

Bounds on Code Parameters

Hanwen Yao

January 16, 2025

1 Singleton bound

Theorem 1. For a length- n code $C \subseteq \{0, 1, \dots, q-1\}^n$ with size q^k and distance d , we have

$$k \leq n - d + 1$$

Proof. If we list all q^k codewords of C in a table, by the pigeonhole principle, there are two codewords $\underline{c}_1, \underline{c}_2$ that agree on the first $k - 1$ locations. Thus

$$d \leq d(\underline{c}_1, \underline{c}_2) \leq n - (k - 1)$$

□

Example 1. Among binary codes, repetition codes and single parity-check codes achieve the singleton bound.

The repetition code of length n has parameters $k = 1$ and distance $d = n$, so we have

$$n = n - 1 + 1 \quad \Rightarrow \quad d = n - k + 1$$

The single parity-check code of length n has parameters $k = n - 1$ and $d = 2$, so

$$2 = n - (n - 1) + 1 \quad \Rightarrow \quad d = n - k + 1$$

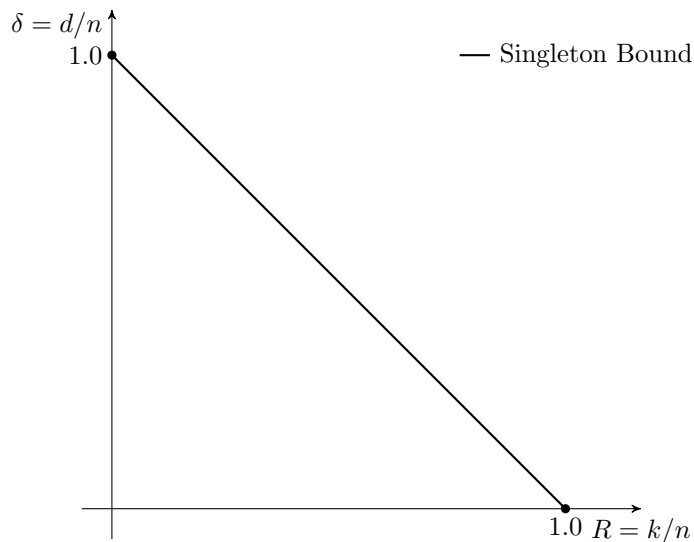
In fact, it can be shown that

Exercise 1. Repetition codes and single parity-check codes are the only non-trivial binary linear codes that achieve the singleton bound. (Hint: study the systematic generator matrix)

The family of codes that meet the Singleton bound is called maximum distance separable (MDS) codes. We will later see an algebraic family of codes called Reed-Solomon codes which achieve the Singleton bound and have dimension $n - d + 1$ and distance d , but its alphabet size grows with the block length.

Consider a family of codes with $\frac{k}{n} \rightarrow R$ and $\frac{d}{n} \rightarrow \delta$, then the Singleton bound tells us that asymptotically

$$R \lesssim 1 - \delta$$



2 Hamming bound

In discrete math, a packing bound constrains the maximum number of non-overlapping objects. The Hamming bound is the natural packing bound for Hamming balls in Hamming metric space.

Theorem 2. For a binary code C of block length n and distance d , let $t = \lfloor \frac{d-1}{2} \rfloor$, then

$$|C| \leq \frac{2^n}{\sum_{i=0}^t \binom{n}{i}}$$

Proof. Define the Hamming ball of radius t around a codeword \underline{c} as

$$B(\underline{c}, t) = \{\underline{w} \in \{0, 1\}^n : d(\underline{w}, \underline{c}) \leq t\}$$

It contains all the points in $\{0, 1\}^n$ within Hamming distance t from the codeword \underline{c} . The volume of this Hamming ball can be computed as

$$|B(\underline{c}, t)| = \sum_{i=0}^t \binom{n}{i}$$

If a code C has distance d , since $d \geq 2t + 1$, we know the Hamming balls of radius t around all the codewords must be disjoint. Therefore, the sum of their volumes has to be smaller or equal to the volume of the entire space $\{0, 1\}^n$:

$$|C| \sum_{i=0}^t \binom{n}{i} \leq 2^n$$

□

The codes that achieve Hamming bound are called perfect codes.

Example 2. $[2^m - 1, 2^m - m - 1, 3]$ binary Hamming codes are perfect.

Hamming code can be defined via its parity-check matrix of size $m \times (2^m - 1)$ that consists of all binary columns of length m except the all-zero column. Hamming codes have distance $d = 3$, so the Hamming ball of radius $t = 1$ around its codewords each have volume

$$|B(\underline{c}, 1)| = 1 + n = 2^m$$

Then since

$$|C| = 2^k = 2^{2^m - m - 1} = \frac{2^{2^m - 1}}{2^m} = \frac{2^n}{1 + n},$$

we know all those disjoint Hamming balls cover the whole space, so Hamming codes achieve the Hamming bound.

It has been shown that the following list contains all the binary perfect codes:

- Trivial codes consisting of just one codeword, or the whole space, or the repetition code
- Hamming code
- The $[23, 12, 7]$ Golay code

For $p \in [0, 1]$, define the binary entropy function

$$H_2(p) = p \log_2 \frac{1}{p} + (1 - p) \log_2 \frac{1}{1 - p}$$

Then the volume of a Hamming ball of radius t in $\{0, 1\}^n$ can be bounded¹ as

$$\frac{1}{\sqrt{8n\varepsilon(q - \varepsilon)}} \cdot 2^{H_2(\varepsilon)n} \leq \sum_{i=0}^t \binom{n}{i} \leq 2^{H_2(\varepsilon)n}$$

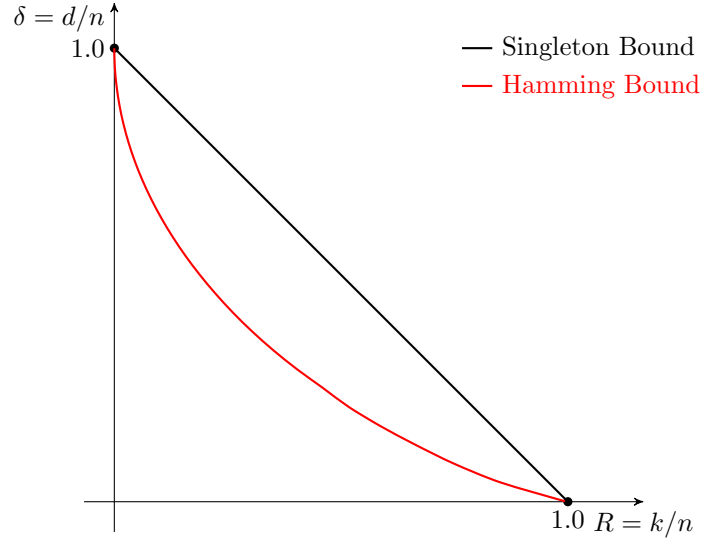
for all integers $n > t \geq 1$ with $\varepsilon = t/n \leq 1/2$. Then for $n \rightarrow \infty$, we have

$$\sum_{i=0}^t \binom{n}{i} \simeq 2^{nH_2(t/n)}$$

Therefore, asymptotically, Hamming bound tells us

$$R = \frac{k}{n} \leq 1 - \frac{1}{n} \log_2 \left(\sum_{i=0}^t \binom{n}{i} \right) \simeq 1 - H_2 \left(\frac{d}{2n} \right)$$

¹see e.g. Ash, Robert B. *Information Theory* (1990, p.121) or Flum, Jörg; Grohe, Martin *Parameterized Complexity Theory* (2006, p.427)



3 Plotkin bound

Lemma 1. For an $[n, k]$ binary linear code C of distance d , we have

$$d \leq \frac{n2^{k-1}}{2^k - 1}$$

Proof. List all codewords in a $2^k \times n$ matrix. We count the number of 1s in this matrix by rows and get

$$(|C| - 1)d \leq \sum_{\substack{\underline{x} \in C \\ \underline{x} \neq 0}} \text{wt}(\underline{x})$$

If we count the number of 1s in this matrix by column, then each column is either all-zero or has half number of 1s by linearity. Therefore

$$\sum_{\substack{\underline{x} \in C \\ \underline{x} \neq 0}} \text{wt}(\underline{x}) \leq n2^{k-1}$$

Hence

$$(|C| - 1)d = (2^k - 1)d \leq n2^{k-1}$$

□

From this lemma, we see that for a family of binary linear codes with $k \rightarrow \infty$, the relative distance δ is bounded above by one-half:

$$\delta = \frac{d}{n} \leq \frac{2^{k-1}}{2^k - 1} \lesssim \frac{1}{2}$$

This is much tighter than the Singleton bound. We now combine Lemma 1 with a shortening argument to derive an asymptotic upper bound for all rates $R > 0$.

Theorem 3. *For a family of binary linear codes with $\lim_{n \rightarrow \infty} \frac{k}{n} = R > 0$, we have*

$$R \leq 1 - \frac{2d}{n} + o(1)$$

Proof. For a code C in this family with large enough n , from the lemma we have

$$n \geq 2d - \frac{d}{2^{k-1}} \geq 2d - 1$$

So $n - 2d + 2$ is a positive integer. Shorten the code C at $(n - 2d + 2)$ positions to C' , then C' has the parameters

$$\begin{aligned} n' &= n - (n - 2d + 2) = 2d - 2 \\ k' &\geq k - (n - 2d + 2) \\ d' &\geq d \end{aligned}$$

Apply lemma 1 on C' and we get

$$d \leq d' \leq \frac{n' \cdot 2^{k'-1}}{2^{k'} - 1} = \frac{(d-1)2^{k'}}{2^{k'} - 1}$$

Therefore $2^{k'} \leq d$, which after we take the \log_2 on both sides gives

$$k - (2n - 2d + 2) \leq k' \leq \log_2 d$$

So

$$R = \frac{k}{n} \leq 1 - \frac{2d}{n} + \frac{\log_2 d}{n} + \frac{2}{n} = 1 - \frac{2d}{n} + o(1)$$

□

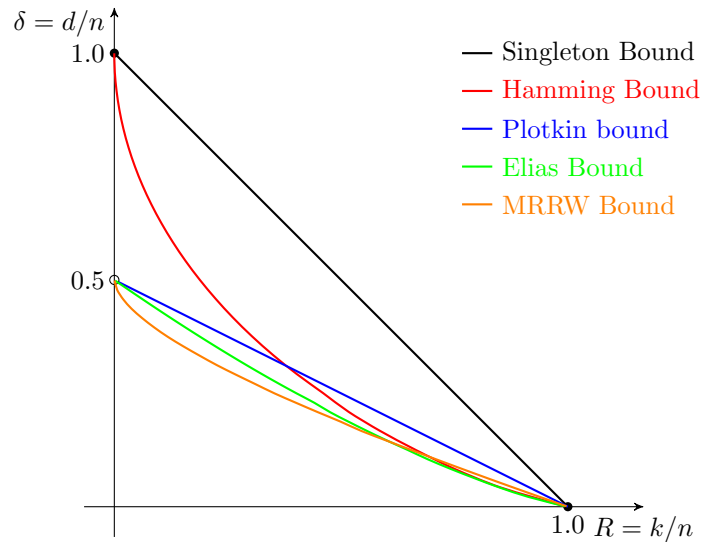
Here, we plot the curves for the Singleton bound, Hamming bound, and Plotkin bound, along with two other tighter upper bounds on the asymptotic achievable rates versus relative distances.

4 Gilbert-Varshamov bound

In discrete math, a covering bound constrains the minimum number of objects required to cover a space. The Gilbert-Varshamov bound is a covering bound for Hamming balls in Hamming metric space.

Theorem 4. *There exists a binary code C of length n and distance d with size*

$$|C| \geq \frac{2^n}{\sum_{i=0}^{d-1} \binom{n}{i}}$$



Proof. Proved by greedy construction of a code in the space $\{0, 1\}^n$ given distance d . Start with any codeword, and keep on adding codewords that have a distance of at least d from all previously chosen codewords, until we can proceed no longer. The process will not terminate as long as the Hamming balls of radius $d - 1$ around every codeword do not cover the whole space, which is true if

$$|C| \sum_{i=0}^{d-1} \binom{n}{i} < 2^n$$

Therefore, this greedy process will produce a code satisfying

$$|C| \geq \frac{2^n}{\sum_{i=0}^{d-1} \binom{n}{i}}$$

□

There also exist linear codes of size given by the Gilbert-Varshamov bound:

Exercise 2. *There exists a binary linear code C of length n , distance d with size*

$$|C| \geq \frac{2^n}{\sum_{i=0}^{d-1} \binom{n}{i}}$$

Hint: greedily construct an $(n - k) \times n$ parity check matrix where every $d - 1$ columns are linearly independent.

Asymptotically, GV bound tells us there exist families of binary codes with

$$R \geq 1 - H_2 \left(\frac{d}{n} \right)$$

