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:

\[
R \leq 1 - f(q) = 1 - \frac{q}{2} \log q - \text{lower order terms}
\]

\[
= O(\varepsilon) \quad \text{if } q = \frac{1}{2} - \varepsilon \quad \varepsilon \to 0
\]

\[
\lim_{q \to 1} \frac{f(q)}{q} = \infty
\]
Existential: Gilbert / Varshamov

\[ R \geq 1 - g(\delta) \]

\[ = 1 - \delta \log \delta - \text{linear order terms} \]

\[ \approx \delta \rightarrow 0 \]

\[ = \Omega(\epsilon^2) \quad \text{if } \delta = \frac{1}{2} - \epsilon \]

\[ \delta \epsilon \rightarrow 0 \]

Constructive: Forney / Justesen

\[ R \geq 1 - h(\delta) \]

\[ = 1 - \sqrt{\delta} \log \delta - \cdots \approx \delta \rightarrow 0 \]

\[ = \Omega(\epsilon^3) \quad \text{if } \delta = \frac{1}{2} - \epsilon \]

\[ \epsilon \rightarrow 0 \]
Some Other Effects

BCH: \( d = \text{fixed or } n^{o(1)} \quad (q=2) \)
\[
    n - k \leq \frac{d}{2} \log n \quad \text{(matches Hamming)}
\]

RS: \( q = n \)
\[
    n = kd - 1 \quad \text{(matches Singleton)}
\]

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

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.

\[ R \leq O(\varepsilon^2 \log \frac{1}{\varepsilon}) \text{ if } \delta = \frac{1}{2} - \varepsilon \]

\[ \varepsilon > 0 \]

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

- Overview of general proofs.
Balanced Codes

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

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

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

$$\uparrow \quad \uparrow$$

Usual Dist.  Special
Criterion  New Stuff.

\[ \text{Theorem: } R = \frac{\log |C|}{n} \text{ satisfies} \]

\[ R = O(\epsilon^2 \log \frac{1}{\epsilon}) \]
Proof [Alon]:

Idea:
- Write \( C \) as a \( K \times n \) matrix with entries being \( \pm \frac{1}{\sqrt{n}} \).

\[ C \cdot C^T \text{ is a } K \times K \text{ matrix with } \]
- diagonals being 1
- \( |\text{off-diagonal}| \leq \epsilon \) \( \Rightarrow \) Contradiction
- rank \( \leq n \)
Lemma 1: \( M \) is a \( K \times K \) matrix with diagonal 1 and \( | \text{off-diagonal} | \leq \varepsilon \)

\[ \Rightarrow \text{rank} (M) \geq \frac{K}{1 + (K-1)\varepsilon^2} \]

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 \( MM \) = \( \lambda_1^2, \lambda_2^2, \ldots, \lambda_K^2 \)
In our case

\[ \sum \lambda_i = k \]

\[ \sum \lambda_i^2 \geq \frac{k^2}{\text{rank}(M)} \quad \text{[Cauchy-Schwarz]} \]

\[ \sum \lambda_i^2 = \text{Trace}(mm^T) \]

\[ = \sum m_{ij}^2 \]

\[ \leq k + K(K-1)\epsilon^2 \]

\[ \Rightarrow \text{rank}(M) \geq \frac{k}{1 + (K-1)\epsilon^2} \quad \Box \]

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

\[ C = O(1) \quad \& \quad K \to \infty \]

\[ (\text{rank} \geq \frac{1}{\epsilon^2}; \text{ we want rank} \geq \frac{1}{\epsilon^2 \log K}) \]
Will work with matrix $M^{(t)}$ whose entries are simply $(M_{ij})^t$.

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

Then $\text{rank}(M^{(t)}) \leq \left(\frac{r + t}{t}\right)$ [Better]

**Proof:** Let $v_1, \ldots, v_r$ span columns of $M$.

Let $V_j^{(k_1, \ldots, k_r)} = V_{j_1}^{k_1} \cdot V_{j_2}^{k_2} \cdots V_{j_r}^{k_r}$

Then $V_1^{(0,0,0,0)}, V_1^{(0,0,1,0)}, \ldots, V_1^{(0,0,0,1)}$

(set of vectors $V_j^{(k_1, \ldots, k_r)} \forall k_i \leq t$)

span $M^{(t)}$

($\sum \alpha_i v_i$ is in span of above)
**Lemma 3:** A $K \times K$ matrix $M$ is diagonal with diagonal $\mathbf{1}$ and at most $\varepsilon$ off-diagonal entries.

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

**Proof:** Let $\varepsilon$ be such that $\varepsilon^2 = \frac{1}{K-1}$.

\[ \Rightarrow \varepsilon = \log \left( \frac{K}{\varepsilon^2} \right) = \frac{1}{\log \frac{1}{\varepsilon^2}} \]

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

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

\[ \Rightarrow \left( \frac{\varepsilon}{\varepsilon} \right)^t \geq K^{1/2} = \varepsilon^2 \Rightarrow t \geq \varepsilon^2 \cdot \log \frac{1}{\varepsilon} \cdot \log K \]

[Theorem follows]
MacWilliams Identity / Linear Programming Bound

**Definition**: Weight Distribution of a code

\[ C \subseteq \{0,1\}^n = (B_0, \ldots, B_n) \]

\[ B_i = \# \text{ codewords of } C \text{ of weight } i. \]

**MacWilliams Identity** for linear \( C \) & dual \( C^\perp \)

\[ B_i(C^\perp) \text{ can be computed from } B_0(C), \ldots, B_n(C) \text{ linearly.} \]

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

(or something like that).
**LP Bound** : **Upper Bound**

\[ B_0(c) + \cdots + B_n(c) \quad \text{s.t.} \]

\[ B_0(c) = 1 \]

\[ B_i(c) = 0 \quad i = 1, \ldots, d-1 \]

**Linear Constraints**

\[ B_j(c^+) = 1 \quad j = 0 \]

\[ B_j(c^-) \geq 0 \quad j > 0 \]

Amazingly .... gives a bound;

works for all codes (non-linear too !)

\[ R \leq O(C^2 \log \frac{1}{\epsilon}) \ldots \]
Recent Proof:

[Friedman + Tillich] linear case

[Narovon + Samorodnitsky] general case

Open Question:

- Asymptotic of $f(q)$ s.t. $R = 1 - S - f(q)$.
- Ternary codes with $n - k \leq \frac{d}{2} \log_2 n$
- Codes of rate $R \geq 1 - \frac{e}{2} \log \frac{1}{s}$. binary $S \to 0$
- Construct codes of rate $R = \omega(\varepsilon^3)$ [8 = $\frac{1}{2} - \varepsilon$]

$R = 1 - o(\sqrt{s \log \frac{1}{\delta}})$ [8 = 70]