STOC8 LECTURE 09 3/5/2008

Note Title 3/4/2008

TODAY: CONCLUDING COMBINATORICS OF CODES.

- ALON'S BOUND FOR "BALANCED" CODES
- Summary

Review of what we've seen so far

Three basic bounds [q=2]

Negative: PLOTKIN / HAMMING-ELIAS-BASSALYGO:

$$R \leq 1 - f(\delta)$$

$$= 1 - \frac{\delta}{2}\log\delta - \text{lesser order terms} \quad \text{as } \delta \to 0$$

$$= O(\epsilon) \text{ if } \delta = 1/2 - \epsilon \text{ \& } \epsilon \to 0$$

Existential: GILBERT / VARSHAMOV

$$R \geq 1 - g(\delta)$$

$$= 1 - \delta \log \delta - \text{lower order terms}$$

as $\delta \to 0$

$$= \Omega(\epsilon^2) \quad \text{if } \delta = \tfrac{1}{2} - \epsilon$$

& $\epsilon \to 0$

CONSTRUCTIVE: FORNEY / JUSTESEN

$$R \geq 1 - h(\delta)$$

$$= 1 - \sqrt{\delta} \log \delta - \ldots \quad \text{as } \delta \to 0$$

$$= \Omega(\epsilon^3) \quad \text{if } \delta = \tfrac{1}{2} - \epsilon$$

$\epsilon \to 0$

Some Other Effects

BCH : $d =$ fixed or $n^{o(1)}$ $(q=2)$

$n-k \le \frac{d}{2} \log n$ (matches Hamming)

RS : $q = n$

$n = k+d-1$ (matches Singleton)

AG : $n-k \le d + \frac{n}{\sqrt{q}-1}$ (beats GV)

So algebraic codes are better than random, except if $q = 2, 3$ (or something small), and $\frac{d}{n} = \delta > 0$.

General feeling:

Eventually Constructive = Existential

Existential vs. Negative unclear...

Except: Can improve Negative.

E.g. $R \leq O\left(\epsilon^2 \log \frac{1}{\epsilon}\right)$ if $\delta = \frac{1}{2} - \epsilon$

$\epsilon \to 0$.

Today: - A simple proof for when codes are "balanced".

- Overview of general proofs.

Balanced Codes

$C \subseteq \{0,1\}^n$ is ϵ-balanced if

$\forall x, y \in C \quad x \neq y$,

$$\left(\frac{1-\epsilon}{2}\right) n \leq \Delta(x,y) \leq \left(\frac{1+\epsilon}{2}\right) n$$

Usual dist. Criterion

Special New Stuff.

Theorem: $R \triangleq \frac{\log|C|}{n}$ satisfies

$$R = O\left(\epsilon^2 \log \frac{1}{\epsilon}\right)$$

Proof [ALON] :

Idea : • Write C as $K \times n$ matrix with entries being $\pm \frac{1}{\sqrt{n}}$.

• $C \cdot C^T$ is $K \times K$ matrix with

$$\left.\begin{array}{l} \bullet \text{ diagonals being } 1 \\ \bullet \ |\text{off-diagonal}| \leq \epsilon \\ \bullet \text{ rank} \leq n \end{array}\right\} \text{Contradictory}$$

Lemma 1: M is $K \times K$ matrix with diagonal 1 & $|\text{off-diagonal}| \leq \epsilon$

$$\Rightarrow \text{rank}(M) \geq \frac{K}{1 + (K-1)\epsilon^2}$$

Linear Algebra Review:

① Real, symmetric matrix M has K eigenvalues $\lambda_1 \ldots \lambda_K$

② $\text{Rank}(M) = K - \#\{i \mid \lambda_i = 0\}$

③ $\sum \lambda_i = \sum M_{ii} = \text{Trace}(M)$

④ Eigenvalues of $M \cdot M = \lambda_1^2, \lambda_2^2 \cdots \lambda_K^2$

In our case

- $\sum \lambda_i = K$

- $\sum \lambda_i^2 \geq \frac{K^2}{\text{rank}(M)}$ [Cauchy-Schwartz]

- $\sum \lambda_i^2 = \text{Trace}(M \cdot M^T)$

$$= \sum_{i,j} m_{ij}^2$$

$$\leq K + K(K-1)\epsilon^2$$

$$\Rightarrow \text{rank}(M) \geq \frac{K}{1+(K-1)\epsilon^2} \quad \boxtimes$$

But doesn't seem to give much .... if

$\epsilon = O(1)$ & $K \to \infty$

(rank $\geq \frac{1}{\epsilon^2}$; we want rank $\geq \frac{1}{\epsilon^2} \log K$)

Will work with matrix $M^{(t)}$ whose entries are simply $(M_{ij})^t$.

Lemma 2: If $\text{rank}(M) \leq r$

then $\text{rank}(M^{(t)}) \leq \binom{r+t}{t}$ [Better than r^t]

Proof: Let $V_1 \ldots V_r$ span columns of M.

Let $V_j^{(k_1, \ldots, k_r)} = V_j^{k_1} \cdot V_j^{k_2} \ldots V_j^{k_r}$

Then $V^{(00000)}, V^{(00001)}, \ldots V^{(t0000)}$

(set of vectors $V^{(k_1 \ldots k_r)}$ $\sum k_i \leq t$)

span $M^{(t)}$

($\left(\sum \alpha_i V_i\right)^t$ is in span of above)

<u>Lemma 3</u>: M is $K \times K$ matrix with diagonal 1 & $|\text{off-diagonal}| \leq \epsilon$

$$\Rightarrow \text{rank}(M) \geq \Omega\left(\frac{1}{\epsilon^2} \cdot \log K \cdot \log \frac{1}{\epsilon}\right)$$

<u>Proof</u>: Let t be s.t. $\epsilon^{2t} \approx \frac{1}{K-1}$

$$\Rightarrow t = \frac{\log(K\quad)}{\log \frac{1}{\epsilon^2}}$$

Thus $\text{rank}(M^t) \geq K/2$ (Lemma 1)

But $\text{rank}(M) \leq \binom{r+t}{t} \approx \left(\frac{r}{t}\right)^t$

$$\Rightarrow \left(\frac{r}{t}\right) \geq K^{1/t} = \frac{1}{\epsilon^2} \Rightarrow r \geq \epsilon^2 \cdot \log \frac{1}{\epsilon} \cdot \log K$$

[The Theorem follows]

MacWilliams Identity / Linear-Programming Bound

Definition: Weight Distribution of a code $C \subseteq \{0,1\}^n$ = $(B_0, \ldots\ldots B_n)$

B_i = # codewords of C of weight i.

MacWilliams Identity for linear C & dual $C^\perp$

$B_0(C^\perp), \ldots \; B_n(C^\perp)$ can be computed from $B_0(C) \ldots \; B_n(C)$ linearly.

$$B_i(C^\perp) = \frac{1}{|C|} \sum_{j=0}^{n} K_j(i) \, B_j(C)$$

(or something like that).

LP Bound : Upper Bound

$B_0(C) + \dots \; B_n(C)$ s.t.

$B_0(C) = 1$

$B_i(C) = 0 \qquad i = 1, \dots, d-1$

linear constraints $\to$ $B_j(C^\perp) = 1 \qquad j = 0$

$\to$ $B_j(C^\perp) \geq 0 \qquad j \geq 0$

Amazingly .... gives a bound;

works for all codes (non-linear too !)

$\Rightarrow \; R \leq O\left(\epsilon^2 \log \frac{1}{\epsilon}\right) \dots$

Recent Proof:

[Friedman + Tillich] linear case

[Navon + Samorodnitsky] general case

Open Questions:

- Asymptotics of $f(q)$ s.t. $R = 1 - \delta - f(q)$.
- Ternary codes with $n-k \leq \frac{d}{2} \log_3 n$
- $\exists$ Codes of rate $R \geq 1 - \frac{\delta}{2} \log \frac{1}{\delta}$. binary $\delta \to 0$
- Construct Codes of rate $R = \omega(\epsilon^3)$ $[\delta = \frac{1}{2} - \epsilon]$

  $R = 1 - o\left(\sqrt{\delta} \log \frac{1}{\delta}\right)$ $[\delta \to 0]$