The Asymptotic information rate function in coding theory

A Thesis

submitted to
Indian Institute of Science Education and Research Pune
in partial fulfillment of the requirements for the
BS-MS Dual Degree Programme

by

Ajinkya Ramdas Gaikwad

Indian Institute of Science Education and Research Pune
Dr. Homi Bhabha Road,
Pashan, Pune 411008, INDIA.

April, 2019

Supervisor: Dr. Krishna Kaipa

Certificate

This is to certify that this dissertation entitled The Asymptotic information rate function in coding theorytowards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Ajinkya Ramdas Gaikwad at Indian Institute of Science Education and Research under the supervision of Dr. Krishna Kaipa, Assistant Professor, Department of Mathematics , during the academic year 2018-2019.

Dr. Krishna Kaipa

Committee:

Dr. Krishna Kaipa

Dr. Anindya Goswami

This thesis is dedicated to my parents and my brother

Declaration

I hereby declare that the matter embodied in the report entitled The Asymptotic information rate function in coding theory are the results of the work carried out by me at the Department of Mathematics, Indian Institute of Science Education and Research, Pune, under the supervision of Dr. Krishna Kaipa and the same has not been submitted elsewhere for any other degree.

Ajinkya Ramdas Gaikwad

Acknowledgments

I am very grateful to my supervisor Prof. Krishna Kaipa for patiently guiding and encouraging me throughout this thesis. I am thankful to my TAC member Prof. Anindya Goswami for his constant support. Finally, I thank my family and friends for all their love and care.

Abstract

Transmission of information through a unreliable communication channel introduces noise in the data. The process of introducing redundancy in this data in order to recover it later is called an error correcting code. Error correcting codes are frequently used for reliable storage in CD's, DVD's, SSD's and in many other cases. A good code must have high error correcting capacity and efficiency which are measured by relative minimum distance and rate of transmission respectively, however these two requirements are mutually conflicting. The trade-off between these quantities is studied through the question of finding the largest size of the code for minimum distance at least d. In general, it is hard to calculate the exact value of this quantity. Therefore, it is important to have good upper and lower bounds for the same. This question become more tractable by defining the asymptotic rate function. It is the asymptotic version of the above quantity which captures the trade-off between these quantities. This in turn raises more mathematical questions about the properties of asymptotic rate function like convexity and differentiability.

The distance enumerator polynomial determines the distance distribution of a code. Binomial moments provide an alternate way to deal with the distance enumerator polynomial which helps in finding bounds on the size of the code. The most intuitive upper bound on the size of the code is sphere packing bound or Hamming bound. The codes attaining this bound are called perfect codes. The problem of existence of perfect codes is a very interesting problem which is still unresolved for case of double error correcting perfect codes. In this thesis, we attempt the above problem by generalizing the differential equation derived by Lloyd in [5] for binary case to find necessary conditions on the existence of perfect codes. We will also present a different proof of the non-linear Mac-Williams identities using binomial moments. The bounds on the size of the code are discussed in Chapter 1 and later we have used them to derive bounds and properties of the asymptotic rate function. We study the recent improvement in Elias bound on the asymptotic rate function due to K. Kaipa. In Chapter 4, We discuss about the open problem of $\cup-$convexity property of the asymptotic rate function.

Contents

Introduction

If we transmit a set of messages through a unreliable or noisy communication channel then some of the data might get corrupted due to noise. To avoid it, we do not send the data in exact form. The main idea of error correcting codes is to add some redundant information to the data which helps us recover the information that was transmitted in the first place without re-transmission. This enables the receiver to correct errors but this is achieved at cost of decrease in rate of transmission.

To translate it mathematically, let us assume that we have a set of messages $\mathcal{M}$. We denote the size of a set $\mathcal{M}$ by M. We also have a set $\mathcal{F}$ of size q. The set $\mathcal{F}$ is called as alphabet set. The set $\mathcal{F}^n$ is a set of words of length n where the letters in the word are coming from a set $\mathcal{F}$. The Hamming distance in space $\mathcal{F}^n$ between any two words x, y is number of positions in which x and y differ. We assume that the set $\mathcal{M}$ is either the full space $\mathcal{F}^k$ or some subset of it where k is a positive integer less than n. Next, we take a one-one map from the set $\mathcal{M}$ to the space $\mathcal{F}^n$. The image of this map is called a code of length n and size M and it is represented as $\mathcal{C}$. Elements in the set $\mathcal{C}$ are called codewords. The reason behind this mapping is to add redundant bits to the data to correct errors in the information due to noise. The minimum distance of a code $\mathcal{C}$ is a least distance between any two distinct elements x, y from a code $\mathcal{C}$. The code $\mathcal{C}$ of length n, minimum distance d and size M over alphabet of size q is denoted as a $[n, M, d]_q$ code $\mathcal{C}$. The two important parameters related to code are relative minimum distance ($\delta_{\mathcal{C}} = d/n$) and rate of transmission ($\mathcal{R}_{\mathcal{C}}$) which determines the error correcting capacity and efficiency of the code. The rate of transmission is defined as

$$\mathcal{R}_{\mathcal{C}} = \frac{\log_q M}{n}$$

For a good code, we need $\delta_\mathcal{C}$ and $\mathcal{R}_\mathcal{C}$ to be high but this is a mutually conflicting requirement. We are interested in the problem of finding the maximum rate of transmission of a code such that the relative minimum distance is at least δ. It depends on the largest size of the code in space $\mathcal{F}^n$. Let $A_q(n, d)$ denotes the maximum value of M such that $[n, M, d]_q$ code exists. The exact formula of $A_q(n, d)$ for general n and d is unknown and therefore it is important to have good upper and lower bounds for it. In Chapter 1, we discuss the bounds and properties of $A_q(n, d)$.

The distance distribution of a code is expressed in the distance enumerator polynomial of a code. The binomial moments provide an alternate way of studying the distance distribution of a code and therefore it help us in finding bounds on the size of the code. In Chapter 2, we study properties of binomial moments and non-linear Mac-Williams identities which led to a linear programming bound on the rate of transmission of a code given in [7] for binary case. Hamming bound or sphere packing bound is the simplest upper bound on the size of the code and it is given as

$$M \leq \frac{q^n}{V_q(n, e)}$$

where $e = \frac{d-1}{2}$ and $V_q(n, r)$ is a size of ball of radius r. The codes attaining the Hamming bound are called perfect codes. The problem of non-existence of perfect codes for $e \geq 3$ is settled for any size of alphabet. Hamming codes and Golay codes are the only known examples of non-trivial perfect codes. The question is unresolved for $e = 2$ case, when size of the alphabet is non-prime power. In 1956, Lloyd gave a strong necessary condition on existence of binary perfect codes for general n and d in [5]. In Chapter 3, we will generalize the differential equation derived by Lloyd for binary case to any size of the alphabet to discuss about the unresolved problem of existence of double error correcting perfect codes for any size of the alphabet.

If a perfect code exists for some n, d and q then in that case we know the exact value of $A_q(n, d)$ but perfect codes exist in only some cases. As we do not know the exact value of $A_q(n, d)$, it is hard to calculate the maximum rate of transmission. This problem becomes tractable by defining the asymptotic rate function $\alpha_q(x)$ which captures the trade-off between $\delta_\mathcal{C}$ and maximum rate of transmission. The asymptotic rate function $\alpha_q(x) : [0, 1] \to [0, 1]$ is given as

$$\alpha_q(x) = \limsup_{n\to\infty} \frac{\log_q A_q(n, xn)}{n}$$

The exact value of $\alpha_q(x)$ is unknown. In Chapter 4, we use bounds and properties of $A_q(n, d)$ to study the asymptotic rate function. We also study the recent improvement in Elias bound due to K. Kaipa in [4]. The question of $\cup-$convexity property of $\alpha_q(x)$ is an open problem. It is discussed at the end of the Chapter 4.

Chapter 1

Error correcting codes in Hamming metric

1.1 Main problem in coding theory

We will begin this section by presenting the terminology used in the following chapters. In the Hamming space $\mathcal{F}^n$, the Hamming metric is defined as follows

Definition 1.1.1. *Hamming distance: The Hamming distance between any two words x, y is given as*

$$d(x, y) = \textit{ Number of positions at which } x \textit{ and } y \textit{ differ}$$

For any code $\mathcal{C}$, we define minimum distance $d(\mathcal{C})$ as

Definition 1.1.2. *Minimum distance: The minimum distance of a code $\mathcal{C}$ is the least distance between any distinct two codewords x and y such that $x, y \in \mathcal{C}$.*

A code $\mathcal{C}$ of size M and minimum distance d in a space $\mathcal{F}^n$ where $|\mathcal{F}| = q$ is represented as $[n, M, d]_q$ code. To make our notations compact, without loss of generality we assume that if $|\mathcal{F}| = q$ then $\mathcal{F}$ is $\mathbb{Z}/q\mathbb{Z}$. We will only use additive structure of $\mathbb{Z}/q\mathbb{Z}$.

Definition 1.1.1. *Hamming weight: For any word x in $\mathcal{F}^n$, Hamming weight of x is non-zero entries in x.*

Hamming distance between any two words x, y can be defined in terms of hamming weight in a following way

$$d(x, y) = \#\text{non-zero entries in } x - y$$

There are two important parameters for a $[n, M, d]_q$ code $\mathcal{C}$ and they are relative minimum distance and rate of transmission.

Definition 1.1.2. *Relative minimum distance: The relative minimum distance for a code $\mathcal{C}$ is d/n and it is represented as $\delta_{\mathcal{C}}$.*

For a code $\mathcal{C}$, the error correcting capacity is measured by relative minimum distance. For a good code, we need high error correcting capacity and therefore high relative minimum distance.

Definition 1.1.3. *Rate of transmission: The rate of transmission of a code $\mathcal{C}$ is*

$$R_{\mathcal{C}} = \frac{\log_q M}{n}$$

Rate of a transmission is a rate at which information is transmitted. For a good code, we need a rate of transmission to be high. The important question in coding theory is that for a given code with fixed minimum distance d in space $\mathcal{F}^n$, what is the maximum rate of transmission can be achieved. This question can be answered by counting the maximum size of a code in space $\mathcal{F}^n$ with minimum distance d.

Definition 1.1.4. *$A_q(n, d)$ represents the maximum value of M such that a $[n, M, d]_q$ code exists.*

It is very hard to calculate the exact value of $A_q(n, d)$ for general n and d. So, the next best thing to do is to calculate good upper and lower bounds.

1.2 Bounds on the largest size of the code

In this section, we state and prove bounds on the size of the code. We also study basic properties of $A_q(n,d)$ as a function of n and d.

Lemma 1.2.1. *Hamming bound: Let $\mathcal{C}$ be any $[n,M,d]_q$ code then*

$$M \leq \frac{q^n}{V_q(n, \lfloor \frac{d-1}{2} \rfloor)} \tag{1.1}$$

where $V_q(n,r)$ is a volume of ball of radius r in space $\mathcal{F}^n$.

Proof. If we draw a ball of radius $\lfloor \frac{d-1}{2} \rfloor$ around every codeword then we observe that these balls are disjoint because least distance between any two codewords is d. If any two balls with centre c_1 and c_2 have a intersection then we can easily prove that the $d(c_1,c_2) < d$ which is a contradiction. □

Hamming bound is also known as sphere packing bound. Codes which attain equality in Hamming bound are called perfect codes. Hamming bound was based on the fact that the diameter of balls was less than d. This bound can be generalized by defining anticodes.

Definition 1.2.1. *Anticode of diameter D is any subset $\mathcal{L}$ of $\mathcal{F}^n$ such that distance between any two elements in $\mathcal{L}$ is at most D.*

Lemma 1.2.2. *Anticode bound: Let $\mathcal{C}$ be any $[n,M,d]_q$ code and $\mathcal{L}$ be any anticode of diameter $d-1$ then*

$$M \leq \frac{q^n}{|\mathcal{L}|} \tag{1.2}$$

Proof. Let us assume that set $\mathcal{L}$ contains origin. If not then we can translate it such that it contains origin. For every $c \in \mathcal{C}$, consider a set $\{c + \mathcal{L} : c \in \mathcal{C}\}$. We will show that any two sets of the form $\{c_1 + \mathcal{L}\}$ and $\{c_2 + \mathcal{L}\}$ are disjoint using contradiction. Let us assume there's exist c_1, l_1, c_2, l_2 such that $c_1 + l_1 = c_2 + l_2$. Therefore using $c_1 - c_2 = l_2 - l_1$, we get

$$d(c_1,c_2) = wt(c_1 - c_2) = wt(l_2 - l_1) = d(l_2 - l_1) \leq d - 1$$

which is a contradiction. It implies that $M|\mathcal{L}| \leq q^n$. □

Lemma 1.2.3. *Singleton bound: For a $[n, M, d]_q$ code $\mathcal{C}$,*

$$M \leq q^{n-d+1} \tag{1.3}$$

Proof. We note that $\mathcal{L} = \mathcal{F}^{d-1} \times \{\bar{0}\}$ a subset of $\mathcal{F}^n$ is an aniticode of diameter $d-1$. Using it in lemma (1.2.2) proves the result. □

Lemma 1.2.4. *Plotkin bound: For a $[n, M, d]_q$ code $\mathcal{C}$,*

$$M \leq \left\lfloor \frac{1}{1-(\theta/\delta)} \right\rfloor \tag{1.4}$$

provided $\delta > \theta$ where $\theta = 1 - q^{-1}$ and $\delta = d/n$.

Proof. We know that the minimum distance between any codewords is d. It implies that the average distance between pair of distinct codewords is also at least d.

$$dM(M-1) \leq \sum_{(c,c') \in \mathcal{C} \times \mathcal{C}} d(c, c') \tag{1.5}$$

Right hand size of the above inequality can be written as

$$\sum_{(c,c') \in \mathcal{C} \times \mathcal{C}} d(c, c') = nM^2 - \sum_{i=1}^{n} \sum_{v \in \mathcal{F}} \#\{(c, c\prime) \in C \times C : c_i = c_i' = v\}$$

$$= nM^2 - \sum_{i=1}^{n} \sum_{v \in \mathcal{F}} n(i, v)^2$$

where $n(i, v) = \#\{c \in \mathcal{C} : c_i = v\}$. We also note that $\sum_{v \in \mathcal{F}} n(i, v) = M$. Using Cauchy-Schwartz inequality, we can write

$$\frac{\sum_{v \in \mathcal{F}} n(i, v)^2}{q} \geq \left[\frac{\sum_{v \in \mathcal{F}} n(i, v)}{q}\right]^2 = \frac{M^2}{q^2}$$

Therefore, we get

$$\sum_{(c,c')\in\mathcal{C}\times\mathcal{C}} d(c,c') \leq nM^2 - \frac{nM^2}{q} = nM^2\theta \tag{1.6}$$

Using (1.5) and (1.6), We can write

$$d(M-1) \leq nM\theta$$

It implies that $\frac{1}{M} \geq 1 - \frac{\theta}{\delta}$. we know that $\theta < \delta$, so taking reciprocal and using the fact that M is an integer, we get

$$M \leq \left\lfloor \frac{1}{1-(\theta/\delta)} \right\rfloor$$

□

Lemma 1.2.5. *Gilbert-Varshamov bound: Let $\mathcal{C}$ be a code of minimum distance d in $\mathcal{F}^n$ of size $A_q(n,d)$ then*

$$A_q(n,d) \geq \frac{q^n}{V_q(n,d-1)} \tag{1.7}$$

Proof. Let us assume that $\exists\ x \in \mathcal{F}^n \backslash \mathcal{C}$ such that $d(x,c) \geq d \ \ \forall c \in \mathcal{C}$. If we add word x to the code then the new code $\mathcal{C} \cup \{x\}$ will be of size greater than $A_q(n,d)$ with minimum distance d which is a contradiction. It implies that for every word in $\mathcal{F}^n$ there's exists at least one codeword in $\mathcal{C}$ such that $d(x,c) \leq d-1$. Therefore, if we draw a ball of radius $d-1$ around every codeword then the set of balls will exhaust the whole space but note that the balls are not necessarily disjoint. This argument proves the Gilbert-Varshamov bound. □

Next, we prove some basic properties of $A_q(n,d)$.

Theorem 1.2.6. 1. *$A_q(n,d)$ is an increasing function of n and decreasing function of d.*

$$A_q(n,d) \geq A_q(n-1,d) \geq A_q(n,d+1)$$

2. *$A_q(n, {}_{\geq}d) = A_q(n,d)$ where $A_q(n, {}_{\geq}d)$ is the largest size of the code with minimum at least d and length n.*

Proof. For the first part let us assume that we have a code $\mathcal{C}$ of size $A_q(n-1,d)$ of minimum distance d in space $\mathcal{F}^{n-1}$. Consider a new code $\mathcal{C}'$ where we add an extra coordinate to every codeword such that $\mathcal{C}' = \{(c,0) : c \in \mathcal{C}\}$. The new code $\mathcal{C}'$ is a $[n, A_q(n-1,d), d]_q$ code. Using the definition of $A_q(n,d)$, we get the first inequality. For the second inequality, let us assume that we have a $[n, A_q(n,d+1), d+1]$ code $\mathcal{C}''$. There must exists a pair of codewords $(c,c') \in \mathcal{C}'' \times \mathcal{C}''$ such that $d(c,c') = d+1$. Consider the i-th coordinate such that $c_i \neq c_i'$ and define a projection map $\phi : \mathcal{F}^n \to \mathcal{F}^{n-1}$ such that i-th coordinate is dropped. The image of the code under the map ϕ generates new $[n-1, A_q(n,d+1), d]$ code. Again using the definition of $A_q(n,d)$, we get the required result.

For the second part, we know that $A_q(n,d) \leq A_q(n,_{\geq} d)$. To prove the opposite inequality let us assume a $[n, A_q(n,d{+}t), d{+}t]$ code. Using the above idea of projection mapping and adding an extra fixed coordinate, we can generate a new $[n, A_q(n,d+t), d+t-1]$ code. Repeating the process t times gives a $[n, A_q(n,d+t), d]$ code. It implies that $A_q(n,d) \geq A_q(n,_{\geq} d)$. □

The next lemma is generalization of anticode bound lemma (1.2.2) and it is called Bassalygo-Elias lemma.

Lemma 1.2.7. *Bassalygo-Elias lemma:*

$$A_q(n,d) \leq \frac{q^n A_q(n,d;\mathcal{L})}{|\mathcal{L}|} \tag{1.8}$$

where $|\mathcal{L}|$ is size of the set $\mathcal{L}$ and $A_q(n,d;\mathcal{L})$ represents largest possible size of code of length n and minimum distance at least d.

Proof. Consider a $[n, M, d']_q$ code $\mathcal{C}$ such that d' is at least d. For every fixed word $v \in \mathcal{F}^n$, we get a new code $(\mathcal{C}+v) \cap \mathcal{L}$ such that it is contained in $\mathcal{L}$ and minimum distance is at least d. It implies that $A_q(n,d;\mathcal{L})$ must be greater than or equal to average of $\#\{(\mathcal{C}+v)\cap\mathcal{L}\}$ over $v \in \mathcal{F}^n$. Therefore, we can write the following

$$q^n.A_q(n,d;\mathcal{L}) \geq \sum_{v\in\mathcal{F}^n} |(\mathcal{C}+v)\cap\mathcal{L}| \tag{1.9}$$

We can rewrite RHS as

$$\sum_{v\in\mathcal{F}^n} |(\mathcal{C}+v)\cap\mathcal{L}| = \sum_{c\in\mathcal{C}}\sum_{v\in\mathcal{F}^n}\begin{cases}1, & \text{if } c+v\in\mathcal{L}\\ 0, & \text{otherwise}\end{cases}$$

For every fixed $c\in\mathcal{C}$, there are $|\mathcal{L}|$ number of $v's$ such that c can be translated to a word in $\mathcal{L}$. It implies that

$$\sum_{v\in\mathcal{F}^n} |(\mathcal{C}+v)\cap\mathcal{L}| = |\mathcal{C}||\mathcal{L}|$$

Substituting above value in equation (1.9) and using theorem(1.2.6), we get

$$\frac{A_q(n,d;\mathcal{L})q^n}{|\mathcal{L}|} \geq M \geq A_q(n,d)$$

□

Lemma 1.2.8. *Cross-sectional bound:*

$$A_q(n,d) \leq A_q(n-t,d)q^t \tag{1.10}$$

where $t\in\{0,1,\dots,n\}$. If $n-t<d$ then we take $A_q(n-t,d)=1$.

Proof. If we substitute $\mathcal{L}=\mathcal{F}^{n-t}\times\{\bar{0}\}\subset\mathcal{F}^n$ in the Bassalygo-Elias lemma then we get the required result. □

In the next chapter, We will discuss about binomial moments which help us calculate bounds on the size of the code.

Chapter 2

Binomial Moments

In this chapter, we will discuss about properties of binomial moments which is an alternate description for distance enumerator polynomial. We will also present a different proof of non-linear Mac-Williams identities using binomial moments. The non-linear Mac-Williams identities led to an linear programming upper bound on the rate of transmission of the code given in [7] for binary case.

2.1 Relation to distance enumerator polynomial

We will begin this section by defining a distance enumerator polynomial for a $[n, M, d]_q$ code $\mathcal{C}$.

Definition 2.1.1. *The distance enumerator polynomial for code $\mathcal{C}$ is defined as*

$$W_{\mathcal{C}}(Z) = \sum_{i=0}^{n} A_i Z^i \tag{2.1}$$

where $A_i = \frac{1}{M}\#\{(c_1, c_2) \in \mathcal{C} \times \mathcal{C} : d(c_1, c_2) = i\}$.

Let $I_{r,n}$ is a set of all possible r positions out of n positions. It is given as $I_{r,n} = (i_1, i_2, \ldots, i_r) : 1 \leq i_1 < i_2 < \cdots < i_r \leq n\}$. The r-th binomial moment is denoted by $\gamma_r(\mathcal{C})$. We have mentioned two equivalent forms of $\gamma_r(\mathcal{C})$ in the definition.

Definition 2.1.2. *The r-th binomial moment for a code $\mathcal{C}$ is*

$$\gamma_r(\mathcal{C}) = \frac{1}{M\binom{n}{r}} \sum_{I \in I_{r,n}} \#\{(c_1, c_2) \in \mathcal{C} \times \mathcal{C} : c_1|_I = c_2|_I\} \tag{2.2}$$

$$\gamma_r(\mathcal{C}) = \frac{1}{\binom{n}{r}} \sum_{j=0}^{n} \binom{n-j}{r} A_j \tag{2.3}$$

This gives a linear transformation from $A_r's$ to $\gamma_r's$. Consider a matrix $T \in GL(n+1, \mathbf{Z})$ such that the elements in T are given as $T_{ij} = \binom{n+1-i}{n+1-j}$. We can write a relation between $A_r/\binom{n}{r}$ and γ_r in a following way

$$T \begin{bmatrix} A_n/\binom{n}{n} \\ A_{n-1}/\binom{n}{n-1} \\ \vdots \\ A_0/\binom{n}{0} \end{bmatrix}_{(n+1)\times 1} = \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_n \end{bmatrix}_{(n+1)\times 1}$$

such that, we get

$$\gamma_r = \sum_{j=0}^{n-r} \frac{\binom{n-r}{j} A_j}{\binom{n}{j}} \tag{2.4}$$

Inverse of the linear transformation T maps $\gamma_r's$ to $A_r's$. The elements of T^{-1} are given as $(T^{-1})_{ij} = (-1)^{i+j}\binom{n+1-i}{n+1-j}$. This gives a expression for $A_r/\binom{n}{r}$ in terms of $\gamma_r's$.

$$\frac{A_{n-r}}{\binom{n}{n-r}} = \sum_{k=r}^{n} (-1)^{k+r} \binom{n-r}{n-k} \gamma_k \tag{2.5}$$

Some of the properties of binomial moments are deeply motivated from Mac-Williams identities. The Mac-Williams identities are discussed in the next section.

2.2 Mac-Williams identities

Definition 2.2.1. *Linear codes: Any k dimensional subspace of a n dimensional vector space over any field* F *of size q is called $[n,k]_q$ linear code.*

A common way to construct a $[n,k]_q$ linear code $\mathcal{C}$ is by giving a linear map from a space F^k to space F^n. The matrix G corresponding to the linear map is a full rank matrix of size $k \times n$. The range space of G is the linear code $\mathcal{C}$ and the matrix is called generator matrix. The same can be achieved by taking a full rank matrix H of size $(n-k) \times n$ and the kernel space of the matrix will be the linear code $\mathcal{C}$. In this case, H is called parity check matrix. The code generated by parity check matrix H is called dual of a code $\mathcal{C}$ and it is denoted as $\mathcal{C}^{\perp}$.

Definition 2.2.2. *Dual code: A dual of a $[n,k]_q$ linear code $\mathcal{C}$ is a $n-k$ dimensional subspace of* F^n *denoted by $\mathcal{C}^{\perp}$ such that* $\mathcal{C}^{\perp} = \{\bar{x} \in \mathrm{F}^n : <\bar{x}, \bar{c}>= 0 \ \forall \ \bar{c} \in \mathcal{C}\}$.

Linear codes have an interesting property that corresponds to their distance distribution known as distance invariant property.

Definition 2.2.3. *Distance invariant code: A code is said to be distance invariant if the number of codewords at distance k from a fixed codeword depends only on k and not on the fixed codeword.*

Lemma 2.2.1. *Linear codes are distance invariant.*

Proof. Let us assume that $A_k(\mathcal{C}; x) = \#\{c \in \mathcal{C} : d(c,x) = k\}$. As the linear combination of codewords is again a codeword implies that $c - x = c'$ is also a codeword. It implies that $A_k(\mathcal{C}; x) = \#\{c' \in \mathcal{C} : wt(c') = k\}$ which is independent of x. □

Let us assume that we have a $[n,k]_q$ linear code $\mathcal{C}$ with distance enumerator polynomial

$$W_{\mathcal{C}}(Z) = \sum_{i=0}^{n} A_i Z^i$$

For linear codes, it is easy to observe that A_i in distance enumerator polynomial become number of codewords of weight i due to distance invariant property. The distance enumerator polynomial of dual of a code $\mathcal{C}$ is represented as

$$W_{\mathcal{C}^\perp}(Z) = \sum_{i=0}^{n} A_i^\perp Z^i$$

Theorem 2.2.2. *For linear codes, Mac-Williams identities gives a transformation from $W_{\mathcal{C}}(Z)$ to $W_{\mathcal{C}^\perp}(Z)$ in a following way*

$$W_{\mathcal{C}^\perp}(Z) = \frac{(1+(q-1)Z)^n}{\#\mathcal{C}} W_{\mathcal{C}}\left(\frac{1-Z}{1+(q-1)Z}\right) \tag{2.6}$$

Proof. We will prove the Mac-Williams identities using binomial moments. For $[n,k]_q$ linear code $\mathcal{C}$, definition of binomial moment can be modified to

$$\gamma_r(\mathcal{C}) = \frac{1}{\binom{n}{r}} \sum_{I \in I_{r,n}} \#\{c \in \mathcal{C} : c|_I = 0\} \tag{2.7}$$

Let G be the generator matrix of code $\mathcal{C}$. If $I \in I_{r,n}$ then G_I represents a submatrix of size $k \times r$. We have a one-one map $\phi : \mathrm{F}^k \to \mathrm{F}^n$ given as $\phi(x) = xG$. Assume that rank(G_I)=r_I.

$$\#\{c \in \mathcal{C} : c|_I = 0\} = \#\{x \in \mathrm{F}^k : xG_I = 0\} = \text{size of } \ (kerG_I^T) = q^{k-r_I}$$

Now, for $I \in I_{r,n}$ assume that $I^\perp \in I_{n-r,n}$ such that $I^\perp$ is complement of I.

$$\#\{c \in \mathcal{C}^\perp : c|_{I^\perp} = 0\} = \#\{v \in \mathrm{F}^n : Gv = 0 \ \text{ and } \ v|_{I^\perp} = 0\} = \#(ker(G_I)) = q^{r-r_I}$$

From above two equations and (2.7), we observe that $\gamma_r(\mathcal{C}^\perp) = q^{n-r-k}\gamma_{n-r}(\mathcal{C})$.

Next, we know that if we expand polynomial $f(Z)$ around $Z=1$ then $f(Z) = \sum_i a_i (Z-1)^i$ where $a_i = \frac{1}{i!}\frac{d^i f}{dZ^i}\big|_{Z=1}$. Taking $f(Z) = Z^n W_{\mathcal{C}}(\frac{1}{Z})$, we get

$$a_i = \frac{1}{i!}\frac{d^i}{dZ^i} Z^n W_{\mathcal{C}}(1/Z)\bigg|_{Z=1} = \sum_{j=0}^{n} \binom{n-j}{i} A_j = \binom{n}{i}\gamma_i(\mathcal{C})$$

Using the expansion of $Z^n W_{\mathcal{C}^\perp}(1/Z)$, we have

$$Z^n W_{\mathcal{C}^\perp}(1/Z) = \sum_{i=0}^{n} \binom{n}{i} \gamma_i(\mathcal{C}^\perp)(Z-1)^i = \sum_{i=0}^{n} \binom{n}{i} q^{n-i-k} \gamma_{n-i}(\mathcal{C})(z-1)^i$$

Substituting $1/Z = X$ and adjusting the summation, we get

$$W_{\mathcal{C}^\perp}(X) = \frac{(qX)^n}{q^k} \sum_{j=0}^{n} \binom{n}{j} \gamma_j(\mathcal{C}) \left(\frac{1-X}{qX} \right)^{n-j}$$

$$= \frac{(1-X)^n}{q^k} \sum_{j=0}^{n} \binom{n}{j} \gamma_j(\mathcal{C}) \left(\frac{qX}{1-X} \right)^{j}$$

$$= \frac{(1-X)^n}{q^k} \sum_{j=0}^{n} \binom{n}{j} \gamma_j(\mathcal{C}) \left(\frac{1+(q-1)X}{1-X} - 1 \right)^{j}$$

$$= \frac{(1-X)^n}{q^k} \left(\frac{1+(q-1)x}{1-x} \right)^n \sum_{j=0}^{n} \binom{n}{j} \gamma_j(\mathcal{C}) \left(\frac{1+(q-1)X}{1-X} - 1 \right)^{j}$$

$$= \frac{(1+(q-1)X)^n}{\#\mathcal{C}} W_{\mathcal{C}} \left(\frac{1-X}{1+(q-1)X} \right)$$

□

In the next section, we will use Mac-Williams identities to derive properties of binomial moments.

2.3 Properties of Binomial Moments

Let $\mathcal{C}$ be any non-trivial $[n, M, d]_q$ code (Size of a code is strictly greater than one).

Lemma 2.3.1. *$\gamma_r(\mathcal{C}) \geq 1$ with equality if and only if $n - d(\mathcal{C}) < r \leq n$.*

Proof. Notice that in the definition of $\gamma_r(\mathcal{C})$ in (2.2), For every $I \in I_{r,n}$ there's exist at least M pairs of codewords of the form (c, c) such that $\{(c, c) \in \mathcal{C} \times \mathcal{C} : c|_I = c|_I\}$ which proves the first part. For the second part, let us assume that $\gamma_r(\mathcal{C}) > 1$. This is true if and only if for some $I \in I_{r,n}$ there exists at least one pair of distinct codewords such that $\{(c_1, c_2) \in \mathcal{C} \times \mathcal{C} : c_1|I = c_2|I\}$. Two distinct codewords can have same entries on at most $n-d$ positions otherwise distance between them will be less than d. It implies that $r \leq n - d$. □

Lemma 2.3.2. *$\gamma_0 > \gamma_1 > \cdots > \gamma_{n-d} > 1$.*

Proof. We define a new quantity β_r in a following way

$$\beta_r \coloneqq (\gamma_r - \gamma_{r+1}) - \binom{n-r-1}{d-1}(\gamma_{n-d} - 1)$$

where $0 \leq r \leq n - d - 1$. It is enough to prove that $\beta_r > 0$. Using value of γ_i from equation (2.3) we can rewrite β_r as

$$\beta_r = \sum_{i=0}^{n-d-r-1} \frac{A_{n-r-i}\binom{n-r-1}{i}}{\binom{n}{n-r-i}}$$

We already know that $A_r^{\prime}s$ are non-negative which implies that $\beta_r > 0$.

□

Lemma 2.3.3. *$\gamma_i(\mathcal{C}) \geq M/q^i$. Also there exists a unique integer $d^{\perp}(\mathcal{C})$ such that $\gamma_i(\mathcal{C}) = M/q^i$ if and only if $i \leq d^{\perp}(\mathcal{C}) - 1$. In case of linear code, we note that $d^{\perp}(\mathcal{C}) = d(\mathcal{C}^{\perp})$.*

Proof. For every $I = \{1 \leq i_1 < i_2 < \cdots < i_r \leq n\}$, we define a map $\phi_I : \mathcal{C} \to \mathcal{F}^r$ such that $\phi_I(c) = \{c_{i_1}, c_{i_2}, \ldots, c_{i_r}\}$. Using a definition (2.2), we can write

$$M\binom{n}{r}\gamma_r(\mathcal{C}) = \sum_{I \in I_{r,n}} \sum_{v \in \mathcal{F}^r} m(I, v)^2$$

where $m(I,v) = \#\{c \in \mathcal{C} : \phi_I(c) = v\}$. Using the Cauchy-Schwartz inequality

$$\frac{\sum_{v \in \mathcal{F}^r} m(I,v)^2}{q^r} \geq \left[\frac{\sum_{v \in \mathcal{F}^r} m(I,v)}{q^r}\right]^2$$

Using the fact that $\sum_{v \in \mathcal{F}^r} m(I,v) = M$, we get the first part. For the second part, we note that

$$\gamma_r(\mathcal{C}) = \frac{M}{q^r} \text{ if and only if } \; m(I,v) = \frac{M}{q^r} \;\; \forall I \in I_{r,n} \text{ and } \; \forall v \in \mathcal{F}^r.$$

If for a fixed $I \in I_{r,n}$ and $\forall v \in \mathcal{F}^n$, we have $\#\{c \in \mathcal{C} : \phi_I(c)\} = \frac{M}{q^r}$ then it implies that for $J \in I_{r-1,n}$ and $J \subset I$, we get $\#\{c \in \mathcal{C} : \phi_J(c)\} = \frac{M}{q^{r-1}}$.

This proves that for a fixed integer $d^\perp(\mathcal{C})$, $\gamma_r(\mathcal{C}) = \frac{M}{q^r}$ if and only if $r < d^\perp(\mathcal{C})$. □

Definition 2.3.1. *We define $\gamma_i^\perp(\mathcal{C}) = \frac{\gamma_{n-i} q^{n-i}}{M}$. We also define $A_i^\perp$ using Mac-Williams identities in a following way*

$$\sum_{i=0}^{n} A_i^\perp(\mathcal{C}) Z^i = W_{\mathcal{C}^\perp}(Z) := \frac{1}{M} W_{\mathcal{C}}\left(\frac{1-Z}{1+(q-1)Z}\right) \tag{2.8}$$

If a code $\mathcal{C}$ is linear then we observe that $\gamma_i^\perp(\mathcal{C}) = \gamma_i(\mathcal{C}^\perp)$.

The $A_i's$ in the distance enumerator polynomial are non-negative for any code $\mathcal{C}$. It follows from the Delsarte's inequalities in [2] that $A_i^\perp{}'s$ are also non-negative. However, it is not clear from definition (2.3.1). Here, we state a theorem by Delsarte.

Theorem 2.3.4. *Non-linear Mac-Williams identities: For any $[n, M, d]_q$ code $\mathcal{C}$, we have*

$$A_i^\perp(\mathcal{C}) \geq 0 \tag{2.9}$$

where $i \in \{0, 1, \ldots, n\}$.

We provide an alternate proof for Delsarte's inequalities and it is based on the fact that variance of random variable is non-negative.

Definition 2.3.2. *For each non-empty $I \subset \{1, \ldots, n\}$ we define a random variable*

$$X_I : \mathcal{F}^n \to \mathbb{R} \ \ \textit{defined by} \ \ X_I(v) = q^{|I|} m(I, v_I)/M.$$

where $m(I, v_I) = \#\{c \in \mathcal{C} : c_I = v_I\}$. Here, $\mathcal{F}^n$ is the probability space with each $v \in \mathcal{F}^n$ being equally likely.

Lemma 2.3.5. *For $I \in I_{r,n}$ random variable X_I, we have* $\mathrm{E}(X_I) = 1$ *and* $\mathrm{Var}(X_I) = \frac{q^r m_I}{M^2} - 1$ *where $m_I = \{(c, c') \in \mathcal{C} \times \mathcal{C} : c_I = c'_I\}$*

Proof. $\mathrm{E}(X_I)$ is given as

$$\mathrm{E}(X_I) = \frac{1}{q^n} \sum_{v \in \mathcal{F}^n} q^{|I|} m(I, v_I)/M = \frac{q^r}{q^n M} \sum_{v \in \mathcal{F}^n} m(I, v_I) = \frac{q^{r-n}}{M} \sum_{v \in \mathcal{F}^r} q^{n-r} m(I, v) = \frac{1}{M} \sum_{v \in \mathcal{F}^r} m(I, v) = 1$$

As we know $\mathrm{E}(X_I)$, we only need to calculate $\mathrm{E}(X_I^2)$.

$$\mathrm{E}(X_I^2) = \frac{1}{q^n} \sum_{v \in \mathcal{F}^n} q^{2r} m(I, v_I)^2 / M^2$$

$$= \frac{q^{2r}}{M^2 q^n} \sum_{v \in \mathcal{F}^n} m(I, v_I)^2$$

$$= \frac{q^{2r}}{M^2 q^n} \sum_{v \in \mathcal{F}^r} q^{n-r} [\#\{c \in \mathcal{C} : c_I = v\}]^2$$

$$= \frac{q^r}{M^2} \#\{(c, c') \in \mathcal{C} \times \mathcal{C} : c_I = c'_I\} = \frac{q^r m_I}{M^2}$$

Therefore, we get the $\mathrm{var}(X_I)$ as given in the lemma. □

Theorem 2.3.6. *For $1 \leq r \leq n$.*

$$\mathrm{var}\left(\sum_{\{J : |J| \leq r\}} (-1)^{r-|J|} \binom{n-|J|}{n-r} X_J \right) = \sum_{\{J : |J| \leq r\}} (-1)^{r-|J|} \binom{n-|J|}{n-r} \mathrm{var}(X_J) = A_r^{\perp}.$$

Proof. Let $Z_r : \mathcal{F}^n \to \mathbb{R}$ be the random variable defined by:

$$Z_r = \sum_{\{J:|J|\leq r\}} (-1)^{r-|J|} \tbinom{n-|J|}{n-r} X_J$$

we can write the $\mathrm{var}(Z_r)$ in a following way

$$\mathrm{var}(Z_r) = \sum_{\{J:|J|\leq r\}} \sum_{\{K:|K|\leq r\}} (-1)^{r-|J|} \tbinom{n-|J|}{n-r} \tbinom{n-|k|}{n-r} \mathrm{cov}(X_J, X_K)$$

Using lemma (2.3.5), we can calculate that $\mathrm{E}(X_K X_J) = \mathrm{E}(X^2_{(K\cap J)})$ which implies that $\mathrm{cov}(X_J, X_K) = \mathrm{var}(X_{J\cap K})$. Let $1 \leq s \leq r$ and let $L \in I_{s,n}$. We can split the summation over K in a following way

$$\mathrm{var}(Z_r) = \sum_{s=1}^{r} \sum_{L\in I_{s,r}} \mathrm{var}(X_L) \sum_{\{J:J\supset L\}} (-1)^{s+|J|} \binom{n-|J|}{n-r} \sum_{K':K'\cap J=\emptyset} (-1)^{|K'|} \binom{n-|K'|-s}{n-r}$$

We note that

$$\sum_{K':K'\cap J=\emptyset} (-1)^{|K'|} \binom{n-|K'|-s}{n-r} = \sum_{i=0}^{r-s} (-1)^i \binom{n-|J|}{i} \binom{n-i-s}{n-r}$$

Next, we will show that

$$\sum_{i=0}^{r-s} (-1)^i \binom{n-|J|}{i} \binom{n-i-s}{n-r} = \delta_{r,j} \tag{2.10}$$

Let $n-s=m$, $r-s=\rho$ and $j-s=l$. The sum then becomes

$$\sum_{i=0}^{\rho} (-1)^i \binom{m-l}{i} \binom{m-i}{m-\rho}$$

We further simplify this by setting $m-l=\mu$

$$\sum_{i=0}^{\mu+l} (-1)^i \binom{\mu}{i} \binom{\mu-i+l}{\rho-i} \tag{2.11}$$

We have,

$$\binom{\mu-i+l}{\rho-i} = \sum_{t=0}^{l} \binom{\mu-i}{\rho-i-t}\binom{l}{t}$$

Using this, (2.11) becomes.

$$\sum_{i=0}^{\mu+l}(-1)^i\binom{\mu}{i}\sum_{t=0}^{l}\binom{\mu-i}{\rho-i-t}\binom{l}{t}$$

Switching the order of summation and multiplying and dividing by $(p-t)!$, we get

$$\sum_{t=0}^{\infty}\sum_{i=0}^{l}(-1)^i\binom{\rho-t}{i}\binom{\mu}{\rho-t}\binom{l}{t} \tag{2.12}$$

Note that

$$\sum_{i=0}^{l}(-1)^i\binom{\rho-t}{i} = (1-1)^{\rho-t}$$

which is 0 except when $\rho = t$. Therefore, we get

$$\sum_{t=0}^{\infty}\delta_{\rho,t}\binom{\mu}{\rho-t}\binom{l}{t} = \binom{l}{\rho} = \binom{j-s}{r-s} = \delta_{r,j}$$

Because $j \le r$, $\binom{j-s}{r-s}$ is 0 whenever $j \ne r$ and 1 when $j = r$. Therefore, we get

$$\mathrm{var}(Z_r) = \sum_{s=1}^{r}\sum_{L\in I_{s,r}}\mathrm{var}(X_L)(-1)^{s+r}\tbinom{n-s}{r-s} = \sum_{\{J:|J|\le r\}}(-1)^{r-|J|}\tbinom{n-|J|}{n-r}\mathrm{var}(X_J)$$

To show the second equality in the theorem, we use $\mathrm{var}(X_J) = -1 + \frac{q^{|J|}m_J}{M^2}$ to get

$$\sum_{\{J:|J|\le r\}}(-1)^{r-|J|}\tbinom{n-|J|}{n-r}\mathrm{var}(X_J) = \sum_{j=1}^{r}(-1)^{r-j}\binom{n-j}{n-r}\binom{n}{j}(\gamma_{n-j}^{\perp}-1)$$

Using equation (2.5) for dual of the code, We have

$$\binom{n}{r}\sum_{j=1}^{r}(-1)^{r-j}\binom{r}{j}(\gamma_{n-j}^{\perp}-1) = A_r^{\perp} \tag{2.13}$$

□

Lemma 2.3.7. *For a code $\mathcal{C}$,*

$$\frac{1}{q} \leq \frac{\gamma_i}{\gamma_{i-1}} < 1 \tag{2.14}$$

for $i \leq n-d$.

Proof. Using the properties of binomial moments for dual of the code, we get

$$\gamma^{\perp}_{n-i} \geq \gamma^{\perp}_{n-i+1}$$

with equality if and only if $i \leq d^{\perp}-1$. Using the definition (2.3.1), we get

$$\frac{\gamma_i q^i}{M} \geq \frac{\gamma_{i-1} q^{i-1}}{M}$$

For $i \leq n-d$, we get

$$\frac{1}{q} \leq \frac{\gamma_i}{\gamma_{i-1}} < 1$$

□

Chapter 3

Perfect codes

In this chapter, we will discuss the unresolved problem of existence of double error correcting perfect codes by generalizing the differential equation derived by Lloyd for binary case to any size of the alphabet. By solving the differential equation for $d = 5$ and $q = 2$, we prove that the only double error correcting binary perfect code is a repetition code. we derive the distance enumerator polynomial for double error correcting perfect codes for any size of the alphabet. Later using the distance enumerator polynomial, we will recover some of the necessary conditions on existence of perfect codes given by Reuvers and Olof Heden in [8, 3]. At the end of this chapter, we present an alternate proof of Lloyd's theorem.

3.1 Generalized differential equation

Let us start by recalling the Hamming bound for a $[n, M, d]_q$ code $\mathcal{C}$. It states that

$$M \leq \frac{q^n}{V_q(n, \lfloor \frac{d-1}{2} \rfloor)}$$

where $V_q(n, r)$ is a volume of ball of radius r in space $\mathcal{F}^n$.

Definition 3.1.1. *Perfect code: If a code $\mathcal{C}$ attains equality in Hamming bound then it is called perfect code.*

To generalize the differential equation by Lloyd for any q, We will follow the discussion in [5]. Let us assume that there exists a perfect code $\mathcal{C}$ with alphabet size q and length n in $\mathcal{F}_q^n$ with balls are of radius e. There are total q^n words in the space. We partition our words based on their distance from codeword. The distance of a word from a perfect code is the least distance from a set of codewords i.e. $d(x, \mathcal{C}) = \min_{\forall c \in \mathcal{C}} d(x, c)$. Let $\nu_{j,s}$ represents number of words at distance j from the perfect code which are of weight s. Since every word lies inside one and only one ball around codeword, we can write

$$\nu_{0,s} + \nu_{1,s} + \cdots + \nu_{e,s} = \binom{n}{s}(q-1)^s$$

Multiplying both sides by x^s and taking sum over s, we get

$$G(x) + G_1(x) + G_2(x) + \cdots + G_e(x) = [1 + (q-1)x]^n$$

where $G_j(x) = \sum_{s=0}^{n} \nu_{j,s} x^s$. We need to $G_j(x)$ in terms of G(x) which is a distance enumerator polynomial of the code. Suppose codeword w is of weight s i.e. w consist of s non-zero terms and n-s zeroes in some order. Number of words at distance j from w is $\binom{n}{j}(q-1)^j$. we choose m non-zero entries out of s and j-m zero entries out of n-s. Now out of m non-zero entries we turn k entries into zero's and m-k are still non zero entries. It implies that out of $\binom{n}{j}(q-1)^j$ words which are at distance j exactly

$$\sum_{m=0}^{\infty} \binom{s}{m}\binom{n-s}{j-m}\binom{m}{k}(q-1)^{j-m}(q-2)^k$$

are of weight s+k+j-2m. We observe that

$$\binom{n}{j}(q-1)^j = \sum_{m=0}^{\infty} \binom{s}{m}\binom{n-s}{j-m}(q-1)^{j-m}(q-1)^m$$

which also can be written as

$$\binom{n}{j}(q-1)^j = \sum_{m=0}^{\infty} \binom{s}{m}\binom{n-s}{j-m}(q-1)^{j-m}\left[\sum_{k=0}^{m} \binom{m}{k}(q-2)^k\right]$$

This ensures that we are not missing any word which is at distance j. Now we can write

$G_j(x)$ as

$$G_j(x) = \sum_{s=0}^{n} \nu_s \Big[\sum_{m=0}^{\infty} \sum_{k=0}^{m} \binom{s}{m} \binom{n-s}{j-m} (q-1)^{j-m} \binom{m}{k} (q-2)^k \Big] x^{s+j+k-2m}$$

Now we observe that

$$[x + (q-2)xy + y]^s [1 + (q-1)xy]^{n-s} =$$

$$\sum_{j=0}^{\infty} y^j \Big[\sum_{m=0}^{\infty} \sum_{k=0}^{m} \binom{s}{m} \binom{n-s}{j-m} (q-1)^{j-m} \binom{m}{k} (q-2)^k \Big] x^{s+j+k-2m}$$

Multiply above equation on both sides by ν_s and taking summation over s we get

$$\sum_{s=0}^{\infty} \nu_s [x + (q-2)xy + y]^s [1 + (q-1)xy]^{n-s} =$$

$$\sum_{s=0}^{\infty} \sum_{j=0}^{\infty} y^j (\nu_s) \Big[\sum_{m=0}^{\infty} \sum_{k=0}^{m} \binom{s}{m} \binom{n-s}{j-m} (q-1)^{j-m} \binom{m}{k} (q-2)^k \Big] x^{s+j+k-2m}$$

Substituting value of $G_m(x)$, we get

$$\sum_{s=0}^{\infty} \nu_s [1 + (q-1)xy]^n \left[\frac{x + (q-2)xy + y}{1 + (q-1)xy} \right]^s = \sum_{m=0}^{\infty} G_m(x) y^m$$

Now, we put $z = \frac{x+(q-2)xy+y}{1+(q-1)xy}$ in above equation to get

$$G(z) \left[\frac{(1-x)(1+(q-1)x)}{1+(q-2)x-(q-1)xz} \right]^n = \sum_{m=0}^{\infty} G_m(x) \left[\frac{z-x}{1+(q-2)x-(q-1)xz} \right]^m$$

Multiplying above equation by $[1 + (q-2)x - (q-1)xz]^{j-1}$ on both sides we get

$$G(z)\left[\frac{(1-x)(1+(q-1)x)}{1+(q-2)x-(q-1)xz}\right]^n [1+(q-2)x-(q-1)xz]^{j-1} = \sum_{m=0}^{\infty} G_m(x)\left[\frac{[z-x]^m}{1+(q-2)x-(q-1)xz]^{m-j+1}}\right]$$

Now, we will partially differentiate above equation j times w.r.t. z at $z=x$. We observe that RHS will contribute only when m=j. So RHS will be

$$G_j(x)\frac{\partial^j}{\partial z^j}\left[\frac{[z-x]^j}{1+(q-2)x-(q-1)xz}\right]\Big|_{z=x} = G_j(x)\left[\frac{j!}{1+(q-2)x-(q-1)x^2}\right]$$

Now, on LHS we will get

$$[(1-x)(1+(q-1)x)]^n \frac{\partial^j}{\partial z^j}\frac{G(z)}{[1+(q-2)x-(q-1)xz]^{n-j+1}}\Big|_{z=x}$$

Equating RHS and LHS, we get

$$G_j(x) = \frac{[(1-x)(1+(q-1)x)]^{n+1}}{j!}\frac{\partial^j}{\partial z^j}\frac{G(z)}{[1+(q-2)x-(q-1)xz]^{n-j+1}}\Big|_{z=x}$$

$$G_j(x) = \sum_{p=0}^{j}\binom{n-p}{j-p}\frac{[(1-x)(1+(q-1)x)]^p[(q-1)x]^{j-p}}{p!}\frac{d^pG(x)}{dx^p}$$

Now taking summation over j=0 to e , we get

$$\sum_{p=0}^{e}\frac{[(1-x)(1+(q-1)x)]^p}{p!}\sum_{r=0}^{e-p}\binom{n-p}{r}[(q-1)x]^r\frac{d^pG(x)}{dx^p} = [1+(q-1)x]^n \tag{3.1}$$

This is a generalize differential equation given by Llyod for binary case to any q. Solving the above differential equation for $q = e = 2$, we get the following result.

Theorem 3.1.1. *The only double error correcting binary perfect code is repetition code* $\mathcal{C} = \{(0,0,0,0,0),(1,1,1,1,1)\}$

Proof. Using the equation (3.1) for $e = q = 2$, we get

$$\sum_{p=0}^{2} \frac{[1-z^2]^p}{p!} \sum_{r=0}^{2-p} \binom{n-p}{r} z^r \frac{d^p W(z)}{dz^p} = [1+(q-1)z]^n \tag{3.2}$$

Next, we substitute $L_r = (1-z^2)^r \frac{d^r}{dz^r}$ and calculate L_2 in terms of $L_1 = L$.

$$L^2 = \left[(1-z^2)\frac{d}{dz}\right]\left[(1-z^2)\frac{d}{dz}\right] = (1-z^2)\frac{d}{dz}\left[(1-z^2)\frac{d}{dz}\right]$$

$$= (1-z^2)\left[(-2z)\frac{d}{dz} + (1-z^2)\frac{d^2}{dz^2}\right] = -2zL \ + \ L_2$$

Using the above equation, we can rewrite the differential equation in a following way

$$(L^2 + 2(1+nz)L + 2(1+nz+\tfrac{n(n-1)}{2}z^2))W(z) = 2(1+z)^n \tag{3.3}$$

We consider a linear operator $\mathbb{T}'$ on vector space $\mathbb{V}$ of smooth functions from $\mathbb{R}$ to $\mathbb{R}$. Let $\mathbb{T}' : C^\infty(-1,1) \to C^\infty(\mathbb{R})$ is given as

$$(\mathbb{T}'f)(z) = f(\tanh x) \tag{3.4}$$

This is a invertible linear transformation. If $D = \frac{d}{dx}$ is a differential operator on $\mathbb{V}$ then we note that $\mathbb{T}'^{-1} \circ D \circ \mathbb{T}' = L$. we can rewrite equation (3.3) as

$$\begin{aligned}(D^2 + 2(1+n\tanh(x))D + 2(1+n\tanh(x)+\tfrac{n(n-1)}{2}\tanh^2(x)) \cdot W(\tanh(x))\\ = 2(1+\tanh(x))^n \quad (3.5)\end{aligned}$$

We note that

$$e^{\int_0^x 1+n\tanh(t)\,dt} = e^x \cosh^n(x)$$

Substituting $H(x) = e^x \cosh^n(x) W(\tanh x)$, we can rewrite the equation (3.5) in a following way

$$[D^2 - (n-1)]H = 2e^{x(1+n)} \tag{3.6}$$

We will solve the above differential equation by symbolic method.

$$H = (D^2 - (n-1))^{-1}(2e^{x(1+n)})$$

$$= (D - \sqrt{n-1})^{-1}(\tfrac{e^{x(1+n)}}{\sqrt{n-1}}) - (D + \sqrt{n-1})^{-1}(\tfrac{e^{x(1+n)}}{\sqrt{n-1}})$$

Now for a scalar λ, we have $(D-\lambda)f = g$ implies $f = \int g(t)e^{(x-t)\lambda}\,dt$ and hence

$$(D-\lambda)^{-1}g = c\,e^{\lambda x}g + \int dt\, e^{(x-t)\lambda}\, g \qquad \text{where } c \text{ is a constant.}$$

Therefore

$$\sqrt{n-1}\,H(x) = \int_0^x dt\, e^{(x-t)\sqrt{n-1}}\, e^{t(1+n)}$$
$$- \int_0^x dt\, e^{-(x-t)\sqrt{n-1}}\, e^{t(1+n)} + ae^{x\sqrt{n-1}} + be^{-x\sqrt{n-1}}$$

for some constants a, b determined by $H(0) = H'(0) = 1$.

$$\sqrt{n-1}\,H(x) = \tfrac{e^{x(1+n)} - e^{x\sqrt{n-1}}}{1+n-\sqrt{n-1}} - \tfrac{e^{x(1+n)} - e^{-x\sqrt{n-1}}}{1+n+\sqrt{n-1}} + ae^{x\sqrt{n-1}} + be^{-x\sqrt{n-1}}$$

We can rewrite the above equation as

$$H(x) = \tfrac{n(n+1)}{n^2+n+2}\cosh(x\sqrt{n-1}) + \tfrac{n\sqrt{n-1}}{n^2+n+2}\sinh(x\sqrt{n-1}) + \tfrac{2}{n^2+n+2}e^{x(1+n)} \tag{3.7}$$

We assume that $k = \sqrt{n-1}$ (not necessarily an integer). Additionally, let $a = 1 + \binom{k}{2}$ and $b = 1 + k + \binom{k}{2}$. Note that $b - a \geq 2$ since $n \geq d = 5$. We also observe that $a + b = n + 1$ and $2ab = (1 + n + \binom{n}{2}) = V_2(n, 2) = 2^n/M$. $W(z) = H(\tanh^{-1}(z))e^{-\tanh^{-1}(z)n}(\tanh^{-1}(z))$

can be written as follows.

$$W(z) = \tfrac{(1+z)^{a-1}}{4ab}\left[2(1+z)^b + n(1-z)^a\big(b(1+z)^{b-a} + a(1-z)^{b-a}\big)\right]$$

Using Mac-Williams transform on $W_{\mathcal{C}}(z)$, we get the following expression for $W_{\mathcal{C}^\perp}(z)$

$$W^\perp(z) = 1 + \tfrac{n}{2}(bz^a + az^b) \tag{3.8}$$

This implies that a, b are non-negative integers. we also know that $\frac{2^n}{V_2(n,2)} = M$ where M is size of the code and therefore integer. It implies that a, b are of the form $2^A, 2^B$ respectively.

$$2ab = 1 + \frac{n(n+1)}{2} = 1 + \frac{(a+b)(a+b-1)}{2}$$

This gives

$$2^{A+B+1} = 1 + (2^A + 2^B - 1)(2^{A-1} + 2^{B-1}) = 1 + 2^{A-1}(2^A + 2^B - 1)(1 + 2^{B-A})$$

For $A > 1$, RHS is odd and LHS is even which is a contradiction. Therefore we substitute $A = 1$ in above equation to get

$$2^{1+B} = 1 + (1 + 2^B)(1 + 2^{B-1}) = 1 + 3 \cdot 2^{B-2} + 4^{B-1}$$

This implies $B = 2$. So, we get $a = 2$ and $b = 4$. Substituting values of a, b in $W_{\mathcal{C}}(z)$, we get the distance enumerator polynomial of a perfect binary code

$$W_{\mathcal{C}}(z) = 1 + z^5 \tag{3.9}$$

It is easy to observe that the above polynomial is a distance enumerator polynomial of a repetition code $\mathcal{C} = \{(0,0,0,0,0), (1,1,1,1,1)\}$. □

3.2 Distance enumerator polynomial

In this section, we will solve the differential for $e = 2$ and general q case to get the distance enumerator polynomial of double error correcting perfect code. The differential equation for

$e = 2$ and general q is given as

$$\sum_{p=0}^{2} \frac{[(1-x)(1+(q-1)x)]^p}{p!} \sum_{r=0}^{2-p} \binom{n-p}{r} [(q-1)x]^r \frac{d^p G(x)}{dx^p} = [1+(q-1)x]^n$$

To simplify the polynomial, we define

$$L_r = (1-x)^r (1+(q-1)x)^r \frac{d^r}{dx^r}$$

We calculate relation between L_2 and $L_1 = L$ in a following way

$$L^2 = (1-x)(1+(q-1)x)\frac{d}{dx}\Big[(1-x)(1+(q-1)x)\frac{d}{dx}\Big]$$

$$= (1-x)(1+(q-1)x)\Big[((q-2)-2(q-1)x)\frac{d}{dx} + (1-x)(1+(q-1)x)\frac{d^2}{dx^2}\Big]$$

$$= ((q-2)-2(q-1)x)L + L_2$$

We rewrite the differential equation in terms of L to get

$$\left((1+n(q-1)x + \binom{n}{2}(q-1)^2x^2) + (1-\tfrac{q}{2}+1+n(q-1)x)L + \frac{L^2}{2}\right)W(x) = (1+(q-1)x)^n$$

We consider a linear operator $\mathbb{T}$ on vector space $\mathbb{V}$ of smooth functions from $\mathbb{R}$ to $\mathbb{R}$ which is a generalization of the linear operator $\mathbb{T}'$ defined in (3.4). Let $\mathbb{T} : C^\infty(\frac{-1}{q-1}, 1) \to C^\infty(\mathbb{R})$ such that

$$(\mathbb{T}f)(x) = f(\frac{e^{qy}-1}{e^{qy}+q-1}) \tag{3.10}$$

This is a invertible linear transformation. If $D = \frac{d}{dy}$ is a differential operator on $\mathbb{V}$ then we note that $\mathbb{T}^{-1} \circ D \circ \mathbb{T} = L$. we can rewrite differential equation as

$$\Bigg((1+n(q-1)\mathbb{T}x+\tbinom{n}{2}(q-1)^2(\mathbb{T}x)^2)$$
$$+(1-\tfrac{q}{2}+1+n(q-1)\mathbb{T}x)\tfrac{d}{dy}+\tfrac{1}{2}\tfrac{d^2}{dy^2}\Bigg)(\mathbb{T}W)=(\tfrac{qe^{qy}}{e^{qy}+q-1})^n$$

To convert the above differential equation with constant coefficients, we define

$$\gamma \coloneqq \tfrac{3}{2}+(n-2)(1-\tfrac{1}{q})$$

and let

$$\phi(y)=\exp\left(\int_0^y q\gamma+n\tfrac{-q(q-1)\exp(-qt)}{1+(q-1)\exp(-qt)}dt\right)=\exp(qy\gamma)\left(\tfrac{1+(q-1)\exp(-qy)}{q}\right)$$

We assume that $H(y)=(\mathbb{T}W)y\ \phi(y)$. We can rewrite the differential equation as

$$\frac{d^2H}{dy^2}-\left(\frac{q\mu}{2}\right)^2 H=2\exp(qy\gamma) \tag{3.11}$$

where $\mu=\sqrt{1+\frac{4(q-1)(n-2)}{q^2}}$. Now, we have a second order differential equation with constant coefficients with initial conditions $H(0)=1$ and $H'(0)=2-\frac{2}{q}$. To make our notations simpler, we assume that $a=\gamma-\frac{\mu}{2}$ and $b=\gamma+\frac{\mu}{2}$. We also note that $V_q(n,2)=\frac{q^2ab}{2}$. It implies that a,b are non-zero numbers. Solving the differential equation, we get the following unique solution

$$H(y)=\frac{2\exp((a+b)qy/2)}{q^2ab}+\frac{\exp((b-a)qy/2)}{2(b-a)}(b-a-1+\tfrac{4}{q}-\tfrac{4}{q^2a})$$
$$+\frac{\exp((a-b)qy/2)}{2(b-a)}(\tfrac{4}{q^b}+b-a-\tfrac{4}{q}+1)$$

Now, substituting the value of $H(y)$ we get

$$\frac{(\mathbb{T}W)y}{(1+(q-1)\mathbb{T}x)^n} = \frac{2}{q^2ab} + \frac{\exp(-aqy)}{2(b-a)}(b-a+2-\tfrac{q}{2}-\tfrac{1}{a}) + \frac{\exp(-bqy)}{2(b-a)}(\tfrac{1}{b}+b-a+\tfrac{q}{2}-2)$$

Next, we substitute $Z = e^{-qy}$. Using the Mac-Williams transformation, we replace $W(Z)$ in terms of $W^{\perp}(Z)$ to get the following expression for $W^{\perp}(Z)$

$$W^{\perp}(z) = 1 + \frac{nq(q-1)}{4}\left[z^a b\left(1+\frac{2/q-1}{b-a}\right) + z^b a\left(1-\frac{2/q-1}{b-a}\right)\right] \tag{3.12}$$

This implies that a, b are positive integers. Using Mac-Williams transformation, we get the following distance enumerator polynomial for double error correcting perfect codes

$$V_q(n,2)W_c(z) =$$

$$(1+(q-1)z)^n + \frac{nq(q-1)}{4}\left[(1+(q-1)z)^{n-a}(1-z)^a b\left(1+\frac{2/q-1}{b-a}\right)+\right.$$

$$\left.(1+(q-1)z)^{n-b}(1-z)^b a\left(1-\frac{2/q-1}{b-a}\right)\right]$$

where $a+b = \frac{4-q+2n(q-1)}{q}$ and $b-a = \sqrt{1+\frac{4(q-1)(n-2)}{q^2}}$.

Using the fact that a, b are positive integers, we will prove the result due to Reuvers given in [8]. Let us assume that q is odd. The fact that $b-a$ is integer implies that q^2 divides $n-2$. Next, we can write

$$a+b-3 = 2q(q-1)(\tfrac{n-2}{q^2}) \tag{3.13}$$

As we know that a and b divides $2q^n$, we observe that 3 is the only common prime factor between a and b. We consider the case $\gcd(q,6) = 1$. In this case, we get a, b to be coprime. Let us assume that q is of the form $q = p_1^{r_1}p_2^{r_2}$ where p_1 and p_2 are primes such that $p_2 \equiv 1 \mod p_1$. As the numbers a, b are coprime they must be of the form $a = 2.p_1^{\alpha_1}, b = p_2^{\alpha_2}$ or

$a = p_1^{\alpha_1}, b = 2.p_2^{\alpha_2}$ where $\alpha_1, \alpha_2 \in \mathbb{N}$. In the first case, we observe that RHS of equation (3.13) is divisible by p_1. It implies that

$$\begin{aligned} 2.p_1^{\alpha_1} + p_2^{\alpha_2} - 3 &\equiv 0 \mod p_1 \\ -2 &\equiv 0 \mod p_1 \end{aligned}$$

which is a contradiction. We can get contradiction for second case in a similar way. It implies that there do not exist any perfect code for alphabet of size $q = p_1^{r_1} p_2^{r_2}$ where p_1 and p_2 are primes such that $p_2 \equiv 1 \mod p_1$.

The following lemma is a generalization of the result by Olof Heden for $e = 1$ case given in [3].

Lemma 3.2.1. *If there exist a q-ary perfect code $\mathcal{C}$ of minimum distance d and length n then $V_q(n, e)$ divides $q^{n-d^{\perp}(\mathcal{C})+1}$ where $d^{\perp}(\mathcal{C})$ is the smallest positive power of Z in $W^{\perp}(Z)$ with non-zero coefficient.*

Proof. We will recall a lemma (2.3.3) proved in second section. It state that $\gamma_i(\mathcal{C}) \geq M/q^i$. Also there exists a unique integer $d^{\perp}(\mathcal{C})$ such that $\gamma_i(\mathcal{C}) = M/q^i$ if and only if $i \leq d^{\perp}(\mathcal{C}) - 1$. Taking $i = d^{\perp}(\mathcal{C}) - 1$, we get that $\frac{M}{q^{d^{\perp}(\mathcal{C})-1}}$ is an integer. In case of perfect code, we have $M = \frac{q^n}{V_q(v,e)}$. It implies that $\frac{q^{n-d^{\perp}+1}}{V_q(n,e)}$ is an integer. □

3.3 Alternate proof of Lloyd's theorem

In order to state Lloyd's theorem, we will first define a set of orthogonal polynomials known as Krawtchouk polynomials.

Definition 3.3.1. *For $k \in \mathbb{Z}^+$, we define Krawtchouk polynomial $K_k(x)$ by*

$$K_k(x; n, q) := K_k(x) := \sum_{j=0}^{k} (-1)^j \binom{x}{j} \binom{n-x}{k-j} (q-1)^{k-j},$$

where

$$\binom{x}{j} := \frac{x.(x-1)\dots(x-j+1)}{j!} \qquad x \in \mathbb{R}$$

We also define Lloyd's polynomial,

$$\psi_j(n,x) := \sum_{k=0}^{j} K_k(n,x) = K_j(x-1,n-1;q)$$

The second equality is given in [9, p. 17].

Theorem 3.3.1. *Lloyd's theorem: If a binary perfect e-error correcting code exists, then $\psi_e(n,x)$ has e distinct zeroes among the integers $1,2,\dots,n$ where $e = \frac{d-1}{2}$.*

Proof. We can rewrite the differential equation (3.5) in a following way

$$O_{n,e}W(x) = (1+(q-1)x)^n$$

where $O_{n,e}$ is an linear differential operator given in a following way

$$O_{n,e} = \sum_{i=0}^{e} \frac{(1-x)^i(1+(q-1)x)^i}{i!}(1+(q-1)x)^{n-i}_{|e-i}\ D_x^i$$

where

$$(1+(q-1)x)^{n-i}_{|e-i} = \sum_{j=0}^{e-i}\binom{n-i}{j}[(q-1)x]^j$$

and $D_x = \frac{d}{dx}$. Next, we will show that $O_{n,e}$ satisfies the following recurrence relation.

Lemma 3.3.2.

$$eO_{n,e} = O_{n,e-1}\{O_{n,1} - (e-1)(q-2)\} - (q-1)(n-e+1)O_{n,e-2}$$

Proof. We note that $O_{n,1} = 1 + n(q-1)x + (1-x)(1+(q-1)x)D_x$. To make the notation simpler, we denote

$$\frac{(1-x)^i(1+(q-1)x)^i}{i!} = b_i(x)$$

Our strategy to prove above recurrence relation is to compare coefficients of $x^m b_k(x) D_x^k$ for $k \in \{0, 1, \ldots, e\}$ and $m \in \{0, 1, \ldots, e-k\}$ on both sides of the equation. The coefficient of $x^m b_k(x) D_x^k$ in $eO_{n,e}$ is given as

$$\text{Coefficient of } x^m b_k(x) D_x^k = e\binom{n-k}{m}(q-1)^m$$

The RHS of equation (2.51) can be written as

$$O_{n.e-1}\{O_{n,1} - (e-1)(q-2)\} - (q-1)(n-e+1)O_{n,e-2} = C_1 + C_2 + C_3 + C_4$$

where C_1, C_2, C_3 and C_4 are given as

$$\begin{aligned}
C_1 &= 1 - (e-1)(q-2)) \sum_{i=0}^{e-1} b_i(x)(1+(q-1)x)^{n-i}_{|e-1-i}\ D_x^i \\
C_2 &= n(q-1) \sum_{i=0}^{e-1} b_i(x)(1+(q-1)x)^{n-i}_{|e-i-1}[xD_x^i + iD_x^{i-1}] \\
C_3 &= \ -(q-1)(n-e+1) \sum_{i=0}^{e-2} b_i(x)(1+(q-1)x)^{n-i}_{|e-2-i}\ D_x^i \\
C_4 &= \ \sum_{i=0}^{e-1} b_i(x)(1+(q-1)x)^{n-i}_{|e-1-i}\{(1-x)(1+(q-1)x)D_x^{i+1} \\
&\qquad\qquad + i[(q-2) - 2(q-1)x]D_x^i + \tbinom{i}{2}[-2(q-1)D_x^{i-1}]\}
\end{aligned}$$

The coefficient of $x^m b_k(x) D_x^k$ on RHS is given as

$$\begin{aligned}\text{Coefficient of } x^m b_k(x)D_x^k =&(q-1)^m\binom{n-k}{m}\Bigg[1-(e-1)(q-2)+\frac{nm}{n-k-m+1}\\ &+(q-1)(n-k-m)+(q-2)m-\frac{m(m-1)}{n-k-m+1}+\frac{k(n-k+1)}{n-k-m+1}\\ &-\frac{2km}{n-k-m+1}+k(q-2)-(q-1)(n-e+1)\Bigg]\\ &=e(q-1)^m\binom{n-k}{m}\end{aligned}$$

□

We use linear operator $\mathbb{T}$ defined in (3.10). We can rewrite the differential equation as

$$(\mathbb{T}\circ O_{n,e}\circ\mathbb{T}^{-1})\mathbb{T}W(x)=\left[\frac{qe^{qy}}{e^{qy}+q-1}\right]^n$$

$$(\mathbb{T}\circ O_{n,e}\circ\mathbb{T}^{-1})W\left(\frac{e^{qy}-1}{e^{qy}+q-1}\right)=\left[\frac{qe^{qy}}{e^{qy}+q-1}\right]^n$$

Using the Mac-Williams transformation above equation can be written as

$$(\mathbb{T}\circ O_{n,e}\circ\mathbb{T}^{-1})\frac{W^{\perp}(e^{-qy})}{1+(q-1)e^{-qy}}=V_q(n,e)\left[\frac{e^{qy}}{e^{qy}+q-1}\right]^n$$

Rearranging the terms in above equation, we get

$$\left[(1+(q-1)e^{-qy})^n\circ\mathbb{T}\circ O_{n,e}\circ\mathbb{T}^{-1}\circ\tfrac{1}{(1+(q-1)e^{-qy})^n}\right]W^{\perp}(e^{-qy})=V_q(n,e) \tag{3.14}$$

Let us denote $P_{n,e}=(1+(q-1)e^{-qy})^n\circ\mathbb{T}\circ O_{n,e}\circ\mathbb{T}^{-1}\circ\frac{1}{(1+(q-1)e^{-qy})^n}$. The recurrence relation for $P_{n,e}$ can be written as

$$eP_{n,e}=P_{n,e-1}\{P_{n,1}-(e-1)(q-2)\}-(q-1)(n-e+1)P_{n,e-2}$$

The term $P_{n,1}$ is given as

$$P_{n,1} = 1 + n(q-1)\frac{e^{qy}-1}{e^{qy}+q-1} + D_y + \frac{nq(q-1)}{e^{qy}+(q-1)}$$

$$= D_y + (n-1)(q-1) + q$$

Therefore, we get

$$eP_{n,e} = P_{n,e-1}\{D_y + (n-1)(q-1) + q - (e-1)(q-2)\} - (q-1)(n-e+1)P_{n,e-2}$$

Now, we recall the Lloyd's polynomial defined in equation (2.32). The recurrence relation satisfied by Krawtchouk polynomial is given in [9, p. 16].

$$(k+1)K_{k+1}(x) = \{k + (q-1)(n-k) + qx\}K_k(x) - (q-1)(n-e+1)K_{k-1}(x)$$

From above relation, we easily get the recurrence relation satisfied by ψ_e and it is given as

$$e\psi_e(-T/q) = \{e + (q-1)(n-e+1) + T\}\psi_{e-1}(-T/q) - (q-1)(n-e+1)\psi_{e-2}(-T/q)$$

It is easy to observe that $P_{n,e}(T) = \psi_e(-T/q)$ and we also have $\psi_e(0) = V_q(n,e)$. Therefore, we get

$$\psi_e\left(\frac{-1}{q}\frac{d}{dy}\right)(W^{\perp}(e^{-qy}) - 1) = 0$$

We know that $W^{\perp}(e^{-qy})$ is a polynomial in variable e^{-qy}. If we substitute

$$W^{\perp}(e^{-qy}) - 1 = \sum_{i=1}^{n} A_i^{\perp} e^{-qyi}$$

in the above differential equation then we get

$$\sum_{i=1}^{n} A_i^{\perp}\psi_e(i)e^{-qyi} = 0$$

Next, we can prove that $f_i(y) = e^{-qyi}$ for $i \in \{1, \ldots, n\}$ are linearly independent functions using Wronskian method. The matrix in the Wronskian method is Vandermonde matrix and therefore we get

$$\begin{vmatrix} e^{-qy} & e^{-2qy} & \cdots & e^{-nqy} \\ (-q)e^{-qy} & (-2q)e^{-2qy} & \cdots & (-nq)e^{-nqy} \\ (-q)^2 e^{-qy} & (-2q)^2 e^{-2qy} & \cdots & (-nq)^2 e^{-nqy} \\ \vdots & \vdots & & \vdots \\ (-q)^{n-1} e^{-qy} & (-2q)^{n-1} e^{-2qy} & \cdots & (-nq)^{n-1} e^{-nqy} \end{vmatrix} \neq 0 \tag{3.15}$$

It implies that $A_i^{\perp}\psi_e(i) = 0 \; \forall \; i \in \{1, 2, \ldots, n\}$. If $A_i^{\perp} \neq 0$ then it implies that $\psi_e(i) = 0$. Let λ_i for $i \in \{1, 2, \ldots, e\}$ are roots of $\psi_e(x)$ and there are $s^{\perp}$ number of powers in $W^{\perp}(z)$ with non-zero coefficient except constant. The roots of $\psi_e(x)$ are distinct and lie in interval $[0, n]$ is given in [9, p. 17]. Let the powers in the $W^{\perp}(z)$ with non-zero coefficient be l_j for $j \in \{1, 2, \ldots, s^{\perp}\}$. It implies that

$$\{l_1, l_2, \ldots, l_{s^{\perp}}\} \subset \{\lambda_1, \lambda_2, \ldots, \lambda_e\}$$

therefore, we have $s^{\perp} \leq e$. It is also proved in [6, p. 175] that $s^{\perp} \geq e$. It implies that $s^{\perp} = e$. As there are exactly e non-zero coefficients in $W^{\perp}(z)$ except the constant term, we get the result by Lloyd for any q that $\psi_e(x)$ must have e distinct integral roots in interval $[0, n]$. □

If a perfect code exists for parameters n, d and q then in that case we can calculate the exact value of $A_q(n, d)$ but perfect codes exist in only few cases. As we do not know the exact value of $A_q(n, d)$, it is hard to calculate the maximum rate of transmission for for general n and d. This question becomes more tractable by defining a function known as the asymptotic rate function. The asymptotic rate function is discussed in more details in the next chapter.

Chapter 4

Asymptotic rate function

In this chapter, we will study about the maximum rate of transmission of a code asymptotically as a function of relative minimum distance. We want both of these quantities to be high for codes. As we have discussed in Chapter 1, these are mutually conflicting requirements. The trade-off between these quantities is captured in a function called asymptotic rate function. The asymptotic rate function is defined as follows

Definition 4.0.1. *Asymptotic rate function for codes is a function* $\alpha_q : [0,1] \to [0,1]$ *defined by:*

$$\alpha_q(\delta) = \limsup_{n\to\infty} \frac{\log_q A_q(n, \delta n)}{n} \tag{4.1}$$

The exact value of $\alpha_q(\delta)$ is unknown because calculating exact value of $A_q(n,d)$ is very hard. To study the nature of asymptotic rate function, we will use the bounds and properties of $A_q(n,d)$ discussed in Chapter 1.

4.1 Basic properties of asymptotic rate function

In this section, we will discuss about the special case of convexity of $\alpha_q(x)$ along with basic properties. We will also prove the continuity of $\alpha_q(\delta)$.

Lemma 4.1.1. $\alpha_q(0) = 1$ *and* $\alpha_q(1) = 0$.

Proof. It is easy to observe that $A_q(n, 0) = q^n$. It implies that $n^{-1} \log_q A_q(n, 0) = 1$. Taking $\limsup_{n\to\infty}$ we get the first result. For the second part, we will first prove that $A_q(n, n) = q$. Using Singleton bound (1.2.3), we get $A_q(n, n) \leq q^{n-n+1} = q$. Now, size of a repetition code is q with minimum distance n in $\mathcal{F}^n$. It implies that $A_q(n, n) = q$ and therefore $n^{-1} \log_q A_q(n, n) = 1/n$. Taking $\limsup_{n\to\infty}$ we get the second result. □

Lemma 4.1.2. *$\alpha_q(x)$ is a decreasing function of x.*

Proof. As we know that $A_q(n, d)$ is a decreasing function of d. It implies that

$$n^{-1} \log_q A_q(n, xn) \geq n^{-1} \log_q A_q(n, yn) \quad \text{for} \quad x \leq y$$

Taking $\limsup_{n\to\infty}$ we get $\alpha_q(x) \geq \alpha_q(y)$ for $x \leq y$. □

Next, we will calculate the asymptotic version of singleton bound.

Lemma 4.1.3. $\alpha_q(x) \leq 1 - x$

Proof. The singleton bound can be written as follows

$$A_q(n, xn) = A_q(n, \lceil xn \rceil) \leq q^{n-\lceil xn \rceil+1} \tag{4.2}$$

It implies that

$$\alpha_q(x) = \limsup_{n\to\infty} \frac{\log_q A_q(n, xn)}{n} \leq \limsup_{n\to\infty} \frac{n - \lceil xn \rceil + 1}{n} = 1 - x \tag{4.3}$$

□

Lemma 4.1.4. $\frac{1-\alpha_q(x)}{x}$ *is a decreasing function of x.*

Proof. Let us take x, y such that $0 \leq x \leq y \leq 1$. We need to show that

$$\alpha_q(x) \leq 1 - \left[\frac{x}{y}(1 - \alpha_q(y))\right]$$

Now, substitute $1 - x/y = \tau$. The above inequality can be written as follows

$$\alpha_q(x) \leq 1 - (1-\tau)\Big[1 - \alpha_q\Big(\frac{x}{1-\tau}\Big)\Big]$$

$$= 1 - \Big[1 - \alpha_q\Big(\frac{x}{1-\tau}\Big) - \tau + \tau\alpha_q\Big(\frac{x}{1-\tau}\Big)\Big]$$

$$= \tau + (1-\tau)\alpha_q\Big(\frac{x}{1-\tau}\Big) \quad \text{where} \;\; 0 \leq x \leq 1-\tau \leq 1$$

We will show that the above inequality is a special case of convexity of $\alpha_q(x)$. The convexity of $\alpha_q(x)$ is an open question. Convexity condition is as follows

$$\alpha_q(\lambda x + (1-\lambda)y) \leq \lambda\alpha_q(x) + (1-\lambda)\alpha_q(y) \tag{4.4}$$

Fixing $x = 0$ and $\lambda = \tau$, we get

$$\alpha_q\Big(0 + (1-\tau)\frac{x}{1-\tau}\Big) \leq \tau\alpha_q(0) + (1-\tau)\alpha_q\Big(\frac{x}{1-\tau}\Big)$$

$$\alpha_q(x) \leq \tau + (1-\tau)\alpha_q\Big(\frac{x}{1-\tau}\Big)$$

So, proving the above inequality is equivalent to prove the special case of convexity. Cross sectional bound on $A_q(n,d)$ for a sequence $d_n = xn$ and $t_n = \lceil \tau n \rceil$ gives

$$A_q(n, xn) \leq A_q(n - t_n, xn)q^{t_n}$$

$$= A_q\Big(n - t_n, (n-t_n)\frac{xn}{n-t_n}\Big)q^{t_n}$$

$$= A_q\Big(n - t_n, (n-t_n)\frac{x}{1-t_n/n}\Big)q^{t_n}$$

We have $t_n/n \geq \tau$ and $A_q(n,d)$ is a decreasing function of d implies that

$$A_q(n, xn) \leq A_q\Big(n - t_n, (n-t_n)\frac{x}{1-\tau}\Big)q^{t_n} \tag{4.5}$$

To get the asymptotic version of the above equation we will multiply and divide by $n - t_n$ to get

$$n^{-1} \log_q A_q(n, xn) \leq \Big(\frac{n - t_n}{n}\Big) \frac{\log_q A_q\Big(n - t_n, (n - t_n)\frac{x}{1-\tau}\Big)}{n - t_n} + \frac{t_n}{n}$$

We can observe that

$$\limsup_{n\to\infty} \frac{\log_q A_q\Big(n - t_n, (n - t_n)\frac{x}{1-\tau}\Big)}{n - t_n} \leq \limsup_{n\to\infty} \frac{\log_q A_q\Big(n, (n)\frac{x}{1-\tau}\Big)}{n}$$

$$= \alpha_q(\frac{x}{1-\tau})$$

because sequence on left hand side is a subsequence of sequence on right hand side. Taking $\limsup_{n\to\infty}$ and using the above inequality, we get

$$\alpha_q(x) \leq (1 - \tau)\alpha_q(\frac{x}{1-\tau}) + \tau \tag{4.6}$$

□

Using lemma (4.1.4) and lemma (4.1.2), we will now show that $\alpha_q(x)$ is a continuous function.

Lemma 4.1.5. *$\alpha_q(x)$ is a continuous function of x.*

Proof. We have to show that

$$\lim_{x\to x_0} \alpha_q(x) = \alpha_q(x_0)$$

Let us take y such that $y < x_0$. Using lemma (3.1.4), we can write

$$\frac{1 - \alpha_q(y)}{y} \geq \frac{1 - \alpha_q(x_0)}{x_0}$$

$$\alpha_q(y) \leq 1 - \frac{y}{x_0}(1 - \alpha_q(x_0))$$

As the function $\alpha_q(x)$ is a decreasing function, we get

$$\alpha_q(y) \geq \alpha_q(x_0)$$

Therefore, we get

$$\alpha_q(x_0) \leq \alpha_q(y) \leq 1 - \frac{y}{x_0}(1 - \alpha_q(x_0))$$

It implies that

$$\lim_{y \to x_0^-} \alpha_q(y) = \alpha_q(x_0)$$

Similarly for $x > x_0$, we get

$$\alpha_q(x_0) \geq \alpha_q(x) \geq 1 - \frac{x}{x_0}(1 - \alpha_q(x_0))$$

Therefore, we get

$$\lim_{x \to x_0^+} \alpha_q(x) = \alpha_q(x_0)$$

□

Using Plotkin bound, we get another linear bound on $\alpha_q(x)$ and it also determines the exact value of $\alpha_q(x)$ in a particular interval.

Lemma 4.1.6.

$$\alpha_q(x) \leq 1 - \frac{x}{\theta} \qquad \textit{if } x \in [0, \theta]$$

$$\alpha_q(x) = 0 \qquad \textit{if } x \in [\theta, 1]$$

Proof. Using Plotkin bound (1.2.4) for a $[n, M, d]_q$ code $\mathcal{C}$, we get

$$M \leq \lfloor \frac{1}{1 - (\theta/\delta)} \rfloor$$

where $\theta = 1 - q^{-1}$ and $\delta = d/n$ provided that $\delta > \theta$. To get the asymptotic version, we take $d = xn$.

$$n^{-1}\log_q A_q(n, xn) = n^{-1}\log_q A_q(n, \lceil xn \rceil) \le n^{-1}\log_q(1 - (n\theta/\lceil xn \rceil))$$

Now taking $\limsup_{n\to\infty}$ on both sides, we get

$$\alpha_q(x) \le \limsup_{n\to\infty} n^{-1}\log_q(1 - (n\theta/\lceil xn \rceil))$$

For $x > \theta$ we get $\alpha_q(x) \le 0$, but we already know that $\alpha_q(x) \ge 0$. It implies that $\alpha_q(x) = 0$ when $x > \theta$. We also get $\alpha_q(\theta) = 0$ due to continuity property of $\alpha_q(x)$. Using lemma (4.1.4) in interval $x \in [0, \theta]$, we have

$$\frac{1 - \alpha_q(x)}{x} \ge \frac{1 - \alpha_q(\theta)}{\theta} = \frac{1}{\theta}$$

$$\alpha_q(x) \le 1 - \frac{x}{\theta} \quad .$$

□

4.2 Non-linear upper bounds on $\alpha_q(x)$

In this section, we discuss the asymptotic version of Hamming and Gilbert-Varshamov bound. We will also state Elias bound which is a better upper bound than Hamming and Plotkin bound for $\alpha_q(x)$ for a low relative minimum distance. Let us begin this section by defining Shannon-Hilbert q-ary entropy function.

Definition 4.2.1. *Shannon-Hilbert q-ary entropy function is a function $H_q : [0,1] \to [0,1]$*

$$H_q(x) = -\left((1-x)\log_q(1-x) + x\log_q\left(\frac{x}{q-1}\right)\right) \tag{4.7}$$

The importance of the above function can be understood from the next two theorems.

Theorem 4.2.1. *Let $S_q(n, r)$ denote the size of sphere of radius r in space $\mathcal{F}^n$ where $|\mathcal{F}| = q$*

and a sequence of numbers t_n such that $\lim_{n\to\infty} t_n/n = t$ then

$$\lim_{n\to\infty} \frac{\log_q(S_q(n,t_n))}{n} = H_q(t) \tag{4.8}$$

Theorem 4.2.2. *Let x_n be a sequence of numbers such that $\lim_{n\to\infty} x_n/n = x$ and $x_n/n \in [0,1]$ then*

$$\lim_{n\to\infty} \frac{\log_q(V_q(n,x_n))}{n} = \begin{cases} H_q(x), & \text{if } 0 \le x \le \theta \\ 1, & \text{if } \theta \le x \le 1 \end{cases} \tag{4.9}$$

The first theorem is a standard theorem in coding theory. The second theorem is a result of first theorem which can be proved using the following relation between $S_q(n,r)$ and $V_q(n,r)$.

$$V_q(n,r) = S_q(n,0) + S_q(n,1) + \cdots + S_q(n,r)$$

Next, we will calculate the asymptotic version of Hamming and Gilbert-Varshamov bound using the above theorem.

Theorem 4.2.3. *Asymptotic version of Hamming bound:*

$$\alpha_q(x) = \alpha_H(x) \le 1 - H_q(x/2) \quad \text{for } \ x \in [0,1]$$

Proof. First, we note that

$$A_q(n,xn) = A_q(n,\lceil xn \rceil)$$

To get the asymptotic version of Hamming bound, We will take $d = xn$.

$$A_q(n,xn) \le \frac{q^n}{V_q(n, \frac{\lceil xn \rceil - 1}{2})}$$

Let us assume that $x_n = \frac{\lceil xn \rceil - 1}{2}$. It is easy to observe that $x_n/n \to x/2$ as $n \to \infty$ and also $0 \leq x/2 \leq \theta$. Using the theorem (4.2.2), we get

$$\alpha_q(x) = \limsup_{n\to\infty} \frac{\log_q A_q(n, xn)}{n} \leq 1 - \limsup_{n\to\infty} \frac{\log_q V_q(n, xn)}{n} = 1 - H_q(x/2)$$

□

Theorem 4.2.4. *Asymptotic version of Gilbert-Varshamov bound:*

$$\alpha_q(x) \geq 1 - H_q(x) \quad \textit{for} \;\; x \in [0, \theta] \tag{4.10}$$

Proof. To get the asymptotic version of Gilbert-Varshamov bound, we take $d = xn$.

$$A_q(n, xn) \geq \frac{q^n}{V_q(n, \lceil xn \rceil - 1)}$$

Let us consider a sequence $x_n = \lceil xn \rceil - 1$. We get that $\lim_{n\to\infty} \frac{\lceil xn \rceil - 1}{n} = x$. As we have assumed that $0 \leq x \leq \theta$, using the theorem (4.2.2) we get

$$\alpha_q(x) = \limsup_{n\to\infty} \frac{\log_q A_q(n, xn)}{n} \geq 1 - \limsup_{n\to\infty} V_q(n, x_n) = 1 - H_q(x)$$

□

To discuss about the Elias bound, we will first recall Bassalygo-Elias lemma given in the Chapter 1. It states

$$A_q(n, d) \leq \frac{q^n A_q(n, d; \mathcal{L})}{|\mathcal{L}|}$$

where $|\mathcal{L}|$ is size of the set $\mathcal{L}$ and $A_q(n, d; \mathcal{L})$ represents largest possible size of code of length n and minimum distance at least d. If we choose a sequence of particular balls $\mathcal{L}_n$ as $\mathcal{L}$ then the above inequality asymptotically gives the Elias bound and it is represented as $\alpha_E(x)$.

Theorem 4.2.5. *Elias bound:*

$$\alpha_q(x) \leq \alpha_E(x) = \alpha_H(2\theta(1 - \sqrt{1 - x/\theta})) \tag{4.11}$$

where $\theta = 1 - q^{-1}$ *and* $x \in [0, \theta]$.

The following graph shows comparison of Elias, Hamming and Plotkin bound for $q = 8$ in Fig. 4.1

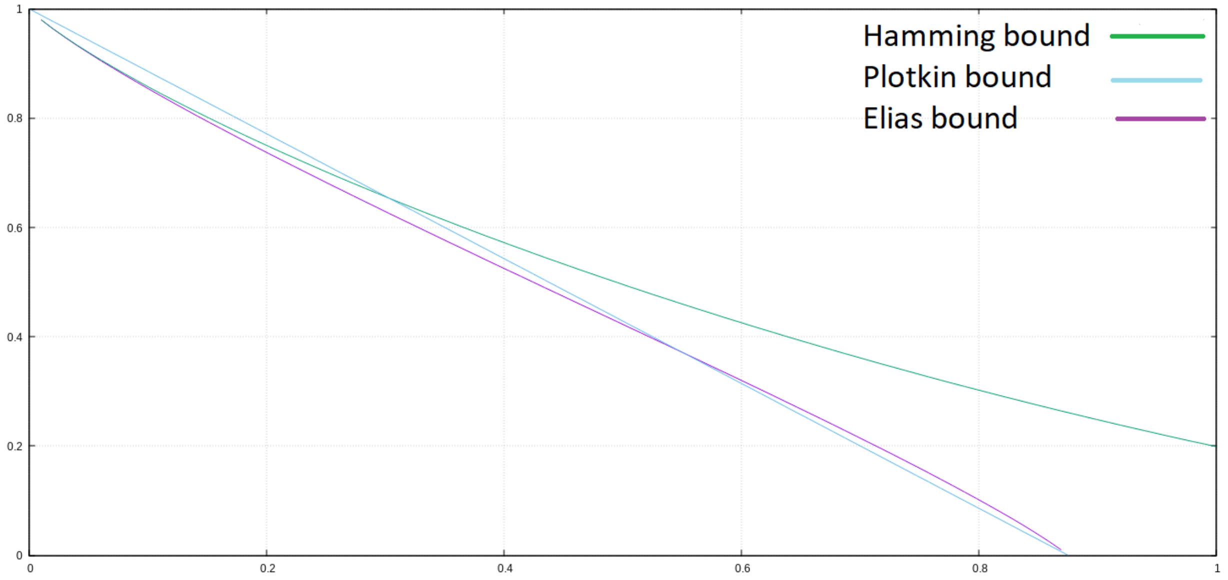

Figure 4.1: The Elias, Hamming and Plotkin bounds shown in magenta, green and blue respectively.

4.3 Convexity property of $\alpha_q(x)$

In this section, we discuss about improvement in the asymptotic version of Hamming and Elias bound due to K. Kaipa. We also discuss about the open question on $\cup-$convexity of $\alpha_q(x)$.

Let us recall the anticode bound stated in the Chapter 1. It states that for any $[n, M, d]_q$ code $\mathcal{C}$ and $\mathcal{L}$ be any anticode of diameter $d - 1$, we have

$$M \leq \frac{q^n}{|\mathcal{L}|}$$

To get good upper bound on the size of the code, we must choose the anticode with maximum size and radius $d - 1$. Let the maximum size of an anticode of radius $d - 1$ and length n is represented by $A_q^*(n, d - 1)$. Therefore, we can rewrite the anticode bound in a following way

$$A_q(n,d) \leq \frac{q^n}{A_q^*(n,d-1)}$$

The asymptotic version of the above bound is

$$\alpha_q(x) \leq 1 - \alpha_q^*(x) \tag{4.12}$$

where $\alpha_q^*(x)$ is given by

$$\alpha_q^*(x) = \liminf_{n\to\infty} \frac{\log_q A_q(n,xn)}{n}$$

For $\alpha_q^*(x)$, we have a following (in)equality [1] due to Ahlswede and Khachatrian in [1].

$$\alpha^*(x) \geq \left\{ \begin{array}{cc} H_q(x/2) & 0 \leq x \leq 2/q \\ 1-(1-x)H_q(1) & 2/q \leq x \leq 1 \end{array} \right. .$$

Using the above expression in asymptotic version of anticode bound, we get a new upper bound and it is called as hybrid Hamming-Singleton bound. This upper bound improves both Hamming and Singleton bound. It is given by

$$\alpha(x) \geq \alpha_{HS}(x) = \left\{ \begin{array}{cc} 1-H_q(x/2) & 0 \leq x \leq 2/q \\ (1-x)H_q(1) & 2/q \leq x \leq 1 \end{array} \right. .$$

We can observe that the hybrid Hamming-Singleton bound is equal to Hamming bound in the interval $[0,2/q]$. The function in interval $[2/q,1]$ is a tangent to $\alpha_H(x)$ at $x=2/q$. The improvement of Hamming and Singleton bound for $q=8$ is compared in a Fig. 4.2

The improvement in Elias bound for non-binary codes is achieved by replacing the set $\mathcal{L}$ in the Bassalygo-Elias lemma by a sequence of specific anticodes $\mathcal{L}_n$ to asymptotically give improvement in Elias bound. The new upper bound is called as hybrid Elias-Plotkin bound. It improves both Elias and Plotkin bound.

[1] It was proved by Ahlswede and Khachatrian that this is actually an equality. For our purposes, however, we only need the inequality and it is much easier to prove.

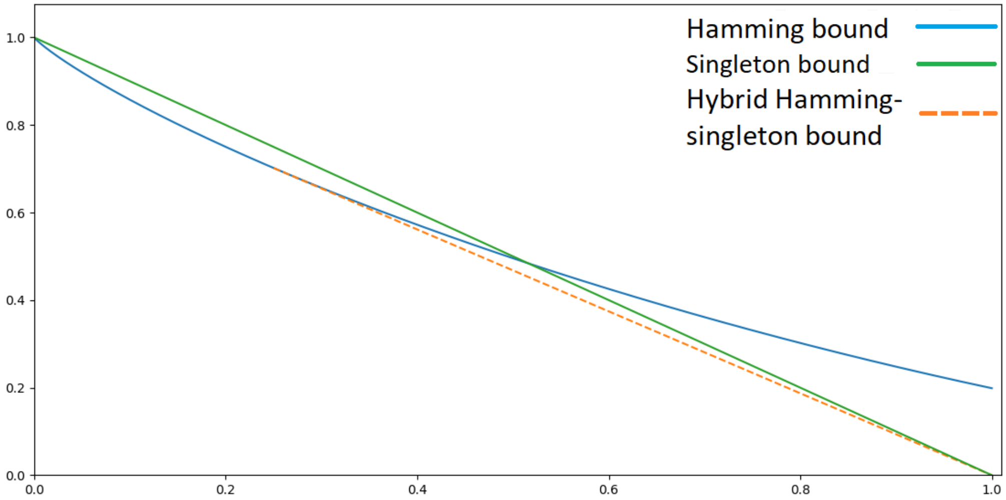

Figure 4.2: Improvement in Singleton and Hamming bound

$$\alpha_{EP}(x) = \begin{cases} 1 - H_q(\theta - \sqrt{\theta^2 - x\theta}) & 0 \leq x \leq \frac{2q-3}{q(q-1)} \\ (\theta - x)\frac{(q-1)H_q(1)}{q-2} & \frac{2q-3}{q(q-1)} \leq x \leq \theta \end{cases}$$

The proof of the hybrid Elias-Plotkin bound is given in [4]. We observe that the hybrid Elias-Plotkin bound is equal to Elias bound in the interval $[0, \frac{2q-3}{q(q-1)}]$. The improvement in the interval $[\frac{2q-3}{q(q-1)}, \theta]$ is a tangent to $\alpha_E(x)$ at $x = \frac{2q-3}{q(q-1)}$. Hybrid Elias-Plotkin bound is a correction to the non-convex part of Elias bound. The improvement of Elias bound for $q = 8$ is compared in Fig. 4.3

The convexity condition on $\alpha_q(x)$ is given as follows

$$\alpha_q(tx + (1-t)y) \leq t\alpha_q(x) + (1-t)\alpha_q(y)$$

The above property holds for $x = 0$. It is proved in lemma(3.1.3). In the paper [4] it has been conjectured by K. Kaipa that the above conditions also holds for $x = \theta$. If we assume this conjecture is true then the improvement in Elias and hamming bound can easily be seen.

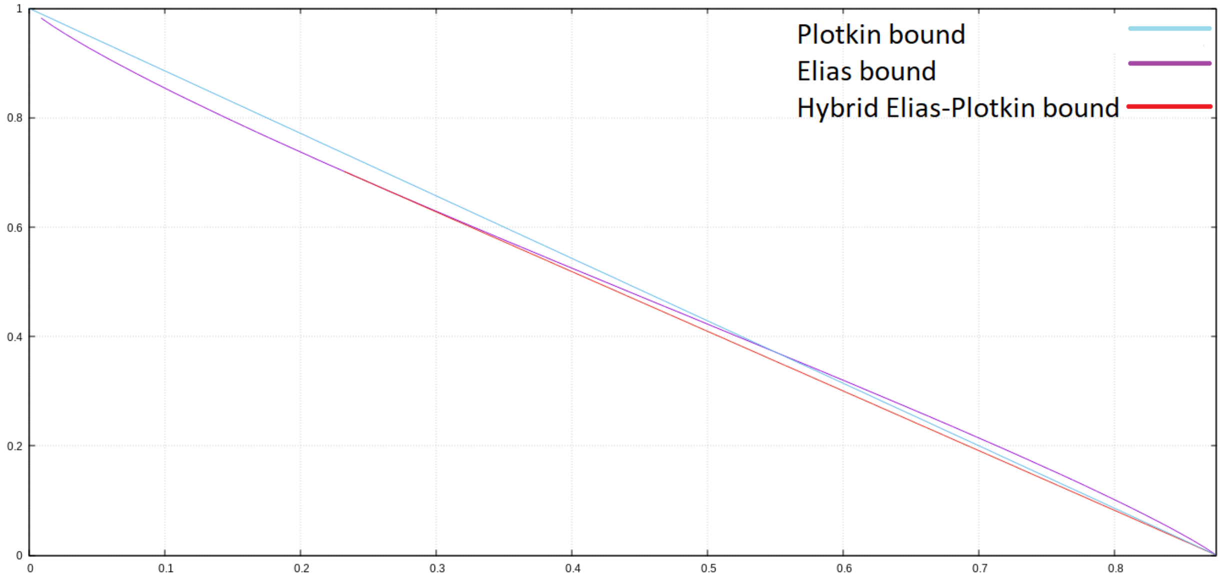

Figure 4.3: Improvement in Elias and Plotkin bound

Chapter 5

Results and Conclusion

In coding theory, we require codes with high error correcting capacity and efficiency which are determined by parameters relative minimum distance and rate of transmission respectively. As these requirements are mutually conflicting, we have studied their trade-off by asking the question that what is the largest size of the code $A_q(n, d)$ for minimum distance d. As the exact value of $A_q(n, d)$ is hard to calculate, in Chapter 1, we have discussed the bounds and properties of $A_q(n, d)$. The most intuitive upper bound for the size of the code is Hamming bound. The codes attaining this bound are called perfect codes. The problem of existence of perfect codes is a very interesting problem and it is unresolved for $e = 2$ case. We attempted this problem in Chapter 3. Our work in Chapter 3 is motivated from the differential equation derived by Lloyd for binary case. In this thesis, we have generalized the differential equation by Lloyd for binary case to any size of the alphabet. By solving the differential equation for $e = 2$ case, we provide the expression for distance enumerator polynomial of double error correcting perfect codes. In this work, we also present an alternate proof for Lloyd's theorem for any size of the alphabet. We have recovered and generalized some necessary conditions on existence of perfect codes by Reuvers and Olof Heden for any e. As the necessary condition by Lloyd is not enough to solve the $e = 2$ case, we will need more stronger necessary conditions to resolve the problem.

The question of finding the maximum rate of transmission became more tractable by defining asymptotic rate function which captures the trade-off between these quantities. As we do not know the exact value of $A_q(n, d)$, the exact value of asymptotic rate function is unknown. Understanding the properties of asymptotic rate function is a very important problem in coding theory. In Chapter 4, we studied the properties and bounds on asymptotic rate function by finding the asymptotic version of bounds on $A_q(n, d)$ discussed in Chapter 1. The work done in [4] by K. Kaipa gives insights into the special case of an open problem of $\cup-$convexity property of the asymptotic rate function which is consider as simpler problem compared to the problem of finding the exact value of the asymptotic rate function or the question to check if the asymptotic rate function is differentiable or not.

Bibliography

[1] Rudolf Ahlswede and Levon H. Khachatrian. The diametric theorem in hamming spaces-optimal anticodes. *Adv. Appl. Math.*, 20(4):429–449, May 1998.

[2] P. Delsarte. An algebraic approach to the association schemes of coding theory. *Philips Res. Rep. Suppl.*, (10):vi+97, 1973.

[3] Olof Heden and Cornelis Roos. The non-existence of some perfect codes over non-prime power alphabets. *Discrete Math.*, 311(14):1344–1348, 2011.

[4] Krishna Kaipa. An improvement of the asymptotic Elias bound for non-binary codes. *IEEE Trans. Inform. Theory*, 64(7):5170–5178, 2018.

[5] S. P. Lloyd. Binary block coding. *Bell System Tech. J.*, 36:517–535, 1957.

[6] F. J. MacWilliams and N. J. A. Sloane. *The theory of error-correcting codes. II.* North-Holland Publishing Co., Amsterdam-New York-Oxford, 1977. North-Holland Mathematical Library, Vol. 16.

[7] Robert J. McEliece, Eugene R. Rodemich, Howard Rumsey, Jr., and Lloyd R. Welch. New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities. *IEEE Trans. Information Theory*, IT-23(2):157–166, 1977.

[8] Henricus Franciscus Hubertus Reuvers. *Some non-existence theorems for perfect codes over arbitrary alphabets.* Technische Hogeschool Eindhoven, Eindhoven, 1977. Written for conferment of a Doctorate in Technical Sciences at the Technische Hogeschool, Eindhoven, 1977.

[9] J. H. van Lint. *Introduction to coding theory*, volume 86 of *Graduate Texts in Mathematics.* Springer-Verlag, Berlin, third edition, 1999.