CSE 533: Error-Correcting Codes **(Autumn 2006)**

Lecture 5: Bounds on Codes

October 13, 2006

Lecturer: Venkatesan Guruswami *Scribe: Ben Birnbaum*

1 Introduction

We now turn to a different approach to achieve reliable communication over BSC_p. Namely, we focus on answering the question posed in Lecture 2: Are there infinite families of codes (n_i, k_i, d_i) of increasing n_i such that $k_i/n_i \geq R + o(1)$ and $d_i/n_i \geq \delta + o(1)$ for some constants R and δ? Specifically, given a fixed distance δ, what is the highest rate that we can achieve with a code? In this lecture, we will answer the first question affirmatively, and we will provide upper and lower bounds on the rate achievable by a code for a fixed δ.

Before we begin to answer these questions, we note the connection between finding codes of high distance and achieving reliable communication over BSC_p in the following theorem.

Theorem 1.1. *If there exists an infinite family of codes $C_i = (n_i, k_i, d_i)$ with rate R and relative distance $\delta = 2p + \varepsilon$ for some constants $p > 0$ and $\varepsilon > 0$, then we can use C_i to communicate over BSC_p with an exponentially decreasing error probability in n_i.*

Proof. For any codeword in C_i, the number of errors introduced by the BSC_p channel is a random variable obeying the binomial distribution with parameters n_i and p. By a Chernoff bound, the probability that at least $\delta n_i/2 = (p + \varepsilon/2)n_i$ errors occurs is $2^{-O(\varepsilon^2 n_i)}$. Because the distance of C_i is δn_i, we can only decode a word incorrectly if at least $\delta n_i/2$ errors occur. This implies that the probability of decoding incorrectly is $2^{-O(\varepsilon^2 n_i)}$. □

Remark 1.2. *The converse of this statement can also be shown. Namely, a necessary condition of having an exponentially small error probability on BSC_p is that the code have positive relative distance.*

2 Some Simple Upper Bounds

We have seen that to achieve reliable communication on BSC_p it suffices to have a code family with a relative distance of at least $2p + \varepsilon$ for some constant $\varepsilon > 0$. A natural question to ask, then, is what is the best rate that we can achieve with relative distance $2p + \varepsilon$? By Theorem 1.1 and the converse of Shannon's Theorem proved in Lecture 3, we know that a code with a relative distance of $2p + \varepsilon$ cannot have a rate higher than $1 - H(p + \varepsilon)$. We show next that this result is essentially an asymptotic version of the volume bound from Lecture 2.

Theorem 2.1. *For a binary code family C with rate R and relative distance δ,*

$$R \leq 1 - H(\delta/2) + o(1) \ ,$$

where H is the binary entropy function *defined by $H(x) = x \log_2 \frac{1}{x} + (1-x) \log_2 \frac{1}{1-x}$.*

Proof. Recall the volume bound, which states that $|C| \operatorname{Vol}_2 \left(n, \left\lfloor \frac{d-1}{2} \right\rfloor\right) \leq 2^n$, where $\operatorname{Vol}_2(n,d)$ is the number of points in a ball of Hamming radius d in $\{0,1\}^n$. Since $k = \log_2 |C|$, we can take logarithms to yield

$$k + \log_2 \operatorname{Vol}_2 \left(n, \lfloor (d-1)/2 \rfloor\right) \leq n \ .$$

It was argued earlier that $\operatorname{Vol}_2(n,d) = 2^{H(d/n)n + o(n)}$. Thus, $k + H\left(\delta/2\right) n + o(n) \leq n$, and hence $R = k/n \leq 1 - H\left(\delta/2\right) + o(1)$. □

If we could find a code having rate $1 - H(\delta/2)$ for all positive values of δ, then by Theorem 2.1, we would have an optimal code. However, we will prove later in this lecture that this is not possible for all δ. In other words, the bound provided in Theorem 2.1 is not tight. We prove next an upper bound on the rate called the Singleton bound.

Theorem 2.2. *For any $(n,k,d)_q$ code, $k \leq n - d + 1$.*

Proof. Let Σ be an alphabet of size q, and let C be an (n,k,d) code over Σ. We need to show that $k = \log_q |C| \leq n - d + 1$. Consider the map $f : C \to \Sigma^{n-d+1}$ defined by $f(c_1, c_2, \ldots, c_n) = (c_1, c_2, \ldots, c_{n-d+1})$. Map f is injective because any two codewords $c \neq c'$ must differ in at least d positions and thus $f(c) \neq f(c')$. Hence $|C| \leq |\Sigma^{n-d+1}| = q^{n-d+1}$ and the theorem follows. □

Remark 2.3. *The asymptotic version of this bound states that $R \leq 1 - \delta + o(1)$.*

Remark 2.4. *For $q = 2$, the Singleton bound is subsumed by the volume bound, but when the volume bound is extended for larger q, this is not always the case.*

3 A Lower Bound

At this point, we have proved some upper bounds, so we ask now what kind of lower bounds on rate we can prove. In this section, we prove that there exists a family of codes having both positive rate and positive distance. The bound we show, called the Gilbert-Varshamov bound, is stated precisely in the following theorem.

Theorem 3.1. *For $0 \leq \delta \leq \frac{1}{2}$, there exists a (not necessarily linear) family of binary codes C having rate R and relative distance δ such that $R \geq 1 - H(\delta) + o(1)$.*

To prove this theorem, we will show that it is essentially the asymptotic version of the following lemma for $q = 2$.

Lemma 3.2. *There exist $(n, k, d)_q$ (not necessarily linear) codes with*

$$q^k \geq \frac{q^n}{\mathrm{Vol}_q(n, d-1)} \ . \tag{1}$$

Proof. Consider the code C over Σ (where $|\Sigma| = q$) defined by the following construction:

1. Set $C \leftarrow \emptyset$.
2. While there exists a word $x \in \Sigma^n$ such that $d(x, c) \geq d$ for all $c \in C$, set $C \leftarrow C \cup \{x\}$.

Clearly, code C has distance d. It also has the property that there are no points in Σ^n that are at a distance of more than $d-1$ from the nearest codeword. Therefore, if we create a ball of radius $d-1$ centered at each codeword, then the entire space Σ^n must be covered. Hence $|C| \, \mathrm{Vol}_q(n, d-1) = q^k \, \mathrm{Vol}_q(n, d-1) \geq q^n$. □

With Lemma 3.2 proved, we are now ready to prove Theorem 3.1.

Proof of Theorem 3.1. In (1), set $q = 2$, take the base 2 logarithm of both sides, and use our usual estimate $\mathrm{Vol}_2(n, d) = 2^{H(\delta)n + o(n)}$. □

Remark 3.3. *The reason that δ is restricted to be no greater than $\frac{1}{2}$ in the statement of Theorem 3.1 is that the estimate for $\mathrm{Vol}_2(n, d)$ using the entropy function is only valid for $\delta \leq \frac{1}{2}$.*

If we extend our definition of the entropy function H for alphabets of size greater than 2, then we can generalize the GV bound for $q \geq 2$.

Definition 3.4. *The* q-ary entropy function H_q *is defined by*

$$H_q(x) = x \log_q(q-1) + \frac{H(x)}{\log_2 q} \ .$$

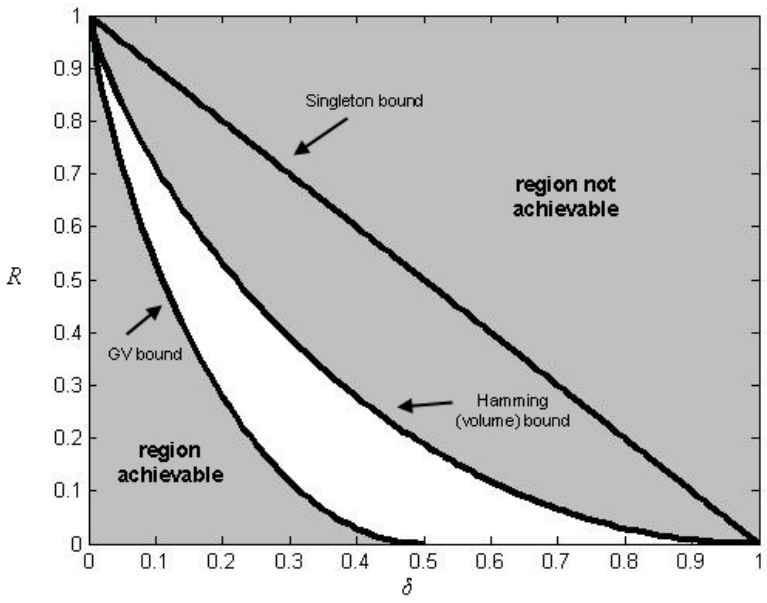
Remark 3.5. *Just as the definition of H is motivated by estimating the volume of a sphere in $\{0,1\}^n$, the definition of H_q is motivated by estimating the volume of a sphere in $\{1, 2, \ldots, q\}^n$. In particular, it can be shown that $\mathrm{Vol}_q(n, \alpha n) = q^{H_q(\alpha)n + o(n)}$.*

The GV bound for all values of q is given in the following theorem, whose proof is left as an exercise.

Theorem 3.6. *For all $q \geq 2$, and for $0 \leq \delta \leq 1 - \frac{1}{q}$, there exists a (not necessarily linear) family of q-ary codes C having rate R and relative distance δ such that $R \geq 1 - H_q(\delta) + o(n)$.*

In Figure 1, we summarize the asymptotic bounds on the rate achievable by binary codes that we have derived so far. We have shown that codes with rates "in the ballpark" of the volume bound exist for $\delta \leq \frac{1}{2}$. However, we do not have any lower bounds yet for the region $\frac{1}{2} \leq \delta \leq 1$. In the next section, we will show that we cannot achieve rates anywhere near the volume bound in this region. In fact, we will prove an upper bound stating that there are no codes with asymptotically positive rate that have a relative distance greater than $\frac{1}{2}$.

Figure 1: Rate bounds for binary codes



4 The Plotkin Bound

The following upper bound, called the Plotkin bound, states that it is impossible to find an infinite family of codes that has a constant rate $R > 0$ for $\delta \geq \frac{1}{2}$.

Theorem 4.1. *Let C be a binary code of distance d and block length n. Then*

1. *if $d = \frac{n}{2}$, then $|C| \leq 2n$, and*
2. *if $d > \frac{n}{2}$, then $|C| \leq (2d)/(2d-n) = (2\delta)/(2\delta - 1)$.*

Remark 4.2. *Asymptotically, Theorem 4.1 states that a binary code family of relative distance greater than or equal to $\delta/2$ has $O(n)$ codewords. Hence $k = O(\log n)$, and $R = k/n \to 0$ as $n \to \infty$. This is what we mean when we say that Theorem 4.1 implies that an infinite family of codes with $\delta \geq \frac{1}{2}$ cannot have a positive rate.*

Remark 4.3. *Although the Plotkin bound is stated in terms of binary codes, its proof can be extended to show a similar bound for $\delta \geq 1 - \frac{1}{q}$ for alphabets of size q such that $|q| \geq 2$. The proof of this is left as an exercise.*

Before we prove Theorem 4.1, we prove a corollary that provides a new bound for $0 \leq \delta \leq \frac{1}{2}$.

Corollary 4.4. *For $0 \leq \delta \leq \frac{1}{2}$, there cannot exist an infinite family of binary codes having rate R and relative distance δ such that $R > 1 - 2\delta + o(1)$.*

Proof. Let C be an $(n,k,d)_2$ code. We show that a consequence of the Plotkin bound is that $2^k \leq 2^{n-2d+1} \cdot 2d$, which implies that $R = \frac{n}{k} \leq 1 - 2\delta + o(1)$ as desired. To do this, we use a trick similar to the one we used in the proof of Theorem 2.2. Consider the map $f : C \to \{0,1\}^{n-2d+1}$ defined by $f(c_1, c_2, \ldots, c_n) = (c_1, c_2, \ldots, c_{n-2d+1})$. For a given $x \in \{0,1\}^{n-2d+1}$, let

$$C_x = \left\{ (c_{n-2d+2}, c_{n-2d+3}, \ldots, c_n) \in \{0,1\}^{2d-1} \ : \ f(c_1, c_2, \ldots, c_n) = x \right\} \ .$$

In other words, if D_x is the set of all codewords in C having x as the prefix, then C_x is the set of suffixes of size $2d-1$ of elements of D_x. Since elements in C_x correspond to codewords in C that have the same prefix of size $n-2d+1$, and since each pair of codewords in C are at a distance of at least d from each other, we know that each pair of elements in C_x must be at a distance of at least d from each other. In other words, C_x is a code of block length $2d-1$ and distance d for all x. Thus by Theorem 4.1, C_x has at most

$$\frac{2d}{2d-(2d-1)} = 2d$$

codewords. This implies that no more than $2d$ codewords in C have the same $x \in \{0,1\}^{n-2d+1}$ as a prefix. Since there are 2^{n-2d+1} possible such x, we conclude that $|C| = 2^k \leq 2^{n-2d+1} \cdot 2d$. □

The Plotkin bound is usually proved combinatorially, but we will prove it in a nice way that relies on the following geometric lemma.

Lemma 4.5. *Let $v_1, v_2, \ldots, v_m$ be unit vectors in $\Re^n$. Then*

1. *if $\langle v_i, v_j \rangle \leq 0$ for all $i \neq j$ then $m \leq 2n$, and*
2. *if there exists a constant $\varepsilon > 0$ such that $\langle v_i, v_j \rangle \leq -\varepsilon$ for all $i \neq j$, then $m \leq 1 + \frac{1}{\varepsilon}$.*

Proof. The proof of the first part is left as an exercise. Note, however, that the bound in the first part is easily reached by taking a basis of $\Re^n$ and its negations.

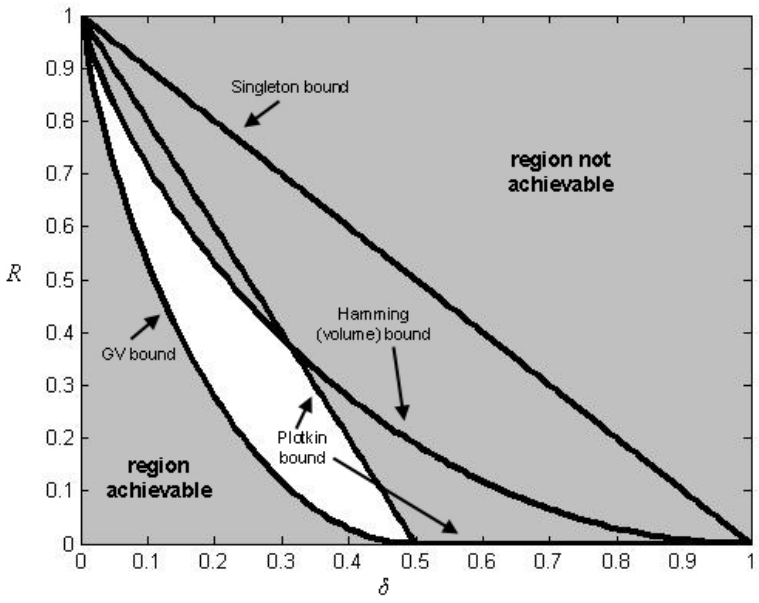
To prove the second part of the lemma, note that

$$\begin{aligned} 0 \leq \|v_1 + v_2 \cdots + v_m\|^2 &= \sum_i \|v_i\|^2 + 2 \sum_{i<j} \langle v_i, v_j \rangle \\ &\leq m + 2\frac{m(m-1)}{2} \cdot -\varepsilon \\ &= m - m(m-1)\varepsilon \ , \end{aligned}$$

which implies that $m \leq 1 + \frac{1}{\varepsilon}$. □

Proof of Theorem 4.1. Consider the given code C. In the proof it will be useful to us to define the map $f : \{0,1\}^n \to \Re^n$ by $f(x_1, x_2, \ldots, x_n) = \frac{1}{\sqrt{n}} \left((-1)^{x_1}, (-1)^{x_2}, \ldots, (-1)^{x_n} \right)$. Clearly, this map is injective. Also, note that for any c and c' in C, if the i^{th} digits of c and c' are the same, then the i^{th} term in the expansion of $\langle f(c), f(c') \rangle$ is $1/n$ and otherwise, the i^{th} term in the expansion of $\langle f(c), f(c') \rangle$ is $-1/n$. Hence, $\langle f(c), f(c') \rangle = 1 - \frac{2}{n}\Delta(c,c') \leq 1 - \frac{2d}{n}$. We can now use these ideas to prove both parts of Theorem 4.1:

Figure 2: Revised rate bounds for binary codes



1. Suppose that $d = n/2$. Then for all $c \neq c'$, $\langle f(c), f(c') \rangle \leq 0$. By the first part of Lemma 4.5 and because f is injective, we conclude that $|C| \leq 2n$.

2. Now suppose that $d > n/2$. Then for all $c \neq c'$, $\langle f(c), f(c') \rangle \leq 1 - \frac{2d}{n} = -\left(\frac{2d-n}{n}\right) < 0$. Now by the second part of Lemma 4.5 and because f is injective, we conclude that

$$|C| \leq 1 + \frac{1}{\frac{2d-n}{n}} = \frac{2d}{2d-n} \ .$$

□

Remark 4.6. *We have seen that the Hadamard code presented in Lecture 2 has 2^r codewords of block length 2^r, and it's distance is 2^{r-1} for all integers $r \geq 1$. Thus we have a code of n codewords that has distance $n/2$. If we create a new code from the Hadamard code by adding the compliment of each codeword, it is not hard to show that the distance is still $n/2$. The existence of this code proves tightness for the first bound given in Theorem 4.1.*

We summarize all of the bounds we have proved in Figure 2. By using some simple ideas, we have achieved fairly tight upper and lower bounds on the rate achievable for any value of δ.