems B_n and B'_n at the end of the protocol. For any good protocol for classical communication, the bound $\Pr\{M' \neq M\} \leq \varepsilon$ applies. Consider the following chain of inequalities (these steps are essentially the same as those in (58)–(61))):

$$\begin{aligned}
\log|\mathcal{M}| &= H(M) && (60)\\
&= I(M;M') + H\left(M|M'\right) && (61)\\
&\leq I(M;M') + 1 + \varepsilon \log|\mathcal{M}| && (62)\\
&\leq I(M;B_nB'_n)_{\omega^{(n)}} + 1 + \varepsilon \log|\mathcal{M}|\,, && (63)
\end{aligned}$$

where the last mutual information is with respect to the state in (71). This chain of inequalities follows for the same reason as those in (58)–(61), with the last step following from quantum data processing. Continuing, we have

$$\begin{aligned}
I(M;B_nB'_n)_{\omega^{(n)}} &= I(M;B_n|B'_n)_{\omega^{(n)}} + I(M;B'_n)_{\omega^{(n)}} && (64)\\
&\leq I(MB'_n;B_n)_{\omega^{(n)}} + I(M;B'_n)_{\omega^{(n)}} && (65)\\
&\leq I(R^{(n)}MB'_n;B_n)_{\omega^{(n)}} + I(M;B'_n)_{\omega^{(n)}}. && (66)
\end{aligned}$$

The first equality is the chain rule for mutual information. The first inequality follows because $I(M;B_n|B'_n) = I(MB'_n;B_n) - I\left(B'_n;B_n\right) \leq I(MB'_n;B_n)$. The second inequality follows from quantum data processing. Now, given that the mutual information $I(R^{(n)}MB'_n;B_n)$ is with respect to the state in (71) and this state has the following form

$$\mathcal{N}_{A_n\to B_n}(\phi_{RA_n}), \tag{67}$$

where ϕ_{RA_n} is some pure state and R is some system not going into the channel (here identified with $R^{(n)}MB'_n$), we can optimize over all such inputs to find that

$$I(R^{(n)}MB'_n;B_n)_{\omega^{(n)}} \leq I(\mathcal{N}), \tag{68}$$

where $I(\mathcal{N})$ is the quantum mutual information of the channel. So this means that

$$\begin{aligned}
I(M;B_nB'_n)_{\omega^{(n)}} &\leq I(\mathcal{N}) + I(M;B'_n)_{\omega^{(n)}} && (69)\\
&\leq I(\mathcal{N}) + I(M;B_{n-1}B'_{n-1})_{\omega^{(n-1)}}. && (70)
\end{aligned}$$

where the last inequality follows from quantum data processing (the system B'_n results from applying the $n-1$ decoder to the systems $B_{n-1}B'_{n-1}$). At this point, we realize that the above chain of steps (77)–(83) can be applied to $I(M;B_{n-1}B'_{n-1})_{\omega^{(n-1)}}$, so we iterate this sequence until we go all the way back to the beginning of the protocol. Putting everything together, we get the following upper bound on any achievable rate C for classical communication with quantum feedback:

$$C - \delta \leq I(\mathcal{N}) + \frac{1}{n} + \frac{\varepsilon}{n}\log|\mathcal{M}|\,, \tag{71}$$

which becomes $C \leq I(\mathcal{N})$ as $n\to\infty$ and $\varepsilon,\delta\to 0$. □

Corollary 6. *The capacity of a quantum channel with unlimited entanglement and classical feedback is equal to the entanglement-assisted classical capacity of $\mathcal{N}$.*

Proof. This result follows because $I(\mathcal{N})$ is a lower bound on this capacity (simply by avoiding use of the classical feedback channel). Also, $I(\mathcal{N})$ is an upper bound on this capacity because the entanglement and classical feedback channel can simulate an arbitrarily large quantum feedback channel via teleportation, and the above theorem gives an upper bound of $I(\mathcal{N})$ for this setting. $\square$