

Foreword

This chapter is based on lecture notes from the course CSE 545– Error-Correcting Codes: Combinatorics, Algorithms and Applications taught by Atri Rudra at University at Buffalo, SUNY.

This version is dated **February 29, 2012**. For the latest version, please go to

<http://www.cse.buffalo.edu/atri/courses/coding-theory/book/>

The material in this chapter is supported in part by the National Science Foundation under CAREER grant CCF-0844796. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF).



©Atri Rudra, 2012.

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. To view a copy of this license, visit

<http://creativecommons.org/licenses/by-nc-nd/3.0/> or send a letter to Creative Commons, 444 Castro Street, Suite 900, Mountain View, California, 94041, USA.

Chapter 4

What Can and Cannot Be Done-I

In this chapter, we will try to tackle Question 2.5.1. We will approach this trade-off in the following way:

If we fix the relative distance of the code to be δ , what is the best rate R that we can achieve?

Note that an upper bound on R is a *negative* result, while a lower bound on R is a *positive* result.

In this chapter, we will consider only one positive result, i.e. a lower bound on R called the Gilbert-Varshamov bound in Section 4.2. In Section 4.1, we recall a negative result that we have already seen– Hamming bound and state its asymptotic version to obtain an upper bound on R . We will consider two other upper bounds: the Singleton bound (Section 4.3, which gives a good upper bound for large enough alphabets (but not binary codes) and the Plotkin bound (Section 4.4).

4.1 Asymptotic Version of the Hamming Bound

We have already seen an upper bound in Section 1.7 due to Hamming. However, we had stated this as an upper bound on the dimension k in terms of n, q and d . We begin by considering the trade-off between R and δ given by the Hamming bound. Recall that Theorem 1.7.2 states the following:

$$\frac{k}{n} \leq 1 - \frac{\log_q \text{Vol}_q \left(\left\lfloor \frac{d-1}{2} \right\rfloor, n \right)}{n}$$

Recall that Proposition 3.3.3 states following lower bound on the volume of a Hamming ball:

$$\text{Vol}_q \left(\left\lfloor \frac{d-1}{2} \right\rfloor, n \right) \geq q^{H_q \left(\frac{\delta}{2} \right) n - o(n)},$$

which implies the following asymptotic version of the Hamming bound:

$$R \leq 1 - H_q \left(\frac{\delta}{2} \right) + o(1).$$

See Figure 4.1 for a pictorial description of the Hamming bound for binary codes.

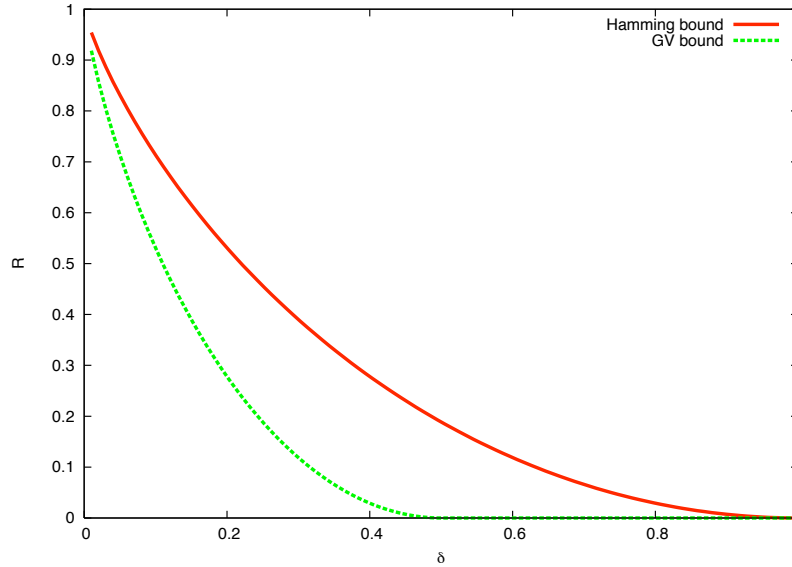


Figure 4.1: The Hamming and GV bounds for binary codes.

4.2 Gilbert-Varshamov Bound

Next, we switch gears we prove our first non-trivial lower bound on R in terms of δ . (In fact this is the only positive result on the R vs δ tradeoff question that we will see in this book.) In particular, we will prove the following result:

Theorem 4.2.1 (Gilbert-Varshamov Bound). *Let $q \geq 2$. For every $0 \leq \delta < 1 - \frac{1}{q}$, and $0 < \varepsilon \leq 1 - H_q(\delta)$, there exists a code with rate $R \geq 1 - H_q(\delta) - \varepsilon$, and relative distance δ .*

The above result was proved for general codes by Edgar Gilbert ([12]) and for linear codes by Rom Varshamov ([46]). Hence, the bound is called the Gilbert-Varshamov bound. The bound is generally referred to as the GV bound. For a pictorial description of the GV bound for binary codes, see Figure 4.1. We will present the two proofs in Sections 4.2.1 and 4.2.2 respectively.

4.2.1 Gilbert Construction

Gilbert proved Theorem 4.2.1 by the following greedy construction (where $d = \delta n$): start with the empty code C and then keep on adding vectors not in C that are at Hamming distance at least d from all the existing codewords in C . Algorithm 5 presents a formal description of the algorithm and Figure 4.2 illustrates the first few executions of this algorithm.

We claim that Algorithm 5 terminates and the C that is output has distance d . The latter is true by step 2, which makes sure that in Step 3 we never add a vector \mathbf{c} that will make the distance of C fall below d . For the former claim, note that, if we cannot add \mathbf{v} at some point, we cannot add it later. Indeed, since we only add vectors to C , if a vector $\mathbf{v} \in [q]^n$ is ruled out in a

Algorithm 5 Gilbert's Greedy Code Construction

INPUT: n, q, d OUTPUT: A code $C \subseteq [q]^n$ of distance d

- 1: $C \leftarrow \emptyset$
 - 2: WHILE there exists a $\mathbf{v} \in [q]^n$ such that $\Delta(\mathbf{v}, \mathbf{c}) \geq d$ for every $\mathbf{c} \in C$ DO
 - 3: Add \mathbf{v} to C
 - 4: RETURN C
-

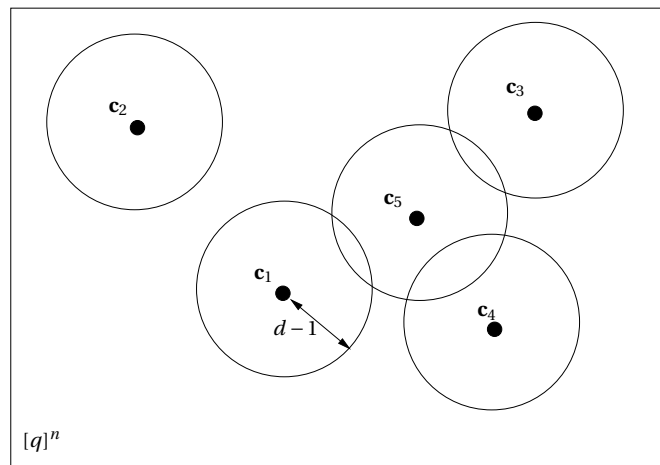


Figure 4.2: An illustration of Gilbert's greedy algorithm (Algorithm 5) for the first five iterations.

certain iteration of Step 2 because $\Delta(\mathbf{c}, \mathbf{v}) < d$, then in all future iterations, we have $\Delta(\mathbf{v}, \mathbf{c}) < d$ and thus, this \mathbf{v} will never be added in Step 3 in any future iteration.

The running time of Algorithm 5 is $q^{O(n)}$. To see this note that Step 2 in the worst-case could be repeated for every vector in $[q]^n$, that is at most q^n times. In a naive implementation, for each iteration, we cycle through all vectors in $[q]^n$ and for each vector $\mathbf{v} \in [q]^n$, iterate through all (at most q^n) vectors $\mathbf{c} \in C$ to check whether $\Delta(\mathbf{c}, \mathbf{v}) < d$. If no such \mathbf{c} exists, then we add \mathbf{v} to C otherwise, we move to the next \mathbf{v} . However, note that we can do slightly better– since we know that once a \mathbf{v} is “rejected” in an iteration, it’ll keep on being rejected in the future iterations, we can fix up an ordering of vectors in $[q]^n$ and for each vector \mathbf{v} in this order, check whether it can be added to C or not. If so, we add \mathbf{v} to C , else we move to the next vector in the order. This algorithm has time complexity $O(nq^{2n})$, which is still $q^{O(n)}$.

Further, we claim that after termination of Algorithm 5

$$\bigcup_{\mathbf{c} \in C} B(\mathbf{c}, d-1) \supseteq [q]^n.$$

This is because if not, then there exists a vector $\mathbf{v} \in [q]^n \setminus C$, such that $\Delta(\mathbf{v}, \mathbf{c}) \geq d$ and hence \mathbf{v} can be added to C . However, this contradicts the fact that Algorithm 5 has terminated. Therefore,

$$\left| \bigcup_{\mathbf{c} \in C} B(\mathbf{c}, d-1) \right| \geq q^n. \quad (4.1)$$

It is not too hard to see that

$$\sum_{\mathbf{c} \in C} |B(\mathbf{c}, d-1)| \geq \left| \bigcup_{\mathbf{c} \in C} B(\mathbf{c}, d-1) \right|,$$

which by (4.1) implies that

$$\sum_{\mathbf{c} \in C} |B(\mathbf{c}, d-1)| \geq q^n$$

or since the volume of a Hamming ball is translation invariant,

$$\sum_{\mathbf{c} \in C} \text{Vol}_q(d-1, n) \geq q^n.$$

Thus, we have

$$\begin{aligned} |C| &\geq \frac{q^n}{\text{Vol}_q(d-1, n)} \\ &\geq \frac{q^n}{q^{nH_q(\delta)}} \\ &= q^{n(1-H_q(\delta))}, \end{aligned} \quad (4.2)$$

as desired. In the above, (4.2) follows from the fact that

$$\begin{aligned} \text{Vol}_q(d-1, n) &\leq \text{Vol}_q(\delta n, n) \\ &\leq q^{nH_q(\delta)}, \end{aligned} \quad (4.3)$$

where the second inequality follows from the upper bound on the volume of a Hamming ball in Proposition 3.3.3.

It is worth noting that the code from Gilbert's construction is not guaranteed to have any special structure. In particular, even storing the code can take exponential space. We have seen in Proposition 2.3.1 that linear codes have a much more succinct representation. Thus, a natural question is:

Question 4.2.1. *Do linear codes achieve the $R \geq 1 - H_q(\delta)$ tradeoff that the Gilbert construction achieves?*

Varshamov showed that indeed this is the case and we look at his construction next.

4.2.2 Varshamov Construction

Now we turn to the result due to Varshamov who showed that a random linear code, with high probability, lies on the GV bound. The Varshamov construction is a use of the probabilistic method (Section 3.2).

By Proposition 2.3.4, we are done if we can show that there exists a $k \times n$ matrix \mathbf{G} of full rank (for $k = (1 - H_q(\delta) - \epsilon)n$) such that

$$\text{For every } \mathbf{m} \in \mathbb{F}_q^k \setminus \{\mathbf{0}\}, wt(\mathbf{mG}) \geq d.$$

We will prove the existence of such a \mathbf{G} by the probabilistic method. Pick a random linear code by picking a random $k \times n$ matrix \mathbf{G} where each of kn entries is chosen uniformly and independently at random from \mathbb{F}_q . Fix $\mathbf{m} \in \mathbb{F}_q^k \setminus \{\mathbf{0}\}$. Recall that by Lemma 3.1.12, for a random \mathbf{G} , \mathbf{mG} is a uniformly random vector from \mathbb{F}_q^n . Thus, we have

$$\begin{aligned} Pr[wt(\mathbf{mG}) < d] &= \frac{Vol_q(d-1, n)}{q^n} \\ &\leq \frac{q^{nH_q(\delta)}}{q^n} \end{aligned} \tag{4.4}$$

where (4.4) follows from (4.3). Thus, by the union bound (Lemma 3.1.5)

$$\begin{aligned} Pr[\exists \mathbf{m}, wt(\mathbf{mG}) < d] &\leq q^k q^{-n(1-H_q(\delta))} \\ &= q^{-\epsilon \cdot n}, \end{aligned}$$

where the equality follows by choosing $k = (1 - H_q(\delta) - \epsilon)n$. Since $q^{-\epsilon n} \ll 1$, by the probabilistic method, there exists a linear code C with relative distance δ . We're almost done except we also need to argue that the code C has dimension at least $k = (1 - H_q(\delta) - \epsilon)n$. To show this we need to show the chosen generator matrix \mathbf{G} has full rank. Note that this is a non-zero probability that a uniformly matrix \mathbf{G} does not have full rank. There are two ways to deal with this. First, we

can show that with high probability a random \mathbf{G} does have full rank, so that $|C| = q^k$. However, the proof above has already shown that, with high probability, the distance is greater than zero, which implies that distinct messages will be mapped to distinct codewords and thus $|C| = q^k$. In other words, C does indeed have dimension k , as desired

Discussion. We now digress a bit to discuss some consequences of the proofs of the GV bound.

We first note that Varshamov's proof shows something stronger than Theorem 4.2.1: *most* linear codes (with appropriate parameters) meet the Gilbert-Varshamov bound.

Varshamov's original proof actually picks a random linear code by picking a random $(n - k) \times n$ parity check matrix. We leave it as exercise to work out the details.

Finally, we note that Theorem 4.2.1 requires $\delta < 1 - \frac{1}{q}$. An inspection of Gilbert and Varshamov's proofs, show that the only reason the proof required that $\delta \leq 1 - \frac{1}{q}$ was because it is needed for the volume bound (recall the bound in Proposition 3.3.3)– $\text{Vol}_q(\delta n, n) \leq q^{H_q(\delta)n}$ – to hold. It is natural to wonder if the above is just an artifact of the proof or, for example,

Question 4.2.2. *Does there exists a code with $R > 0$ and $\delta > 1 - \frac{1}{q}$?*

We will return to this question in Section 4.4.

4.3 Singleton Bound

We will now change gears again and prove an upper bound on R (for fixed δ). We start by proving an upper bound due to Richard C. Singleton, which not surprisingly is called the Singleton bound [40].

Theorem 4.3.1 (Singleton Bound). *For every $(n, k, d)_q$ code,*

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

Proof. Let $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_M$ be the codewords of an $(n, k, d)_q$ code C . Note that we need to show $M \leq q^{n-d+1}$. To this end, we define \mathbf{c}'_i to be the prefix of the codeword \mathbf{c}_i of length $n - d + 1$ for every $i \in [M]$. See Figure 4.3 for a pictorial description.

We now claim that for every $i \neq j$, $\mathbf{c}'_i \neq \mathbf{c}'_j$. For the sake of contradiction, assume that there exists an $i \neq j$ such that $\mathbf{c}'_i = \mathbf{c}'_j$. Note that this implies that \mathbf{c}_i and \mathbf{c}_j agree in all the first $n - d + 1$ positions, which in turn implies that $\Delta(\mathbf{c}_i, \mathbf{c}_j) \leq d - 1$. This contradicts the fact that C has distance d . Thus, M is the number of prefixes of codewords in C of length $n - d + 1$, which implies that $M \leq q^{n-d+1}$ as desired. \square

Note that the asymptotic version of the Singleton bound states that $k/n \leq 1 - d/n + 1/n$. In other words,

$$R \leq 1 - \delta + o(1).$$

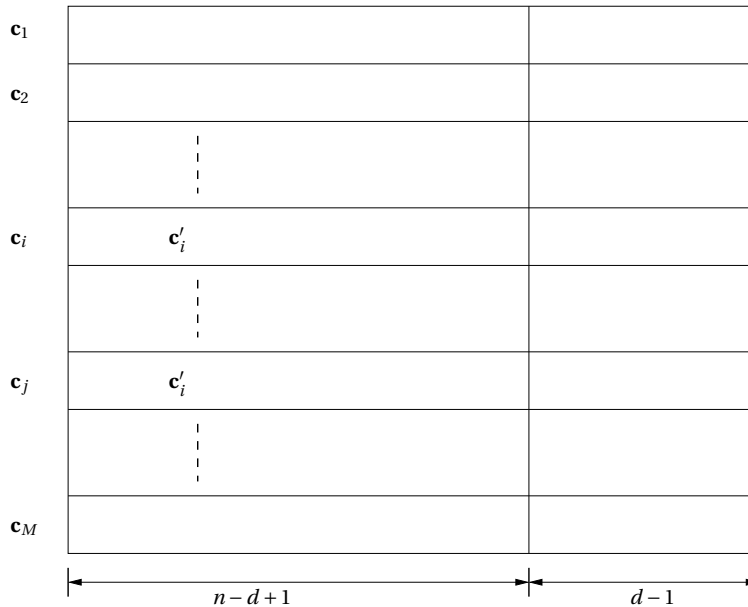


Figure 4.3: Construction of a new code in the proof of the Singleton bound.

Figure 4.4 presents a pictorial description of the asymptotic version of the Singleton bound. It is worth noting that the bound is *independent* of the alphabet size. As is evident from Figure 4.4, the Singleton bound is worse than the Hamming bound for binary codes. However, this bound is better for larger alphabet sizes. In fact, we will look at a family of codes called Reed-Solomon codes in Chapter 5 that meets the Singleton bound. However, the alphabet size of the Reed-Solomon codes increases with the block length n . Thus, a natural follow-up question is the following:

Question 4.3.1. *Given a fixed $q \geq 2$, does there exist a q -ary code that meets the Singleton bound?*

We'll see an answer to this question in the next section.

4.4 Plotkin Bound

In this section, we will study the Plotkin bound, which will answer Questions 4.2.2 and 4.3.1. We start by stating the bound.

Theorem 4.4.1 (Plotkin bound). *The following holds for any $C \subseteq [q]^n$ with distance d :*

1. If $d = (1 - \frac{1}{q})n$, $|C| \leq 2qn$.

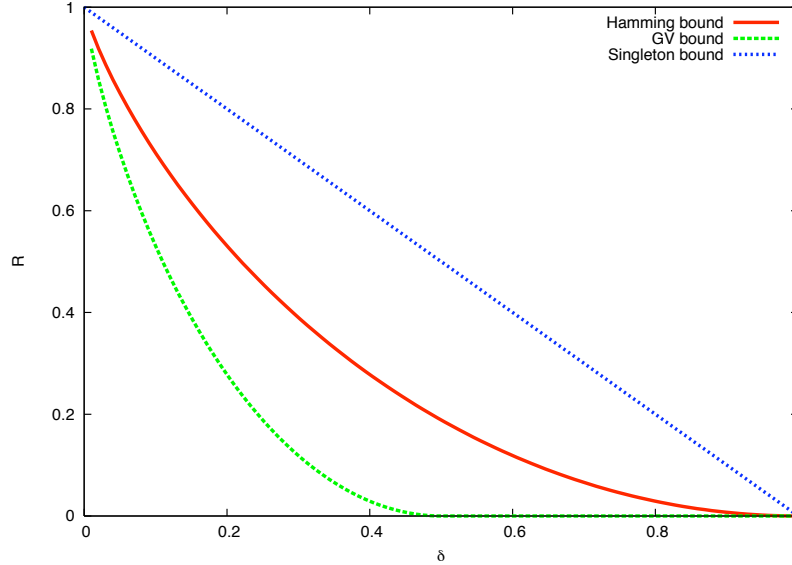


Figure 4.4: The Hamming, GV and Singleton bound for binary codes.

2. If $d > (1 - \frac{1}{q})n$, $|C| \leq \frac{qd}{qd - (q-1)n}$.

Note that the Plotkin bound implies that a code with relative distance $\delta \geq 1 - \frac{1}{q}$, must necessarily have $R = 0$, which answers Question 4.2.2 in the negative.

Before we prove Theorem 4.4.1, we make couple of remarks. We first note that the upper bound in the first part of Theorem 4.4.1 can be improved to $2n$ for $q = 2$, which is tight. (The proof is left as an exercise.) Second, it can be shown that this bound is tight– this again is left as an exercise. (Hint: Consider the code obtained by adding the “complement” of each codeword in the Hadamard code to the Hadamard code.) Third, the statement of Theorem 4.4.1 gives a trade-off only for relative distance greater than $1 - 1/q$. However, as the following corollary shows, the result can be extended to work for $0 \leq \delta \leq 1 - 1/q$. (See Figure 4.5 for an illustration for binary codes.)

Corollary 4.4.2. For any q -ary code with distance δ , $R \leq 1 - (\frac{q}{q-1})\delta + o(1)$.

Proof. The proof proceeds by shortening the codewords. We group the codewords so that they agree on the first $n - n'$ places, where $n' = \lfloor \frac{qd}{q-1} \rfloor - 1$. In particular, for any $\mathbf{x} \in [q]^{n-n'}$, define

$$C_{\mathbf{x}} = \{(c_{n-n'+1}, \dots, c_n) \mid (c_1 \dots c_N) \in C, (c_1 \dots c_{n-n'}) = \mathbf{x}\}.$$

Define $d = \delta n$. For all \mathbf{x} , $C_{\mathbf{x}}$ has distance d as C has distance d .¹ Additionally, it has block length $n' < (\frac{q}{q-1})d$ and thus, $d > (1 - \frac{1}{q})n'$. By Theorem 4.4.1, this implies that

$$|C_{\mathbf{x}}| \leq \frac{qd}{qd - (q-1)n'} \leq qd, \tag{4.5}$$

¹If for some \mathbf{x} , $\mathbf{c}_1 \neq \mathbf{c}_2 \in C_{\mathbf{x}}$, $\Delta(\mathbf{c}_1, \mathbf{c}_2) < d$, then $\Delta((\mathbf{x}, \mathbf{c}_1), (\mathbf{x}, \mathbf{c}_2)) < d$, which implies that the distance of C is less than d (as by definition of $C_{\mathbf{x}}$, both $(\mathbf{x}, \mathbf{c}_1), (\mathbf{x}, \mathbf{c}_2) \in C$).

where the second inequality follows from the facts that $d > (1 - 1/q)n'$ and that $qd - (q - 1)n'$ is an integer.

Note that by the definition of $C_{\mathbf{x}}$:

$$|C| = \sum_{\mathbf{x} \in [q]^{n-n'}} |C_{\mathbf{x}}|,$$

which by (4.5) implies that

$$|C| \leq \sum_{\mathbf{x} \in [q]^{n-n'}} qd = q^{n-n'} \cdot qd \leq q^{n - \frac{q}{q-1}d + o(n)}.$$

In other words, $R \leq 1 - \left(\frac{q}{q-1}\right)\delta + o(1)$ as desired. \square

Note that Corollary 4.4.2 implies that for any q -ary code of rate R and relative distance δ (where q is a *constant* independent of the block length of the code), $R < 1 - \delta$. In other words, this answers Question 4.3.1 in the negative.

Let us pause for a bit at this point and recollect the bounds on R versus δ that we have proved till now. Figure 4.5 depicts all the bounds we have seen till now (for $q = 2$). The GV bound is the best known lower bound till date. For larger (but still constant) values of q , better upper bounds than the GV bound are known. In particular, for any prime power $q \geq 49$, there exist linear codes, called *algebraic geometric* (or AG) codes that outperform the corresponding GV bound. AG codes out of the scope of this book. One starting point could be the following [26]. Better upper bounds are known and we will see one such trade-off (called the Elias-Bassalygo bound) in Section 7.7.

We now turn to the proof of Theorem 4.4.1, for which we will need two more lemmas.

The first lemma deals with vectors over real spaces. We quickly recap the necessary definitions. Consider a vector \mathbf{v} in \mathbb{R}^n , that is, a tuple of n real numbers. This vector has (Euclidean) norm $\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}$, and is a unit vector if and only if its norm is 1. The inner product of two vectors, \mathbf{u} and \mathbf{v} , is $\langle \mathbf{u}, \mathbf{v} \rangle = \sum_i u_i \cdot v_i$. The following lemma gives a bound on the number of vectors that can exist such that every pair is at an obtuse angle with each other.

Lemma 4.4.3 (Geometric Lemma). *Let $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m \in \mathbb{R}^n$ be non-zero vectors.*

1. *If $\langle \mathbf{v}_i, \mathbf{v}_j \rangle \leq 0$ for all $i \neq j$, then $m \leq 2n$*
2. *Let \mathbf{v}_i be unit vectors for $1 \leq i \leq m$. Further, if $\langle \mathbf{v}_i, \mathbf{v}_j \rangle \leq -\epsilon < 0$ for all $i \neq j$, then $m \leq 1 + \frac{1}{\epsilon}$*

The proof of the Plotkin bound will need the existence of a map from codewords to real vectors with certain properties, which the next lemma guarantees.

Lemma 4.4.4 (Mapping Lemma). *Let $C \subseteq [q]^n$. Then there exists a function $f : C \rightarrow \mathbb{R}^{nq}$ such that*

1. *For every $\mathbf{c} \in C$, $\|f(\mathbf{c})\| = 1$.*

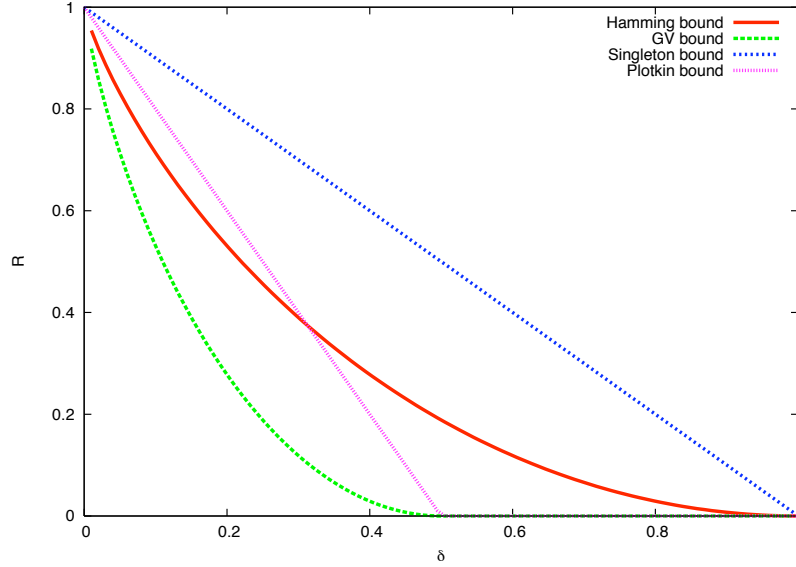


Figure 4.5: The current bounds on the rate R vs. relative distance δ for binary codes. The GV bound is a lower bound on rate while the other three bounds are upper bounds on R .

$$2. \text{ For every } \mathbf{c}_1 \neq \mathbf{c}_2 \in C, \langle f(\mathbf{c}_1), f(\mathbf{c}_2) \rangle = 1 - \left(\frac{q}{q-1}\right) \left(\frac{\Delta(\mathbf{c}_1, \mathbf{c}_2)}{n}\right).$$

We defer the proofs of the lemmas above to the end of the section. We are now in a position to prove Theorem 4.4.1.

Proof of Theorem 4.4.1 Let $\{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_m\} = C$. For all $i \neq j$,

$$\langle f(\mathbf{c}_i), f(\mathbf{c}_j) \rangle \leq 1 - \left(\frac{q}{q-1}\right) \frac{\Delta(\mathbf{c}_i, \mathbf{c}_j)}{n} \leq 1 - \left(\frac{q}{q-1}\right) \frac{d}{n}.$$

The first inequality holds by Lemma 4.4.4, and the second holds as C has distance d .

For part 1, if $d = (1 - \frac{1}{q})n = \frac{(q-1)n}{q}$, then for all $i \neq j$, $\langle f(\mathbf{c}_i), f(\mathbf{c}_j) \rangle \leq 0$ and so by the first part of Lemma 4.4.3, $m \leq 2nq$, as desired.

For part 2, $d > \left(\frac{q-1}{q}\right)n$ and so for all $i \neq j$, $\langle f(\mathbf{c}_i), f(\mathbf{c}_j) \rangle \leq 1 - \left(\frac{q}{q-1}\right) \frac{d}{n} = -\left(\frac{qd - (q-1)n}{(q-1)n}\right)$ and, since $\varepsilon \stackrel{\text{def}}{=} \left(\frac{qd - (q-1)n}{(q-1)n}\right) > 0$, we can apply the second part of Lemma 4.4.3. Thus, $m \leq 1 + \frac{(q-1)n}{qd - (q-1)n} = \frac{qd}{qd - (q-1)n}$, as desired \square

4.4.1 Proof of Geometric and Mapping Lemmas

Next, we prove Lemma 4.4.3.

Proof of Lemma 4.4.3. We begin with a proof of the first result. The proof is by induction on n . Note that in the base case of $n = 0$, we have $m = 0$, which satisfies the claimed inequality $m \leq 2n$.

In the general case, we have $m \geq 1$ non-zero vectors $\mathbf{v}_1, \dots, \mathbf{v}_m \in \mathbb{R}^n$ such that for every $i \neq j$,

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle \leq 0. \quad (4.6)$$

Since rotation, translation and scaling all the vectors by the same amount does not change the sign of the inner products, w.l.o.g. we can assume that $\mathbf{v}_m = \langle 1, 0, \dots, 0 \rangle$. (The formal proof of this claim is left as an exercise.) For $1 \leq i \leq m-1$, denote the vectors as $\mathbf{v}_i = \langle \alpha_i, \mathbf{y}_i \rangle$, for some $\alpha_i \in \mathbb{R}$ and $\mathbf{y}_i \in \mathbb{R}^{n-1}$. Now, for any $i \neq 1$, $\langle \mathbf{v}_1, \mathbf{v}_i \rangle = 1 \cdot \alpha_i + \sum_{j=2}^m 0 = \alpha_i$. However, note that (4.6) implies that $\langle \mathbf{v}_1, \mathbf{v}_i \rangle \leq 0$, which in turn implies that

$$\alpha_i \leq 0. \quad (4.7)$$

Next, we claim that at most one of $\mathbf{y}_1, \dots, \mathbf{y}_{m-1}$ is can be the all zeroes vector, $\mathbf{0}$. If not, assume w.l.o.g., that $\mathbf{y}_1 = \mathbf{y}_2 = \mathbf{0}$. This in turn implies that

$$\begin{aligned} \langle \mathbf{v}_1, \mathbf{v}_2 \rangle &= \alpha_1 \cdot \alpha_2 + \langle \mathbf{y}_1, \mathbf{y}_2 \rangle \\ &= \alpha_1 \cdot \alpha_2 + 0 \\ &= \alpha_1 \cdot \alpha_2 \\ &> 0, \end{aligned}$$

where the last inequality follows from the subsequent argument. As $\mathbf{v}_1 = \langle \alpha_1, \mathbf{0} \rangle$ and $\mathbf{v}_2 = \langle \alpha_2, \mathbf{0} \rangle$ are non-zero, this implies that $\alpha_1, \alpha_2 \neq 0$. (4.7) then implies that $\alpha_1, \alpha_2 < 0$. However, $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle > 0$ contradicts (4.6).

Thus, w.l.o.g., assume that $\mathbf{v}_1, \dots, \mathbf{v}_{m-2}$ are all non-zero vectors. Further, note that for every $i \neq j \in [m-2]$, $\langle \mathbf{y}_i, \mathbf{y}_j \rangle = \langle \mathbf{v}_i, \mathbf{v}_j \rangle - \alpha_i \cdot \alpha_j \leq \langle \mathbf{v}_i, \mathbf{v}_j \rangle \leq 0$. Thus, we have reduced problem on m vectors with dimension n to an equivalent problem on $m-2$ vectors with dimension dimension $n-1$. If we continue this process, we can conclude that every loss in dimension of the vector results in twice in loss in the numbers of the vectors in the set. Induction completes the proof.

We now move on to the proof of the second part. Towards that end, define $\mathbf{z} = \mathbf{v}_1 + \dots + \mathbf{v}_m$. Now consider the following sequence of relationships:

$$\|\mathbf{z}\|^2 = \sum_{i=1}^m \|\mathbf{v}_i\|^2 + 2 \sum_{i < j} \langle \mathbf{v}_i, \mathbf{v}_j \rangle \leq m + 2 \cdot \binom{m}{2} \cdot (-\varepsilon) = m(1 - \varepsilon m + \varepsilon).$$

The inequality follows from the facts that each \mathbf{v}_i is a unit vector and the assumption that for every $i \neq j$, $\langle \mathbf{v}_i, \mathbf{v}_j \rangle \leq -\varepsilon$. As $\|\mathbf{z}\|^2 \geq 0$,

$$m(1 - \varepsilon m + \varepsilon) \geq 0.$$

Thus, we have $m \leq 1 + \frac{1}{\varepsilon}$, as desired. □

Finally, we prove Lemma 4.4.4.

Proof of Lemma 4.4.4. We begin by picking a map $\phi : [q] \rightarrow \mathbb{R}^q$ with certain properties. Then we apply ϕ to all the coordinates of a codeword to define the map $f : \mathbb{R}^q \rightarrow \mathbb{R}^{nq}$ that satisfies the claimed properties. We now fill in the details.

We begin by defining the map $\phi : [q] \rightarrow \mathbb{R}^q$. For every $i \in [q]$, define

$$\phi(i) = \left\langle \frac{1}{q}, \frac{1}{q}, \dots, \underbrace{\frac{-(q-1)}{q}}_{i^{\text{th}} \text{ position}}, \dots, \frac{1}{q} \right\rangle,$$

that is all but the i th position in $\phi(i) \in \mathbb{R}^q$ has a value of $1/q$ and the i th position has value $-(q-1)/q$.

Next, we record two properties of ϕ that follow immediately from its definition. For every $i \in [q]$,

$$\phi(i)^2 = \frac{(q-1)}{q^2} + \frac{(q-1)^2}{q^2} = \frac{(q-1)}{q}. \quad (4.8)$$

Also for every $i \neq j \in [q]$,

$$\langle \phi(i), \phi(j) \rangle = \frac{(q-2)}{q^2} - \frac{2(q-1)}{q^2} = -\frac{1}{q}. \quad (4.9)$$

We are now ready to define our final map $f : C \rightarrow \mathbb{R}^{nq}$. For every $\mathbf{c} = (c_1, \dots, c_n) \in C$, define

$$f(\mathbf{c}) = \sqrt{\frac{q}{n(q-1)}} \cdot \langle \phi(c_1), \phi(c_2), \dots, \phi(c_n) \rangle.$$

(The multiplicative factor $\sqrt{\frac{q}{n(q-1)}}$ is to ensure that $f(\mathbf{c})$ for any $\mathbf{c} \in C$ is a unit vector.)

To complete the proof, we will show that f satisfies the claimed properties. We begin with condition 1. Note that

$$\|f(\mathbf{c})\|^2 = \frac{q}{(q-1)n} \cdot \sum_{i=1}^n \|\phi(i)\|^2 = 1,$$

where the first equality follows from the definition of f and the second equality follows from (4.8).

We now turn to the second condition. For notational convenience define $\mathbf{c}_1 = (x_1, \dots, x_n)$ and $\mathbf{c}_2 = (y_1, \dots, y_n)$. Consider the following sequence of relations:

$$\begin{aligned} \langle f(\mathbf{c}_1), f(\mathbf{c}_2) \rangle &= \sum_{\ell=1}^n \langle f(x_\ell), f(y_\ell) \rangle \\ &= \left[\sum_{\ell: x_\ell \neq y_\ell} \langle \phi(x_\ell), \phi(y_\ell) \rangle + \sum_{\ell: x_\ell = y_\ell} \langle \phi(x_\ell), \phi(y_\ell) \rangle \right] \cdot \left(\frac{q}{n(q-1)} \right) \\ &= \left[\sum_{\ell: x_\ell \neq y_\ell} \left(\frac{-1}{q} \right) + \sum_{\ell: x_\ell = y_\ell} \left(\frac{q-1}{q} \right) \right] \cdot \left(\frac{q}{n(q-1)} \right) \end{aligned} \quad (4.10)$$

$$= \left[\Delta(\mathbf{c}_1, \mathbf{c}_2) \left(\frac{-1}{q} \right) + (n - \Delta(\mathbf{c}_1, \mathbf{c}_2)) \left(\frac{q-1}{q} \right) \right] \cdot \left(\frac{q}{n(q-1)} \right) \quad (4.11)$$

$$\begin{aligned}
&= 1 - \Delta(\mathbf{c}_1, \mathbf{c}_2) \left(\frac{q}{n(q-1)} \right) \left[\frac{1}{q} + \frac{q-1}{q} \right] \\
&= 1 - \left(\frac{q}{q-1} \right) \left(\frac{\Delta(\mathbf{c}_1, \mathbf{c}_2)}{n} \right),
\end{aligned}$$

as desired. In the above, (4.10) is obtained using (4.9) and (4.8) while (4.11) follows from the definition of the Hamming distance. \square