6.897 Algorithmic Introduction to Coding Theory October 2, 2002

Lecture 8

Lecturer: Madhu Sudan *Scribe: Laura Serban*

1 Overview

- Construction of Justesen Codes
- Plotkin Bound
- Hamming-Elias-Bassalygo Bound
- Johnson Bound

2 What we have achieved so far

So far, we have examined three types of codes:

- **Atomic Codes**: Hamming, Hadamard, Reed-Solomon, and Reed-Muller codes
- **Random Codes**
- **Concatenated Codes**: Forney and Justesen Codes

The last two types of codes are asymptotically good. However, while the Random Codes are non-constructible, the Concatenated Codes provide us with an explicit way of producing asymptotically good codes. Further in the course we will focus on decoding algorithms for some of the codes examined so far, and on applications of coding theory to other areas of Computer Science.

Since we have failed to give an explicit construction for the Justesen Codes in the previous lectures, for completion, we will take up that task here.

3 Construction of Justesen Codes

Consider $n = q = 2^l$ and the field $\mathbb{F}_{2^l}$. For any message $m = (c_0, c_1, \ldots, c_{k-1}) \in \mathbb{F}_{2^l}^k$ define the polynomial $p(x) = \sum_{i=1}^{i=k} c_i x^i$. Encode m as the concatenation of $(p(\alpha), \alpha p(\alpha))$ for each $\alpha \in \mathbb{F}_{2^l}$. The message length is kl bits since each message is represented as k elements of the field F_{2^l}. The blocklength is $2l2^l$ since $p(\alpha)$ is l bits long and the pairs $(p(\alpha), \alpha p(\alpha))$ are concatenated for all 2^l elements of F_{2^l}. The minimum distance is asymptotic to $(2^l - k)l$, i.e. $c(2^l - k)l$ where c is a global constant. We will not prove this result here. Thus, Justesen codes are $(l2^{l+1}, kl, c(2^l - k)l)_2$ codes. Surprisingly, by appending a low-performace Reed-Solomon code with a translation of itself, we obtain a code with remarkably good asymptotic properties.

4 Plotkin Bound

To motivate our first bound, let us re-examine the bounds we have so far. On the one hand, we have the Singleton and Hamming upper bounds on codes which show that $R \leq 1 - \delta$ and $R \leq 1 - H(\delta/2)$, respectively, the latter dominating the former. On the other hand, the Gilbert-Varshamov bound shows that there exist random codes with rate $R \geq 1 - H(\delta)$. The Plotkin bound addresses the largest gap between these bounds. More precisely, it rules out the existence of codes of positive rate and minimum relative distance larger than $\frac{1}{2}$.

Theorem 1 (Plotkin) *Let C be a code with relative minimum distance δ and rate R.*

Plotkin 1 *If $\delta = \frac{1}{2} + \epsilon$, then $|C| \leq \frac{1}{2\epsilon} + 1$. That is, if the relative minimum distance of the code exceed $\frac{1}{2}$, there are only a constant number of allowable codewords, i.e. the rate of the code tends to zero as the blocklength becomes large.*

Plotkin 2 *If $\delta = \frac{1}{2}$, then $|C| \leq 2n$.*

Plotkin 3 *The rate of the code satisfies $R \leq 1 - 2\delta$.*

Observations

- The Simplex Code meets the first Plotkin bound within a constant factor.
- The Hadamard Code meets the second Plotkin bound tightly.

Next, we will prove the first Plotkin bound. The last two bounds are left as an exercise.

Proof Idea: Embed the Hamming space into the Euclidean space. Specifically, define a map from $\{0,1\}$ to $\mathbb{R}$ such that 0 corresponds to 1, and 1 correspond -1, respectively. We can then inductively define a map from $0,1^n$ to $\mathbb{R}^n$. Therefore any object in the Hamming Space $\{0,1\}^n$ can be represented by a vector in the Euclidean Space. The embedding function has the following easy and useful properties.

Fact 2 *For any $x \in C \subset \{0,1\}^n$, the corresponding vector in the n-dimensional Euclidean space v_x has length $\sqrt{n}$, i.e. $||v_x||^2 = n$.*

Fact 3 *For any $x, y \in C \subset \{0,1\}^n$, with Hamming distance $\Delta(x,y)$ the inner product of the corresponding vectors in the n-dimensional Euclidean space, v_x and v_y is $< v_x, v_y >= n - 2\Delta(x,y)$.*

Thus two codewords that are far from each other in the Hamming space have a small inner product. In particular, if two codewords differ in more than half of the positions the angle between their corresponding vectors in Euclidian space is obtuse.

Let us now translate the first version of the Plotkin Bound to the Euclidean space. All vectors corresponding are scaled to length one.

Fact 4 (Plotkin 1 Euclidean Space) *Given k vectors $v_1, v_2, \ldots, v_k \in \mathbb{R}^n$ of unit length such that $< v_i, v_j >\leq -\alpha$, $k \leq 1 + \frac{1}{\alpha}$. To obtain the first version of the Plotkin Bound take $\alpha = 2\epsilon$.*

We will give two proofs of the fact above. One is intuitive, but fairly involved, the other is short and simple, but doesn't provide much intuition. In the proofs belown require only that the length of each vector v_i is at most one, rather than exactly 1.

Proof [Proof 1 - Intuitive] Without loss of generality let $v_1 = (1, 0, \ldots, 0)$. Since $< v_1, v_i >\leq \alpha$, v_i must be of the form $v_i = (-\alpha_i, u_i)$ where $\alpha_i \leq \alpha$ and $u_i \in \mathbb{R}_{n-1}$, for any i, $2 \leq i \leq k$. Then, for any i, j such that $2 \leq i, j \leq k$. Note that:

$$\begin{aligned} < u_i, u_j > &= < v_i, v_j > -\alpha_i\alpha_j \\ &\leq < v_i, v_j > -\alpha^2 \\ &\leq -\alpha - \alpha^2 = -\alpha(1+\alpha) \end{aligned}$$

So the u_i's form a set of $k-1$ vectors of length at most 1 such that for the inner product of any two is at most $-\alpha(1+\alpha)$. We repeat the argument above for this new set of vectors. Note that at every round the number of vectors decreases by one. Continuing this argument for more than $\frac{1}{\alpha} + 1$ steps we construct a set of vectors of length at most 1 such that the inner product of any two of them is less than -1. This is clearly impossible. Therefore, we should be left with zero vectors after at most $\frac{1}{\alpha} + 1$

rounds, whcih implies that $k \leq \frac{1}{\alpha} + 1$.

We can actually do even better by observing that $< u_i, u_i > \leq 1 - \alpha_i^2 \leq 1 - \alpha^2$. We can scale each u_i up by a factor of $\frac{1}{\sqrt{1-\alpha^2}}$ while still maitaining the required properties of the set. The new sets of vectors thus constucted will have inner products that are less than $-\alpha$ by a larger quantity.

The Plotkin Bound is tight. To see that in Euclidean space reverse engineer the inductive proof above to construct a set of vectors that satisfies the bound tightly. In the Hamming space, one can proof tightness by examples of specific codes that achieve the bound.

■

Proof [Proof 2]

Let $z = v_i + v_2 + \ldots v_k$. Recall that $< v_i, v_i > \leq 1$ since the length of v_i is at most one, and $< v_i, v_j > \leq -\alpha$ for all $1 \leq i, j \leq k$. Compute

$$\begin{aligned} < z, z > &= \sum_{i,j} < v_i, v_j > \\ &= \sum_{i} < v_i, v_i > + \sum_{i \neq j} < v_i, v_j > \\ &\leq k + k(k-1)(-\alpha) = k(1 + (1-k)\alpha) \end{aligned}$$

But $< z, z > \geq 0 \Rightarrow k(1 + (1-k)\alpha) \geq 0 \Rightarrow k \leq 1 + \frac{1}{\alpha}$.

■

Next we state the Euclidean version of the second Plotkin Bound.

Fact 5 (Plotkin 2 Euclidean Space) *Given k vectors $v_1, v_2, \ldots, v_k \in \mathbb{R}^n$ of length at most one such that $< v_i, v_j > \leq -\alpha$, then $k \leq 2n + 1$. If the vectors are non-zero, the bound becomes $k \leq 2n$.*

The proof of this fact is similar to the proof for the first Plotkin about. A rigurous analysis can be found in the *Lecture Notes from 2001.*

5 Hammilton-Elias-Bassalygo Bound

The current state of affairs for upper bounds involves both the Hamming and the Plotkin bounds. More specifically, for small values of the δ, the Hamming bound is better than the Plotkin bound, while the Plotkin bound is stronger for larger values of the relative minimum distance. Our next upper bound is always stronger than the Hamming bound although it is very close to the Hamming bound for small values of δ. Before we introduce it, we define a new notion of error-correcting codes, which will be later refered to as "list-decodable" codes.

Definition 6 ((t,l)-error-correcting codes) *A code C is (t, l) error-correcting if for every vector $x \in \{0,1\}^n$, the $|Ball(x, t) \bigcap C| \leq l$.*

That is, if at most t errors happen, the intial codeword could any l given codewords. Note that this a generalization of the Hamming notion of error-correcting codes. In particular, an error-correcting code in the Hammming sense (all the codes we have encountered so far) is $(t, 1)$ error-correcting.

Theorem 7 (Hammilton-Elias-Bassalygo Bound) *If C is (t, l) error-correcting then*

$$|C| \leq \frac{l 2^n}{\mathrm{Vol}_2(t, n)} \approx l 2^{n(1 - H(\frac{t}{n}))}$$

Proof Draw a ball of radius t around each codeword. Each point lies in at most l balls. Otherwise, a point $x \in \{0,1\}^n$ the ball of radius t around that x would contain more than l codewords, contradicting the definition of a (t, l) error-correcting code. Since there are only 2^n distinct points in the Hamming space $|C|\text{Vol}_2(t, n) \leq l2^n$ which yields the bound. ■

In order to get asymptotic results, we would like to relate the (relative) minimum distance of a code to its l-error-correcting radius for $l > 1$.

Theorem 8 (Johnson Bound) *A $(n, k, \delta n)_2$ code is $(t, O(n))$ error-correcting where $\frac{t}{n} = \frac{1}{2}(1-\sqrt{1-2\delta})$.*

Our first sanity check confirms that $\frac{\delta}{2} \leq \frac{1}{2}(1-\sqrt{1-2\delta}) \leq \delta$. We next attempt to give some intuition as to why the quantity $1-\sqrt{1-2\delta}$ is appropriate. Construct a non-linear code that has large minimum distance, but has one Hamming ball of radius t containing $\exp(n)$ codewords. This ensures that the code is not (t, l) error-correcting. Fix a ball of radius t around the origin in the Hamming space of dimension n, and define a random code inside it. We can show that we can pick exponentially many codewords inside this ball before any two become too close to each other.