CSE522/Math581EB Coding Theory **Fall 2001**

Lecture 7: October 30

Lecturer: Elchanan Mossel *Scribe: Ashish Sabharwal*

7.1 Bounds for General Codes (continued)

7.1.1 (t, l) Error-correcting codes and Johnson Bound

Definition 7.1 *A* (t, l) error-correcting code *is a code such that any ball* $B(y, t)$ *contains at most* l *codewords.*

Proposition 7.2 *A* (t, l) *error-correcting code with parameters* $[n, k, d]_2$ *satisfies* $2^k \text{Vol}(t, n) \leq l2^n$.

Proof: Consider 2^k balls of radius t around each codeword. The sum of the volumes of the balls is $2^k \text{Vol}(t, n)$. Each point in the balls can however appear in only l balls. Therefore the sum of the volumes must be bounded by $l2^n$. ■

Lemma 7.3 *If* $y, z_1, \ldots, z_m \in \mathbb{R}^n$ *satisfy* $\forall i,\ \langle y, z_i \rangle > 0$, *and* $\forall i, j,\ \langle z_i, z_j \rangle \leq 0$, *then* $m \leq n$.

Proof: Assume $m > n$. It follows that $z_1, \ldots, z_m$ are linearly dependent. Hence there exist non-empty T, S and positive α_k such that $\sum_{i \in T} \alpha_i z_i = \sum_{j \in S} \alpha_j z_j = z$ (say) and $T \cap S = \phi$. We will obtain a contradiction by showing that both $z \neq 0$ and $z = 0$ follow from this.

First consider $\langle y, z \rangle = \sum_i \alpha_i \langle y, z_i \rangle$. From the conditions of the Lemma, this quantity is positive and hence $z \neq 0$. Now consider $\langle z, z \rangle = \langle \sum_{i \in T} \alpha_i z_i, \sum_{j \in S} \alpha_j z_j \rangle = \sum_{i \in T, j \in S} \alpha_i \alpha_j \langle z_i, z_j \rangle$. From the conditions of the Lemma, this quantity is at most 0 and therefore, being the inner product of a vector with itself, exactly 0. Hence $z = 0$. This completes the proof. ■

Theorem 7.4 (Johnson bound [Johnson1,Johnson2]) *An* $(n, k, \delta n)$ *code is* $(\tau n - 1, n)$ *error-correcting, where* $\tau = \frac{1}{2}(1 - \sqrt{1 - 2\delta})$.

Proof: Let $y \in \{0, 1\}^n$ and $x_1, \ldots, x_m$ be the codewords in $B(\tau n - 1, y)$. Let $\bar{y}$ be the vector in $\mathbb{R}^n$ corresponding to y, where each 1 becomes a -1 and each 0 becomes a 1, and the vector is scaled so that $||\bar{y}|| = 1$. Similarly define $\bar{x}_i$'s.

Since y and x_i (for any i) disagree in less than τ fraction of places, $\langle \bar{y}, \bar{x}_i \rangle > 1 - 2\tau$. Also $\langle \bar{x}_i, \bar{x}_j \rangle \leq 1 - 2\delta$ since any two codewords differ in at least a δ fraction of places. We will now translate these vectors in $\mathbb{R}^n$ and apply Lemma 7.3.

Look at vectors $(1 - \alpha)\bar{y}, \bar{x}_1 - \alpha\bar{y}, \ldots, \bar{x}_m - \alpha\bar{y}$. To apply Lemma 7.3 on these vectors, we need $\langle \bar{x}_i - \alpha\bar{y}, \bar{x}_j - \alpha\bar{y} \rangle \leq 0$. This quantity is at most $(1 - 2\delta) - 2\alpha(1 - 2\tau) + \alpha^2$, which is minimized when $\alpha = 1 - 2\tau$ (obtained by setting the derivative to 0). Substituting $1 - 2\tau$ for α, we get $(1 - 2\delta) - (1 - 2\tau)^2$, which is 0 for $\tau = \frac{1}{2}(1 - \sqrt{1 - 2\delta})$.

The second condition of the Lemma is easily verified for $\alpha = 1 - 2\tau$: $\langle (1 - \alpha)\bar{y}, \bar{x}_i - \alpha\bar{y} \rangle = (1 - \alpha)\langle \bar{x}_i, y \rangle - \alpha(1 - \alpha) > (1 - \alpha)((1 - 2\tau) - \alpha) = 0$. Applying Lemma 7.3 now, m, the number of codewords in $B(\tau n - 1, y)$, is at most n. ■

The above proof is due to [ABZ]

7.1.2 Elias-Bassalygo Bound

We have seen that Hamming bound implies the asymptotic bound on codes, $R+H(\delta/2) \leq 1$. We now obtain better asymptotic bound using Johnson bound.

Theorem 7.5 (Elias-Bassalygo bound [Bassalygo]) *For all codes, $R + H(\frac{1}{2}(1 - \sqrt{1-2\delta})) \leq 1$.*

Proof: Start with any $(k, n, \delta n)$ code. Let $\tau = \frac{1}{2}(1 - \sqrt{1-2\delta})$. From Johnson bound, a $(k, n, \delta n)$ code is also $(\tau n - 1, n)$ error correcting. Put a ball of radius τn around each of the 2^k codewords. The sum of the volumes of these balls is $2^k Vol(\tau n, n)$. On the other hand, each point in these balls can be in at most n balls. Hence

$$\begin{aligned} 2^k Vol(\tau n, n) &\leq n2^n \\ \Rightarrow k + nH(\frac{\tau n - 1}{n}) &\leq n + \log n \\ \Rightarrow R + H(\frac{\tau n - 1}{n}) &\leq 1 + \frac{\log n}{n} \end{aligned}$$

Therefore, asymptotically $R + H(\tau) \leq 1$. ■

Note that this is a tighter bound than $R + H(\delta/2) \leq 1$ whenever $\delta/2 < \frac{1}{2}(1 - \sqrt{1-2\delta}) \leq \frac{1}{2}$, which is true for all $0 < \delta \leq 1$.

7.2 MacWilliams identities

Lemma 7.6 *If C is an $[n, k, d]_q$ linear code and $x \notin C^\perp$ (where $C^\perp$ is the dual code of C), then for all $a \in \mathbb{F}_q$,*

$$|\{c \in C : \langle c, x\rangle = a\}| = q^{k-1}.$$

*(*Remark*: For $x \in C^\perp$, $\langle c, x\rangle = 0$ for all $c \in C$)*

Proof: Let G be the generator matrix of the code C.

$$\begin{aligned} |\{c \in C : \langle c, x\rangle = a\}| &= |\{y \in \mathbb{F}_q^k : \langle yG, x\rangle = a\}| \\ &= |\{y \in \mathbb{F}_q^k : yGx^\top = a\}| \\ &= |\{y \in \mathbb{F}_q^k : y(Gx^\top) = a\}| \\ &= q^{k-1} \text{ since } x \notin C^\perp \text{ implies } Gx^\top \text{ is a nonzero vector} \end{aligned}$$

■

7.2.1 Notation

We now define a few quantities that will be used throughout the section and prove some basic facts about them.

Definition 7.7 *For an $[n,k,d]_2$ linear code C and $x \in \mathbb{F}_2^n$, define*

1. $A_i = \{c \in C : wt(c) = i\}$, *and*
2. $A_C(x) = \sum_{i=0}^{n} A_i x^i$.

Clearly, for a code with distance d, $A_0 = 1, A_1 = A_2 = \ldots = A_{d-1} = 0$ and $A_d \neq 0$. The rest of the A_i's give even more information about the code and allow us to get better bounds for linear codes.

Definition 7.8 *For $b = b_1 b_2 \ldots b_n, x, y \in \mathbb{F}_2^n$ and a linear code C, define*

1. $W_{b_i}(x,y)$ *equal x if $b_i = 0$ and y if $b_i = 1$,*
2. $W_b(x,y) = \prod_{i=1}^{n} W_{b_i}(x_i, y_i)$, *and*
3. $W_C(x,y) = \sum_{b \in C} W_b(x,y)$.

Proposition 7.9 *For n, b, x and y as in Definition 7.8,*

$$W_b(x+y, x-y) = \sum_{b' \in \mathbb{F}_2^n} (-1)^{\langle b, b' \rangle} W_{b'}(x,y).$$

Proof: Let us first verify the proposition for $n = 1$. By definition, $W_b(x+y, x-y)$ is equal to $x+y$ if $b = 0$ and $x-y$ if $b = 1$. Therefore

$$\begin{aligned} W_b(x+y, x-y) &= (-1)^{\langle b,0 \rangle} x + (-1)^{\langle b,1 \rangle} y \\ &= \sum_{b' \in \mathbb{F}_2} (-1)^{\langle b, b' \rangle} W_{b'}(x,y) \end{aligned}$$

For general n,

$$\begin{aligned} W_b(x+y, x-y) &= \prod_{i=1}^{n} W_{b_i}(x_i + y_i, x_i - y_i) \\ &= \prod_{i=1}^{n} \sum_{b' \in \mathbb{F}_2^n} (-1)^{\langle b_i, b_i' \rangle} W_{b_i}(x_i, y_i) \\ &= \sum_{b' \in \mathbb{F}_2^n} (-1)^{\langle b_i, b_i' \rangle} \prod_{i=1}^{n} W_{b_i'}(x_i, y_i) \\ &= \sum_{b' \in \mathbb{F}_2^n} (-1)^{\langle b_i, b_i' \rangle} W_{b'}(x,y) \end{aligned}$$

■

Proposition 7.10 *For C, x and y as in Definition 7.8,*

$$W_{C^\perp}(x,y) = \frac{1}{|C|} W_C(x+y, x-y).$$

Proof:

$$\begin{aligned} W_C(x+y, x-y) &= \sum_{c\in C} W_c(x+y, x-y) \\ &= \sum_{c\in C}\sum_{b\in\mathbb{F}_2^n} (-1)^{\langle c,b\rangle} W_b(x,y) \\ &= \sum_{b\in\mathbb{F}_2^n} W_b(x,y) \sum_{c\in C} (-1)^{\langle c,b\rangle}. \end{aligned}$$

By Lemma 7.6 it follows that if $b \in C^\perp$, then $\sum_{c\in C}(-1)^{\langle c,b\rangle} = |C|$, while if $b \notin C^\perp$, then $\sum_{c\in C}(-1)^{\langle c,b\rangle} = 0$. Therefore the last sum equals

$$|C| \sum_{b\in C^\perp} W_b(x,y) = |C| W_{C^\perp}(x,y),$$

as needed. ■

Proposition 7.11 *For C, x and y as in Definition 7.8,*

$$x^n A_C(y/x) = W_C(\underbrace{x,\ldots,x}_{n}, \underbrace{y,\ldots,y}_{n}).$$

Proof:

$$\begin{aligned} W_C(\underbrace{x,\ldots,x}_{n}, \underbrace{y,\ldots,y}_{n}) &= \sum_{b\in C} y^{wt(b)} x^{n-wt(b)} \\ &= x^n \sum_{b\in C} (y/x)^{wt(b)} \\ &= x^n \sum_{i=0}^{n} A_i (y/x)^i \\ &= x^n A_C(y/x) \end{aligned}$$

■

7.2.2 MacWilliams Identities

Theorem 7.12 (MacWilliams Identities [mWilliams]) *For a linear code C,*

$$A_{C^\perp}(y) = \frac{1}{|C|}(1+y)^n A_C\left(\frac{1-y}{1+y}\right)$$

Proof:

$$\begin{aligned} A_{C^\perp}(y) &= W_{C^\perp}(\underbrace{1,\ldots,1}_{n}, \underbrace{y,\ldots,y}_{n}) \\ &= \frac{1}{|C|} W_C(1+y,\ldots,1+y,1-y,\ldots,1-y) && \text{(Proposition 7.10)} \\ &= \frac{1}{|C|}(1+y)^n A_C\left(\frac{1-y}{1+y}\right) && \text{(Proposition 7.11)} \end{aligned}$$

■

7.2.3 The linear programming bounds

MacWilliams identities relate the weight distribution of a code and its dual. Suppose we want to find a linear code C with distance d. Let A_i be the weight distribution for the code C and B_i the weight distribution for the dual $C^{\perp}$. Then we get $A_0 = 1, A_2 = \ldots, A_{d-1} = 0$ and $A_i \geq 0$ for all i. Moreover, we know that all of the i it holds that $B_i \geq 0$. We thus get a linear programming problem where we want to maximize $\sum_{i=1}^{n} A_i$ (the number of codewords).

This methods has been used in order to obtain the best known bounds for linear codes [mERRW]. A result by [Delsarte] implies that this bound is valid for all codes.

References

[1] E. Agrell, A. Vardy and K Zegger (200) Upper bounds for constant-weight codes. *IEEE Trans. Inform. Theory* **46**, 2373–2395.

[2] L. A. Bassalygo (1965) New upper bounds for error-correcting codes. (Russian) *Problemy Peredači Informacii* **1 vyp. 4**, 41–44.

[3] P. Delsarte (1973) An algebraic approach to the association schemes of coding theory. *Philips Res. Rep. Suppl.* **No. 10**.

[4] R. J. McEliece, E. R. Rodemich, H. Rumsey and L. R. Welch. (1977) New upper bounds on the rate of a code via the Delsarte-MacWilliams i nequalities. *IEEE Trans. Information Theory* **IT-23** (1977), 157–166.

[5] S. M. Johnson (1963) Improved asymptotic bounds for error-correcting codes. *IEEE Trans. Information Theory* **IT-9**, 198–205.

[6] S. M. Johnson (1962) A new upper bound for error-correcting codes. *IRE Trans.* **IT-8**, 203–207.

[7] J. MacWilliams (1963) A theorem on the distribution of weights in a systematic code. *Bell System Tech. J.* **42**, 79–94.