INFORMATION TRANSMISSION THROUGH A QUANTUM INFORMATION CHANNEL

Thesis Submitted To

The Department of Electrical and Electronic Engineering,
BRAC University
in partial fulfillment of the requirements of the award of the degree of
Bachelor of Science in Electrical and Electronic Engineering

By

Tareq Ahmed

Department of Electrical and Electronic Engineering
BRAC University
December, 2017

DECLARATION

I hereby declare that the thesis titled Information transmission through a Quantum information channel submitted to Department of Electrical and Electronic Engineering, BRAC UNIVERSITY for the fulfillment of degree in Bachelors of Science Electrical and Electronic Engineering. Any information used from other sources has been acknowledged in the reference section.

____

Candidate
Tareq ahmed
ID:15221011

____

Certified
Dr. Mahbub Alam Majumdar
Professor
Department of Computer Science and Engineering
Brac University

Abstract

We consider the transmission of classical information over a quantum channel. The channel is expressed by an alphabet of quantum states. For an example, we can express quantum states by photon polarization. For transmitting information, we use specified set of probabilities. If we find that the receiver is unable to make separate measurement on the received letter then we have to use Holevo theorem. From this theorem we see that, in such case, the amount of information per letter we are sending cannot be larger than the Von Neumann entropy H of the letter ensemble. It happens most of the time that, the actual amount of information which will be transmitted is less than H. However if we use block coding scheme which has options to choose code words that respects the priori probabilities of the letter states then we find a different situation. In this case the receiver distinguishes whole words rather than individual letters. In this way, the information transmitted per letter can be made arbitrarily close to H. Block coding scheme helps us to find clear information of theoretical interpretation of Von Neumann entropy in quantum mechanics. We use this experiment in superdense coding and we consider this extension to noisy channels.

Acknowledgement

Firstly, I express my deepest thank to my supervisor Dr. Mahbub Alam Majumdar for giving me a chance to work with him. Before I start my thesis I have less knowledge about my topic. As I started my work I found great interest in this topic. Without his suggestion it maynot be possible for me to know about this interesting topic. Specially, the new term like unitary operator, Von Neumann entropy theorem, Holevo bound, quantum entanglement theory, Bell states have increased my interest in this topic.

Secondly, I also thank Muhammad Lutfor Rahman and Mohammad Mosaddidur Rahman. They both are the faculties of MNS Department of BRAC UNIVERSITY. Specially Lutfor sir gave me many valuable advices and suggested me to read some books which are related to my thesis topic. Really without his help it was really very difficult for me to complete my thesis work.

Finally, I am again grateful to my supervisor for giving me opportunity to work with this interesting topic and also his cooperation to complete my thesis. I wish sir will keep me with him in his further research activities.

Contents

Chapter 1

Introduction

Noisy channel coding theorem is the main outcome of Shannons classical theory of information. From this research, a high reliable process came out which is measuring the capacity of a noisy channel. From this measurement, it can be easily determined what will be the maximum rate of classical information which can transmit reliably through the channel. A quantum source can emit unknown quantum state. Our purpose is to convey this state through the channel to some receivers. The main obstacle we face in this case is that, the channel has some noises. The noise prevents transmission of information through the channel properly. Our main target is to increase the efficiency of transmitting information over a given quantum channel. We can recognize a channel for transmitting a signal reliably on this condition if the sequence of block coding and block decoding system can be found that acquire perfect fidelity within the limit of large block size. A new concept has arised for a perfect quantum information channel which is channel capacity theorem. From this theorem we can get an effective procedure which explicit algorithm to evaluate the channel capacity. There are two parts in this theorem, one part creates an upper bound for transmitting information properly and the other part gives the instruction for coding and decoding scheme to attain that bound.

This paper has two main sections. In first sections, we showed that increasing of efficiency of the general process for proving upper founds on the capacity of a noisy quantum channel. These techniques are applied in various different classes of quantum noisy channel problem. In other section, we showed some new innovations that quantum mechanics introduces into the noisy channel problem.

Chapter 2

Quantum information channel

The transmission and manipulation of information is one of the prime concerns of quantum information theory. The storing system of information must be considered as quantum mechanical system. One of the vital question in quantum information theory that raised by physicist, what is the limit of perfection to transmit information within a given set of resources. Comparing to the classical information fact, we find two very different questions in quantum theory. However, quantum state can be conveyed by a sender himself. In the quantum state the sender find an unknown state of quantum system. The goal of the sender is that, the receiver will end up with a similar system in quantum state. To solve this case, recently a coding theorem is proved by scientists. On the other hand, quantum state may be used for conveying classical information in the sequence of zeros and ones. When the quantum states are not orthogonal to each other (which states are used by a sender and receiver) then we notice an interesting thing. In this case, it becomes very difficult to distinguish from each other perfectly by the receiver. This type of problem gives us a chance to a new information theoretic-interpretation of the Von Neumann entropy of an ensemble of states. We can use non- orthogonal quantum states in a variety of context to transmit information. To avoid eavesdropping, non-orthogonal signals are used intentionally in the field of quantum cryptography. Besides, in the field of quantum level for transmitting signal ,such as weak coherent pulses in an optical fiber, any ambiguity between signals may be more a matter of non-orthogonally(e.g., overlapping pulses) than classical noise. Although our analysis applies to all quantum system, most of the time we will imagine the signals to be non-orthogonal photon polarization state

We can imagine a situation that a sender, Alice wants to transmit classical information to a receiver, Bob. To do this job Alice use quantum mechanical communication channel (for example, the polarization of photon). Alice will try to prepare the channel in various quantum states. Following this process Alice will represent her messages. On the other hand, the receiver Bob will measure the channel and try to recover the information. Messages sent by Alice will be limited without error. The quantum mechanics of the channel will do this task. As we mentioned before unless the signal state used by Alice are orthogonal, it will become very difficult for Bob to distinguish accurately between the signals. As a result the accuracy of Bob to recover the messages of Alice without any error will be limited. Bob will complete this work with the help of quantum mechanics of the channel. In reality, it is almost impossible for any Decoding observable to recover the whole information content of the message in the quantum signal source. The maximum information that we can be able to recover in a measurement performed on the system M that conveys that message. For this reason, it is more reasonable to consider accessible mutual information is the proper measurement of recovered information. For a pair of random variables of mutual information X and Y is defined to be

$$I(X:Y) = H(X) - H(X|Y) \tag{2.1}$$

Here H is the Shannon entropy, which is a function of the probabilities p(xi) of the possible values of X:

$$H(X) = -\sum_i p(x_i)\log_2 p(x_i) \tag{2.2}$$

Here we interpret plog2p as taking the value zero when pH(XuY) is the expected entropy of X once one knows the value of Y. That is,

$$H(X|Y) = \sum_j p(y_j)\Big[-\sum_i p(x_i|y_j)\log_2 p(x_i|y_j)\Big] \tag{2.3}$$

The mutual information is the amounts of information about X that we find by determining the value of Y in the classical information theory .We justify our focus in the crucial theorem which we find from classical information theory. The justification on I(X: Y) is this: if we find a communication channel which has mutual information I(X: Y) between the input signal X and the received output Y than the meaning of sufficiently redundant coding is that the channel can be used to send information not existing I(X:Y) binary digits per use of the channel with arbitrarily low probability of error. In the field of quantum context, if we denote it by B, the outcome of a measurement of an observable on M , the actually I(A:B) measure the information about the message source A that we find from the measurement of the observable. For this reason, it also measures the number of binary digit which can be conveyed per signal when the receiver works with this observable.

Gordon and levitin stated a theorem about quantum information. This theorem was first proved by Holevo. The theorem tells that, the amount of information which is accessible to Bob is limited .This limitation is occurred by the entropy of ensemble of signal states. We can imagine a situation where Alice represents each message (with a priori probability pa). Usually this state is a mixed state. Now for an observable, we can say that Bob will take the opportunity to measure information. The mutual information I(A:B) between Alices input A and Bobs measurement outcome B is bounded by

$$I(A:B) \leq H(\rho) - \sum_a p_a H(\rho_a) \tag{2.4}$$

Where rho=(the average density matrix for ensemble of signals) and H()=-Tr(here we find the von Neumann entropy of the density matrix). If this condition true that, the signal states a all pure states and the second term on the right vanishes, then we can simply say that

$$I(A:B) \leq H(\rho) \tag{2.5}$$

Holeveo noted some special situations when message are sent to Bob. In this situation I(A:B) does not approach H(P) very closely for any choice of Bobs decoding observable. Although this theorem gives an upper bound on the amount of information which is accessible to Bob, this upper bound is not always strong.

Two physicists Peres and Wooter researched more about this example. Alice wants to send a photon in one of three linear polarization states (called letter states). These three states are separated by 1200 .Three states are equally useful for acquiring information. This signal ensemble has a Von Neumann entropy H(p) of 1.000 bit. On the other hand, we see that Bobs optimal decoding observable creates mutual information I(A:B) of .585 bit. There will be nine possible states if we use two photons. The Von Neumann entropy is 2.000 bits and Bob can easily find the optimal mutual information which is 1.170 bits. Alice try to send two photons, but she works only with three states. It is noted that, the individual letter states are being used with their original (equal) probabilities in this restricted choice of two photon states. In this case, the ensemble entropy S (p) is only 1.585 bits. But we find the optimal mutual information is 1.369 bits or about .685 bit per photon. In other words, we can restrict the code to a subset of the possible code words. In spite of restricting the code, Alice has a chance to increase the distinguish ability of the code words and increase the information which is conveyed per photon to Bob.[1]

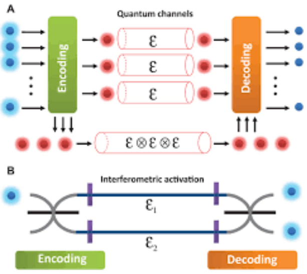

Figure 2.1: Encoding and decoding of quantum channel

From the above example, we can say that, we have some chances for increasing the accessible information per elementary signal by (a) using code composed of several elementary signals and (b) deleting the possible code words in the ensemble; still we are respecting the elementary signal frequency which we have acquired before. For the purpose of distinguishing the real code words, the receiver than chooses a decoding observable optimized. It is noted that, this observable will not be observable for a separate set of measurement on the individual elementary signals at all times. Instead of that will be a joint measurement on the whole set of signals constituting a code word.

This consideration creates an ambiguity which is that whether Alice and Bob will be able to use this strategy for approving the Holevo bound. That is, given a priori ensemble of pure state signals with entropy H (), is it possible for Alice and Bob to choose a set of code words respecting probability of the original signals together with a decoding observable, so that information can be transmitted reliably at a rate approaching H (p) per elementary signal? The answer is yes. Besides, it will be bound that, no such code will be able to transmit signal exceeding H (p). It is now convenient for us to give a precise formulation of our main result. Suppose we are given an ensemble of letter states of an elementary quantum system (it is not need a photon) with a priori probabilities Pa. The letter ensemble has a density matrix

$$\rho = \sum_a p_a \left|a\right\rangle \left\langle a\right| \tag{2.6}$$

With a Von Neumann entropy H(p)=-Trlnp.
A code [(N,l) code] has two things. One is a set of nN code words S˙ij=1,.n where each code word is a sequence (i.e, product) of letter states (but not every time all such sequence of letter states are code words). Another thing is that an a priori probabilities pSi assigned to each code word. The tolerance of the code is defined by

$$\tau = max_a|f_a - p_a| \tag{2.7}$$

Where fa is the overall frequency of occurrence of the letter $\left|a\right\rangle$ among the Nl letters of all the code words, taking into account then a priori probabilities of the code words. That is

$$f_a = \frac{1}{l}\sum_{i=1}^{N} Ps_i^n a, s_i \tag{2.8}$$

For construction of given words low tolerance code will use the letters approximately with their given a priori probabilities Pa.
Theorem: Suppose I be the least upper bound on the information per letter transmissible with any code having tolerance ¡.
This theorem gives a precise information-theoretic interpretation of von Neumann entropy in quantum mechanics. To put it in somewhat looser but more familiar terms, the theorem says that if Alice is highly needed to use certain quantum states as signals with certain specified frequencies of occurrence, the number of binary digits she can convey to Bob per particle can be made

arbitrarily close toH()-, but not greater than , the von Neumann entropy of the ensemble of signals as we think before. Before we prove the theorem it is noticeable that comparing our result to the channel capacity theorem for classical information channel we get a clear concept for transmitting information reliably through a noisy channel. The maximum von Neumann entropy we find from our quantum channel has the priori probabilities

$$C := max_{p_a} H(\rho) \tag{2.9}$$

From the above discussion we have found some similarities between classical channel capacity and the above notion of quantum channel capacity. In spite, of these similarities we also find some differences, in particular the origin of the conditional probabilities in eqn (3). Generally we find these probabilities are fixed in the classical setting, we are able to transmit information like letter states without altercation. There will be no difficulty for Bob to decode the message. Now there is a problem to distinguish perfectly non-orthogonal quantum states by any measurement. For decoding bob has enough freedom to choice and conditional probabilities in equation (3).
The proof of the theorem that we have found just before, has some similarities with the classical channel capacity theorem in some respects in that sense, both we found rely on the construction of the letter states to obtain code words and proving of the set of all possible code words to be used in the channel.

In the classical setting, the purpose of proving codes is to increase redundancy. On the other hand, we see that in quantum setting the aim of proving code words is to distinguish of the code word states. It is a matter of interest this concentration and proving do not result in a channel that differs from the original. There is a question arise in both situation that if we allow for repeated transmission of elementary letter states, what will be our minimum rate so that we can convey information. For demonstrating the existence of a code with desired properties both proofs utilize a method of random coding

$$(1-\epsilon)2^{l[H(\rho)-\delta]} < \dim\Lambda < 2^{l[H(\rho)+\delta]} \tag{2.10}$$

The typical subspace L is constructed as follows: The eigenvalues qi of the one-letter density operator r form a "probability distribution for the Eigen states of r, for which the classical Shannon entropy is just the von Neumann entropy H(r). Eigen states of are sequences of r Eigen states. By the weak law of large numbers, we can find a set of typical Eigen states of l which the frequencies of the ρ Eigen states are close to the probabilities qi is the subspace spanned by these typical Eigen states. Suppose we sum the squares of the eigenvalues of l, but restrict ourselves to the typical subspace. Then we get

$$\mathrm{Tr}\prod\nolimits_{\Lambda}(\rho^{l})^{2}\prod\nolimits_{\Lambda} < (\dim\Lambda)(2^{-l[H(\rho)-\delta]})^{2} < 2^{-l[H(\rho)-3\delta]} \tag{2.11}$$

Chapter 3

Decoding Observable

We suppose a situation where Alice is using a code. The words that he uses in the code are long enough for the typical subspace to exist. The words have the properties outlined above. It might be happened for Bob that he is not using all of the possible code words. One of the main problem for Bob is to distinguish a collection of vectors in the Hilbert space H' be a collection of such vectors (possibly not normalized). We consider the operator

$$\Phi = \sum_k |\phi_k\rangle \langle\phi_k| \tag{3.1}$$

Which we find a positive operator whose support is the subspace spanned by the vectors. On this subspace exists and is invertible, so we can form the vectors

$$|\mu_k\rangle = \Phi^{-\frac{1}{2}} |\phi_k\rangle \tag{3.2}$$

It is corresponds to positive operators. These positive operators can easily be shown to be a resolution of the identity on this subspace:

$$\begin{aligned} \sum_k |\mu_k\rangle \langle\mu_k| &= \sum_k \Phi^{-\frac{1}{2}} |\phi_k\rangle \langle\phi_k| \Phi^{-\frac{1}{2}} && (3.3) \\ &= \Phi^{-\frac{1}{2}} (\sum_k |\phi_k\rangle \langle\phi_k|) \Phi^{-\frac{1}{2}} && (3.4) \\ &= \Phi^{-\frac{1}{2}} \Phi \Phi^{-\frac{1}{2}} = 1 && (3.5) \end{aligned}$$

The operators supplemented by a projection onto the subspace perpendicular to the span of it thus arise from the outcome operators of a POM. The vectors specify a particular POM which employs the outcome operators. This is the POM that Bob chooses in order to distinguish among the vectors. This is a reasonable choice. If the vectors are orthogonal and thus completely distinguishable, the resulting measurement does indeed distinguish them perfectly. There is no reliable way of specifying the best observable in general, but this observable will be good enough for our purposes. The vectors have another interesting and for us useful property. Let Sjk be the matrix of inner products among the vectors:

$$S_{jk} = \langle\phi_j|\phi_k\rangle \tag{3.6}$$

If there are N vectors, this is an N*N complex matrix with positive eigenvalues. The μ˙mk vectors are related to the square root of this matrix by

$$(\sqrt{s})_{jk} = \langle\mu_j|\phi_k\rangle \tag{3.7}$$

In fact, this property of the vectors can be taken as an implicit definition for them.

Bob will try to decode Alices message, for this reason he will employ an observable to distinguish between her signal states. But we will find it more useful to suppose that he distinguishes between projections of the signal states into the typical subspace Lthat is, between non-normalized vectors. For this purpose Bob will employ the square root measurement which was described a short time ago. Since we see that, the typical subspace contains almost all of the set of available code words in the sense of the previous section this refinement introduces negligible error, as we shall show. Thus, we can define the matrix Si so that

$$\langle \mu_i | s_j \rangle = \langle \mu_i | \sigma_j \rangle = (\sqrt{S}_{ij}) \tag{3.8}$$

Chapter 4

Probability of error

Alice code contains N code of words. Each code is used with equal frequency. These codes can be used with equal frequency. So we say that, the information content of a single code word is therefore log2N.Each code word is a sequence of l letters which we found from the set of possible letters. (Only for this case, we will disregard the probabilities of those letters in the given ensemble.) Bob will try to devise his decoding observable as we mentioned before. Alice sends the signal $|s_i\rangle$. The probability of sending her signals is 1/N. Bob will try to interpret the signal accurately. It means that, he will obtain the i outcome in his decoding POM-with probability.

$$p(\mu_i|s_i) = \mathrm{Tr}||\mu_i\rangle \langle\mu_i|s_i\rangle \langle s_i| = |\langle\mu_i|s_i\rangle|^2 \tag{4.1}$$

From the equation, we see that the vector is real and non-negative. The average probability of error we find

$$P_E = 1 - \sum_i \frac{1}{N} \langle\mu_i|s_i\rangle^2 = \frac{1}{N} \sum_i (1 - \langle\mu_i|s_i\rangle)(1 + \langle\mu_i|s_i\rangle) \tag{4.2}$$

$$\leq \frac{2}{N} \sum_i (1 - \langle\mu_i|\sigma_i\rangle) \tag{4.3}$$

In terms of the Sijmatrix, this is

$$P_E \leq \frac{2}{N} \sum_i [1 - (\sqrt{S}_{ii})] \tag{4.4}$$

The square root function is bounded below by a parabola: for $x \geq \ 0$,

$$\sqrt{x} \geq \frac{3}{2}x - \frac{1}{2}x^2 \tag{4.5}$$

The matrix S is a matrix with non-negative eigenvalue, so this inequality may be applied to it:

$$\sqrt{S} \geq \frac{3}{2} - \frac{1}{2}S^2 \tag{4.6}$$

This means that, for a complex N vector with components Zk we find

$$\sum_{kl} z_k^* (\sqrt{S})_{kl} z_l \geq \frac{3}{2} \sum_{kl} z_k^* S_{kl} z_l - \frac{1}{2} \sum_{klj} z_k^* S_{kj} S_{jl} z_l \tag{4.7}$$

Chapter 5

Random code

In this section and the coming section we will prove our main result. Our main result will be proved in this section and the next section. We will show that Alice can choose N code words with N sufficiently large so that log2N is approximately lH(r), such that Bob (using the scheme above) has probability of error PE. The probability is nearly equal to zero. In the coming section we will also see that, Alice's code may be choose arbitrarily. Alice code has small tolerance. As we defined in (7) we will also show that, an information rate of H (p) bits per letter cannot be expected to exceed in the limit of vanishing tolerance. We will show that, the existence of this code. To do this task, we will in fact show that the job can be done by most of the code. This means we will measure the average probability of errors. Our target is, by doing this job we will ensemble of random codes of N words. By applying this process, we will generate a random code. Each of the N code words we found is a sequence of l letter states which generated by using the priori probabilities for the letters. We denote an average over random codes by

$$\sum_{kl} z_k^*(\sqrt{S})_{kl} z_l \geq \frac{3}{2}\sum_{kl} z_k^* S_{kl} z_l - \frac{1}{2}\sum_{klj} z_k^* S_{kj} S_{jl} z_l \tag{5.1}$$

Next we take the average of PE over random codes

$$\langle P_E\rangle \leq \frac{2}{N}\sum_i (1-\langle n_i\rangle_c + \frac{1}{2}\sum_{j\neq i}\langle S_{ij}S_{ji}\rangle_c \tag{5.2}$$

Each of the averages in this expression is easy to calculate. The average norm of the ith projected code word is

$$\langle n_i\rangle_c = \langle \mathrm{Tr}(\prod\nolimits_\Lambda |s_i\rangle\langle s_i|\prod\nolimits_\Lambda)\rangle_c \tag{5.3}$$

$$= \mathrm{Tr}(\prod\nolimits_\Lambda \rho^l \prod\nolimits_\Lambda) \geq 1-\epsilon \tag{5.4}$$

For ji, the code words $|s\rangle_j$and $|s\rangle_i$ are independent so that

$$\langle S_{ij}S_{ji}\rangle = \langle\langle s_i|\prod\nolimits_\Lambda|s_j\rangle\,\langle s_j|\prod\nolimits_\Lambda|s_i\rangle\rangle_c \tag{5.5}$$

$$= \langle \mathrm{Tr}(\prod\nolimits_\Lambda|s_i\rangle\langle s_i||s_j\rangle\langle s_j|\prod\nolimits_\Lambda)\rangle_c \tag{5.6}$$

$$= \mathrm{Tr}\prod\nolimits_\Lambda(\rho^l)^2\prod\nolimits_\Lambda \tag{5.7}$$

The j sum yields a factor N-1 thus

$$\langle P_E\rangle_c < 2\epsilon + N\,\mathrm{Tr}\prod\nolimits_\Lambda(\rho^l)^2\prod\nolimits_\Lambda \tag{5.8}$$

We use the properties of typical subspace to obtain

$$\langle P_E\rangle_c < 2\epsilon + N2^{-l[H(\rho)-3\delta]} \tag{5.9}$$

If the average probability of error is below this bound, then Alice and Bob will be able to find some particular code for which

$$P_E < 2\epsilon + N2^{-l[H(\rho)-3\delta]} \tag{5.10}$$

This code can be transmitted at any asymptotic rate of H (p) bits per letter which error probability is very low.
Remark: In fact, we have a chance to do better. We have a system to modify a code with low error probability. Suppose we throw the worst half of the code words in the suitable code. As we found the average probability of error for this code is less than 3, we have

$$\frac{1}{2^{lH(\rho)}}\sum_i [1 - p(\mu_i|s_i)] \leq 3\epsilon \tag{5.11}$$

From the above discussion we can conclude that, at least half of the code words must have conditional probability of error of less than 6 otherwise, these code words have contribution at least 3. to the sum. Thus the reduced code book we have code words. If we throughout half of the code words, then we have a chance to change the rate from H(p) to H(p)-1ł, a negligible difference for large l.

Chapter 6

Letter Frequencies and Channel Capacity

In our previous section, the arrangement we have made is constructive. We were not able to construct a code with low probability of error. We have merely showed that this type of code should be exist. For this reason, we become fail to know its property in details. However, our task is now to show that the code may be chosen having arbitrarily small tolerance. Previously we saw that, each code word is used equally most of the time. So, many times a given letter appears in the list of N code words help us to predict it's frequency of occurrence when the code will be used by Alice and Bob. In the following way, we can apply the weak law of large numbers to the set of codes: if N l is sufficiently large, it might be happened that the set of all random codes divided into two classes: a set of "atypical" codes. This type of codes is generally generated by random coding with small total probability. These codes which has small total probability have very small effect to the overall average probability of error estimated above. In this way $\langle P_E \rangle_c$ must also be very small even if Alice and Bob are given no chance to use "typical" codes-every one of which has letter frequencies matching the a priori probabilities p˙a to within any specified tolerance. Different types of situation might be happened if Alice and Bob are not need to use any particular frequencies. In this case, they are free to adjust them as their own wish in order to maximize the information per letter conveyed by their channel. Similar to section 1 in this case we have the chance to define the channel capacity C of a quantum channel with a particular alphabet.

$$C = max_{p_a} H(\rho) \tag{6.1}$$

From our above argument clear this fact that Alice may communicate with arbitrarily low probability of error up to C bits per letter, to Bob using the letter states $|a\rangle$.

Chapter 7

Super dense coding

Bennet and Wiesner (14) proposed quantum communication scheme. This code is called Super dense coding. We can apply our result in this interesting scheme. We use super dense coding in the quantum entanglement lower systems her enhancing their information security. We now want to give an example how the super dense coding works. At first Alice and Bob share a pair of two state system. We suppose this state is one of the four orthogonal Bell states which is given by,

$$|\Psi^{\pm}\rangle \, \frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle \pm |\downarrow_1\uparrow_2\rangle) \tag{7.1}$$

$$|\Phi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|\uparrow_1\uparrow_2\rangle \pm |\downarrow_1\downarrow_2\rangle) \tag{7.2}$$

The communicating power of single spin is very limited. In reality,we find only one bit of information by transmitting this type of spin. The process is occurred by a simple application of the W holevo theorem. The other situation might be happened. Bob will be able to know two bits of information from Alice. This task is happened by the per-determined entanglement between the pair of spins. Both Alice and Bob have to do a special task for this special condition. Bob will be able to recover perfectly by Alice 4 way choices which encodes a two bit message. Our interest is in a more general situation. Instead of spin, both Alice and Bob will work with N-state quantum

system. At the same time they can possess a considerable supply of them. The purpose of Alice and Bob are to use blocs coding of many independent messages.
Now we suppose that, Alice and Bob are sharing many pairs of system. These systems in some entangles pure state like the Bell states may or may not be maximally entangled. Now our prime concern is, what will be information capacity of these entangled systems for super dense coding. It is very easy for Alice to perform a unitary transformation on her system. After unitary transformation, Alice will deliver it to Bob. We might imagine different transformation performed by Alice with different a priory probabilities, contribute a vital rule to ensemble of states for the pair of system. Later, Bob will measure these states. From our theorem it is proved that with the help of tricky coding and Bobs decoding observable Alice can easily convey an amount of information up to the entropy of the ensemble.

Now it is our main concern, that what will be limit of this entropy or what is the maximum capacity of this scheme. Now it is very easy to realize that, it is not possible for the entropy of this system can be larger than He log2N. The reason is that manipulations of Aliice's system do not affect the density matrix pf Bobs system. The total entropy for the pair of system will be always smaller than the entropy of Bobs system which is denoted by Alice. The same thing will be happened with the entropy of the Alices system. From our discussion we can also say that, it is very easy for a particular ensemble of transformation to make the overall entropy equal to He+ log2 N. It is very easy for this type of ensemble of transformation to include all permutations of the Schmidt basis states.

Form the above discussion, we can make a decision that,the channel capacity of the super dense coding scheme is He + log2N. It is a sensible result. We suppose that, at the initial time the pair of a system in a product state He=O and same to the previous state Alice only can send log2N bits in each state.

Chapter 8

Noisy Channels:

In our discussion we have assumed that, Alice sends a letter state la ¿to Bob. The state remain unchanged when it arrives to Bob. But in reality it is not possible of all the times. The channel will often create noise. Besides, the signal wills arrived in mixed state Pa. In that case Alice will take another process to send her message. Instead of ensemble of pure state Alice will use ensemble of mixed state. We do not use ensemble of mixed states in our theorem. In this case we will use the following conjecture.

$$\chi = H(\rho) - \sum_a p_a H(\rho_a) \tag{8.1}$$

To make logical our assumption considering two different ensembles is very helpful. Alice is given the ensemble of mixed states. We consider a random code. The original mixed state ensemble contributes to construct this code. We do not find any physical difference between these two codes as we have used al set of possible transmission. The difference we find in code is that the knowing of Alice is sending the pure state.
We suppose that our main theorem applies to 8 code constructed in this way. It Alice is able to know by which pure signal she is sending her signal, then it will be easy to her to convey up to H (p) bits per letter using a typical 8 code constructed as above. By getting this information Bob will be able to know the specific Eigen states. In reality Alice has the capacity only to know the mixed state. In this way the amount of information that will be

obtained by Bob per letter about this Eigen states is the average entropy of the mixed signal states Pa, that (Pa). The result we found from the additivity of information is that

$$H_(\rho) - \sum_a p_a H(\rho_a) \tag{8.2}$$

It is our failure that the above theorem does not apply to the codes because of the lack of strict independence among the code words. However in general sense we can say that in more elevator arrangement will discuss it more dearly and thereby prove the conjecture.

Chapter 9

Quantum operation:

How can we perfectly define quantum noisy channel? What is the accurate process to describe it mathematically? We will use quantum operation to describe noisy channels. Same as before, we define quantum operation as completely positive maps or super scattering operators. There are many example of a state change in quantum mechanics. Among them writary evolution is one of the simple state example which is experienced by a closed quantum system. The mathematical expression by which we show the trial system state of the system which is related to the initial state by a unitary transformation.
img

$$\rho \to \mathcal{E}(\rho) - U\rho U^\dagger \tag{9.1}$$

Most of the closed quantum systems are described by unitary evolutions. More general state changes are possible by Schrodingers equation in his open quantum system. In this case, we can give the example of noisy quantum channel.
How one can describe a general state change in quantum mechanics. With quantum operation formalism we can give a proper solution. Kraus described the formalism properly. We find in his formalism an input state and an output state. The input state and output state are connected by a map.

$$\rho \to \frac{\mathcal{E}(\rho)}{\mathrm{Tr}[\mathcal{E}(\rho)]} \tag{9.2}$$

We find only one term in the sum A1= U of a unitary transformation. One of the interest things that we find in a class of operation is the trace preserving or non selective operation. Generally we find trace preserving operations where the system is interacted to some environment which is not under observation.
Quantum Operation Theorem representation (theorem for trace preserving quantum operation) Suppose is a trace preserving quantum operation on a system which has a d-dimensional state space. We can able to construct an environment E of at most d2 dimension. In this situation, the system and environment are initially uncorrelated. We also think that, at the beginning the environment is in a pure state $\sigma = \langle s \rangle$. At the same time we also think, there exist a unitary evolution U on system and environment such that

$$\mathcal{E}(\rho) - \mathrm{Tr}_E[U(\rho \otimes \sigma)U^{\dagger}] \tag{9.3}$$

From this theorem we find that any trace preserving quantum operation can always be mocked up as a unitary evolution. The unitary evaluation is connected to an environment with which the system can interact unitarily. Conversely, from this theorem we can say that, any such unitary interactions with an initially uncorrelated environment give rise to a trace preserving quantum operation. Both of those cases are useful in what follows.
Here the state of the system before the interaction with the environment is denoted by Q. On the other hand, the state of the system after the interaction is denoted by Q'. Unless stated otherwise in all case we will follow the convention that Q and are Q'd- dimensional. For the description of noisy channel the only thing, that we concern is the dynamics of Q. For any given quantum operation ϵ we find various expression of in terms of environments and interactions $U\hat{}$QE for our convenient we always think that the beginning state at E is a pure state and regard E as a mathematical artifice. In reality, the actual physical environment Ea may be impure at the beginning

state. The meaning of discrete is that, there are few members at input and output states exist on the channel. By discrete quantum channel we realize that, it has a finite number of Hilbert space dimensions. In the classical case, the meaning of memory less is that the output of the channel is independent of the past, conditioned on knowing the state at the source.
Phrased in the language of quantum operations we assume that there is a quantum operation N which describes the dynamics of the channel. The relation of input p; with the output of the channel is expressed by the following equation.

$$\rho_i \to \rho_o - \mathcal{N}(\rho_i) \tag{9.4}$$

Chapter 10

Entanglement Fidelity:

Now we will review a quantity which is known as the entanglement fidelity. We use this quantity. We defined the entanglement fidelity for a process which is expressed by a quantum operation acting on some initial state p. There are some special significance of entanglement fidelity. Among them when an entanglement fidelity close to one, we find that the process preserve the state well It also indicates that the process preserves the entanglement well. On the other side, we kind a completely different situations when the entanglement fidelity is close to zero. With this situation we find that the ϵ on its entanglement was not well preserved by the operation. So, we can say that entanglement fidelity is the overlap between the initial purification
Psi^{RQ} of the state between it is we send it through the channel with the state of the joint system RQ. After that, we send it through the channel. We see that entanglement fidelity depends only on ρ and ϵ instead of the particular purification
Psi^{RQ} of that is used. If has operation elements Ai, than we find the following expression of entanglement fidelity.

$$F_e(\rho, \mathcal{E}) - \frac{\sum_i |\mathrm{Tr}(A_i \rho)|^2}{\mathrm{Tr}[\mathcal{E}(\rho)]} \tag{10.1}$$

As we find the denominator is 1, so this expression will simplify trace- preserving quantum operation.
We use entanglement fidelity for various reasons to measure our success in

transmitting quantum state. If we become success to send a source Ps, with high entanglement fidelity, then we will be able to send any ensemble for s with high average pure-state fidelity. So, entanglement fidelity is more usual for quantum coherence than average pure-state fidelity. In addition, the ability to preserve entanglement has great importance in applications of quantum coding. For instance in quantum computation we apply error correction in a modular fashion to small portions of quantum computer.

From our previous discussion we can say that if the subspace fidelity is close to one then we can say that the entanglement fidelity is also close to one. This converse is not true at all the time. It means that reliable transmission of subspace has more contribution than transmission of entanglement. Therefore for reliable transmission entanglement fidelity yields capacity at least as great as those obtained when subspace fidelity is used. For convenient, we think that those too capacities are identical.
The main lesson we have learned from this section is that there are many different process of measurement which we can use to quantify how reliably quantum states are transmitted. Different measures may result in different capacities. Which measure we will choose it depends on what resource is most important for the application of interest. As measure of our reliability we use the entanglement fidelity in this course.
One of the most useful inequalities is quantum fano inequality. This inequality rates the entropy exchange and entanglement fidelity. It is

$$S_{\rho,\mathcal{E}} \leq 1 + 2[1 - F_e(\rho, \mathcal{E})] \log_2 d \tag{10.2}$$

From this result we find bound on the change in the entanglement fidelity when the input state is excited. We find that during the proof a co-efficient $\sqrt{F(\rho, \epsilon)}$ was dropped suddenly from the first term on the right hand side of the inequality. In some applications it is sometimes useful applying the inequality with the co-efficient in place.

Chapter 11

Noisy-Channel Coding Revisited

One of the main aim of noisy-channel coding is to choose what source states can be sent with high entanglement fidelity for this reason it becomes our purpose what states PS encoding and decoding operation can be found such that

$$F_e(\rho_s, \mathcal{D} \circ \mathcal{N} \circ \mathcal{C}) \approx 1 \tag{11.1}$$

Shannons noisy-channel coding theorem is one of the best examples of a channel capacity theorem. There are two parts in this theorem. (1) We set a rate on the upper bound so that we can send information reliably through the channel. It must be expected that we can calculate the upper bound perfectly. (2) A reliable system for encoding and decoding exist that comes arbitrarily close to attaining the upper bound found in (1), the system was proved before.

This maximum rate is called channel capacity. In this rate we can send information reliably through the channel. Channel capacity results may be easier to understand if he use the language of error correcting code. We want to protect information against the effects of noise. So we encode the information using an error correcting code, with the encoding operation represented by c. We then subjected it to the noise which is represented by N. At last the encoding is undone by using the decoding operation, D. If we become able to find a good error correcting code then it is easy to find C and D which preserves the information being encoded, The function of a channel capacity theorem is to set up a ultimate achievable limit on the effectiveness of these error correcting codes, for a given noise mode N.

Chapter 12

Mathmatical Formulation of noisy channel coding

So far, noisy-channel coding procedure has been described elaborately, but we did not make all of our definitions mathematically precise. In this subsection we try to give a precise formulation for the most important concepts appearing in our work on noisy-channel coding. The nth density operator refers representing the state of emission from the source. Its other meaning is taking units of time.

We abuse notation usually by omitting explicit mention of the Hilbert spaces Hs and He when we apply this notation in various equation. It is noticeable that the channel has some possibility to have different input and output Hilbert spaces. For our convenient we do not consider that case here, but all the results that we have proved here go through without change.

Suppose a source state Ps and a channel N. The ultimate goal noisy-channel coding is to find a best solution with an encoding C and a d-coding 1) such that Fe1s,J is close to 1. In fact the best solution we means here source state Ps and its entanglement will be transmitted almost perfectly. In reality this situation is almost impossible. However Shannon showed that in the classical content if we consider blocs of output from the source and performs block encoding and decoding, then it might be possible to considerably expand the close of source states Ps for which it is possible. The main goal that we have done quantum mechanical version is to find a sequence of n codes (Cn,Dn).

We imagine that our last previous sequence of codes exist for a given source . In this case, the channel has to transmit reliably. At the same we see the reliable transmission rate of the channel is $R = S()$.
In most part of this paper, we assume as in the previous equation the action of the Channel which is described by the operation is generally trace preserving. The assumption we make here is similar to the physical assumption no classical information about the state of the system or its environments is obtained by an external classical observer. In addition to the environment E has also tremendous usefulness her introducing a reference system Rin the following way. It might be ambiguous to someone that the system is initially part of a larger system RQ from our perspective we can say that the introduction of R emerged as a mathematical device to purify the initial state. According to the dynamics $I_R \otimes \epsilon$ we find from the joint system RQ

$$\rho^{R^{'}Q^{'}} = (I_R \otimes \epsilon)(\rho^{RQ}) \tag{12.1}$$

Where I_R is the identity dynamics has the reference system R. So far, we have discussed the procedures are only applicable to trace preserving quantum operation. Later in the paper we have will also discuss quantum operations which are not traced preserving. This means those types of quantum operation do not satisfy the relation and in general we can say this is not equal to 1. These types of quantum operation we find in the theory of generalized measurements. For each outcome M of a measurement we find an associated quantum operation ϵ_m with an operator sum representation

$$\epsilon_m \rho = \sum_i A_{mi} \rho A_{mi}^{\dagger} \tag{12.2}$$

The probability is

$$Pr_m = \mathrm{Tr}[\epsilon_m(\rho)] = \mathrm{Tr}(\sum_i A_{mi}^{\dagger} A_{mi} \rho) \tag{12.3}$$

It is a matter of interest that, the formulation of quantum measurement based of the projection postulate, taught in most classes on quantum mechanics, is a special case of the quantum operations formation. When we h a single projector Am=Pm in the operator-sum representation m for then we find these types of quantum operation formalism. We find the relation of the formalism of positive operator valued measures (POVMS) to the generalized measurements formalism Em= are the elements at the RUVM that is measured. We can prove a result analogous to the earlier representation there for general operation. The representation theorem expresses the trace preserving quantum operation.

Chapter 13

Entropy Exchange

In this section we will review the definition and some basic results about the entropy exchange. The entropy of a channel, also called the map of entropy is defined as the entropy of the state corresponding to the channel by the Jamiolkowski isomorphism. In briefly the Jamiolkowski isomorphism allows to map many statements about states to statements about channels and vice versa. E.g., a channel is completely positive exactly if the Choi state is positive, a channel is entanglement breaking exactly if the Choi state is separable and so further. We try tofind a clear concept of noisy quantum channel from the discussion of the entropy exchange.
We define the entropy exchange of a quantum operation *epsilon* with input
ρ is
$S_e(\rho, \epsilon) \equiv S(\rho')$ We have found many essentially different information transmission problems in quantum mechanics. We have discussed mainly two problems. One is the transmission system of a discrete set of mutually orthogonal quantum states through the channel. Another one is the transmission of entire subspace of quantum states through the channel. This type of transmission has great contribution to keep all other quantum resources, inducing entanglement. In many applications like quantum computation, cryptography, teleportation we find the contributions of such types of transmission.

Chapter 14

Conclusion

So far, we have proved that, Von Neumann entropy H (p) has an important relation with the capacity of the quantum channel to transmit classical information in an ensemble of pure quantum state. In this case, the quantum channel transmits the states with their given priori probabilities. It is true that, for all non-orthogonal ensemble, one can get sufficient amount of informations by measuring on a signal system which is sharply less than H (p) [6]. At the same time, it is also true that, Von Neumann entropy H (p) is equal to the capacity of quantum channel. One can use this trick to increase the capacity of quantum channel by having the receiver discriminating among whole code words. This process is better than trying to distinguish the individual signal state.

In final, we can say that, entropy has great impact in transmitting information. In communication problem, entropy is the vital factor of the actual channel capacity and at the same time it is not merely an upper bound.

List of Figures

Bibliography

[1] Schumacher, B. (1995). Quantum coding. *Physical Review A*, 51(4):2738.