Quantum Shannon Theory and Glimpses of a Resource Theory

Course project for CS682A - Quantum Computing

Sagnik Bhattacharya

Department of Electrical Engineering, IIT Kanpur

E-mail: sagnikb@iitk.ac.in

Advisor - Prof Dr. Rajat Mittal

Department of Computer Science and Engineering, IIT Kanpur

E-mail: rmittal@iitk.ac.in

Abstract. This report focuses on Quantum Shannon Theory, which extends classical information theory to the quantum domain, and the resource theory of quantum information which gives a powerful way to talk about quantum communication protocols. It also considers formalising the description of quantum channels, and simple communication protocols that can be analysed using arguments that are also applicable to more general situations. It uses resource theoretic arguments to prove the optimality of these protocols. It also discusses typicality, which is an important concept in information theory.

Contents

1. Introduction and Overview

We first go over the extension of noiseless quantum theory to the noisy quantum theory via density matrices. After going over some further generalisations of the concept of density matrices, we reach the concept of a quantum channel, that allows us to formalise evolutions, density matrices etc under one unified representation. Along with this, we cover the Choi-Kraus Theorem, that gives necessary and sufficient conditions for something to be a quantum channel. We then state the purification lemma that allows us to consider noisy channels within the framework of the noiseless theory.

We then see the most basic communication protocols, the unit quantum protocols. Here, we show that the various protocols covered in the course (eg teleportation) are optimal. The first optimality proofs have some motivated but unproven assumptions, but in a first encounter with a 'Shannon-like' proof involving a direct coding theorem and a converse coding theorem, we see we can refine the proofs, thus demonstrating the power of such an argument.

We then introduce classical and quantum typicality, that form the basis for Shannon's coding theorems. We also state two lemmas, called the packing lemma and

the covering lemma that are useful in proofs regarding the optimality of communication protocols.

2. Noisy Quantum Theory and Density Matrices

2.1. Motivation and Definition

An unsaid assumption in an introduction to quantum theory is that all states and unitary operations can be perfectly prepared. However, we know that this is never possible in practice. So, for completeness of the theory, we are motivated to consider the effect of noise in our theory, and to come up with a language to describe this noise. We will see later that this will be useful when we consider the general framework of a quantum channel.

The most useful language to talk about the noisy quantum theory is the language of density matrices. Suppose we have an ensemble of quantum states, which are distributed according to some probability distribution p. That is, the classical probability that we obtain the state $|\psi_x\rangle$ when we 'pick-up' any state from this ensemble is distributed according to some random variable $\mathcal{X}$. Then, the density matrix corresponding to this ensemble is given by

$$\rho \coloneqq \sum_{x \in \mathcal{X}} p_X(x) \, |\psi_x\rangle \langle\psi_x|$$

Of course, the following holds

$$\sum_{x \in \mathcal{X}} p_X(x) = 1$$

Note that this strictly classical probability is different from the probabilistic collapse into basis states given by the measurement postulate.

An example should make the concept of density matrices clear. Suppose we have a bag in which we put 25 each of qubits $|0\rangle$, $|1\rangle$, $|+\rangle$ and $|-\rangle$. Then the density matrix corresponding to such an ensemble is given by

$$\rho = \frac{1}{4} |0\rangle \langle 0| + \frac{1}{4} |1\rangle \langle 1| + \frac{1}{4} |+\rangle \langle +| + \frac{1}{4} |-\rangle \langle -|$$

We generalise our notion of 'state' to mean anything that can be described by a density matrix. Our original definition of state as a unit vector in the Hilbert space is called a 'pure state' in this language. We define the purity of a density matrix to be

$$P(\rho) = Tr\{\rho^2\}$$

It can be easily shown that the purity of a pure state is exactly one, while for 'impure' states, the purity is strictly less than one.

Also, we should note here that there is only one density matrix corresponding to an ensemble, but the converse is not true. However, with each density matrix ρ, we associate a canonical ensemble given by its singular value decomposition, which is unique.

2.2. Properties of Density Matrices

The following are some useful properties of density matrices. These are not very difficult to prove starting from the definitions and known results from the noiseless theory.

- **Unit trace:** $Tr\{\rho\} = 1$.
- **Hermiticity:** ρ is Hermitian.
- **Positive-semidefiniteness:** ρ is positive semi-definite, because the eigenvalues are related to probabilities, that must be real and non-negative.
- **Evolution:** Given a unitary U the evolved density matrix is $\rho' = U\rho U^\dagger$
- **Measurement:** Given a projection operator Π_i the post-measurement state is given by

$$\frac{\Pi_i \rho \Pi_i}{Tr\{\rho \Pi_i\}}$$

In fact, we can define a density matrix as any positive semi-definite matrix that has unit trace. Also, since each of these properties is basis invariant, we see that we can define the density matrix of an ensemble without reference to a particular basis.

2.3. Further Definitions

Now we look at further extensions of the concept of density matrices that let us consider more general states.

2.3.1. Ensemble of Ensembles An ensemble of density matrices. It has a density matrix representation given by

$$\rho = \sum_{i \in \mathcal{X}} p_X(x) \rho_x$$

Note that we can still describe this in terms of density matrices, showing some of the power of the density matrix representation.

2.3.2. Composite States For multiple Hilbert spaces, simple tensor product analogues to the results discussed also hold. The joint density matrix for two independent systems with density matrices ρ and σ respectively is given by $\rho \otimes \sigma$.

Definitions of product states, separable states and entangled states are similar to those used for the composite noiseless case. For example, a density matrix for the composite of two Hilbert spaces A and B is separable if it can be written in the form $\rho_{AB} = \rho_A \otimes \rho_B$ and entangled if it cannot.

A partial trace is also defined, to consider a subset of multiple Hilbert spaces. It is defined (in the case of two composite Hilbert spaces) as

$$Tr_B\{\rho_{AB}\} = \rho_A$$

which 'traces-out' the Hilbert space B.

3. Quantum Channels

We finally reach the definition of quantum channels which allows us to have an axiomatic theory of quantum evolutions, and a characterisation of the same in terms of the Choi-Kraus theorem. First, we need some definitions.

- Let $\mathcal{H}$ be a Hilbert space.
- $\mathcal{L}(\mathcal{H})$ is the space of all linear operators acting on $\mathcal{H}$.
- $\mathcal{D}(\mathcal{H})$ is the space of all density operators acting on $\mathcal{H}$.
- $\mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$ is the space of all linear operators from $\mathcal{H}_A$ to $\mathcal{H}_B$.
- $\mathcal{N}$ is a convex linear map that takes density operators in $\mathcal{H}_A$ to density operators in $\mathcal{H}_B$.
- We can uniquely extend such a map to all linear operators from $\mathcal{H}_A$ to $\mathcal{H}_B$, and we call this extended map a **quantum channel**.

On physical grounds, we postulate that a quantum channel must have certain properties.

- **The channel must be linear**. This follows from the fact that quantum mechanics without measurement is a linear theory.
- **The channel must be completely positive.** $I_R \otimes \mathcal{M}$ preserves positive semi-definite operators on $\mathcal{L}(\mathcal{H})$ for a reference system R of arbitrary size. This is to allow for preservation of semi-definiteness even when the Hilbert space is part of a bi-partite state with another Hilbert space.
- **The channel must preserve trace.** Here we impose the reasonable condition that the channel must take density matrices to density matrices.

3.1. The Choi-Kraus Theorem

Now we can state the Choi-Kraus theorem. A map in $\mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$ is a quantum channel iff it has a Choi Kraus Decomposition given by

$$\mathcal{N}(X_A) = \sum_{l=0}^{d-1} V_l X_A V_l^\dagger$$

where

$$X_A \in \mathcal{L}(\mathcal{H}_A)$$
$$V_l \in \mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$$
$$d \leq dim(H_A) dim(\mathcal{H}_B)$$
$$\sum_{l=0}^{d-1} V^\dagger V_l = I_A$$

One can probably see that using this definition only, it is difficult to check if a given linear map is a quantum channel. This task is made much easier with the introduction of the Choi operator corresponding to a given linear map.

3.1.1. The Choi operator Suppose we are given two isomorphic Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_R$ and another Hilbert space $\mathcal{H}_B$. Let $\{|i\rangle_R\}$ and $\{|i\rangle_A\}$ be orthonormal basis for $\mathcal{H}_A$ and $\mathcal{H}_R$ respectively. Let $\mathcal{N}_{\mathcal{A}\to\mathcal{B}}$ be the given linear map from $\mathcal{L}(\mathcal{H}_A)$ to $\mathcal{L}(\mathcal{H}_B)$. The Choi operator for the given map and bases is given by

$$(id_R \otimes \mathcal{N}_{\mathcal{A}\to\mathcal{B}})(|\Gamma\rangle\langle\Gamma|_{RA}) = \sum_{i,j=0}^{dim(\mathcal{H}_\mathcal{A})-1} |i\rangle\langle j|_R \otimes \mathcal{N}_{\mathcal{A}\to\mathcal{B}} |i\rangle\langle j|_A$$

where

$$|\Gamma\rangle_{RA} = \sum_{i=0}^{dim(\mathcal{H}_\mathcal{A})-1} |i\rangle_R \otimes |i\rangle_A$$

That is, a maximally entangled state.

The Choi-operator is positive semi-definite if and only if $\mathcal{N}_{\mathcal{A}\to\mathcal{B}}$ is completely positive, so this gives an easy test to see if a given map is a quantum channel.

The quantum channel gives a powerful way to talk about quantum evolution because all of the following turn out to be special cases of quantum channels.

- state preparation (a classical to quantum channel)
- state appending
- discarding channels
- unitaries (a quantum to quantum channel)
- density matrices
- measurement channel (a quantum to classical channel)

4. Purification[1]

As mentioned in the introduction, the idea of purification allows us to consider noisy quantum theory within the theory of the noiseless theory. Essentially it can be interpreted as saying that we can interpret any noise in the system as arising due to entanglement with the 'environment' - a system we cannnot control/access.

A purification for a density matrix $\rho_A \in \mathcal{D}(\mathcal{H}_A)$ is a **pure bipartite state** $|\psi\rangle_{RA} \in \mathcal{H}_R \otimes \mathcal{H}_A$ with a reference system R which has the property that if we 'trace-out' the reference system, we re-obtain ρ_A. That is,

$$Tr\{|\psi\rangle\langle\psi|_{RA}\} = \rho_A$$

We can show that given a spectral decomposition for ρ_A

$$\rho_A = \sum_x p(x) |x\rangle\langle x|$$

a purification is given by

$$|\psi\rangle_{RA} = \sum_x \sqrt{p(x)} |x\rangle_R |x\rangle_A$$

Also, we can show that all purifications are unitarily related, so we can only work with one of them. One particularly convenient purification is the *canonical purification*:

$$|\psi\rangle_{RA} = (I_R \otimes \sqrt{\rho_A})\,|\Gamma\rangle_{RA}$$

where $|\Gamma\rangle_{RA}$ is as defined in the definition of the Choi operator.

5. Unit Communication Protocols

These are the simplest quantum communication protocols, and therefore admit an explanation in very simple terms. Here, we introduce non-local unit resources, introduce four simple quantum communication protocols, introduce the language of resource theory and finally prove the optimality of the protocols. Also, here we assume that all the protocols and states are noiseless. This is motivated by the aforementioned purification theorem that gives a way to deal with noisy quantum communication in the framework of the noiseless theory.

5.1. Non-local unit resources

We define a resource as **non-local** if it can be used at two points that are spatially separated, or which can be used to communicate between two such points.

We define a resource as **unit** if we can quantify it in terms of units like classical bits, qubits or entangled bits.

A resource which is both non-local and unit is called a non-local unit resource. When faced with communication tasks, we usually consider these resources to be the most expensive and (similar to the query model for quantum computation) we characterise the efficiency of a protocol by the number of times it uses such a resource.

As an example, consider any map of the form $|i\rangle_A \to |i\rangle_B$. This is a noiseless qubit channel, and is a **non-local unit resource**.

The non-local unit resources which we consider are the following:

- **Noiseless qubit channel:** We have already seen that this is of the form $|i\rangle_A \to |i\rangle_B$. It is represented in the resource theory as $[q \to q]$.
- **Classical Information Channel:** We define it in terms of density matrices, and represent it as $[c \to c]$

  $$\begin{cases} |i\rangle\langle i|_A \to |i\rangle\langle i|_B & \text{for i = j} \\ |i\rangle\langle j|_A \to 0 & \text{for i} \neq \text{j} \end{cases}$$

  We can see that a classical channel throws away superpositions of states. In other words, a classical channel cannot maintain arbitrary superpositions of quantum states.
- **Entanglement:** Quantified by an entangled bit, or ebit, defined as

  $$\left|\Phi^+\right\rangle_{AB} \coloneqq \frac{1}{\sqrt{2}}\left(|00\rangle_{AB} + |11\rangle_{AB}\right)$$

We represent it as $[qq]$. This looks very similar to a Bell pair, but there is a distinction that must be made clear. We call a Bell pair an entangled bit if and only if one qubit of the pair is spatially separated from the other. Without this, the Bell pair is not a non-local resource.

5.2. Unit quantum protocols

Before we move on to resource inequalities, we consider four unit quantum communication protocols.

5.2.1. Entanglement Distribution

- **Preparing a Bell state** Alice creates two qubits in the state $|0\rangle_A |0\rangle_{A'}$. She applies Hadamard to the first qubit to get the state $\frac{1}{\sqrt{2}} (|0\rangle_A + |1\rangle_A) |0\rangle_{A'}$. She then performs CNOT with A as control and A' as target to get $\frac{1}{\sqrt{2}} (|00\rangle_{AA'} + |11\rangle_{AA'})$ which is a Bell state.
- **Creating an entangled bit** She uses the noiseless qubit channel to send the qubit A' to Bob, thereby creating the state $\frac{1}{\sqrt{2}} (|00\rangle_{AB} + |11\rangle_{AB})$.

Re-emphasising a point already made earlier, note that we have only called it an ebit after the transfer takes place.

5.2.2. Elementary Coding This is a really simple protocol that uses a quantum channel to achieve classical communication.

- Alice has a classical bit which she wants to transmit. Depending on what the bit is, she creates a $|0\rangle$ or a $|1\rangle$ qubit. Note that this does not have to be a $|0\rangle$ or $|1\rangle$, any pair of orthonormal states will work.
- She uses a noiseless qubit channel to transfer this to Bob.
- Bob measures it to find which qubit was sent, and can then recover the classical bit Alice started with.

5.2.3. Super-dense coding This gives a way to transmit two classical bit using one use of the quantum channel.

- Alice and Bob share an entangled bit.
- Alice applies one of $\{I, X, Z, XZ\}$ to her share.
- She transmits her share to Bob using a quantum channel.
- Bob measures in $\{|0\rangle, |1\rangle\}$ basis, and recovers which unitary Alice applied. Note that this can be done because the possible states that may be obtained in this protocol are all orthogonal.

Since the code-space consists of 4 codes, two bits of classical information can be transmitted.

5.2.4. Teleportation This requires two uses of a noiseless classical channel to transmit a quantum state. This has been covered in the course, so we skip the protocol.

5.3. Resource Inequalities

First, consider the following argument. We know that a classical channel cannot maintain arbitrary superpositions of quantum states. Therefore, a classical channel *cannot* simulate a quantum channel. However, the elementary coding protocol gives a way of accomplishing the opposite task, ie, using a quantum channel to simulate a classical channel.

Hence, we can make the statement that a quantum channel is more powerful than a classical channel, and write the inequality

$$[q \to q] \geq [c \to c]$$

This is an example of a resource inequality. Written in this form, it means that *there exists a protocol that uses the resource on the right to generate the resource on the left.* We can write similar resource inequalities for the other protocols.

- **Entanglement generation** $[q \to q] \geq [qq]$
- **Super-dense coding** $[qq] + [q \to q] \geq 2\,[c \to c]$
- **Teleportation** $[qq] + 2\,[c \to c] \geq [q \to q]$

We have a protocol that uses a quantum channel to generate entanglement. Can we do the opposite? The answer is no. Suppose such a task were possible. Since the measurement of one pair of an entangled bit instantaneously changes the state of the other pair, we would have then been able to communicate superluminally between two spatially separated points, which is not allowed by the General Theory of Relativity. Therefore, we can say that the quantum channel is in this sense 'stronger' than the other two channels.

5.4. Optimality

A natural question that arises now is whether the protocols that we have seen are optimal. Let us consider this question for each protocol in turn.

5.4.1. Elementary Coding We ask if the following resource inequality and the corresponding protocol is possible, with $E > 1$.

$$[q \to q] \geq E\,[c \to c]$$

The Holevo bound states that it is not possible to transmit more than n bits of classical information by transmitting n qubits and maintain perfect recoverability. Hence, this is not possible.

5.4.2. Entanglement Generation We ask if the following resource inequality is possible, with $E > 1$.

$$[q \to q] \geq E\,[q \to q]$$

The answer again is no. The proof is by contradiction. Let's assume we can have such a protocol. We can combine it with the teleportation protocol and assume free classical communication to achieve the following resource inequality:

$$[q \to q] \geq E\,[q \to q]$$

We can then keep repeating this and achieve unbounded amount of quantum communication, which is impossible.

If we do not assume free classical communication we get an extra $E\,[c \to c]$ on the LHS. The argument for ignoring that factor is the following. Since we are dealing with a noiseless protocol, we would want our qubits to be transmitted perfectly. Now, to perfectly locate a qubit on the Bloch sphere, we need infinite classical information for the two angles on the sphere. So, quantum communication is much more expensive than classical communication and hence can be neglected.

5.4.3. Super-dense coding Can we have the resource inequality

$$[q \to q] + [qq] \geq 2C\,[c \to c]$$

with $E > 1$? The answer is again no.

- Suppose we have such a protocol and infinite entanglement available.
- Then we can have the resource inequality

  $$[q \to q] + \infty\,[qq] \geq 2E\,[c \to c] + \infty\,[qq]$$

  because the protocol uses a finite amount of entanglement.
- We can chain it with teleportation to obtain

  $$2E\,[c \to c] + \infty\,[qq] \geq E\,[q \to q] + \infty\,[qq]$$

- Therefore, we have a protocol that can be described by

  $$[q \to q] + \infty\,[qq] \geq E\,[q \to q] + \infty\,[qq]$$

  We get this inequality by noticing that the RHS of the first inequality and the LHS of the previous inequality are the same.
- We can repeat this to obtain

  $$[q \to q] + \infty\,[qq] \geq E^k\,[q \to q] + \infty\,[qq]$$

  Where k can be any finite number. Thus we end up generating unbounded quantum communication using an infinite source of entanglement. However we know that entanglement only cannot give us a quantum channel, so we reach a contradiction.

The argument for using infinite entanglement in the proof is that since we have shown that the protocol does not work even after assuming infinite entanglement, it also will not work if we assume entanglement is finite.

5.5. Better Optimality Proof

In the proofs we have already seen, we have assumptions like free classical communication or availability of infinite entanglement. The question that arises now is, can we do better? Yes, we can, and the proof follows.

We define a 3-dimensional space, the axes of which are labelled C, Q and E. Each point in the space corresponds to a protocol involving unit resources. If a protocol generates a resource X, then the sign of X in the coordinates of the point corresponding to that protocol is positive, and vice versa.

For example, Superdense coding corresponds to the point $x_{SD} \coloneqq (2, -1, -1)$. Entanglement generation corresponds to the point $x_{EG} \coloneqq (0, -1, 1)$. Similarly for teleportation, $x_T \coloneqq (-2, 1, -1)$.

5.5.1. Unit Resource Achievable Region, $\widetilde{C_U}$ Next, we define the Unit Resource Achievable Region, $\widetilde{C_U}$. These are all the points obtained by linear combinations of the different protocols.

The space in the C, Q, E space corresponding to $\widetilde{C_U}$ obeys the following equations:

$$C + Q + E \leq 0 \tag{1}$$
$$Q + E \leq 0 \tag{2}$$
$$C + 2Q \leq 0 \tag{3}$$

This can be shown by noting that

$$\begin{bmatrix} C \\ Q \\ E \end{bmatrix} = \begin{bmatrix} -2 & 2 & 0 \\ 1 & -1 & -1 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix}$$

gives all achievable triples along with the fact that the α, β, γ cannot be negative (since we cannot repeat a protocol a negative number of times. Here, α, β, γ can be any real number. To see this, note that we can repeat each protocol multiple times to get to points that have α, β, γ as integers. We can then extend the definition to rational numbers, and since every real number can be approximated by a rational number with as high an accuracy as required, we can extend it to real numbers as well.

5.5.2. Unit Resource Capacity region C_U This is the set of all points in the C, Q, E space that have corresponding protocols that can implement them. That is, all points in this region satisfy the following resource inequality:

$$0 \geq C\,[c \to c] + Q\,[q \to q] + E\,[qq]$$

5.5.3. C_U and $\widetilde{C_U}$ are equal We have a theorem that states that the above two regions are equal, that is, we can show that

$$\widetilde{C_U} = C_U$$

The direct coding theorem, which states that $\widetilde{C_U} \subseteq C_U$ follows easily from the definitions.

The converse coding theorem is more difficult. The proof uses the following lemmas, which we have argued to be true at various points.

- It is impossible to get classical or quantum communication from entanglement alone.
- Only classical communication cannot generate quantum communication or entanglement.
- Holevo bound, ie, we cannot transmit more than one bit of classical information by transmitting one qubit only.

Using these, we can consider each octant in the C, Q, E space and show that the above assumptions and the definition for the point belonging in C_U to show that it also lies in $\widetilde{C_U}$.

6. Classical and Quantum Typicality

6.1. Classical Typicality

We first consider classical typicality. We define an i.i.d. (independent and identically distributed) sequence obtained by sampling from a probability distribution to be *typical* if the entropy of the sequence (called the sample entropy) is close to the entropy of the probability distribution. This is called *weak typicality*. We refer to the probability distribution as the **classical information source.**

This is formalised in the following definitions:

Given a sequence x^n drawn from a probability distribution $p_X(x)$, its **sample entropy** is defined as:

$$\bar{H}(x^n) := -\frac{1}{n} log(p_{X^n}(x^n))$$

where

$$p_{X^n}(x^n) = \prod_{i=1}^{n} p_X(x_i)$$

We formalise the idea of 'sufficiently close' in the following definition of a **typical sequence**. A sequence is called δ-typical if its sample entropy $\bar{H}(x^n)$ is δ close to the entropy $H(X)$ of the source random variable X, and the set of all such sequences is called the **typical set**.

This idea is extremely useful in the information theory of Shannon. Let us first see some properties of this set.

- **Unit probability** As n becomes large, it becomes more and more likely that the sequence is a typical sequence ie the typical set has asymptotic probability equal to 1.
- **Exponentially smaller cardinality** The cardinality is exponentially smaller than the cardinality of the set of all sequences, for all random variables excluding the uniform random variable.

- **Asymptotic Equipartition Property** Each element of the set is approximately equally probable in the asymptotic limit.

The proofs can be found in any book on Information Theory. I refer to [3].

6.2. Quantum Typicality

For the quantum case, we first need to define a **quantum information source** as a device that randomly emits pure states from a Hilbert space $\mathcal{H}_A$ of finite dimension. We can write it as a density matrix.

$$\rho_A \coloneqq \sum_y p_Y(y) \left|\psi_y\right\rangle \left\langle\psi_y\right|_A$$

where the states $|\psi_y\rangle$ need not be orthonormal. Now, we can write a spectral decomposition of the density matrix, as

$$\rho_A = \sum_x p_X(x) \left|\psi_x\right\rangle \left\langle\psi_x\right|_A$$

Now, the $|\psi_x\rangle$ do form an orthonormal basis and $p(x)$ is still a probability distribution.

We define the quantum entropy to be

$$H(A)_\rho \coloneqq -Tr\{\rho_A log \rho_A\}$$

and this can easily be shown to equal the classical Shannon entropy of the probability distribution $p(x)$.

Now we can extend the ideas of classical typicality to the quantum case. The sequence of states can be thought of as being described by a density matrix that is the tensor product of n copies of the density matrix that we have already seen. The Hilbert space of the sequence is then denoted as $\mathcal{H}_{A^n}$. It is spanned by the states $|x^n\rangle_{A^n}$ which can be thought of as particular realisations of the density matrix realisation. The quantum **typical subspace** is then the space spanned by all the $|x^n\rangle_{A^n}$ such that the corresponding classical sequences are δ-typical as per the definition already seen. A projector onto this typical subspace is called the **typical projector**.

The distinguishability problem of quantum states gives an important difference between the classical typical set and the quantum typical subset. If the quantum information source outputs states that are orthogonal (and hence perfectly distinguishable) and the dimensionalities of the support of the classical random variable and that of the Hilbert space are equal, then the sizes of the typical set and the typical subset will be equal. However, if that is not the case, then the quantum typical subspace will have exponentially smaller cardinality as compared to the classical typical subset.

The same properties hold for the typical subspace, namely

- unit probability in the asymptotic limit
- exponentially smaller cardinality compared to the full Hilbert space $\mathcal{H}_{A^n}$
- asymptotic equipartition

7. The packing lemma

The packing lemma arises when considering the problem of encoding classical information into a Hilbert space while maintaining distinguishability of the encoded messages. It is important in several quantum encoding and communication protocols.

We have encountered this problem of encoding classical information in a Hilbert space before when we considered the optimality of the elementary coding protocol and said that there is a limit to the amount we can encode (the statement of the Holevo bound).

To get to the packing lemma we need to first define an **ensemble**. Let $\mathcal{X}$ be a set of cardinality $|\mathcal{X}|$. Suppose the elements of the set are x. Let X be a random variable with probability density function $p_X(x)$. Now, suppose that for each x there is an associated quantum state σ_x such that $\sigma_x \in \mathcal{D}(\mathcal{H})$ (the set of all density operators in the Hilbert space $\mathcal{H}$. Then, the expected density operator corresponding to this ensemble is given by

$$\sigma = \sum_{x \in \mathcal{X}} p_X(x)\sigma_x$$

We can pick a subset $\mathcal{C}$ of $\mathcal{X}$ and consider each element of the subset x to correspond to a message belonging to a message set $\mathcal{M}$. So, if Alice wants to send a message $m \in \mathcal{M}$ to Bob, she chooses the corresponding σ_x and sends the same to Bob. Also, we assume that she chooses a message from the set uniformly at random. Now we want that Bob should be able to distinguish these states. The packing lemma gives conditions for this distinguishability.

Let there be given an ensemble as defined above, projectors Π and $\{\Pi_x\}_{x \in X}$ that are projection operators for subspaces of $\mathcal{H}$, and these projectors satisfy the constraints

$$Tr\{\Pi\sigma_x\} \geq 1 - \epsilon$$
$$Tr\{\Pi_x\sigma_x\} \geq 1 - \epsilon$$
$$Tr\{\Pi_x\} \leq d$$
$$\Pi\sigma\Pi \leq \frac{\Pi}{D}$$

where $0 \leq \epsilon \leq 1$, $D > 0$ and $0 < d < D$. Let $\mathcal{M}$ be a set of size $|\mathcal{M}|$ with elements m. Now, we generate a random code $\mathcal{C} = \{C_m\}_{m \in \mathcal{M}}$ like this : each random variable C_m takes a value in $\mathcal{X}$ and corresponds to message m, but with a distribution that depends only on $p_X(x)$ and not on m. Then, the packing lemma gives a construction of a POVM $\{\Lambda_m\}_{m \in \mathcal{M}}$ such that the expected (with respect to $\mathcal{C}$) average probability (with respect to $\mathcal{M}$) of getting the correct state (for Bob) is high. The bound is that it is greater than

$$1 - 2(\epsilon + 2\sqrt{\epsilon}) - 4|\mathcal{M}|\frac{d}{D}$$

Here, $\frac{d}{D}$ is small, $|\mathcal{M}| \ll \frac{D}{d}$ and ϵ is small.

This lemma is used as a very crucial part of the Schumacher Compression protocol.

8. Future Work

I would like to cover more of the paper [4]. Also, I recently came across some very interesting work that uses category theory to unify and generalise many of the results that I have encountered. In particular, the whole subject of quantum evolution can be seen as a special case of symmetric monoidal categories and one can also find general frameworks that all resource theories must satisfy - work that was partly influenced by the success of the resource theory of quantum communication. I would like very much to learn about these areas in future.

What I learned from the Project

I originally wanted to read the whole of Devetak, Harrow and Winter's paper [4]. It turned out that to understand the later parts of the paper, knowledge of quantum Shannon theory is essential, and to understand that the knowledge of both the density matrix formalism of quantum mechanics and classical information theory is necessary. To learn about these topics I referred to [1], [2] and [3]. Doing this project has given me a good idea of both classical and quantum information theory, and also introduced me to a lot of the density matrix formalism.

Acknowledgements

I am grateful to Prof Dr. Rajat Mittal for giving me this opportunity to work on this project and for making the fundamentals of noiseless quantum theory clear in the course so that I could easily extend my knowledge into the other areas. I am also grateful to him and all my fellow students in the course for attending both the mid-term and end-term presentations, and asking great questions that increased my understanding of the material. Also, I would like to acknowledge the efforts of my friend Prannay Khosla (Deptt. of Computer Science and Engineering, IIT Kanpur), for being a dummy audience for both my mid-term and end-term presentations and for this report. He asked great questions that clarified several points in all of the above and bettered my understanding too. Acknowledgements also to Shruti Joshi and Homanaga Bharadhwaj (Deptt. of Electrical Engineering, IIT Kanpur) for going through this document and pointing out errors.

References

[1] Nielsen, Michael A. and Chuang, Isaac L., Quantum Computation and Quantum Information: 10th Anniversary Edition, 2011, Cambridge University Press

[2] Wilde, Mark M., From Classical to Quantum Shannon Theory, arXiv:1106.1445 [quant-ph]

[3] Cover, Thomas M. and Thomas, Joy A., Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing), 2006

[4] I. Devetak, A.W. Harrow, A. Winter, A Resource Framework for Quantum Shannon Theory, arXiv:quant-ph/0512015