

Verification of CQ Coding Theorem using MATLAB

Archana Chaudhary
Department of Electronics and Communication
Engineering
Netaji Subhas University of Technology
New Delhi, India

Harish Parathasarathy
Department of Electronics and Communication
Engineering
Netaji Subhas University of Technology
New Delhi, India

Abstract— In a communication system, whenever data is transferred it has to be received as it is on the input side but the data gets corrupted due to channel noise because of which the information or data will be transmitted incorrectly. So, as to minimize this problem codes have been formulated with the aim to reduce the noise by minimizing the probability of error in a communication system. In a quantum communication system with classical alphabet and there strings encoded into quantum states, even if the channel is noiseless, uncertainly is inherently introduced into the decoding process by the quantum state matrices because of the method by which probability enter into quantum mechanics via the inner product between state matrices and detection operators.

Keywords— Kronecker (Kron), Quantum bits (Qubits), Classical-Quantum coding theorem (CQ coding theorem)

I. INTRODUCTION

A. History of quantum information theory

The beginning of quantum information started during the twentieth century, when the classical information was being developed. Quantum mechanics was first created by Heisenberg in the form of non-commuting matrices that represent observables, then extended by Schrödinger who created wave mechanics as an alternate description of quantum mechanics. Finally, Dirac unified the two pictures by moving that they are equivalent in that both lead to the same evolution for average values of observables which are the physical quantities of real interest. Max Born introduced probability into quantum theory by appropriately interpreting the Schrödinger wave function. This is how uncertainty inherently enters into quantum mechanics via state matrices used to transmit data. The state/density matrices that encode classical data into quantum data inherently define the uncertainty introduced in the CQ channel.

There main focus was on error probability and channel capacity. Though this research work was going on yet unfortunately, the number of researchers in the field has been rapidly declining in the past years. This line of research came to decline in the early 1980s. One of the reason behind the decline was the fact that coding theorem of Shannon and his quantum counterparts stated only existence of codes which gave low error probability that data was transmitting at not too higher rate. They did not give any constructive algorithm for the generators of such optimal codes.

Information theory is an application of Probability theory. Claude Shannon single handedly made a great contribution to information theory. In fact by providing the noiseless and noisy coding theorem by introducing concepts such as entropy, mutual information and channel capacity, Shannon created the subject of information theory. Due to Shannon's contribution in information theory many Quantum information theorists call Quantum information as Quantum Shannon theory. The main aim of information theory is data compression and to transmit information efficiently from sender to the receiver.

B. Classical and quantum information theory

In a number of respects, quantum information theory varies dramatically from classical information theory. Bits are the most fundamental unit in conventional information theory or classical information theory and they are represented as 0 or 1, whereas qubits are the most fundamental unit in quantum information theory. We represent qubits as $|0\rangle$, or $|1\rangle$. In a classical information system, Shannon entropy is used to assess uncertainty, but in a quantum system, Von Neumann entropy is used to measure entropy. We use $H = -\sum P(x) \log P(x)$ to measure entropy in classical system and in quantum von Neumann entropy uses a density operator called rho and to measure entropy we use $S = -\text{Tr}(\rho \log \rho)$.

C. Quantum bits

A quantum bit, also known as a qubit, is the most fundamental unit of a quantum system. The qubits are a structure of two-levels. The representation of qubits is different from that of bits. In qubits $|0\rangle$ this represent one of the state of quantum system .In the representation of qubits we are using a vertical bar in the left side of 0 and a angle bracket in the right side which shows that we use Dirac notation. $|1\rangle$ it represents the another possible state of the qubit.

So the qubit can be mapped in the following way:

$$0 \rightarrow |0\rangle, 1 \rightarrow |1\rangle$$

Because superposition states are possible, we cannot represent the states $|0\rangle$ and $|1\rangle$ with there classical bits 0 and 1's because Boolean algebra superposition states are not permitted. We instead will define the row and column vector for our qubits as it is beneficial to find the states are represented as vectors $|0\rangle$ and $|1\rangle$:

$$|0\rangle \equiv \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |1\rangle \equiv \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

In linear algebra, these are the column vectors and these column vectors are called as “kets”. In the similar we can also define the row vectors these row vectors are called as “bras”.

The row vectors corresponding to the column vectors $|0\rangle$ and $|1\rangle$ are as follows:

$$\langle 0| \equiv [1 \ 0] \quad , \quad \langle 1| \equiv [0 \ 1]$$

D. Quantum Entanglement

The difficulties of sending a classical message over a quantum channel and estimating a classical analogue for every quantum state by passing it through a channel. These are not problems that are inherently quantum-specific. Quantum extensions of classical issues do occur, though the non-locality of quantum mechanics is the primary source of these quantum extensions difficulties whereas Quantum mechanics, on the other hand, is more than simply a theory in a non-commuting format. There are many protocols in quantum system that have no classical counterpart. The advantage of using entanglement is that we examined primarily the effects. Entanglement quantification a uniquely quantum phenomenon emerges from composite quantum systems.

Schrodinger was the first to notice that if there are more than one quantum systems it can entangled and coined the term after observing some of the strange side effects of this process.

For this to understand we will start by considering a simple, unentangled states and we are considering these states to be named as Alice and Bob, who might share the state. So as to see how an unentangled state differs from an entangled state we will assume that they distribute the states as $|0\rangle_A |0\rangle_B$ where we can say that Alice is having qubit in a system A and while Bob has a qubit in system B. By looking at this we can easily say in which system Alice is and in which system bob is, there is no difficulty in this scenario. The development of shared randomness is our first application of entanglement. We the probability distribution for two binary values X_A and X_B , which are two random variables, is defined as one bit of shared randomness:

$$p_{X_A; X_B}(x_A; x_B) = 1/2 \delta(x_A; x_B)$$

where δ the Kronecker delta function is denoted. Let us assume that Alice is owing a random variable represented as X_A and while Bob owns a random variable represented as X_B . As a consequence of probability being $1/2$ they will either both have a zero or neither has a zero. They will be alternatively, in possession of one. We refer to a single occurrence of shared randomness as a resource $[cc]$ implying that a little amount of shared randomness is a quiet, conventional resource shared by two people.

E. Entropy and Information

Entropy can be used to calculate the degree of uncertainty in a system. Both classical and quantum information theories can be used to investigate entropy. It employs the formula $H(X) = P \log_2 P$

Classical information: Claude Shannon's information concepts serve as the foundation for classical information. Bits of binary strings are used to store the smallest unit of classical information. A capable bit is defined as any system with two states.

Shannon entropy: Shannon entropy is a measure of how much information can be extracted by measuring the values of a random variable. Another way of putting it is that it is a means of looking at a system's uncertainty despite the measurement. Shannon entropy can be thought of as the average quantity of information connected with the occurrences x_1, \dots, x_n , associated with the probability distribution $P(x_1), P(x_2), \dots, P(x_n)$ represented in bits: $H(X) = H[P(x_1), P(x_2), \dots, P(x_n)] = -\sum P(x_i) \log_2 P(x_i)$

Von Neumann entropy : Von Neumann entropy is similar to Shannon entropy. The von Neumann entropy is used for Theory of quantum information In classical information theory, Shannon entropy is used. It is used to find the entropy of a quantum system and to characterise the information or uncertainty in a quantum state. We employ density operators in this, and the expression for entropy is $S(\rho) = -\text{Tr}(\rho \log \rho) = -\sum \lambda_i \log \lambda_i$, where λ_i are the eigenvalues of ρ . In quantum information, the Von Neumann entropy functions similarly to the Shannon entropy in conventional information.

F. Shannon Result

Shannon not only studied about noiseless coding theorem but he have also studied about the noisy coding theorem. According to the study work of Shannon there are two types of channel. These are:

- Noisy Channel: Noisy channel is the type of channel that actually introduces the error in the transmission and due to this error's we are unable to transfer the information correctly at the receiver. So, we started introducing the redundancies in the system so as to overcome the errors because if the errors are less we can transmit the data correctly without error. So the way of reducing error in the system is called Channel Coding.

- Noise-free Channel: Noiseless or Noise-free is the channel that does not have error in the transmission. Even if there are any error in the noise free channel it can be removed by multiple ways. In this we can even decompress the data at the receiver and this is called as Source coding.

Below figure is the model of a communication system,

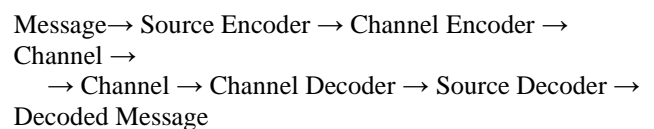


Fig .1. COMMUNICATION SYSTEM MODEL

For source coding, we can overcome the errors using hamming codes, Huffman code, Shannon Fano coding. These are the different types of error correcting codes that are used in source coding.

G. Shannon noisy coding theorem

The Shannon noiseless coding theorem have multiple ways by which the errors can be reduced while the

Shannon noisy coding theorem do not have any specific way to overcome the errors. So the sub-optimal codes like convolution codes are used to overcome the errors. The Shannon noisy coding theorem can be used for various uses like it is used for the purpose of describing the efficiency and it is also used for data corruption. The theorem can be used in both communication as well as for the purpose of data storage.

Shannon introduced a different way to follow the noise. In this the symbols that are the input to the channel are actually obtained from one of the input alphabet called X and the channel what it does is it further transfer it to one of the output alphabet, the alphabet which is at the output is represented as Y . So thus from the above discussion we can write the relation among all of them

X belongs $x \rightarrow$ channel \rightarrow y belong to Y

In this coding theorem we will consider discrete memoryless channel. The discrete memoryless channels are those who do not consider symbols it means the system will be independent. In Shannon theorem we are having two types of rate one is transmission rate which is represented by C and the other one is the fixed rate which is represented by R. To achieve small probability of error we need to satisfy the condition $R < C$, if we are satisfying this condition we can not only reduce the probability but we can also transmit the data error free. But if the condition is not satisfied then we cannot reduce the probability of error to a very small value.

II. PROBLEM FORMULATION

A. CQ coding problem

The Fig1. Shows the formulation of CQ coding problem. In this paper we discussing about the problem that arises if the probability of error is not very small as it is very much important to reduce the error in the system. In the CQ coding problem there are N messages that are encoded into binary strings ($\Phi_k(1), \dots, \Phi_k(n)$), where $k = 1, 2, \dots, N$. 0 bit is encoded as the mixed state $w(0)$. 1 bit is encoded as the mixed state $w(1)$. Since the CQ channel is Discrete Memoryless Channel, the kth message is encoded as the state

$$w(\Phi_k) = \sum_j^n w(\Phi_k(j))$$

Detection operators $\{Y_1, \dots, Y_k\}$ are to be constructed satisfying $0 \leq Y_j$, $\sum_{j=1}^k Y_j = I$. If the kth message is transmitted , then it is correctly decoded with the probability

$Tr(w(\phi_k)Y_k)$. The average error probability of decoded decision is

$$p_r(E) = \frac{1}{N} \sum_{k=1}^N Tr(w(\phi_k)(I - Y_k)).$$

We choose the code $C = \{\Phi_1, \dots, \Phi_N\}$ and decision operator Y_1, \dots, Y_N so that probability of error is minimum. We show via simulation studies using code search, that the minimum error probability is very small if $\log N/n < C$ and large when $\log N/n > C$ confirming the cq coding theorem, namely $\min Pr(E) \rightarrow 0$ as $n \rightarrow \infty$ if \lim

$\log N/n < C$ and $\min Pr(E) \rightarrow 1$ as $n \rightarrow \infty$ if $\lim \log N/n > C$ where

$$C = \min \left[\sum_{x=0,1} P(x) H(w(x)) - H \left(\sum_{x=0,1} P(x) w(x) \right) \right]$$

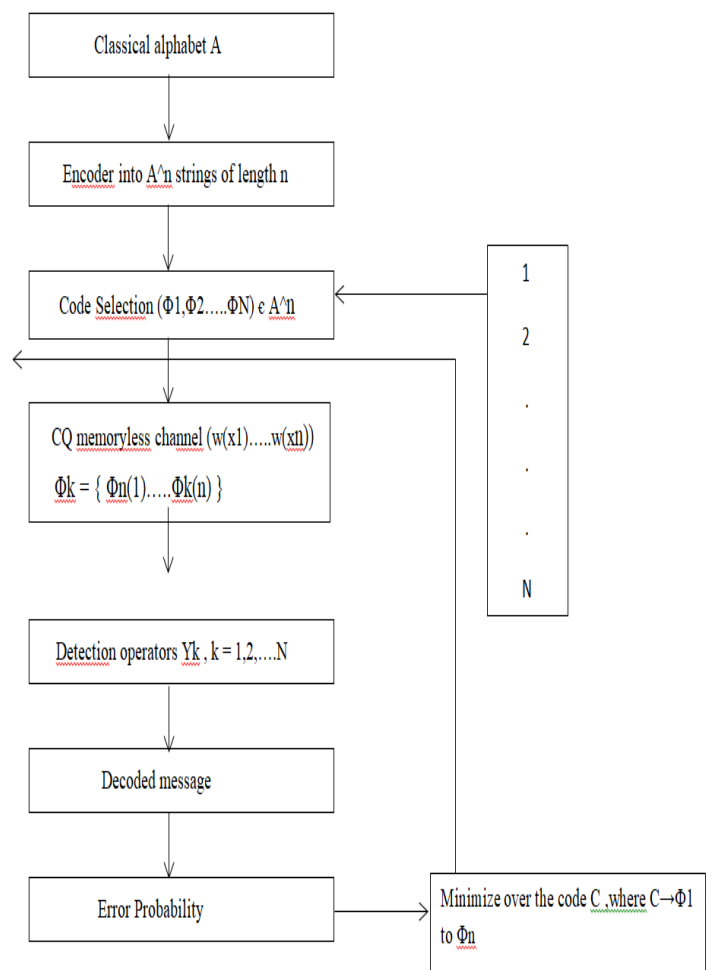


FIG. 2. FORMULATION OF CQ CODING PROBLEM

III. SIMULATION STUDIES

A. Algorithm

- Generate two density matrices W_0 and W_1 , ie, two positive definite matrices of unit trace, both 2×2 matrices

- Generate p_0, p_1 positive real numbers such that $p_0+p_1=1$.
- Choose a positive integer n large enough and generate the density matrix $W_p=p_0.W_0+p_1.W_1$ and $W_{pn}=W_p^{\otimes n}$, namely Kronecker product of W_p with itself n times.
- Generate all the 2^n sequences (i_1, i_2, \dots, i_n) of length n with $i_1, i_2, \dots, i_n=1, 2$.
- Calculate the density matrices $W(\phi(i)) = \text{kron}(W_{i_1}, \dots, W_{i_n}), i=1, 2, \dots, 2^n$.
- Calculate for each binary sequence for different values of n and obtain $\phi(i)$ for every sequences individually by obtaining the eigen decomposition, the projection operator $P(i) = \{W(\phi(i))\}^{\otimes N} W_p$, $i=1, 2, \dots, 2^n$ where N is a positive integer so that $\log(N)/n < C$ where $C = H(W_{pn}) - p_0 H(W_0) - p_1 H(W_1)$.
- Choose a subset E of the set of above 2^n sequences, consisting of N sequences, ie, E is a code of size N .
- Calculate $Y(i) = (\sum_{j \in E} P(j))^{-1/2} \cdot P(i) \cdot (\sum_{j \in E} P(j))^{-1/2}$, $i=1, 2, \dots, N$ for each $\phi(i)$ in E . Note that the elements of E are indexed by $\phi(1), \dots, \phi(N)$.
- Calculate the error probability for the coder-decoder $(\phi(i): i=1, 2, \dots, N, Y(i), i=1, 2, \dots, N)$ as

$$p_r(E) = \frac{1}{N} \sum_{k=1}^N \text{Tr}(w(\phi_k)(I - Yk))$$

- Repeat steps 7,8,9 by varying the set E of 2^n sequences and record the least error probability.
- Repeat 6,7,8,9 by taking $\log(N)/n > C$.
- Show that when $\log(N)/n < C$, the least error probability is much much smaller than when $\log(N)/n > C$.

B. Simulation results

In this research paper we have made a code according to the algorithm given above and results are obtained in numerical form, they are not in the form of graph or figures. So we have shown the results in the below given tables as the results obtained are numerical values. The results are obtained by varying the values of N as shown in the table I, table II, table III and table IV.

To perform the code we will consider different binary values and will obtain the error probabilities in each case. Firstly we are considering N to be equal to 5 it is defining the state of the coding theorem. As the value of N is 5 we will have code word in pairs of 5 and as n is also equal to 5 we will take five binary code words at a time as it is mentioned in the table I. W_0 and W_1 are the density matrices let $W_0=[0.5 \ 0; 0 \ 0.5]$, $W_1=[0.99 \ 0.02; 0.02 \ 0.01]$. The error probabilities will be obtained for every code word individually and will be compared among each other and we will choose the one which is smallest. Thus, the error probabilities obtained are listed below:

TABLE I: RESULT I

For $N=5, n=5, \text{Capacity}=0.6577$ (obtained)

Code word	Error Probability
00000-00100	0.7929
00101-01001	0.7687
01010-01110	0.8558
01111-10011	0.8600
10100-11000	0.7272
11001-11101	0.9096

- In the next case we are considering N to be equal to four. As the value of N is 4 we will have code word in pairs of four and as n is also equal to 4 we will take four binary code words at a time and as the maximum binary value that we can achieve in pairs of four is from 0000-1111 we will consider them in this case. The error probabilities in each case is mentioned in the table II. The density matrices will remain same in this case as above:

TABLE II: RESULT II

For $N=4, n=4, \text{Capacity}=0.5090$ (obtained)

Code word	Error Probability
0000-0011	0.8179
0100-0111	0.8818
1000-1011	0.8224
1100-1111	0.6252

- To check error probability in third case we are considering N to be equal to three. As the value of N is 3 we will have code word in pairs of three and as n is also equal to 3 we will take three binary code words at a time, it will range from 000 to 111 but as we need to make pair of 3 and in last case it is not possible so we have considered only two cases. The error probabilities for both the cases will be as it is mentioned in the table III. The density matrices will remain same in this case as above:

TABLE III: RESULT III

For $N=3, n=3, \text{Capacity}=0.3602$ (obtained)

Code word	Error Probability
000-010	0.8287
011-101	0.5895

- In the last case we will take N to be equal to two and will take the code words in pair of two. The obtained error probability are listed below in the

REFERENCES

Table IV. The density matrices will remain same in this case also. From the obtained error probabilities we will choose the one with the least error probability as we want our error to be as small as possible because the system with low error is considered to be the best. Thus, the error probabilities obtained in case 4 are:

TABLE IV: RESULT IV
For N=2, n=2, Capacity=0.3602 (obtained)

Code word	Error Probability
00-01	0.6295
10-11	0.6068

- [1] Charles H. Bennet , Peter W. Shor (1998), "Quantum information theory" , IEEE Transactions on Information Theory, VOL. 44 , NO. 6, OCTOBER 1998.
- [2] Masahito Hayashi , ed (2006) Quantum Information Theory : Mathematical Foundation DOI 10.1007/978-3-662-49725-8 , Japan , Springer-Verlag Berlin Heidelberg.
- [3] Dan C. Marinescu , Gabriela M. Marinescu , "Classical and Quantum information" Academic Press (2011).
- [4] Ranjan Bose , "Information theory, coding and cryptography", McGraw Hill Education(2008).
- [5] Shu Lin/Daniel J.Costello, Jr." Error control coding: Fundamentals and applications Prentice-Hall International, Hemel Hempstead, Herts., U.K., 1982.
- [6] Mark M.Wilde , "Quantum Information Theory" , DOI.10.1017/978131680997 , Cambridge University Press , February 2017.

IV. DISCUSSION

We have verified the channel coding theorem for two states for binary alphabet it can be extended upto M-ary alphabet in the similar way. In order to verify it for longer strings we need a supercomputer because the search involves so much. We cannot do it using ordinary computer, with the current facilities we are not able to address that problem currently.

V. CONCLUSION

In Noisy coding theorem , the channel itself is noisy in nature because of which the data gets corrupted. In this paper we have minimized this problem by obtaining the smallest value of error probability after comparing it with different values of N. For a single value of N we took multiple binary sequences as a pair of two, three, four and five depending on the value of N. Based on the calculation the least value is obtained for N = 4 , the error probability is 0.6252. Further, it is compared with the capacity so as to satisfy the condition $\log N/n < C$ which is achieved in both the cases for N=4 and N=5 but as we need to choose the smallest error probability so we selected 0.6252 as the error probability . So, thus we have successfully reduce the error in the channel.

ACKNOWLEDGMENT

Archana Chaudhary is grateful to Professor Harish Parthasarathy for his time to time guidance, support at various stages in the task of developing this project.