18.409 Applied Extremal Combinatorics	October 22, 1998
Lecture 11	
Lecturer: Daniel A. Spielman	*Scribe: Rados Radoicic*

1 Introduction

Def 1: The **distance** $d(x_1, x_2)$ between two codewords x_1 and x_2 is the number of positions at which they disagree.

If all the codewords are distance d apart, then one can correct $\frac{d}{2} - 1$ errors. So, if we want to make a good error-correcting code, it's reasonable to make all codewords to be far apart. The question that naturally arises from here is:

Question: How many codewords of the certain length can one create, so they are all far apart?
This is one of the main combinatorial problems arising from the code theory. We'll study the following function:

$$A(n,d) \stackrel{def}{=} \max_m \{\exists \text{ code } C : \{0,1\}^m \to \{0,1\}^n \text{ of minimum distance } d\},$$

where **minimum distance** d means that any 2 codewords in C are at least d apart.

We can define an asyptotic version of this function:

$$A(\delta) \stackrel{def}{=} \lim_{n\to\infty} \frac{A(n,\delta n)}{n}$$

$A(\delta)$ is actually a maximal rate such that all codewords are distance δn apart. This later version is more of a combinatorial nature, than practical.
What can we say about this function?
In the rest of this lecture, we'll try to examine the behaviour of $A(\delta)$. We will look only at random linear codes, which will appear to have rather large minimum distance.

2 Crash Course in Linear Codes

Def2: A **linear** code is a code $C : \{0,1\}^{rn} \rightarrow \{0,1\}^n$ whose codewords $C(\{0,1\}^{rn})$ form a vector space over 0,1. A **generator** matrix for a linear code C is a $rn \times n$ $0-1$ matrix whose rows are a basis for the set of codewords.

So, if C is a linear code, then the set of codewords of C is exactly the set of all linear combinations of the rows of its generator matrix with 0 and 1 as coefficients. Therefore, there are 2^{rn} codewords, all specified by the generator matrix. As we can see, linear codes, unlike the nonlinear counterparts, have the efficiency of specification.
Encoding linear codes is easy: just multiply a message vector of length n by the generator matrix. Therefore, the complexity of encoding is the complexity of the multiplication of a column vector by a matrix.
However, there is nothing nontrivial known for **decoding** the linear codes, and the best known decoding still takes exponential time.
Since a code C is linear iff for all codewords $w_1, w_2 \in C : w_1 + w_2 \in C$, we have that the **zero-word** (denote it with $\bar{0}$) is always a codeword (simply $\bar{0} = x + x \in C$).

Def3: The **weight** of a codeword is the number of $1's$ which appear in it. The **minimum relative distance** δ of a code C is

$$\delta \stackrel{def}{=} \frac{1}{n}(\min_{x \neq y \in C(\{0,1\}^{rn})}\{d(x,y)\})$$

For a linear code $d(C(x), C(y)) = d(C(x = y), \bar{0})$, so the minimum weight across all codewords is equal to δn.
One popular example of linear codes is Hamming code.
A random linear code is a linear code which is obtained from a random generator matrix, that is each entry of the matrix is indepently chosen to be 0 or 1, being one or the other with probability 1/2. It turns out that a random linear code will, with high probability, have a minimum relative distance; we see this in the next section.

3 Gilbert-Varshamov Bound

Theorem 1. (Gilbert,Varshamov) $A(\delta) \geq 1 - H(\delta)$

proof: Consider a random linear code. Assume that the rows of its generator matrix are linearly independent, in which case there are 2^{rn} codewords (when the rows are not independent, this number is smaller). Consider a probability that there exists a codeword of weight $\leq \delta n$. We want to show that that probability is relatively small with the right choice of r. What's

the probability that the first row of a generator matrix has weight $\leq \delta n$? Obviously, $Pr[\text{first row has weight} \leq \delta n] = \frac{1}{2^n}\sum_{i=0}^{\delta n}\binom{n}{i}$.
Since any linear combination of rows at the generator matrix (different from $\bar{0}$) is also uniformly randomly chosen just like the first row, we have that:

(1) $Pr[\exists \text{ a non-zero codeword of weight} \leq \delta n] \leq 2^{rn}.\frac{\sum_{i=0}^{\delta n}\binom{n}{i}}{2^n} \sim 2^{n(r+H(\delta)-1)}$

If $1-r > H(\delta)$, then the RHS of (1) is less than 1, and consequently there must exist some linear code at rate r in length n in which every nonzero codeword has weight $> \delta n$ $\Box$

It is not known whether this lower bound on $A(\delta)$ can be improved.

4 Weak Upper Bounds

Theorem 2. (Singleton bound) $A(\delta) \leq 1-\delta$

proof: Take a set of codewords, focus on the first coordinate of the codewords, and define 2 sets: $C|_{x_1=i}$ is the set of codewords whose first coordinate is i, $i \in 0, 1$. Since these 2 sets form a disjoint union of the set of all codewords, one of these sets contains at least 1/2 of all the codewords. So, w.l.o.g assume that set is $C|_{x_1=0}$. Now, we ignore the first coordinate in all the codewords of $C|_{x_1=0}$ and end up with the code of length $n-1$, with $\geq 2^{rn-1}$ codewords and minimum distance $\geq d$. From here we get that $A(n-1,d) \geq A(n,d)-1$. We repeat the above operation $n-d$ times and end up with

$$A(d,d) \geq A(n,d)-(n-d).$$

Since you can fit only 2 codewords of length d at distance d (all-one word and zero-word), we get that

$$n-d+1 \geq A(n,d).$$

Asymptotiucally, from here we get that $A(\delta) \leq 1-\delta$ $\Box$

Theorem 3(Volume bound) $A(\delta) \leq 1-H(\delta/2)$

proof:The proof is based simply on volume considerations, If δ is the minimum relative distance between codewords, then we can put disjoint balls of radius $\delta/2$ around each of the 2^{rn} codewords. Since each of these balls contains approximately $2^{nH(\delta/2)}$ points, we obtain the following inequality:

$$2^{rn}2^{nH(\delta/2)} \leq 2^n,$$

which implies that $r \leq 1 - H(\delta/2)$ □

These bounds are quite weak. For $\delta = 1/2$, the volume bound only gives us that $A(1/2) < 1 - H(1/4) \approx 0.188$, leaving open the possibility that a set of $2^{0.188n}$ codewords could have a minimum relative distance 1/2. However, it's in fact impossible for a set to contain more than n codewords and still have minimum relative distance 1/2. Therefore, we need stronger upper bounds, which we present in the following two sections.

5 Plotkin Bound

Theorem 4 $A(1/2 + \varepsilon) = 0$ for $\forall \varepsilon > 0$.

proof: The proof proceeds by counting $\sum_{i<j} d(w_i, w_j)$, the sum of all pairwise distances between codewords, in two different ways. Since the minimum distance between codewords is δn, we have that:

$$(2)\ \sum_{i<j} d(w_i, w_j) \geq \binom{2^{rn}}{2} \delta n$$

Now consider a $2^{rn} \times n$ matrix whose rows are the codewords. Let a_i be the fraction of rows which have a 0 in the i-th position (i.e in the i-th column). A contribution to the total pairwise distance is made whenever two words differ in position i; consequently, the bits in the i-th position contribute $a_i(1 - a_i)2^{2rn}$ to the total. Summing over all columns, we find that:

$$(3)\ \sum_{i<j} d(w_i, w_j) = \sum_{i=1}^{n} a_i(1 - a_i)2^{2rn}$$

The RHS of (3) is maximized when $a_i = 1/2$. Combining this with (2), we get:

$$\frac{n2^{2rn}}{4} \geq \binom{2^{2rn}}{2} \delta n = \frac{2^{2rn}(2^{2rn} - 1)\delta n}{2}$$

Simplifying, we find that $\delta \geq 2^{rn}(\delta - 1/2)$. Now, if $\delta > 1/2$, we can divide through by $(\delta - 1/2)$ and obtain $\frac{\delta}{\delta - 1/2} \geq 2^{rn}$. So, if δ is a constant greater than 1/2, then 2^{rn} is bounded above by a constant, which is as it should be. Furthermore, we have that $r \to 0$ as $n \to 0$ for fixed $\delta > 1/2$. It turns out that the Plotkin bound at $\delta = 1/2$ can be extended linearly back to $\delta = 0$.(see picture 1.) □

6 Elias Bound

This is the second upper bound known until now. The only better bound is the linear-programming (LP) bound, which we will talk about later in the semester.

Theorem 5. $A(\delta) \leq 1 - H(\frac{1}{2} - \sqrt{\frac{1}{4} - \frac{\delta}{2}})$

proof: Consider a randomly chosen ball of radius ρn (such a ball is randomly chosen by choosing randomly a word and all the codewords within ρn distance). We will determine ρ later in the proof. The expected number of codewords in the ball is $2^{rn}\frac{\sum_{i=0}^{\rho n}\binom{n}{i}}{2^n} \sim 2^{n(r+H(\rho)-1)}$. If the center of the ball is a word x, then we can recenter the ball to the origin ($\hat{0}$ word) simply by adding x to each codeword in the ball ($x + x = \bar{0}$). Therefore, there exists a ball B with $2^{n(r+H(\rho)-1)}$ codewords, such that each codeword has weight at most ρn and the minimum distance $\geq \delta n$. Denote $M = 2^{n(r+H(\rho)-1)}$. As in the proof of Plotkin bound, we count $\sum_{i<j} d(w_i, w_j)$ in two different ways. Since the minimum distance between codewords is δn, we have that

$$\sum_{i<j} d(w_i, w_j) \geq \binom{M}{2}\delta n \ (4).$$

Now, consider a $M \times n$ matrix S whose rows are the codewords from ball B. Let a_i be the fraction of rows which have a 1 in the i-th position (i.e in the i-th column). As in section 5, we obtain:

$$\sum_{i<j} d(w_i, w_j) = \sum_{i=1}^{n} a_i(1 - a_i)M^2 \ (5)$$

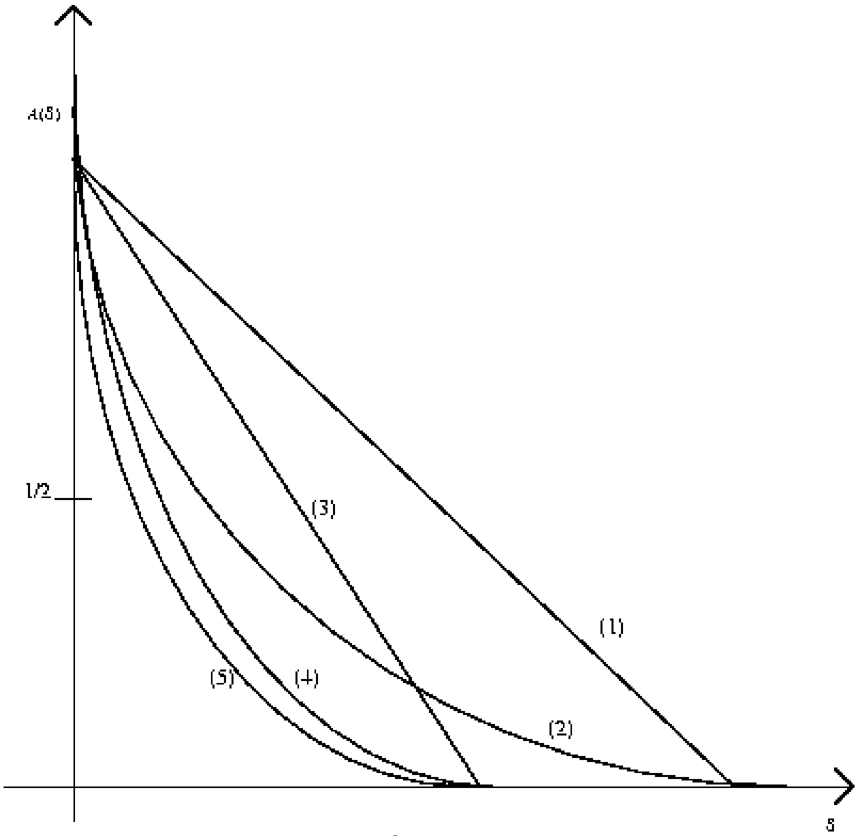
Now, since the weight of every codeword in the matrix S is at most ρn, we get that the total number of ones in S is $\leq M\rho n$, counting by rows. Counting the ones in S by columns, we get: $\sum a_i M \leq M\rho n$ (6). From (6) we conclude that the RHS of (5) is maximized when $a_i = \rho$. Now, from (4) and (5) we get:

$$\binom{M}{2}\delta n \leq M^2(\rho - \rho^2)n \Leftrightarrow M = 2^{n(r+H(\rho)-1)} \leq \frac{\delta}{\delta - 2(\rho - \rho^2)} = \text{const}$$

If $\delta > 2(\rho - \rho^2)$, then $r \leq 1 - H(\rho)$.
Therefore $r \leq 1 - H(\frac{1}{2} - \sqrt{\frac{1}{4} - \frac{\delta}{2}})$ $\square$

7 Conclusion

The bounds for $A(\delta)$, shown in this lecture, can be summarized in the following graph:



Picture 1: (1) Singleton bound
(2) Volume bound
(3) Plotkin bound
(4) Elias bound
upper bounds
(5) Gilbert-Varshamov bound