特　別　研　究　報　告

題　目

Sphere Packing Bound and Gilbert-Varshamov Bound for b-Symbol Read Channels

b-シンボル読出通信路における球充填限界とG-V限界

指　導　教　員

Prof. Toru Fujiwara

報　告　者

SONG SEUNGHOAN

February 14, 2017

Department of Information and Computer Sciences
School of Engineering Science
Osaka University

Sphere Packing Bound and Gilbert-Varshamov Bound for b-Symbol Read Channels

b-シンボル読出通信路における球充填限界とG-V限界

Supervised by
Prof. Toru Fujiwara

Presented by
SONG SEUNGHOAN

February 14, 2017

Department of Information and Computer Sciences
School of Engineering Science
Osaka University

Sphere Packing Bound and Gilbert-Varshamov Bound for b-Symbol Read Channels

SONG SEUNGHOAN

Abstract

Recent rapidly growing storage technologies have led to the development of high-density storage devices. However when information is written with high resolution write devices and read by lower resolution devices, it possibly occurs that individual symbols of data are not distinguished at a read but several symbols are read at a read. Over the hypothesis that the number of symbols read at once is two, a channel model called symbol-pair read channel is proposed by Cassuto and Blaum. Pair-metric is defined for the error-correction in symbol-pair read channel and sphere-packing bound, Gilbert-Varshamov(G-V) bound, and asymptotic G-V bounds were derived for symbol-pair read channel. Also, from analysis of asymptotic G-V bounds, Cassuto and Blaum showed that there exist strictly higher rate codes over symbol-pair read channel compared to those over conventional channels with Hamming metric. As a generalization of symbol-pair read channel, Yaakobi *et.al.* proposed b-symbol read channel whose number of concurrently read symbols is b and defined b-symbol metric for error-correction over this channel model.

In this research, we derive sphere-packing bound, G-V bound, and asymptotic G-V bounds for b-symbol read code. Based on a calculation of the size of b-symbol sphere and ball, sphere packing bound and G-V bound are obtained. Then we show asymptotic G-V bounds by approximation and asymptotic methods of G-V bound. Finally, the relationship between asymtotic b-symbol G-V bounds and the number b of read symbols is analyzed.

We confirmed that the derived bounds for 1-symbol read channels coincide with the bounds for conventional channels with Hamming metric. Also, from the analysis of asymptotic b-symbol G-V bounds proved in this research, we showed that the number b of read symbols achieving the highest bound is determined dependently on fractional minimum distance $\delta = \frac{d}{n}$. Furthermore, the existence of b-symbol codes with strictly higher rates is shown as b becomes larger.

Keywords

error-correcting code, codes for storage media, symbol-pair read channel, b-symbol read channel, sphere packing bound, Gilbert-Varshamov bound, asymptotic G-V bound

平成 29 年度 特別研究報告

b-シンボル読出通信路における球充填限界と G-V 限界

SONG SEUNGHOAN

内容梗概

近年, 記憶装置の発展により, きわめて高密度な記憶装置が開発されている. しかし記憶装置が高密度であるものの読出装置の精度が低い場合, 一度の読出しで一つのシンボルではなく複数個のシンボルが同時に読み出される現象が起こり得る. 同時に読み出されるシンボル数が二つである前提で, 2010 年, Cassuto らによってシンボルペア読出通信路が提案された. シンボルペア読出通信路での誤り訂正のためにペア距離空間が定義され, ペア距離空間での球充填限界, Gilbert-Varshamov(G-V) 限界, 漸近的 G-V 限界が Cassuto らによって導出されている. また, 漸近的 G-V 限界を分析し, シンボルペア読出通信路上では, ハミング距離を用いる通信路上でより, 高い符号レートを持つ符号が存在することが示された.

シンボルペア読出通信路を一般化し, 隣接した b 個のシンボルが同時に読み出される b-シンボル読出通信路が, 2016 年, Yaakobi らによって提案された. しかし, b-シンボル読出通信路における球充填限界, G-V 限界, 漸近的 G-V 限界はまだ示されていない.

本研究では, b-シンボル読出通信路における球充填限界, G-V 限界, 漸近的 G-V 限界を導出する. まず, b-シンボル距離の求め方を提案し, それを用いて b-シンボル球面や b-シンボル球のサイズを導出することで, 球充填限界と G-V 限界を示す. また, G-V 限界の漸近的な近似により, 漸近的 G-V 限界を導出する. 最後に, 漸近的 G-V 限界がシンボル読み出し数 b によって, どのように変化するか分析する.

導出した限界に対して, $b = 1$ のときの球充填限界, G-V 限界, 漸近的 G-V 限界が, ハミング距離を用いた通信路の球充填限界, G-V 限界, 漸近的 G-V 限界と一致することを確認できた. また, この研究で証明した漸近的 G-V 限界に対して, 最も高い符号レートを持つ符号が求まる b は, fractional minimum distance $\delta = \frac{d}{n}$ に依存して決まることがわかった. 加えて, b を大きくすれば, より高いレートを持つ b-シンボル符号が存在することがわかった.

主な用語

誤り訂正符号, 記憶媒体向け符号, シンボルペア符号, b-シンボル読出通信路, 球充填限界, Gilbert-Varshamov(G-V) 限界, 漸近的 Gilbert-Varshamov(G-V) 限界

Contents

1 Introduction

One of the fundamental objective of information theory is to construct reliable communications process. Due to the existence of many sorts of errors in information process, error correction during information process is indispensable for reliable communications. In data transmition, data are sent to the receiver, errors occur until the data arrive to the receiver, and then errored data are corrected by receiver. In storage devices, written data are considered to be sent data, errors occur while reading the written data, and the reader corrects errors of data. As an integrated framwork explaining both of data transmition and storage device reading, information theory introduce a framework called a *channel model* which inputs data to be sent and outputs data corrupted by errors.

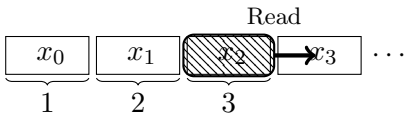
Traditional approaches in information theory are based on an unit of data, called *symbol*. Sent and received data is expressed as a sequence of individual symbols. And in most channels, the symbols for sent data and received data is assumed to be the same units. Based on this assumption, researches related to codes and bounds for channels have been progressed.

However, even though the information theory accumulated significant achievements, the recent rapidly growing hardware technologies brought out a new problem related to symbols. With the development of high density data storage devices, it happens that the performance of writing devices and reading devices varies, which lead to several kinds of decoding error. One specific form of these errors is the phenomenon that several symbols are read at once because the resolution of the reading device is lower than that of writing device. In this case, limited to the current channel models, the decoding process should be seperated into two parts; seperating multiple symbols read at once to each symbol, and recovering read errors. It is expected to improve the efficiency of decoding by integrating two decoding process.

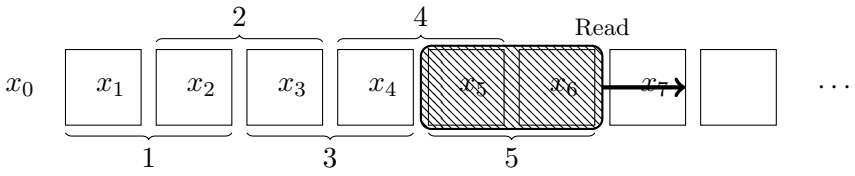
The framework called *symbol-pair read channel* was proposed by Cassuto and Blaum [2], for this purpose in case of the number of concurrently read symbols is two. In symbol-pair read channel, consecutive two symbols, called *pair-symbol*, are read as a one unit but it is hypothesized that the former symbol and the latter symbol can be distinguish after reading the pair-symbol. For example, if 101 is the input of the channel and the channel makes no error, the output of symbol-pair read channel would be $[10, 01, 11]$. In this framework, at least one error of pair-symbol is considered a *pair-error*.

The fundamental structures of symbol-pair channels and codes were proposed in [2] and [3]. The definition of pair-metric, pair-error correctability, code construction and decoding methods, and lower and upper bounds for codesizes were answered. Based on these results, in [6], [7], and [8], maximum distance seperable codes for symbol-pair read channels were studied and an MDS codes with certain parameters were found. Decoding algorithms were also studied over symbol-pair read channel: decoding for cyclic codes [5] and [9], syndrome

(A) : Channel without resolution deficiency of reader



(B) : Symbol-pair read channel



(C) : b-symbol read channel($b = 4$)

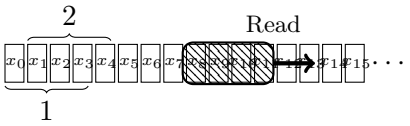


Figure 1: Read process of channels

decoding [10], algebraic decoding of BCH codes [11], and linear programming decoding for binary linear codes [12]. Symbol-pair read channel was generalized to *b-symbol read channel* [4] where $b \geq 2$ consecutive symbols were read as a unit, instead of two symbols. In [4], basic notations and code constructions for b-symbol read channels were introduced.

One of the primary goal of this report is deriving sphere packing bound, Gilbert-Varshamov bound, and finally, asymptotic Gilbert-Varshamov bound for b-symbol read channels. These three bounds for symbol-pair read channel is obtained in [2]. It should be denoted that the derivation of these bounds for b-symbol read channel in this report is basically based on [2].

Then we analyzes the relationship between b and the asymptotically achievable rate. In [2], Cassuto and Blaum showed that symbol-pair codes achieve higher asymptotic Gilbert-Varshamov bound for relative minimum distance $\delta < 0.27$ compared to codes for Hamming metric. Later in [3], Cassuto and Litsyn improved the result in [2] and showed the enhanced asymptotic Gilbert-Varshamov bound for symbol-pair codes is strictly higher than asymptotic Gilbert-Varshamov bound for Hamming metric for all the relative minimum distance δ. In terms of b-symbol read channels, we conjectured the bound may grow higher as b grows, but the growth of the bound may have limitation. In this report, we used the method similar to that of [2] and generalized the asymptotic Gilbert-Varshamov bound

for b-symbol codes. As a result, for the bound we showed in this report, we disproved that the larger b is, the more suitable for getting higher bounds. Rather, a finite number b such that achieves highest asymptotic b-symbol G-V bound is determined dependently on the relative minimum distance δ.

The remainder of this report is organized as follows. We state the basics of error-correcting codes with Hamming distance in Section 2 and the basic concepts of b-symbol read channel and b-symbol codes in Section 3. In Section 4, we review sphere packing bound and Gilbert-Varshamov bound with the size of sphere and ball when a metric space is given. In Section 5, sphere packing bound and Gilbert-Varshamov bound for b-symbol read channel will be derived. Finally, in Section 6, we will show asymptotic b-symbol Gilbert-Varshamov bound and give, in terms of the bound we derived, the result that higher b does not nessesarily lead to the higher asymptotic b-symbol Gilbert-Varshamov bound. The report ends with a short conclusion and some open questions in Chapter 7.

2 Error Correcting Codes

In the study of error-correcting codes, the standard scheme has been constructed with Hamming distance. Before we explore b-symbol read channels, it is worth reviewing the standard error-correcting codes which have been studied for several decades. Therefore we introduce elementary concepts about codes, channels and error-correcting process in this section.

In Section 2.1 and 2.2, we will see formal discription of codes and error-correcting process. Though it is possible to integrate the two sections, we seperate these to distinguish the the properties that do not change when the channel model is b-symbol read channel and those that change. In Section 2.1, we focus on the definitions and properties of codes that are applied in the same way to both channel models, and those not in Section 2.2. Lastly, we see how we evaluate asymptotically good codes in Section 2.3.

2.1 Formal Description of Codes

A Simple Code and Error-Correcting Process

To help understanding code and error-correcting process, we introduce a simple example of code called *repetition code*.

Example 2.1. *Consider sending a message in*

$$M = \{00, 01, 10, 11\}.$$

If a bit flip error occur while sending a message, the received message is different from the original message. For example, if we are sending 00 *and a bit flip occured to the fisrt symbol, the received message is* 10*. In this case the receiver cannot discern whether recieved message* 10 *is sent from sender or it is a corrupted message.*

To resolve this problem, consider sending an element of

$$C = \{000000, 000111, 111000, 111111\}$$

instead of $\{00, 01, 10, 11\}$*. Then a bit-flip error can be recovered to the most similar element in* C*. If the sender wants to send a message* 01 *and* 000111 *is sent instead, then a bit-flip error to the first symbol* 100111 *will be successfully recovered to* 000111*.*

Here, the set C *is called a code and the elements of* C *are called codewords.*

Let us get the formal description of codes we previewed in the former example. A code is a set of words to be sent through channels. The concepts in this section are directly applied to codes for b-symbol read channel.

Alphabet : The finite set Σ of possible symbols. Throughout this report, we suppose a symbol 0 is in Σ without losing generality. $|\Sigma|$ is denoted as q.

Word : An element of the set Σ^n.

Code Length : The number n of symbols that a word has.

Code : A set C of words that are sent through channels. It is the subset of Σ^n. With $q = |\Sigma|$, a code C is called a q-ary code. When $|\Sigma| = 2$, we call C is a binary code.

Codeword : A word in a code C.

In Example 2.1, the alphabet Σ is $\{0, 1\}$ and codewords are sequences consist of 0s and 1s that represent messages in $\{00, 01, 10, 11\}$.

Definition 2.1 (Code Rate)**.** *We call the following value $R(C)$ code rate, or rate, of q-ary code C:*

$$R(C) = \frac{\log |C|}{\log \Sigma^n} = \frac{\log_q |C|}{n}. \tag{1}$$

Code rate is the density of the code over the space Σ^n. Code rate indicates how many codewords a code expresses using limited n resources with error-correcting ability. The maximum value of the code rate is 1 when the $C = \Sigma^n$. However, even though higher rate is expected when building codes, it is, in most cases, the biggest challenge for code constructors because the code rate is a trade-off for error corretablilty what we will see in next section.

2.2 Error-Correcting Process for Channels with Hamming Metric

Figure 2: Channel Model

Channel is a structure that inputs a codeword and outputs a vector with a specific kind of error. The decoding process is recovering the output vector with error to a codeword. The error-correcting process succeeds if the recovered codeword is not the same as the input codeword. If the recovered codeword is the same as the input codeword, it is called *decoding error*. If the output vector can not be recovered to any codeword, it is called *decoding failure*. If the set R of possible output vector is decided, the decoding process is considered as a mapping from R to Σ^n or $\Sigma^n \cup \{fail\}$. Therefore designing the mapping is the main goal of error correction and it is reliant on the property of channels which decide the kinds of errors.

In most of noisy memoryless channel models, the channel output vector is an element in Σ^n. Considering the input codewords are also elements of Σ^n, we can evaluate the difference of the output vector and codewords with an evalutaion metric over Σ^n; in

most cases in error correcting codes, Hamming distance which is defined in this section is the evaluation metric. With an evaluation metric over Σ^n, we can find a closest codeword whose difference from the output vector is the least and decode the output to the codeword. This decoding method is called *minimum distance decoding*. Of course, finding the closest codeword from the output vector is complicated in most of decoding processes. Therefore constructing decoding algorithm is one of the main research areas in error-correcting codes.

To describe the minimum distance decoding with Hamming distance, we first should define Hamming metric, an evaluation metric that we think over. In most of coding schemes, Hamming distance is a basic evalutation metric for decoding.

Definition 2.2 (Hamming Distance)**.** *Let $\vec{x}, \vec{y}$ be words in Σ^n and $\vec{a}_i$ denote the i-th coordinate of a word $\vec{a}$. Then Hamming distance function $D_H : \Sigma^n \times \Sigma^n \to \mathbb{N}$ is defined as follows ($\mathbb{N} = \{0, 1, 2, ...\}$):*

$$D_H(\vec{x}, \vec{y}) \triangleq |\{i : \vec{x}_i \neq \vec{y}_i\}|. \tag{2}$$

With Hamming distance as an evaluation metric, we can decode received word $\vec{y}$ to a codeword $\vec{x}' \in C$ from which Hamming distance to $\vec{y}$ is minimum compared to all other codewords in C.

Next, we consider how many errors a code can correct when the minimum distance d_H is determined. Fisrt, for a code $C \subset \Sigma^n$, *minimum distance* d_H is defined as follows:

$$d_H = \min_{\vec{x},\vec{y} \in C, \vec{x} \neq \vec{y}} D_H(\vec{x}, \vec{y}). \tag{3}$$

If the minimum distance d_H is known, error correctability of a code is proposed as the next theorem. To make the next theorem obvious, we denote *the number of errors* as the distance $D_H(\vec{x}, \vec{y})$ where $\vec{x}$ is a sent codeword and $\vec{y}$ is a received word.

Theorem 2.1. *A code C can correct t errors if and only if $d_H \geq 2t + 1$.*

Proof. Suppose $\vec{x} \in C$ is a sent codeword and $\vec{y} \in \Sigma^n$ is the received word with t errors. Then $D_H(\vec{x}, \vec{y}) = t$.

If there exists another codeword $\vec{x}' \in C$ such that $D_H(\vec{x}', \vec{y}) \leq t$, the received word $\vec{y}$ with t errors cannot be corrected because $\vec{y}$ might be miscorrected to $\vec{x}'$ when $D_H(\vec{x}', \vec{y}) < t$, called *decoding error*, or error correction might fail when $D_H(\vec{x}', \vec{y}) = t$, called *decoding failure*. We show that there exists no such a codeword $\vec{x}'$.

If there exists a codeword $\vec{x}' \in C$ such that $\vec{x} \neq \vec{x}'$ and $D_H(\vec{x}', \vec{y}) \leq t$,

$$D_H(\vec{x}, \vec{x}') \leq D_H(\vec{x}, \vec{y}) + D_H(\vec{x}, \vec{y}) \leq 2t. \tag{4}$$

This contradicts to

$$2t + 1 \leq d_H \leq D_H(\vec{x}, \vec{x}'). \tag{5}$$

Therefore there exists no codeword $\vec{x}' \in C$ such that $D_H(\vec{x}', \vec{y}) \leq t$ except for $\vec{x}$. □

2.3 Asymptotically Good Codes

Asymptotic analysis is a traditional method in information theory. It is introduced from the initial stage of information theory [1].

There are trade-offs between the code rate and the minimum distance. If a code has both good rate and good relative distance, the code is considered as an asymptotically good code. To make it formal, we consider a code C with minimum distance d asymptotically good if

$$\lim_{n\to\infty} R(C) > 0 \text{ and } \lim_{n\to\infty} \delta > 0 \tag{6}$$

where $\delta = \frac{d}{n}$.

For an example, a code C generated by three times repetition of each symbol in Σ^n, as a code in Example 2.1, has the rate

$$R(C) = \frac{|\Sigma|^n}{|\Sigma|^{3n}} = |\Sigma|^{-2n} \tag{7}$$

and minimum distance $d_H = 3$. This code C is not considered an asymptotically good code because when n grows asymptotically,

$$\lim_{n\to\infty} R(C) = \lim_{n\to\infty} |\Sigma|^{-2n} = 0, \lim_{n\to\infty} \frac{d_H}{n} = 0. \tag{8}$$

Construction of asymptotically good codes is an important but hard question in coding theory. In terms of existence of asymptotically good codes, it is already proved in channel with Hamming metric [14] and symbol-pair metric [2]. We will prove with b-symbol metric in Section 6.

3 b-Symbol Read Channels

In this section, we review a formal definition and properties of codes over b-symbol read channels proposed in [4]. Note on the difference from Section 2 while reading this section.

3.1 Formal Description of b-Symbol Read Channels

Let us review the previous section before we get to the formal description of b-symbol read channels. The standard methods for error correction is to decode the received vectors sent through the information channel to a codeword. To make it more formal, a finite set of symbols is defined as an *alphabet* Σ and a vector with symbol length n as a *word*. We also define the length of words as *code length* n, a set of words as a *code* $C \subset \Sigma^n$, and words in code C as *codewords*. Then channel can be described as a structure whose input is a codeword in C and the output is an element of R which is called a *channel output set*, and decoder is a mapping $dec : R \to \Sigma^n$. Differently from the ordinary channel whose output set R is Σ^n, b-symbol read channel takes $(\Sigma^b)^n$ as a channel output set. Therefore in this channel model, each channel output is a length n vector of which every coordinate consists of b symbols. Throughout this report, we call the codes over this channel *b-symbol codes*.

Figure 3: b-Symbol Read Channel

To distinguish vectors in $(\Sigma^b)^n$ from vectors in Σ^n, we express vectors in $(\Sigma^b)^n$ with the symbol $\leftrightarrow$, and call these vectors *b-symbol vectors*.

Notation 3.1 (b-Symbol Vector[4])**.** *b-symbol vector $\overleftrightarrow{u} \in (\Sigma^b)^n$ is denoted as follows:*

$$\overleftrightarrow{u} = [u_0^0 u_0^1 ... u_0^{b-1}, u_1^0 u_1^1 ... u_1^{b-1}, ..., u_{n-1}^0 u_{n-1}^1 ... u_{n-1}^{b-1}] \tag{9}$$
$$= [\overleftrightarrow{u}_0 ... \overleftrightarrow{u}_{n-1}] \tag{10}$$

with the notation that $\overleftrightarrow{u}_j = u_j^0 u_j^1 ... u_j^{b-1}$.

For a word $\vec{x} \in \Sigma^n$, we define the *b-symbol read vector* $\pi(\vec{x})$ that corresponds to $\vec{x}$.

Definition 3.1 (b-Symbol Read Vector[4])**.** *Let $\vec{x} = x_0 x_1 ... x_{n-1}$ a word in Σ^n. The b-symbol read vector $\pi(\vec{x})$ is defined as*

$$\pi_b(\vec{x}) \triangleq [x_0 x_1 ... x_{b-1}, x_1 x_2 ... x_b, ..., x_{n-1} x_0 ... x_{b-2}] \tag{11}$$
$$= [\pi_b(\vec{x})_0 ... \pi_b(\vec{x})_{n-1}]. \tag{12}$$

Each coordinate of length n vector $\pi_b(\vec{x})$ is composed of cyclically consecutive b symbols. With an example, it is simply understood how a b-symbol read vector is contructed.

Example 3.1. *If $b = 3$ and $\vec{x} = x_1x_2x_3x_4 = 0110$, then*

$$\begin{aligned}\pi_b(\vec{x}) &= [x_1x_2x_3, x_2x_3x_4, x_3x_4x_1, x_4x_1x_2] \\ &= [011, 110, 101, 001]\end{aligned}$$

From notation 3.1 and definition 3.1, it is trivial that every word in Σ^n has corresponding b-symbol read vectors, but not vice-versa. To make this clear, let a set $\pi_b(X)$ for $X \subset \Sigma^n$ [4] be

$$\pi_b(X) \triangleq \{\pi_b(\vec{x}) | \vec{x} \in X\}. \tag{13}$$

Then for a code C over Σ^n, $\pi_b(C) \subseteq \pi_b(\Sigma^n) \subset (\Sigma^b)^n$. In case that a b-symbol vector $\overleftrightarrow{u}$ is in $\pi_b(\Sigma^n)$, it is called *consistent*. In other words, All consistent vectors have corresponding word in Σ^n.

3.2 b-Symbol Distance

A metric over b-symbol read channel is a suitable evaluation metric for error-correctring as Hamming metric is for channels in Section 2. First, we define b-symbol distance over $(\Sigma^b)^n$.

Definition 3.2 (b-Symbol Distance[4])**.** *Let $\overleftrightarrow{u}, \overleftrightarrow{v}$ be vectors in $(\Sigma^b)^n$. Then, b-symbol distance function $D_b : (\Sigma^b)^n \times (\Sigma^b)^n \to \mathbb{N}$ is defined as follows ($\mathbb{N} = \{0, 1, 2, ...\}$).*

For $\overleftrightarrow{u}, \overleftrightarrow{v} \in (\Sigma^b)^n$,

$$D_b(\overleftrightarrow{u}, \overleftrightarrow{v}) \triangleq |\{i : (\overleftrightarrow{u})_i \neq (\overleftrightarrow{v})_i\}|. \tag{14}$$

For notational convenience, in case that one or both b-symbol vectors are consistent vectors, we express the b-symbol distance with the following notations:

for $\vec{x}, \vec{y} \in \Sigma^n$,

$$D_b(\vec{x}, \vec{y}) \triangleq D_b(\pi_b(\vec{x}), \pi_b(\vec{y})) \tag{15}$$

$$D_b(\overleftrightarrow{u}, \vec{x}) \triangleq (\overleftrightarrow{u}, \pi_b(\vec{x})) \tag{16}$$

$$D_b(\vec{x}, \overleftrightarrow{u}) \triangleq D_b(\pi_b(\vec{x}), \overleftrightarrow{u}). \tag{17}$$

As $\pi_b(\Sigma^n) = \{\pi_b(\vec{x}) | \vec{x} \in \Sigma^n\}$ is proper subset of $(\Sigma^b)^n$, Σ^n is a metric space with a distance function D_b. We confirm Σ^n is metric space.

Theorem 3.1. *Σ^n is a metric space with the distance function $D_b : (\Sigma^n, \Sigma^n) \to \mathbb{N}$.*

Proof. We show that the set Σ^n has the following three properties for all $\vec{x}, \vec{y}, \vec{z} \in \Sigma^n$.

(i) $D_b(\vec{x}, \vec{y}) \geq 0$; equality holds if and only if $\vec{x} = \vec{y}$

(ii) $D_b(\vec{x}, \vec{y}) = D_b(\vec{y}, \vec{x})$

(iii) $D_b(\vec{x}, \vec{y}) \leq D_b(\vec{x}, \vec{z}) + D_b(\vec{z}, \vec{y})$

Proving properties (i),(ii) is trivial from the defintion of the distance function D_b. Property (3) is derived by $[\pi_b(\vec{x})_i \neq \pi_b(\vec{y})_i] \implies [\pi_b(\vec{x})_i \neq \pi_b(\vec{z})_i] \vee [\pi_b(\vec{z})_i \neq \pi_b(\vec{y})_i]$. □

For a code $C \subseteq \Sigma^b$, *minimum b-symbol distance* d_b is defined as follows:

$$d_b = \min_{\vec{x},\vec{y} \in C, \vec{x} \neq \vec{y}} D_b(\vec{x}, \vec{y}). \tag{18}$$

Definition 3.3 (b-Symbol Weight[4])**.** *For* $\overleftrightarrow{u} \in (\Sigma^b)^n$*, b-symbol weight* w_b *is defined as follows* ($\vec{0} = 0^n = \underbrace{0...0}_{n}$) :

$$w_b(\overleftrightarrow{u}) \triangleq D_b(\pi(\vec{0}), \overleftrightarrow{u}). \tag{19}$$

3.3 Error Correctability of Codes over b-Symbol Read Channels

In absence of errors, $\pi_b(\vec{x})$ is received when $\vec{x}$ is sent. On the other hands, if errors exist over the channels, errors over the channels corrupt the original codeword $\vec{x}$ to a b-symbol vector $\overleftrightarrow{u}$. The number of errors is counted as the number of coordinates that b-symbol read vector $\pi_b(\vec{x})$ and b-symbol vector $\overleftrightarrow{u}$ differs.

By the similar way to Theorem 3.1, it is confirmed that $(\Sigma^b)^n$ is a metric space. For $\pi(C) \subset (\Sigma^b)^n$, the decoding over b-symbol read channels is also based on b-symbol distance.

Proposition 3.1. *A code* C *can correct* t *b-symbol errors if and only if* $d_b \geq 2t + 1$*.*

The proof of Proposition 3.1 is the same as Hamming metric case, Theorem 2.1.

4 Sphere Packing Bound and Gilbert-Varshamov Bound

Bounds for codes are indicators for the performance a code can achieve. Lots of bounds in information theory are derived from combinatorial calculation. Sphere packing bound and Gilbert-Varshamov bound are the two bounds for the codesizes. These two code bounds are derived when the metric of the space is determined. In this section, we suppose a metric D is predetermined and then explain these two bounds.

4.1 Sphere and Ball

As the sphere packing bound and Gilbert-Varshamov bound are expressed with the size of the ball, we define the sphere and the ball before. We denote a metric space X with the distance function $D : (X, X) \to \mathbb{N}$ as the metric space (X, D).

Definition 4.1 (Sphere and Ball)**.** *Let (Σ^n, D) be a metric space. Then for a word $\vec{x} \in \Sigma^n$, sphere and ball are defined, respectively as follows:*

$$\mathcal{S}_h(\vec{x}) \triangleq \{\vec{y} \in \Sigma^n | D(\vec{x}, \vec{y}) = h\}, \tag{20}$$

$$\mathcal{B}_r(\vec{x}) \triangleq \{\vec{y} \in \Sigma^n | D(\vec{x}, \vec{y}) \leq r\} = \bigcup_{h=0}^{r} \mathcal{S}_h(\vec{x}). \tag{21}$$

Note that the sizes of sphere and ball, $|\mathcal{S}_h(\vec{x})|$ and $|\mathcal{B}_r(\vec{x})|$ respectively, are independent of the word $\vec{x}$. In metric space with Hamming distance, regardless of the choice of $\vec{x}$,

$$|\mathcal{S}_h(\vec{x})| = \binom{n}{h}(q-1)^h, |\mathcal{B}_r(\vec{x})| = \sum_{h=0}^{r} \binom{n}{h}(q-1)^h. \tag{22}$$

4.2 Sphere Packing Bound

Sphere packing bound is upper bound for code size when error-correctability of the code is given that t or less errors are always correctable.

Theorem 4.1 (Sphere-Packing Bound[14])**.** *Let (Σ^n, D) be a metric space. If a q-ary code $C \subset \Sigma^n$ can correct all t or less errors, then*

$$|C||\mathcal{B}_t(\vec{x})| \leq q^n. \tag{23}$$

We give simple proof of sphere packing bound. When a word $\vec{x}$ is sent, the output vectors with less than t errors are in the radius t ball centered at $\vec{x}$. Considering that code C can corrects all t or less errors, radius t balls centered at codewords should not intersect. The size of the entire radius t balls in space Σ^n without intersections is less than the size of space Σ^n (Figure 4).

(The size of radius t ball) $\times$ (The number of codewords) $\leq$ (The size of space Σ^n)

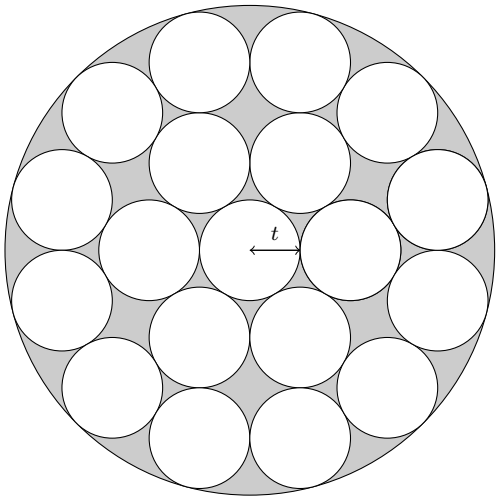


Figure 4: A diagram for obtaining sphere packing bound

4.3 Gilbert-Varshamov Bound

While sphere paking bound proposes upper bound for codesizes, Gilbert-Varshamov(G-V) bound gives lower bound for codesizes.

Theorem 4.2 (Gilbert-Vashamov Bound[14]). *Let (Σ^n, D) be a metric space. There exists a q-ary code $C \in \Sigma^n$ with minimum distance $d = \min_{\vec{x},\vec{y}\in C,\vec{x}\neq\vec{y}} D(\vec{x},\vec{y})$ such that*

$$|C||\mathcal{B}_{d-1}(\vec{x})| \geq q^n. \tag{24}$$

We also give simple proof of Gilbert-Varshamov bound. Considering minimum distance is d, no code word should be in radius $d-1$ ball centered at other codewords. If codewords are not in radius $d-1$ balls centered at other codewords, minimum distance d is kept.

Then we construct a code satisfying Gilbert-Varshamov bound as follows: repeat adding a word outside radius $d-1$ balls centered at other codewords the to code C until there is no word to add in the space Σ^n (Figure 5). If there is no room for adding a word to code, the entire balls cover the space Σ^n and we get G-V bound.

However, differently from the proof of sphere packing bound, the balls can intersect if

4.4 Asymptotic Gilbert-Varshamov Bound

Gilbert-Varshamov bound is reformulated with asymptotic methods. The reformulation is called *asymptotic Gilbert-Varshamov bound*, the bound for existence of certain rate codes when code length n grows infinity.

Asymptotic G-V bounds are great indicator for the performance of channels. We need an indicator for comparing the performance of codes over channels. However, G-V bound

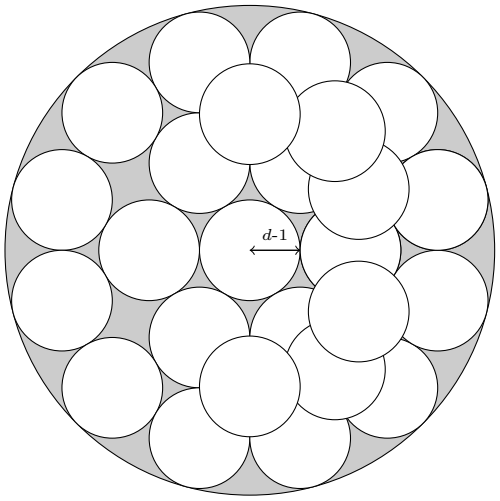


Figure 5: A description of adding a word(center of yellow ball) to code C(set of centers of ball) in the proof of Gilbert-Varshamov bound

is expressed in terms of the code length n. Asymptotic G-V bounds make it possible to compare performance of codes by restricting code length n to infinity Therefore we will compare the performance of codes over channels in terms of asymptotic G-V bounds.

Furthermore, asymptotic G-V bound is written in terms of fractional minimum distance $\delta = \frac{d}{n}$ which is great indicator for error-correctability. G-V bound is written in terms of minimum distance d. The error correctability of two codes with minimum distance d is not same if code length n is different.

Asymptotic Gilbert-Varshamov bound in Hamming metric [14] and symbol-pair metric [2] were expressed with code rate and a special function called *entropy*. We show asymptotic Gilbert-Varshamov bound in b-symbol metric is also derived in Section6. After introducing the definition and an important property of entropy function, we suggest the asymptotic Gilbert-Varshamov Bound in Hamming metric and 2-symbol metric(symbol-pair metric).

4.4.1 Entropy Function

Entropy function is an important function in information theory. It is interpreted as the quantity of uncertainty.

Here, we only reference simple definitions and a property of entropy function because these are sufficient for proof of asymptotic Gilbert-Varshamov bound. For further properties and meaning of entropy, see [14].

Entropy function is defined as below.

Definition 4.2 (Entropy)**.** *Let p be a real number such that $0 \leq p \leq 1$. Then entropy function $H_(p)$ is defined as follows:*

$$H(p) \triangleq p \log_2 \frac{1}{p} + (1-p) \log_2 \frac{1}{1-p}. \tag{25}$$

Entropy function can be extened to q-ary cases.

Definition 4.3 (q-ary Entropy Function)**.** *Let q be an integer and p be a real number such that $q \geq 2$ and $0 \leq p \leq 1$. Then q-ary entropy function $H_q(p)$ is defined as follows:*

$$H_q(p) \triangleq p \log_q (q-1) + p \log_q \frac{1}{p} + (1-p) \log_q \frac{1}{1-p}. \tag{26}$$

In definition 4.2 and 4.3, we set $(0 \log_q 0) = 0$. Entropy function Definition 4.2 is the special case of q-ary entropy function when $q = 2$.

Theorem 4.3 ([13])**.** *Let q be an integer and p be a real number such that $q \geq 2$ and $0 \leq p \leq 1 - \frac{1}{q}$. Then the following inequality holds:*

$$\sum_{j=0}^{pn} \binom{n}{j} (q-1)^j \leq q^{H_q(p)n}. \tag{27}$$

The proof of Theorem 4.3 is in Appendix A.

4.4.2 Asymptotic Gilbert-Varshamov Bound

For proving asymptotic G-V bound, we reformulate G-V bound into the form related with rate $R(C)$. It follows from (24) that

$$R(C) = \frac{\log_q |C|}{n} \geq \frac{1}{n} \log_q \frac{q^n}{|\mathcal{B}_{d-1}(\vec{x})|}. \tag{28}$$

Asymptotic G-V bound is the form that taking limitation $n \to \infty$ to the left and right terms of (28).

$$R(C_\delta) \geq \lim_{n \to \infty} \frac{1}{n} \log_q \frac{q^n}{|\mathcal{B}_{d-1}(\vec{x})|} \tag{29}$$

C_δ is notation for a code with fractional minimum distance $\delta = \frac{d}{n}$. Therefore Gilbert-Varshamov bound is stated by another way: there exists a code C_δ that satisfies (29). The remaining task is calculation of $|\mathcal{B}_{d-1}(\vec{x})|$ whose value depends on the metric of the space. And it is already proved in terms of Hamming metric [14] and 2-symbol metric(symbol-pair metric) [2].

Theorem 4.4 (Asymptotic Gilbert-Varshamov Bound in Hamming Metric[14])**.** *There exist q-ary codes with fractional minimum Hamming distance $0 \leq \delta (= d_H/n) \leq 1 - \frac{1}{q}$ and rate*

$$R(\delta) \geq 1 - H_q(\delta) \tag{30}$$

From [2], binary asymptotic 2-symbol Gilbert-Varshamov bound is proved as follows.

Theorem 4.5 (Binary Asymptotic 2-Symbol Gilbert-Varshamov Bound[2])**.** *There exist binary codes with fractional minimum 2-symbol distance* $\delta(= d_2/n)$ *and rate*

$$R(\delta) \geq 1 - H(\delta/2) - \delta. \tag{31}$$

We generalize asymptotic 2-symbol Gilbert-Varshamov bound to b-symbol Gilbert-Varshamov bound for any integer $b \geq 2$ in Section 6.

5 Sphere Packing Bound and G-V Bound for b-Symbol Read Channels

From this section to the next section, we show the main results of this report: bounds for b-symbol read channels. In this section, we will prove sphere packing bound and Gilbert-Varshamov bound for b-symbol read channels by calculating the size of the sphere and ball over this channel, called b-symbol sphere and b-symbol ball, with combinatorial methods.

We use a notation $\mathcal{S}_h^{(b)}(\vec{x})$ for the size of b-symbol sphere and $\mathcal{B}_r^{(b)}(\vec{x})$ for b-symbol ball over the metric space (Σ^n, D_b).

5.1 Invariance of the Size of b-Symbol Sphere to $\vec{x}$

We suppose $\vec{x} = \vec{0}$ from this section, as the size of sphere $|\mathcal{S}_h(\vec{x})|$ and ball $|\mathcal{B}_r(\vec{x})|$ do not depend on a symbol vector $\vec{x}$. It is proposed by the following Proposition 5.1.

Proposition 5.1. $|\mathcal{S}_h^{(b)}(\vec{x})| = |\mathcal{S}_h^{(b)}(\vec{0})|$.

Proof. From Definition 4.1, $\mathcal{S}_h^{(b)}(\vec{x})$ and $\mathcal{S}_h^{(b)}(\vec{0})$ are written as follows:

$$\mathcal{S}_h^{(b)}(\vec{x}) \triangleq \{\vec{y} \in \Sigma^n | D_b(\vec{x}, \vec{y}) = h\},$$
$$\mathcal{S}_h^{(b)}(\vec{0}) \triangleq \{\vec{y} \in \Sigma^n | D_b(\vec{0}, \vec{y}) = h\}.$$

It is enough to show that there exists one-to-one correspondence between $\mathcal{S}_h^{(b)}(\vec{x})$ and $\mathcal{S}_h^{(b)}(\vec{0})$.

Think of bijection between alphabet Σ and $\Sigma_q = \{0, 1, 2, ..., q\}$ which maps 0 of Σ to 0 of Σ_q. Then we can regard a word in Σ^n as a word in Σ_q^n and addition and subtraction between words in Σ^n are defined as the operation over Σ_q^n.

Then we show there exists one-to-one correspondence between $\mathcal{S}_h^{(b)}(\vec{x})$ and $\mathcal{S}_h^{(b)}(\vec{0})$.

(i) Define $f : \mathcal{S}_h^{(b)}(\vec{x}) \to \mathcal{S}_h^{(b)}(\vec{0})$ as $f(\vec{y}) = \vec{y} - \vec{x} \in \mathcal{S}_h^{(b)}(\vec{0})$. Then f is injective: for $\vec{y}_1, \vec{y}_2 \in \mathcal{S}_h^{(b)}(\vec{x})$, if $\vec{y}_1 \neq \vec{y}_2$, then $\vec{y}_1 - \vec{x} \neq \vec{y}_2 - \vec{x}$. Therefore $|\mathcal{S}_h^{(b)}(\vec{x})| \leq |\mathcal{S}_h^{(b)}(\vec{0})|$.

(ii) Define $g : \mathcal{S}_h^{(b)}(\vec{0}) \to \mathcal{S}_h^{(b)}(\vec{x})$ as $g(\vec{z}) = \vec{z} + \vec{x} \in \mathcal{S}_h^{(b)}(\vec{x})$. Then g is injective: for $\vec{z}_1, \vec{z}_2 \in \mathcal{S}_h^{(b)}(\vec{x})$, if $\vec{z}_1 \neq \vec{z}_2$, then $\vec{z}_1 + \vec{x} \neq \vec{z}_2 + \vec{x}$. Therefore $|\mathcal{S}_h^{(b)}(\vec{0})| \leq |\mathcal{S}_h^{(b)}(\vec{x})|$.

Thus we get $|\mathcal{S}_h^{(b)}(\vec{0})| = |\mathcal{S}_h^{(b)}(\vec{x})|$. □

From Proposition 5.1, we calculate $|\mathcal{S}_h^{(b)}(\vec{0})|$ instead of $|\mathcal{S}_h^{(b)}(\vec{x})|$ throughout Section 5.

5.2 Calculation of b-Symbol Weight

$|\mathcal{S}_h^{(b)}(\vec{0})|$ is the number of vectors whose b-symbol weight is h. We suggest a convenient way for calculating b-symbol weight and this helps the calculation of the size of the b-symbol ball $|\mathcal{S}_h^{(b)}(\vec{0})|$.

In this subsection, we only think binary words.

5.2.1 Calculation of 2-Symbol Weight

The way of calculating 2-symbol weight is mentioned in [2]. Let us review the result of [2] before we generalize this result.

Theorem 5.1 (Calculation of 2-Symbol Weight)**.** *Let w be the number of coordinates that the symbol is 1 and L be the number of runs of 1s in $\vec{y}$. Then 2-symbol weight of $\vec{y} \in \Sigma^n$ is calculated as $w_2(\vec{y}) = w + L$.*

Proof. 2-symbol weight of $\vec{y}$ is the number that the coordinate of $\pi(\vec{y})$ is $11, 10$, or 01. When $\vec{y}$ is $y_0...y_{n-1}$, $\pi(\vec{y})$ is $[y_0y_1, ..., y_ny_0]$ from the defintion. Therefore, if we note on the former symbols in $\pi(\vec{y})$, the number of coordinates 11 or 10 is the same as the number of 1s in $\vec{y}$.

Then the remaining is the number of coordinates that b-symbols are 01. A b-symbol is 01 if and only if it is a b-symbol in the coordinate before a run of 1s in $\vec{y}$. Therefore the number of 01s in $\pi(\vec{y})$ is the same as the number of runs of 1s. □

Example 5.1. *Consider 2-symbol weight of the word $\vec{y} = 001100111$. We arrange $\vec{y}$ and $\pi(\vec{y})$ by each coordinate. We express the coordinates that $\vec{y}$'s symbol is 1 with underline, and those that $\pi(\vec{y})$'s symbol is 01 with double underline.*

$$\begin{array}{ccccccccccc} \vec{y} & = & 0 & \underline{0} & \underline{1} & \underline{1} & 0 & \underline{0} & \underline{1} & \underline{1} & \underline{1} \\ \pi(\vec{y}) & = & 00 & \underline{\underline{01}} & \underline{11} & \underline{10} & 00 & \underline{\underline{01}} & \underline{11} & \underline{11} & \underline{10} \end{array}$$

From above equations, we can confirm that every run of 1s in $\vec{y}$ generates a blue coordinate 01. As a result, 2-symbol weight of $\vec{y}$ is $w_2(\vec{y}) = 5 + 2$.

5.2.2 Calculation of b-Symbol Weight

In Section 5.2.1, 2-symbol weight is calculated in terms of runs of 1s. However, even if we change the definition of L to "the number of runs of 0s in $\vec{y}$", 2-symbol weight $w_2(\vec{y})$ is the same as $w + L$. Noting on how runs of 0s contribute b-symbol weight, we obtain an equation for b-symbol weight as following Theorem 5.2. Before giving the theorem, we define parameters for a word and then see an example.

Definition 5.1. *For a word $\vec{y}$, parameters w, z_k, L, Z is defined as follows:*

$$\begin{cases} w\text{: Hamming weight of } \vec{y}, \\ z_k\text{: the number of runs of 0s whose lengths are } k, \\ L = \sum_{k=b-1}^{n} z_k, \\ Z = \sum_{k=1}^{b-2} kz_k. \end{cases}$$

We will represent b-symbol weight $w_b(\vec{y})$ in terms of w, L and Z.

Example 5.2. *Consider b-symbol weight $w_b(\vec{y})$ when $b = 3$. Suppose $\vec{y}$ is defined as*

$$\begin{array}{rcccccccccccc} \vec{y} = & 1 & 1 & 1 & \underline{0} & \underline{0} & 1 & 1 & 1 & \underline{0} & \underline{0} & \underline{0} & \underline{0} \\ \pi_3(\vec{y}) = & 111 & 110 & 100 & 001 & 011 & 111 & 110 & 100 & \underline{000} & \underline{000} & 001 & 011 \end{array}$$

The coordinates of nonzero 3-symbols in $\pi_3(\vec{y})$ are the coordinates at which the symbol of $\vec{y}$ is 1, the coordinates of length 2 run of 0s, and last two coodinates of length 4 run of 0s. Therefore $w_3(\vec{y})$ is $w + Z + L(b-1) = 6 + 2 + 1(3-1) = 10$.

From Example 5.2, we can understand how symbol 1s and runs of 0s contribute to b-symbol weight. Then, we can represent b-symbol weight in terms of w, L and Z.

Theorem 5.2 (Calculation of b-Symbol Weight)**.** *For a word $\vec{y} \in \Sigma^n$, b-symbol weight $w_b(\vec{y})$ is $w + Z + (b-1)L$.*

Proof. b-symbol weight of $\vec{y}$ is the number of b-symbols which are not 0^b, called *nonzero*. To count how many b-symbols are nonzero, we seperate $\vec{y} = y_1 y_2 ... y_n$ into three sorts of parts: (i) 1, (ii) a run of 0s whose length is at most $b-2$, (iii) a run of 0s whose length is greater than $b-2$.

Then, we count how many b-symbols are counted as nonzero for a part (i),(ii), and (iii).

(i) Suppose y_i is 1. Then the b-symbol $\pi_b(\vec{y})_i$ is nonzero because in $\pi_b(\vec{y})_i = y_i ... y_{i+b-1}$, $y_i = 1$.

(ii) Let the coordinate this run starts be s and ends be e. Then all b-symbols in this run $\pi_b(\vec{y})_s, ..., \pi_b(\vec{y})_e$ are nonzero because the symbol y_{e+1} is 1 and this symbol is included in all b-symbols of this run $\pi_b(\vec{y})_s, ..., \pi_b(\vec{y})_e$.

(iii) Let the coordinate this run starts be s and that ends be e. Then last $b-1$ b-symbols $\pi_b(\vec{y})_{e-(b-2)}, ..., \pi_b(\vec{y})_e$ are nonzero because the symbol y_{e+1} is 1 and this symbol is included in $\pi_b(\vec{y})_{e-(b-2)}, ..., \pi_b(\vec{y})_e$. The other b-symbols $\pi_b(\vec{y})_s, ..., \pi_b(\vec{y})_{e-(b-1)}$ is 0^b. Therefore $b-1$ is added for each run of this sort when calculating b-symbol weight.

From above, it is confirmed that all coordinates are nonzero for a part of (i) and (ii) and $b-1$ coordinates are not for a part of (iii). The number of coordinates of parts in (i) is w and in (ii) is $Z = \sum_{k=1}^{b-2} k z_k$, and the number of parts in (iii) is $L = \sum_{k=b-1}^{n} z_k$. Therefore b-symbol weight $w_b(\vec{y})$ is $w + Z + (b-1)L$. □

b-symbol weight $w_b(\vec{y})$ is also derived as $w + Z' + (b-1)L'$ with the notations $Z' = \sum_{k=1}^{b-1} k z_k$ and $L' = \sum_{k=b}^{n} z_k$ considering that runs of 0s with length $b-1$ are counted same regardless of whether they are treated as parts in (ii) or (iii) in proof of Theorem 5.2.

5.3 The Size of b-Symbol Sphere

Next, for the derivation of the size of b-symbol sphere, we partition b-symbol sphere $\mathcal{S}_h^{(b)}(\vec{0})$ into subsets $\mathcal{S}_h^{(b)}(L, \vec{z})$ with two parameters L and $\vec{z}$.

Definition 5.2 ($\mathcal{S}_h^{(b)}(L, \vec{z})$)**.** *$\mathcal{S}_h^{(b)}(L, \vec{z})$ is a set of words that satisfy the parameters $h, L, \vec{z}$ where h is b-symbol weight, $L = \sum_{k=b-1}^{n} z_k$ and $\vec{z} = (z_1, ..., z_{b-2})$.*

$Z = \sum_{k=1}^{b-2} k z_k$ of vectors in $\mathcal{S}_h^{(b)}(L, \vec{z})$ is determined by $\vec{z}$. And from Theorem 5.2, Hamming weight w of vectors in $\mathcal{S}_h^{(b)}(L, \vec{z})$ is determined as $h - Z - (b-1)L$.

Next we make it explicit the meaning of $\mathcal{S}_h^{(b)}(L, \vec{z})$ with an example.

Example 5.3. *Consider following* $\vec{y} \in \Sigma^{19}$ *when* $b = 4$.

$$\vec{y} = 1\underline{0}111\underline{\underline{000}}1\underline{0}11\underline{00}1\underline{\underline{0000}}$$
$$L = 2, w = 8, \vec{z} = \{z_1, z_2\} = \{2, 1\},$$
$$h = w + Z + (b-1)L = 18.$$

Therefore, $\vec{y}$ *is an element of* $\mathcal{S}_{18}^{(4)}(2, \{2, 1\})$.

$\mathcal{S}_h^{(b)}(\vec{0})$ is partitioned by $\mathcal{S}_h^{(b)}(L, \vec{z})$ as follows: for $1 \leq h < n$,

$$\mathcal{S}_h^{(b)}(\vec{0}) = \bigcup_{(L,\vec{z})\in\mathcal{K}} \mathcal{S}_h^{(b)}(L, \vec{z}),$$

where $\mathcal{K} = \{(L, \vec{z}) = (L, z_1, ..., z_{b-2}) \in \mathbb{N}^{b-1} | L \geq 1, z_i \geq 0, w = h - Z - (b-1)L \geq \sum_{k=1}^{n} z_k\}$. In the set $\mathcal{K}$, the condition of w is derived considering the number of 1s in a vector should be at least the number of runs of 0s. By the condition of w, parameters $z_1, ..., z_{b-2}$ and L are bounded above.

The size of b-symbol sphere is written as follows: for a word $\vec{x} \in \Sigma^n$ and $1 \leq h < n$, the size of b-symbol sphere is

$$|\mathcal{S}_h^{(b)}(\vec{x})| = |\mathcal{S}_h^{(b)}(\vec{0})| = \sum_{(L,\vec{z})\in\mathcal{K}} |\mathcal{S}_h^{(b)}(L, \vec{z})|.$$

Lemma 5.1. $|\mathcal{S}_h^{(b)}(L, \vec{z})|$ *is derived as follows:*
for $1 \leq h < n$,

$$|\mathcal{S}_h^{(b)}(L, \vec{z})| = \frac{n}{w} \binom{w}{L, z_1, \ldots, z_{b-2}} \binom{n-h+L-1}{L-1} (q-1)^w,$$

where $w = h - Z - (b-1)L$. *Here, multinomial coefficient is defined as follows:*

$$\binom{n}{k_1, ..., k_{m-1}} \triangleq \frac{n!}{k_1! \ldots k_{m-1}! (n - \sum_{i=1}^{m-1} k_i)!},$$

where $k_1, \ldots, k_m$ *are nonnegative integers and* n *is an integer at least* $\sum_{i=1}^{m-1} k_i$.

Proof. Define a subset Y of $\mathcal{S}_h^{(b)}(L, \vec{z})$,

$$Y = \{\vec{y} = (y_0 \ldots y_{n-1}) \in \Sigma^n | y_0 \neq 0, \vec{y} \text{ satisfies } h, L, \vec{z}\}.$$

Then, $\mathcal{S}_h^{(b)}(L, \vec{z})$ is rewritten as

$$\mathcal{S}_h^{(b)}(L, \vec{z}) = \{\sigma_i(\vec{y}) | \vec{y} \in Y, 0 \leq i < n\},$$

where $\sigma_i(\vec{x})$ is i-times cyclic right shift of $\vec{x}$.

We seperate the proof by two parts. $|Y|$ is calculated in [Part I] and $|\mathcal{S}_h^{(b)}(L, \vec{z})|$ is in [Part II]. Each part explains the multiplicands $(q-1)^w \binom{w}{L, z_1, \ldots, z_{b-2}} \binom{n-h+L-1}{L-1}$ and $\frac{n}{w}$, respectively.

Suppose $q_i (i = 1, \ldots, w)$ is a nonzero symbol in Σ. Noting on Hamming weight of $\vec{y}$ is w, every vector $\vec{y}$ in the set Y can be written as the following form

$$\vec{y} = q_1 0^{i_1} q_2 0^{i_2} \ldots q_w 0^{i_w}.$$

Here, 0^i is a run of 0s whose length is $i(\geq 0)$ and 0^0 is thought that there exists no symbol.

Part I: The Size of Y.

In [Part I], we calculate the number of ways to determine i_ks and q_ks$(k = 0 \ldots w)$ of

$$\vec{y} = q_1 0^{i_1} q_2 0^{i_2} \ldots q_w 0^{i_w}$$

to satisfy the given variables $L, z_1, \ldots, z_{b-2}$.

For convenience, we extend the definition of z_i, or the number of 0^i, not only for $i \geq 1$ but for $i \geq 0$. Here, z_0 means the number of i_ks that no symbol is inserted between two neighbored q_is. Then, z_0 is determined as $z_0 = w - L - \sum_{i=1}^{b-2} z_i$.

(i) The number of ways to determine $q_1, \ldots, q_w$ is $(q-1)^w$.

(ii) Each i_k takes a value in $\{0, \ldots, n\}$. Define i'_k which takes a value in $\{0, \ldots, b-2, l\}$:

$$i'_k \triangleq \begin{cases} i_k & \text{if } i_k \leq b-2 \\ l & \text{if } i_k > b-2. \end{cases} \tag{32}$$

The process of determining i'_k is the same as distributing w members of i'_k into b distinct groups, of sizes $z_0, z_1, \ldots, z_{b-2}, L$ respectively. The number of ways of this distribution is $\binom{w}{L, z_1, \ldots, z_{b-2}}$.

(iii) The values of i_k where $i'_k = i_k \leq b-2$ are determined in (i). The remianing task is determining the L values of i_k greater than $b-2$ so that code length of $\vec{y} = q_1 0^{i_1} q_2 0^{i_2} \ldots q_w 0^{i_w}$ becomes n.

Of n symbols of $\vec{y}$, the number of nonzero symbols is w and the number of 0s in the runs with length at most $b-2$ is $Z = \sum_{k=1}^{b-2} k z_k$. Therefore $n - (w + Z)$ symbols construct L values of i_k greater than $b-2$. Therefore $\sum_{i_k > b-2} i_k = n - (w + Z)$.

From combinatorics, the number of ways to distribute m identical objects into k distinct groups given that each group contains at least t objects is $\binom{m-(p-1)t-1}{p-1}$. Because each i_k greater than $b-2$ has at least $b-1$ symbols, the number of ways for allocating $n - (w + Z)$ to i_ks greater than $b-2$ is

$$\binom{n - (w+Z) - (b-2)L - 1}{L-1} = \binom{n - h + L - 1}{L-1}.$$

Finally, we get the size of Y:

$$|Y| = (q-1)^w \binom{w}{L, z_1, \ldots, z_{b-2}} \binom{n - h + L - 1}{L-1}.$$

Part II: The Size of $\mathcal{S}_h^{(b)}(L, \vec{z})$

In [Part II], we (i) construct a multiset S_m of cyclically shifted vectors from Y and (ii) calculate the size of $\mathcal{S}_h^{(b)}(L, \vec{z})$ by relieving multiplicity of multiset S_m.

(i) First define multiset S_m:

$$S_m = \{\sigma_i(\vec{y}) | \vec{y} \in Y, 0 \le i < n\}.$$

Multiplicity of elements is allowed in S_m whereas it is prohibited in $\mathcal{S}_h^{(b)}(L, \vec{z})$. The size of S_m is $n|Y|$ because each element of Y generates n shifted vectors.

(ii) Every vector $\vec{y}$ in Y can be written as a repetition of a short vector. Let $\vec{y}_s$ be the shortest vector generating $\vec{y}$ by r times repetition($r \ge 1$). Then all $\vec{y}$ in Y can be written uniquely with $\vec{y}_s$ as below:

$$\vec{y} = (\vec{y}_s)^r = (q_1 0^{i_1} \cdots q_{w_r} 0^{i_{w_r}})^r.$$

In n cyclically shifted vectors of $\vec{y}$, same vector appears r times.

On the other hand, each of following $w_r (= \frac{w}{r})$ vectors, $\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_{w_r} \in Y$, generates same vectors by n cyclic shifts:

$$\begin{aligned}
\vec{y}_1 &= (q_1 0^{i_1} q_2 0^{i_2} \cdots q_{w_r} 0^{i_{w_r}})^r \\
\vec{y}_2 &= (q_2 0^{i_2} q_3 0^{i_3} \cdots q_1 0^{i_1})^r \\
&\vdots \\
\vec{y}_{w_r} &= (q_{w_r} 0^{i_{w_r}} q_1 0^{i_1} \cdots q_{w_r-1} 0^{i_{w_r-1}})^r.
\end{aligned}$$

Therefore every distinct element of S_m appears $w (= r \cdot w_r)$ times. Finally,

$$|\mathcal{S}_h^{(b)}(L, \vec{z})| = \frac{1}{w}|S_m| = \frac{n}{w}|Y|.$$

By results of [Part I] and [Part II], $|\mathcal{S}_h^{(b)}(L, \vec{z})|$ is derived as

$$|\mathcal{S}_h^{(b)}(L, \vec{z})| = \frac{n}{w} \binom{w}{L, z_1, \ldots, z_{b-2}} \binom{n-h+L-1}{L-1} (q-1)^w.$$

□

When $b = 1$, Lemma 5.1 corresponds to the size of Hamming sphere and when $b = 2$, Lemma 5.1 corresponds to the result of [2].

5.4 Sphere Packing Bound and G-V Bound for b-Symbol Read Channels

The size of the b-symbol sphere directly leads to sphere packing bound and Gilbert-Varshamov bound for b-symbol read channels as discussed in Section 4.

The size of the b-symbol ball $\mathcal{B}_r^{(b)}(\vec{x})$ should be calculated before.

Theorem 5.3 (the Size of the b-Symbol Ball)**.** *Let (Σ^n, D_b) be a metric space. Then for a symbol vector $\vec{x} \in \Sigma^n$ and $1 \le r < n$,*

$$\mathcal{B}_r^{(b)}(\vec{x}) = \bigcup_{h=0}^{r} \mathcal{S}_h^{(b)}(\vec{x}), \tag{33}$$

$$|\mathcal{B}_r^{(b)}(\vec{x})| = 1 + \sum_{h=1}^{r} |\mathcal{S}_h^{(b)}(\vec{x})|. \tag{34}$$

Table 1: b-symbol weight distribution($2 \leq b \leq 8$) with code length $n = 20$

h	$\lvert\mathcal{S}_h^{(2)}(\vec{x})\rvert$	$\lvert\mathcal{S}_h^{(3)}(\vec{x})\rvert$	$\lvert\mathcal{S}_h^{(4)}(\vec{x})\rvert$	$\lvert\mathcal{S}_h^{(5)}(\vec{x})\rvert$	$\lvert\mathcal{S}_h^{(6)}(\vec{x})\rvert$	$\lvert\mathcal{S}_h^{(7)}(\vec{x})\rvert$	$\lvert\mathcal{S}_h^{(8)}(\vec{x})\rvert$
1	0	0	0	0	0	0	0
2	20	0	0	0	0	0	0
3	20	20	0	0	0	0	0
4	190	20	20	0	0	0	0
5	340	40	20	20	0	0	0
6	1270	210	40	20	20	0	0
7	2680	380	80	40	20	20	0
8	6585	810	270	80	40	20	20
9	13220	1980	500	160	80	40	20
10	25694	3940	1030	410	160	80	40
11	44820	7780	2080	780	320	160	80
12	72095	14915	4260	1570	710	320	160
13	105100	26820	8180	3120	1380	640	320
14	138250	45930	15420	6120	2750	1330	640
15	161824	74184	28220	11880	5440	2620	1280
16	165490	111915	49965	22560	10680	5210	2590
17	143580	155720	85100	42160	20800	10320	5140
18	100590	195400	138590	77280	40160	20360	10230
19	51680	212180	213420	138600	76800	40000	20320

Theorem 5.4 (b-Symbol Sphere-Packing Bound)**.** *If a q-ary code $C \subset \Sigma^n$ can correct all t or less b-symbol errors, then*

$$|C||\mathcal{B}_t^{(b)}(\vec{x})| \leq q^n. \tag{35}$$

Theorem 5.5 (b-Symbol Gilbert-Vashamov Bound)**.** *There exists a q-ary b-symbol code $C \in \Sigma^n$ with minimum distance d_b such that*

$$|C||\mathcal{B}_{d_b-1}^{(b)}(\vec{x})| \geq q^n. \tag{36}$$

6 Asymptotic b-Symbol Gilbert-Varshamov Bound

6.1 Asymptotic b-Symbol Gilbert-Varshamov Bound

In Section 5, we derived the equation for the size of b-symbol ball. With the size of b-symbol ball, we derive asymptotic b-symbol G-V bound.

Before the proof of asymptotic b-symbol G-V bound, we give multinomial theorem.

Lemma 6.1 (Multinomial Theorem[15])**.**

$$\begin{aligned}&(x_1+\cdots+x_m)^n\\ =&\sum_{(*)}\binom{n}{k_1,...,k_{m-1}}x_1^{k_1}\dots x_{m-1}^{k_{m-1}}x_m^{(n-\sum_{i=1}^m k_i)}\end{aligned}$$

The symbol $(*)$ *means all cases that* $n, k_1, \dots, k_{m-1}$ *satisfy the condition of multinomial coefficients.*

If $x_1 = \cdots = x_m = 1$,

$$(m)^n = \sum_{(*)}\binom{n}{k_1,...,k_{m-1}}. \tag{37}$$

For proving asymptotic G-V bound, we reformulate G-V bound into the form related with rate $R(C)$. It follows from G-V bound that

$$R(C) = \frac{\log_q |C|}{n} \geq \frac{1}{n}\log_q \frac{q^n}{|\mathcal{B}^{(b)}_{d_b-1}(\vec{x})|}. \tag{38}$$

Therefore G-V bound is stated by another way: there exists a b-symbol code C that satisfies (38). Asymptotic G-V bound is obtained by taking limitation $n \to \infty$ to the left and right terms of (38). The remaining task is calculating $|\mathcal{B}_{d_b-1}(\vec{x})|$ and then taking limitation to (38).

Lemma 6.2. *For* $1 \leq r < n$,

$$|\mathcal{B}^{(b)}_r(\vec{x})| < nb^{r+1}q^{nH_q(\frac{r}{bn})+b}.$$

Proof. From Lemma 5.1,

$$\begin{aligned}&|\mathcal{S}^{(b)}_h(\vec{x})|\\ =&\sum_{(L,\vec{z})\in\mathcal{K}}|\mathcal{S}^{(b)}_h(L,\vec{z})|\\ =&\sum_{(L,\vec{z})\in\mathcal{K}}\frac{n}{w}\binom{w}{L,z_1,\dots,z_{b-2}}\binom{n-h+L-1}{L-1}(q-1)^w\end{aligned}$$

where $\mathcal{K} = \{(L,\vec{z}) = (L, z_1, ..., z_{b-2}) \in \mathbb{N}^{b-1} | L \geq 1, z_i \geq 0, w = h - Z - (b-1)L \geq \sum_{k=1}^n z_k\}$.

We get the following inequalities of $|\mathcal{S}_h^{(b)}(\vec{x})|$:

$$
\begin{aligned}
&|\mathcal{S}_h^{(b)}(\vec{x})| \\
=&\sum_{(L,\vec{z})\in\mathcal{K}} \frac{n}{w}\binom{w}{L, z_1, \ldots, z_{b-2}}\binom{n-(h-L+1)}{L-1}(q-1)^w \\
<&\sum_{(L,\vec{z})\in\mathcal{K}} \frac{n}{w}\binom{w}{L, z_1, \ldots, z_{b-2}}\binom{n}{L-1}(q-1)^w \qquad (39)\\
\leq&\sum_{(L,\vec{z})\in\mathcal{K}} n\binom{h}{L, z_1, \ldots, z_{b-2}}\binom{n}{L-1}(q-1)^h \qquad (40)\\
=&n(q-1)^h \sum_{(L,\vec{z})\in\mathcal{K}} \binom{h}{L, z_1, \ldots, z_{b-2}}\binom{n}{L-1} \\
<&n\binom{n}{\lfloor \frac{h}{b} \rfloor}(q-1)^h \sum_{(L,\vec{z})\in\mathcal{K}} \binom{h}{L, z_1, \ldots, z_{b-2}} \qquad (41)\\
<&nb^h\binom{n}{\lfloor \frac{h}{b} \rfloor}(q-1)^h. \qquad (42)
\end{aligned}
$$

(39) is from $h-L+1>0$ (by the inequality $h=w+Z+(b-1)L>L$), (40) is derived from $w\leq h$, and (41) is derived from $L-1\leq \lfloor \frac{h}{b} \rfloor \leq \frac{n}{2}$ which we get from

$$
\begin{aligned}
w = h-Z-(b-1)L \geq L,\\
h \geq h-Z \geq bL,\\
\frac{h}{b} \geq L. \qquad (43)
\end{aligned}
$$

Considering L is an integer, (43) changes to $\lfloor \frac{h}{b} \rfloor \geq L$. (42) is derived by $\mathcal{K}\subset(*)$ and multinomial theorem (37):

$$
\sum_{(L,\vec{z})\in\mathcal{K}} \binom{h}{L, z_1, \ldots, z_{b-2}} < \sum_{(L,\vec{z})\in(*)} \binom{h}{L, z_1, \ldots, z_{b-2}} = b^h.
$$

Then we calculate the size of the b-symbol ball $|\mathcal{B}_r(\vec{x})|$,

$$
\begin{aligned}
|\mathcal{B}_r(\vec{x})| &= 1+\sum_{h=1}^{r}|\mathcal{S}_h(\vec{x})| \\
&< \sum_{h=1}^{r} nb^h\binom{n}{\lfloor \frac{h}{b} \rfloor}(q-1)^h \\
&< nb^r\sum_{h=1}^{r}\binom{n}{\lfloor \frac{h}{b} \rfloor}(q-1)^h.
\end{aligned}
$$

From $\sum_{j=0}^{\lfloor pn \rfloor}\binom{n}{j}(q-1)^j < q^{nH_q(p)}$ for $p\leq 1-\frac{1}{q}$,

$$
\sum_{h=1}^{r}\binom{n}{\lfloor \frac{h}{b} \rfloor}(q-1)^h < b\sum_{j=0}^{\lfloor \frac{r}{b} \rfloor}\binom{n}{j}(q-1)^{j+b} < bq^{nH_q(\frac{r}{bn})+b}.
$$

Therefore

$$|\mathcal{B}_r^{(b)}(\vec{x})| < nb^r \sum_{h=1}^{r} \binom{n}{\lfloor \frac{h}{b} \rfloor} (q-1)^h < nb^{r+1} q^{nH_q(\frac{r}{bn})+b}.$$

□

With Lemma 6.2, we can derive asymptotic b-symbol G-V bound.

Theorem 6.1 (Asymptotic b-Symbol G-V Bound)**.** *For $b \geq 2$ and $0 \leq \delta \leq 1$, there exist q-ary b-symbol codes C_δ with fractional minimum b-symbol distance $\delta(= d_b/n)$ and rate*

$$R(C_\delta) > 1 - H_q(\delta/b) - \delta \log_q b.$$

Proof. To calculate (38), we derive the following inequality:

$$\log_q \frac{q^n}{|\mathcal{B}_{d_b-1}^{(b)}(\vec{x})|} > n - (\log_q n + d_b \log_q b + nH_q\Big(\frac{d_b-1}{bn}\Big) + b).$$

Then, we take limitation to (38) and get asymptotic G-V bound:

$$\begin{aligned} R(C_\delta) &\geq \lim_{n\to\infty} \frac{1}{n} \log_q \frac{q^n}{|\mathcal{B}_{d_b-1}^{(b)}(\vec{x})|} \\ &> 1 - H_q\Big(\frac{\delta}{b}\Big) - \delta \log_q b. \end{aligned}$$

□

When $b = 2$, Theorem 6.1 coincides with asymptotic symbol-pair G-V bound in [2]. If the range of δ is narrowed to $0 \leq \delta \leq 1 - \frac{1}{q}$, we can extend Theorem 6.1 to $b \geq 1$ because asymptotic G-V bound for Hamming metric(Theorem 4.4) coincides with asymptotic 1-symbol G-V bound.

We also note that it is not confirmed whether the asymptotic bound given above is the tight bound or not. If this bound is not the tight form of asymptotic b-symbol G-V bound, there are possiblities to improve this bound. Considering inequalities in the proof, especially (42), are not tight, it is rather probable that asymptotic b-symbol G-V bound proved in this article is not tight. Of course, there are possiblities that this bound is tight. For showing the tightness of this bound, we should prove the RHS inequality of (44) should be equality.

6.2 The Best b for The Asymptotic G-V b-Symbol Bound with respect to Fractional Minimum Distance

In this subsection, we analyze the derived asymptotic b-symbol Gilbert-Varshamov bound. To be specific, we find the symbol-read number b that leads to the best asymptotic b-symbol G-V bound when the size of alphabet q and relative minimum distance δ are fixed.

When we observe the traces of asymptotic b-symbol G-V bounds where $b = 1, 2, 3$ and $q = 2$(Figure 6), it is confirmed the bound for $b = 2$ achieves the highest bound between the two intersections of bounds, $0.19 < \delta < 0.27$. If this tendency that the mid-value achieves the highest bound continues for arbitrary three consecutive bs, the bound for $b = 2$ achieves the highest bound in the range $0.19 < \delta < 0.27$ compared to any other bs, because (2-symbol bound) $>$ (3-symbol bound) $>$ (4-symbol bound) $> \cdots$ in the range. We will show this tendency is true for asymptotic b-symbol G-V bounds for bs larger than a certain number s and explore related properties of asymptotic b-symbol G-V bounds.

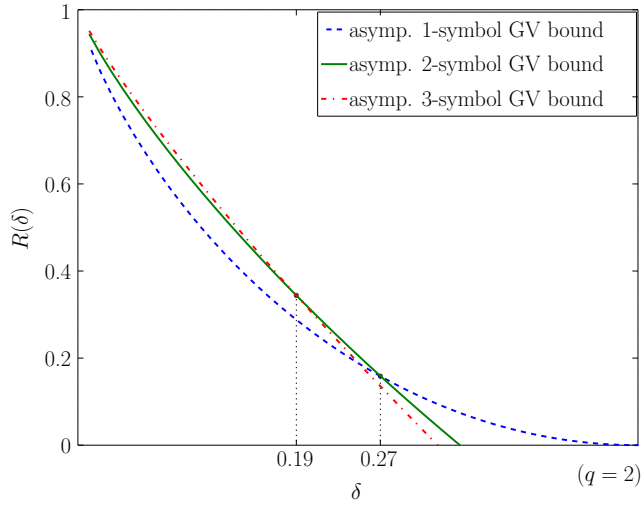


Figure 6: Asymptotic G-V bounds when $b = 1$(dashed),2(solid),and 3(dash-dotted) ($q = 2$).

Let RHS of asymptotic Gilbert-Varshamov bound be a function

$$f_b(\delta) = 1 - H_q(\delta/b) - \delta \log_q b. \tag{44}$$

Theorem 6.2. *For $b \geq 3$ and $0 < \delta \leq 1$, the equation $f_b(\delta) = f_{b-1}(\delta)$ has at most one solution, a unique solution $\delta = \delta_b$ or no solution.*

Proof. Define a function $g_b(\delta)$ on $0 \leq \delta \leq 1$,

$$g_b(\delta) = f_b(\delta) - f_{b-1}(\delta). \tag{45}$$

Then the equation $f_b(\delta) = f_{b-1}(\delta)$ is equivalent to $g_b(\delta) = 0$. We prove the following three equation and inequalities. Here, $\lim_{\delta \to 0^+}$ means right side limit of 0.

$$g_b(0) = 0 \tag{46}$$

$$\lim_{\delta \to 0^+} g_b'(\delta) > 0 \tag{47}$$

$$g_b''(\delta) < 0. \tag{48}$$

The trace $(\delta, g_b(\delta))$ starts at $(0,0)$ from (46) and the gradient at $\delta = 0$ is positive from (47). Considering $g_b(\delta)$ is concave function from (48), the trace of $g_b(\delta)$ is shown like Figure 7. Therefore these directly lead to $g_b(\delta) = 0$ has at most one solution on $0 < \delta \leq 1$.

(i) We get the following from the definition of $f_b(\delta)$,

$$g_b(\delta) = H_q\Big(\frac{\delta}{b-1}\Big) - H_q\Big(\frac{\delta}{b}\Big) + \delta \log_q \frac{b-1}{b}. \tag{49}$$

(ii) From defintion of entropy (26),

$$H_q(x) = x \log_q (q-1) - x \log_q x - (1-x) \log_q (1-x) \tag{50}$$
$$H_q'(x) = \log_q (q-1) - \log_q x + \log_q (1-x) \tag{51}$$
$$H_q''(x) = \frac{1}{x(x-1)}. \tag{52}$$

(iii) Considering $0 \log 0 = 0$ from Definition 4.3 of entorpy function, it is trivial that $g_b(0) = 0$.

(iv) Next, we show $\lim_{\delta \to 0^+} g_b'(\delta) > 0$. $g_b'(\delta)$ is calculated from $g_b(\delta)$:

$$g_b'(\delta) = \frac{1}{b-1} H_q'\Big(\frac{\delta}{b-1}\Big) - \frac{1}{b} H_q'\Big(\frac{\delta}{b}\Big) + \delta \log_q \frac{b-1}{b} \tag{53}$$

In $H_q'(x)$ (51), a non-convergent term when $x \to 0$ is only $(-\log_q x)$. Therefore when $\delta \to 0^+$, non-convergent terms of $g_b'(\delta)$ are

$$-\frac{1}{b-1} \log_q \frac{\delta}{b-1} + \frac{1}{b} \log_q \frac{\delta}{b} \tag{54}$$
$$= \frac{1}{b(b-1)} \log_q \Big(\frac{b-1}{\delta}\Big)^b \Big(\frac{\delta}{b}\Big)^{b-1}. \tag{55}$$

By taking $\delta \to 0^+$ to (55),

$$\lim_{\delta \to 0^+} \frac{1}{b(b-1)} \log_q \Big(\frac{b-1}{\delta}\Big)^b \Big(\frac{\delta}{b}\Big)^{b-1} = \infty. \tag{56}$$

Therefore we get

$$\lim_{\delta \to 0^+} g_b'(\delta) = \infty > 0. \tag{57}$$

(v) Then we show (48).

$$\begin{aligned} g_b''(x) &= \frac{1}{(b-1)^2} H_q'\Big(\frac{\delta}{b-1}\Big) - \frac{1}{b^2} H_q'\Big(\frac{\delta}{b}\Big) \\ &= \frac{1}{\delta(\delta-(b-1))} - \frac{1}{\delta(\delta-b)} \\ &= \frac{1}{\delta(\delta-(b-1))(\delta-b)} \\ &< 0. \end{aligned} \tag{58}$$

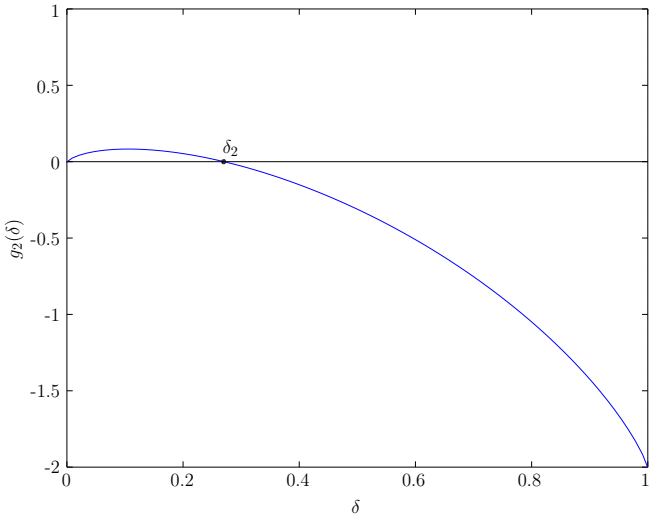


Figure 7: The trace of $g_2(\delta)$ $(q = 2)$. $g_b(\delta)$s have the similar shape in $0 < \delta \leq 1$.

□

For the equation $g_b(\delta) = f_b(\delta) - f_{b+1}(\delta) = 0$, if a solution on $0 < \delta \leq 1$ exists, we denote the unique solution as δ_b. From the trace of the graph of $g_b(\delta)$(Figure 7), we can understand the following properties of $f_b(\delta)$.

(i) if a unique solution $\delta = \delta_b$ exists,

$$f_b(\delta) > f_{b-1}(\delta), \text{ for } 0 < \delta < \delta_b, \tag{59}$$

$$f_b(\delta) \leq f_{b-1}(\delta), \text{ for } \delta_b \leq \delta \leq 1. \tag{60}$$

(ii) if there is no solution, for $0 < \delta \leq 1$,

$$f_b(\delta) > f_{b-1}(\delta). \tag{61}$$

Therefore we understand that for $b \geq 3$, b-symbol codes achieve higher rate codes compared to $(b-1)$-symbol codes in the range (i) $0 < \delta \leq \delta_b$ or (ii) $0 < \delta \leq 1$.

δ_b for $3 \leq b \leq 7$ and $2 \leq q \leq 8$ is shown in Table 6.2

Theorem 6.3. *For $b \geq 3$, when $\delta = \delta_b$ is the unique solution of $f_b(\delta) = f_{b-1}(\delta)$ in the range $0 < \delta \leq 1 - \frac{1}{q}$, then the unique solution $\delta = \delta_{b+1}$ of $f_{b+1}(\delta) = f_b(\delta)$ exists in the range $0 < \delta \leq 1 - \frac{1}{q}$ and*

$$\delta_{b+1} < \delta_b. \tag{62}$$

Table 2: δ_b for $3 \le b \le 7$ and $2 \le q \le 8$.

q	δ_3	δ_4	δ_5	δ_6	δ_7
2	1.9×10^1	1.1×10^1	5.1×10^2	2.3×10^2	1.0×10^2
3	3.6×10^1	2.0×10^1	1.0×10^1	4.6×10^2	2.0×10^2
4	5.0×10^1	3.0×10^1	1.5×10^1	6.8×10^2	3.0×10^2
5	6.3×10^1	3.9×10^1	2.0×10^1	9.1×10^2	4.0×10^2
6	7.3×10^1	4.7×10^1	2.4×10^1	1.1×10^1	5.0×10^2
7	8.3×10^1	5.5×10^1	2.9×10^1	1.3×10^1	5.9×10^2
8	9.2×10^1	6.2×10^1	3.3×10^1	1.6×10^1	6.9×10^2

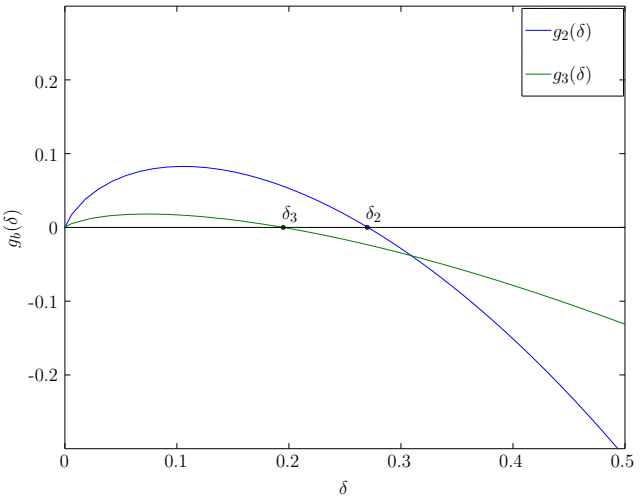


Figure 8: $g_2(\delta)$(solid), $g_3(\delta)$(dashed), and δ_2, δ_3 ($q = 2$). we can confirm $g_3(\delta_2) < 0$ and $\delta_3 < \delta_2$.

Proof. We prove that for $0 < \delta_b \le 1 - \frac{1}{q}$ such that $g_b(\delta_b) = f_b(\delta_b) - f_{b-1}(\delta_b) = 0$,

$$g_{b+1}(\delta_b) = f_{b+1}(\delta_b) - f_b(\delta_b) < 0. \tag{63}$$

From Theorem 6.2 and the shape of graphs $g_b(\delta)$ and $g_{b+1}(\delta)$ shown in Figure 8, (63) leads to the result that δ_{b+1} exists and $\delta_{b+1} < \delta_b$.

$g_{b+1}(\delta_b)$ is written with a function $h_\delta(x) \triangleq H_q\Big(\frac{\delta}{x}\Big)$,

$$g_{b+1}(\delta_b) = H_q\Big(\frac{\delta_b}{b}\Big) - H_q\Big(\frac{\delta_b}{b+1}\Big) + \delta_b \log_q \frac{b}{b+1} \tag{64}$$

$$= h_{\delta_b}(b) - h_{\delta_b}(b+1) + \delta_b \log_q \frac{b}{b+1}. \tag{65}$$

As $0 < \delta_b \le 1 - \frac{1}{q}$, $h_{\delta_b}(x)$ is convex function from Lemma B.1 and thus

$$h_{\delta_b}(b) - h_{\delta_b}(b+1) < h_{\delta_b}(b-1) - h_{\delta_b}(b). \tag{66}$$

Therefore

$$(65) < h_{\delta_b}(b-1) - h_{\delta_b}(b) + \delta_b \log_q \frac{b}{b+1} \tag{67}$$

$$= -\delta_b \log_q \frac{b-1}{b} + \delta_b \log_q \frac{b}{b+1} \tag{68}$$

$$= \delta_b \log_q \frac{b^2}{b^2-1} \tag{69}$$

$$< 0. \tag{70}$$

(68) is derived from $g_b(\delta_b) = h_{\delta_b}(b-1) - h_{\delta_b}(b) + \delta_b \log_q \frac{b-1}{b} = 0$.

Therefore, we get $g_{b+1}(\delta_b) < 0$ and therefore $\delta_b > \delta_{b+1}$. □

Corollary 6.1. *For $b \geq 3$, if the unique solution $\delta = \delta_b$ of $f_b(\delta) = f_{b-1}(\delta)$ on $0 < \delta \leq 1 - \frac{1}{q}$ exists, for all $b' \geq b$, the unique solution $\delta = \delta_{b'}$ of $f_{b'}(\delta) = f_{b'-1}(\delta)$ on $0 < \delta \leq 1 - \frac{1}{q}$ exists. And the following inequalities are satisfied.*

$$\delta_b > \delta_{b+1} > \delta_{b+2} > \delta_{b+3} > \cdots . \tag{71}$$

Corollary 6.2. *Let $s \geq 3$ be the smallest integer such that the solution $\delta = \delta_s$ of $f_s(\delta) = f_{s-1}(\delta)$ on $0 < \delta \leq 1 - \frac{1}{q}$ exists. Then compared to all b-symbol codes with $b \geq 2$,*

(i) $(s-1)$-symbol codes achieve the highest asymptotic G-V bound on $\delta_s < \delta \leq 1 - \frac{1}{q}$.

(ii) for $t \leq s$, t-symbol codes achieve the highest asymptotic G-V bound on $\delta_{t+1} < \delta \leq \delta_t$.

(iii) on the range $1 - \frac{1}{q} < \delta \leq 1$, integers b that achieve the highest asymptotic b-symbol G-V bounds are less than s.

Theorem 6.4. *Let $s \geq 3$ be the smallest integer such that the solution $\delta = \delta_s$ of $f_s(\delta) = f_{s-1}(\delta)$ on $0 \leq \delta \leq 1 - \frac{1}{q}$ exists. Then for a sequence $\{\delta_s, \delta_{s+1}, \delta_{s+2}, \cdots\}$,*

$$\lim_{b \to \infty} \delta_b = 0. \tag{72}$$

Proof. From Corollary 6.1, it is enough to prove the following proposition: for sufficiently small $\epsilon > 0$, there exists B such that $\delta_B < \epsilon$. Instead of $\delta_B < \epsilon$, we show $g_B(\epsilon) < 0$. Then, from the shape of the graph(Figure 8), $\delta_B < \epsilon$.

$H_q(x)$ and $g_b(\delta)$ is written as follows:

$$H_q(x) = x \log_q (q-1) - \log_q (1-x) + x \log_q \Big(\frac{1-x}{x}\Big) \tag{73}$$

$$\begin{aligned} g_b(\delta) =& f_b(\delta) - f_{b-1}(\delta) \\ =& H_q\Big(\frac{\delta}{b-1}\Big) - H_q\Big(\frac{\delta}{b}\Big) + \delta \log_q \frac{b-1}{b} \\ =& \frac{\delta}{b(b-1)} \log_q (q-1) + \log_q \left(\frac{1-\frac{\delta}{b}}{1-\frac{\delta}{b-1}}\right) \\ &+ \log_q \left\{ \left(\frac{\frac{\delta}{b}}{1-\frac{\delta}{b}}\right)^{\frac{\delta}{b}} \left(\frac{1-\frac{\delta}{b-1}}{\frac{\delta}{b-1}}\right)^{\frac{\delta}{b-1}} \Big(\frac{b-1}{b}\Big)^{\delta} \right\} \end{aligned} \tag{74}$$

If we take B such that $\frac{1}{\epsilon} \leq B < \frac{1}{\epsilon} + 1$. Then $\frac{1}{B}$ is written with ϵ and Δ:

$$\frac{1}{B} \leq \epsilon < \frac{1}{B-1} \leq \frac{\epsilon}{1-\epsilon} \tag{75}$$

$$\frac{1}{B} = \epsilon - \Delta. \tag{76}$$

Here, Δ follows inequalities below:

$$\begin{aligned} 0 < \Delta &= \epsilon - \frac{1}{B} \\ &< \frac{1}{B-1} - \frac{1}{B} = \frac{1}{B(B-1)} \\ &< \frac{1}{(B-1)^2} \\ &\leq \left(\frac{\epsilon}{1-\epsilon}\right)^2 \\ &\ll \epsilon. \end{aligned}$$

Then $\frac{1}{B}$ and $\frac{1}{B-1}$ are also approximated:

$$\frac{1}{B} = \epsilon - \Delta \simeq \epsilon, \tag{77}$$

$$\frac{1}{B-1} = \frac{\frac{1}{B}}{1-\frac{1}{B}} \simeq \frac{\epsilon}{1-\epsilon}. \tag{78}$$

Then, $g_b(\delta)$ is approximated:

$$g_b(\epsilon) \simeq \frac{\epsilon^3}{1-\epsilon} \log_q (q-1) + \log_q \left(\frac{1-\epsilon^2}{1-\frac{\epsilon^2}{1-\epsilon}}\right) + \log_q \left\{ \left(\frac{\epsilon^2}{1-\epsilon^2}\right)^{\epsilon^2} \left(\frac{1-\frac{\epsilon^2}{1-\epsilon}}{\frac{\epsilon^2}{1-\epsilon}}\right)^{\frac{\epsilon^2}{1-\epsilon}} (1-\epsilon)^{\epsilon} \right\} \tag{79}$$

$$\simeq \log_q (1-\epsilon)^{\epsilon} < 0. \tag{80}$$

From (79) to (80), we take $\epsilon^3 = 0$ and $\epsilon^2 = 0$ from $\epsilon^3 \ll \epsilon^2 \ll \epsilon$ and use following inequalities:

$$\begin{aligned} &\left(\frac{\epsilon^2}{1-\epsilon^2}\right)^{\epsilon^2} \left(\frac{1-\frac{\epsilon^2}{1-\epsilon}}{\frac{\epsilon^2}{1-\epsilon}}\right)^{\frac{\epsilon^2}{1-\epsilon}} \\ &= \left(\frac{\epsilon^2}{1-\epsilon^2}\right)^{\frac{\epsilon^2}{1-\epsilon}(1-\epsilon)} \left(\frac{1-\frac{\epsilon^2}{1-\epsilon}}{\frac{\epsilon^2}{1-\epsilon}}\right)^{\frac{\epsilon^2}{1-\epsilon}} \\ &= \left(\frac{\epsilon^2}{1-\epsilon^2}\right)^{\frac{-\epsilon^3}{1-\epsilon}} \left(\frac{\epsilon^2}{1-\epsilon^2} \cdot \frac{1-\frac{\epsilon^2}{1-\epsilon}}{\frac{\epsilon^2}{1-\epsilon}}\right)^{\frac{\epsilon^2}{1-\epsilon}} \\ &\simeq 1. \end{aligned}$$

Therefore, for an integer B satisfying $\frac{1}{\epsilon} \leq B < \frac{1}{\epsilon} + 1$, the solution $\delta = \delta_B$ of $f_B(\delta) = f_{B-1}(\delta)$ exists. Finally, from Corollary 6.1, we get that for $b \geq B$, the solution δ_b exists. □

Corollary 6.3. *For all $\delta \in (0, 1]$, there exists a finite integer k such that k-symbol codes achieve the highest asymptotic G-V bound compared to all other b-symbol codes with $b \geq 2$.*

Corollary 6.3 implies that larger b for b-symbol codes does not nessesarily bring out better asymptotic b-symbol G-V bound. In other words, at least in the view of asymptotic b-symbol G-V bound, the most suitable number of read-symbol b is determined dependently on relative minimum distance δ.

On the other hand, we can also interprete these analysis that there exist b-symbol codes with strictly higher rates in $\delta \in (0, \delta_b)$ than best known codes with symbol-read numbers less than b. Therefore, the performance of b-symbol codes are partially improved as b becomes larger in terms of asymptotic G-V bound.

7 Conclusions

In this report we derived sphere packing bound, Gilbert-Varshamov bound, and asymptotic Gilbert-Varshamov bound for b-symbol read channels. We calculated the size of b-symbol sphere and b-symbol ball and these sizes led to the bounds for b-symbol read channels. When $b = 1$, the bounds obtained here coincide with bounds for Hamming metric and when $b = 2$, with symbol-pair codes. We also showed that for the asymptotic G-V bound proved in this report, when relative minimum distance δ and the size of alphabet q are given, there exists a finite symbol-read number b such that asymptotic b-symbol G-V bounds achieve the best rate than other symbol-read numbers. It means that higher b may not guarantee higher rate code for a given relative minimum distance.

There are lots of remaing tasks related to this research.

Most of all, asymptotic b-symbol G-V bounds proved in this article should be improved or the tightness of these bounds should be proved. With the enhanced bounds, the possibilities and limitations of b-symbol read channel will be more clear. For the improvement of asymptotic b-symbol G-V bound, reference [3], which proved the improved asymptotic G-V bound for $b = 2$, might be helpful reference. The analysis of error probability for b-symbol codes is also a task for a comprehensive study. Though a perfect code, a code satisfying sphere packing upper bound, for 2-symbol read code was suggested in [2], it is also unknown whether perfect codes for b-symbol read channel exists or not, Furthermore, it is an open question how to construct a code achieving b-symbol Gilbert-Varshamov bound.

Acknowledgement

First and foremost, I want to thank my advisor, Prof. Toru Fujiwara. Most importantly, he supervised this research with his broad understanding of information theory. He introduced me b-symbol read channel and related research achievements and gave sincere advices from the selection of research theme to the representation of results. Also, for most of difficulties that I have encountered in this research, he suggested me possible answers and related reference materials.

Besides from this research, he helped me understand elementary knowledges related to the study of information theory. Every question I had was solved by his comments and especially my significant interest to information theory was brought out from entirely by his course 'Information and Coding Theory'. Without these motivations, I would not make the results in this report.

I am also grateful for Prof. Yasunori Ishihara, Prof. Naoto Yanai and other members of Fujiwara laboratory, Osaka university. They set aside time for my research presentation and sincerely helped to revise the details of my results.

References

[1] C. E. Shannon, "A mathematical theory of communication," Bell System Technical Journal, vol.25, pp.379-423, 623-656, July, Octoctober 1948.

[2] Y. Cassuto and M. Blaum, "Codes for symbol-pair read channels," IEEE Transactions on Information Theory, vol.57, no.12, pp.8011-8020, December 2011.

[3] Y. Cassuto and S. Litsyn, "Symbol-pair codes: Algebraic constructions and asymptotic bounds," IEEE International Symposium on Information Theory Proceedings, St. Petersburg, Russia, pp.2348-2352, Jul. 2011.

[4] E. Yaakobi, J. Bruck, and P. H. Siegel, "Constructions and decoding of cyclic codes over b-symbol read channels," IEEE Transactions on Information Theory, vol.62, no.4, pp.1541-1551, April 2016.

[5] E. Yaakobi, J. Bruck, and P. H. Siegel, "Decoding of cyclic codes over symbol-pair read channels," IEEE International Symposium on Information Theory Proceedings, Cambridge, MA, USA, pp.2891-2895, Jul. 2012.

[6] Y. M. Chee, H. M. Kiah, and C. Wang, "Maximum distance separable symbol-pair codes," IEEE International Symposium on Information Theory Proceedings, Cambridge, MA, USA, pp.2886-2890, Jul. 2012.

[7] Y. M. Chee, L. Ji, H. M. Kiah, C. Wang, and J. Yin, "Maximum distance separable codes for symbol-pair read channels," IEEE Transactions on Information Theory, vol.59, no.11, pp.7259-7267, November 2013.

[8] X. Kai, S. Zhu, and P. Li, "A construction of new MDS symbol-pair codes," IEEE Transactions on Information Theory, vol.61, no.11, pp.5828-5834, November 2015.

[9] M. Takita, M. Hiromoto, and M. Morii, "A decoding algorithm for cyclic codes over symbol-pair read channels," IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, vol.E98-A, no.12, pp.2415-2422, December 2015.

[10] M. Takita, M. Hiromoto, and M. Morii, "Syndrome decoding of symbol-pair codes," IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, vol.E98-A, no.12, pp.2423-2428, December 2015.

[11] M. Takita, M. Hiromoto, and M. Morii, "Algebraic decoding of BCH codes over symbol-pair read channels: Cases of Two-Pair and Three-Pair Error Correction," IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, vol.E99-A, no.12, pp.2179-2191, December 2016.

[12] S. Horii, T. Matsushima, and S. Hirasawa, "Linear programming decoding of binary linear codes for symbol-pair read channel," IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, vol.E99-A, no.12, pp.2170-2178, December. 2016.

[13] W. W. Peterson and E. J. Weldon Jr., 2nd ed., Error Correcting Codes, Cambridge, MIT Press, USA, 1972.

[14] F. J. MacWilliams and N. J. A. Sloane, The Theory of Error Correcting Codes, North Holland, The Netherlands, 1977.

[15] R. L. Graham, D. E. Knuth, and O. Patashnik, Concrete Mathematics: A Foundation for Computer Science, 2nd ed., Addison-Wesley Publishing Company, USA, 1994.

Appendix

A The Proof of Theorem 4.3

Theorem 4.3 is written as follows.

Let q be an integer and p be a real number such that $q \geq 2$ and $0 \leq p \leq 1 - \frac{1}{q}$. Then the following inequality holds.

$$\sum_{j=0}^{pn} \binom{n}{j} (q-1)^j \leq q^{H_q(p)n}. \tag{81}$$

Proof. We can derive the following inequalities:

$$\begin{aligned}
1 &= \{p + (1-p)\}^n \\
&= \sum_{j=0}^{n} p^j (1-p)^{n-j} \\
&= \sum_{j=0}^{pn} p^j (1-p)^{n-j} + \sum_{j=pn+1}^{n} p^j (1-p)^{n-j} \\
&\geq \sum_{j=0}^{pn} p^j (1-p)^{n-j} \\
&= (1-p)^n \sum_{j=0}^{pn} (q-1)^j \Big(\frac{p}{(1-p)(q-1)} \Big)^j \\
&\geq \Big(\frac{p}{(1-p)(q-1)} \Big)^{pn} (1-p)^n \sum_{j=0}^{pn} (q-1)^j \qquad (82) \\
&= q^{-H(p)n} \sum_{j=0}^{pn} (q-1)^j. \qquad (83)
\end{aligned}$$

Inequality (82) is obtained from $\frac{p}{(1-p)(q-1)} < 1$ which is derived from the condition $0 \leq p \leq 1 - \frac{1}{q}$.

Then from (83),

$$\sum_{j=0}^{pn} \binom{n}{j} (q-1)^j \leq q^{H_q(p)n}. \tag{84}$$

□

B Convex Function

A function $f(x)$ definend on an interval $[a,b]$ is called *convex function* if the following inequality is satisfied for all $x_1, x_2 \in [a,b]$ and $t \in (0,1]$,

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2). \tag{85}$$

Theorem B.1. *$f(x)$ is twice-diffrentiable and convex on the interval $[a,b] \Leftrightarrow f''(x) > 0$ for $a \leq x \leq b$.*

Lemma B.1. *Define a function $h_\delta(x)$ as follows: For $0 \leq \delta \leq 1 - \frac{1}{q}$,*

$$h_\delta(x) \triangleq H_q\Big(\frac{\delta}{x}\Big). \tag{86}$$

Then $h_\delta(x)$ is convex function on $x \geq 2$.

Proof. From Theorem B.1, we show $h''_\delta(x) > 0$ on $x \geq 2$.

First, we get $h'_\delta(x)$ and $h''_\delta(x)$ simply by differentiating $h_\delta(x)$:

$$\begin{aligned} h_\delta(x) &= H_q\Big(\frac{\delta}{x}\Big) \\ &= \frac{\delta}{x}\log_q(q-1) - \frac{\delta}{x}\log_q\frac{\delta}{x} - \Big(1-\frac{\delta}{x}\Big)\log_q\Big(1-\frac{\delta}{x}\Big) \qquad (87) \\ h'_\delta(x) &= \frac{\delta}{x^2\log q}\left\{\log\frac{\delta}{(q-1)(x-\delta)}\right\} \qquad (88) \\ h''_\delta(x) &= -\frac{\delta}{x^3(x-\delta)}\left\{\log_q\Big(\frac{\delta}{(q-1)(x-\delta)}\Big)^{2(x-\delta)}e^x\right\} \qquad (89) \end{aligned}$$

From $0 < \delta \leq 1 - \frac{1}{q}$ and $x \geq 2$,

$$\frac{\delta}{x-\delta} < \frac{1-\frac{1}{q}}{2-\Big(1-\frac{1}{q}\Big)} \leq \frac{q-1}{q+1}, \tag{90}$$

$$\frac{\delta}{(q-1)(x-\delta)} < \frac{q-1}{(q-1)(q+1)} = \frac{1}{q+1} \leq \frac{1}{3}. \tag{91}$$

And we get the inequialities:

$$\begin{aligned} \Big(\frac{\delta}{(q-1)(x-\delta)}\Big)^{2(x-\delta)} &< \Big(\frac{1}{3}\Big)^{2(x-\delta)} \\ &< \Big(\frac{1}{3}\Big)^{2x} \\ &< \Big(\frac{1}{9}\Big)^{x}, \qquad (92) \\ \Big(\frac{\delta}{(q-1)(x-\delta)}\Big)^{2(x-\delta)}e^x &< \Big(\frac{e}{9}\Big)^{x} < 1. \qquad (93) \end{aligned}$$

By applying (93) to (89),

$$h''_\delta(x) = -\frac{\delta}{x^3(x-\delta)}\left\{\log_q\Big(\frac{\delta}{(q-1)(x-\delta)}\Big)^{2(x-\delta)}e^x\right\} > 0. \tag{94}$$

Therefore, $h_\delta(x)$ is convex function on $x \geq 2$. □