Introduction to Coding Theory **CMU: Spring 2010**

Notes 5.2: Proof of the first MRRW bound

February 2010

Lecturer: Venkatesan Guruswami *Scribe: Venkat Guruswami & Srivatsan Narayanan*

1 Fourier analytic proof of the first MRRW bound

We develop a proof of the first MRRW (JPL) bound for binary *linear* codes based on a covering argument. Our main theorem (Theorem 1) states that the dual code $C^\perp$ has a small essential covering radius. From this, we conclude that the size of the dual code $|C^\perp|$ is large, and equivalently, the size of the code C is small.

Theorem 1 (Dual codes have a small "essential covering radius") *Let C be a binary linear code of distance at least d. Then,*

$$\left| \bigcup_{z \in C^\perp} B(z, r) \right| \geq \frac{2^n}{n} \tag{1}$$

for $r = \frac{1}{2}n - \sqrt{d(n-d)} + o(n)$.

Remark 2 *We say that a set S has a covering radius at most r if every point in $\{0,1\}^n$ is inside $B(z,r)$ for some $z \in S$. Theorem 1 asserts that the dual code satisfies a relaxed version of the property: for large enough r, at least a $n^{-O(1)}$ fraction of the points are inside $B(z,r)$ for some $z \in C^\perp$. This is sufficient for establishing an asymptotic bound on the rate of the codes, as shown in Corollary 3.*

Before we prove Theorem 1, we observe that it directly implies an asymptotic upper bound on the rate of codes.

Corollary 3 *Let C be a binary linear code with distance at least $d = \delta n$. Then,*

1. $|C| \leq n\mathrm{Vol}(n, r)$
2. $R(C) \leq h\left(\frac{1}{2} - \sqrt{\delta(1-\delta)}\right) + o(1)$

PROOF: The covering condition (Equation 1) gives us that:

$$|C^\perp| \cdot \mathrm{Vol}(n, r) \geq \left| \bigcup_{z \in C^\perp} B(z, r) \right| \geq \frac{2^n}{n}$$

since the volume of each ball $B(z,r)$ is exactly $\mathrm{Vol}(n,d)$. But, the sizes of C and $C^\perp$ are related as follows (see Exercise 1): $|C^\perp| \cdot |C| = 2^n$. Combining the two observations, we directly obtain a bound on the code size:

$$|C| = \frac{2^n}{|C^\perp|} \leq n\mathrm{Vol}(n,r)$$

To obtain the bound on the rate, we need to translate the above bounds to asymptotic notation. Writing $d = \delta n$,

$$r = n\left(\frac{1}{2} - \sqrt{\delta(1-\delta)} + o(1)\right)$$

so that,

$$|C| \leq n\mathrm{Vol}(n,r) = n2^{nh\left(\frac{1}{2} - \sqrt{\delta(1-\delta)} + o(1)\right)}$$

Taking logarithms, and noting that $(\log n)/n = o(1)$ gives the claimed bound.

Exercise 1 *Show that* $|C| \cdot |C^\perp| = 2^n$.

Hint: What is the relation between the dimensions of a code and its dual? How many points lie on a subspace of dimension k? □

We now need to establish Theorem 1. But, at first we introduce some notation. Let A denote the adjacency matrix of the boolean hypercube; that is, for $x, y \in \{0,1\}^n$, $A_{xy} = 1$ if and only if x and y differ in exactly one coordinate. We now extend the concept of maximum eigenvalue to a set $B \subseteq \{0,1\}^n$.

Definition 4 *For* $B \subseteq \{0,1\}^n$, *define its maximum eigenvalue to be*

$$\lambda_B = \max\left\{ \frac{\langle Af, f\rangle}{\langle f, f\rangle} \;\middle|\; f : \{0,1\}^n \to \mathbb{R}, \mathrm{Supp}(f) \subseteq B \right\}$$

Proposition 5 *The maximum eigenvalue of the whole hypercube is* $\lambda_{\{0,1\}^n} = n$.

In order to establish Theorem 1, we observe that $B(z,r) = z + B(0,r)$; that is, $B(z,r)$ is really a translate of $B(0,r)$ by z. Now, we break the proof down into two parts. First, we show that for *any* $B \subseteq \{0,1\}^n$ with sufficiently large maximum eigenvalue, $\bigcup_{z \in C^\perp}(z+B)$ covers a significant (*i.e.,* $n^{-O(1)}$) fraction of the whole space. We then show that $B(0,r)$ achieves the required maximum eigenvalue, even when r is not too large.

Theorem 6 *Let* C *be a binary linear code with block length* n *and distance* d. *Suppose* $B \subseteq \{0,1\}^n$ *has maximum eigenvalue* $\lambda_B \geq n - 2d + 1$. *Then,*

$$\left| \bigcup_{z \in C^\perp} (z + B) \right| \geq \frac{2^n}{n}$$

Theorem 7 (Hamming balls have a large maximum eigenvalue)

$$\lambda_{B(0,r)} \geq 2\sqrt{r(n-r)} - o(n)$$

Now, our main theorem follows as a corollary of Theorems 6 and 7.

PROOF:(**For Theorem 1**) We just need to pick an r large enough such that

$$\lambda_{B(0,r)} \geq n - 2d + 1$$

This is satisfied provided:

$$2\sqrt{r(n-r)} - o(n) \geq n - 2d + 1$$

Neglecting $o(n)$ terms, we get:

$$2\sqrt{r(n-r)} \geq n - 2d$$

which is true for:

$$r = \frac{1}{2}n - \sqrt{d(n-d)} + o(n)$$

□

Remark 8 *We cannot hope to improve the above bound simply by employing a "better" choice for B. Hamming balls are so-called "Faber-Krahn" minimizers for the Hamming cube. That is, of all sets B with a given volume, Hamming balls have almost optimum value of the maximum eigenvalue: if $|B| = \mathrm{Vol}(n,r)$, then $\lambda_B \leq (1+o(1))\lambda_{B(0,r)}$.*

2 Proof of Theorem 5

We recall the following facts on Fourier transforms from the first installment of these notes. For $f : \{0,1\}^n \to \mathbb{R}$,

$$\begin{aligned} \hat{f}(\alpha) &= \frac{1}{2^n}\sum_x f(x)(-1)^{\alpha\cdot x} \\ f(x) &= \sum_\alpha \hat{f}(\alpha)(-1)^{\alpha\cdot x} \end{aligned}$$

The Parseval identity says that for $f, g : \{0,1\}^n \to \mathbb{R}$,

$$\mathbb{E}_x\left[f(x)g(x)\right] = \sum_\alpha \hat{f}(\alpha)\hat{g}(\alpha)\ .$$

Also, we have a natural correspondence between the Fourier transform of a code and its dual:

$$\widehat{1_C}(\alpha) = \frac{|C|}{2^n}1_{C^\perp}(\alpha) \tag{2}$$

$$\widehat{1_{C^\perp}}(\alpha) = \frac{|C^\perp|}{2^n}1_C(\alpha)$$

The following lemma is a direct consequence of Equation 2.

Lemma 9 *Let C be a binary linear code with distance at least d. Suppose that $\alpha \in \{0,1\}^n$ such that $\alpha \neq 0$ and $\mathrm{wt}(\alpha) < d$. Then,*

$$\widehat{1_{C^\perp}}(\alpha) = 0.$$

We are now equipped to prove the required claim. Let $f_B : \{0,1\}^n \to \mathbb{R}$ be a function with $\mathrm{Supp}(f_B) \subseteq B$, such that

$$\langle Af_B, f_B \rangle = \lambda_B \langle f_B, f_B \rangle$$

(In other words, f_B is a function that maximizes $\langle Af, f\rangle / \langle f, f\rangle$.) Since the set B is understood from the context, for notational ease, we simply set $\lambda = \lambda_B$ and $f = f_B$.

The following exercises make many simplifying observations.

Exercise 2 *For $g : \{0,1\}^n \to \mathbb{R}$, show that $Ag(x) = \sum_{i=1}^n g(x + e_i)$.*

Exercise 3 *Prove the following statements.*

1. *f is nonnegative: for every x, $f(x) \geq 0$.*
2. *For all $x \in B$, we have $Af(x) = \lambda f(x)$.*
3. *For all x, we have $Af(x) \geq \lambda f(x)$. Equivalently, $Af \geq \lambda f$.*

Hint for part (3): *For $x \in B$, equality holds. On the other hand, for $x \notin B$, $f(x) = 0$, and hence the inequality is trivially satisfied. (Note that all the terms in the left hand side are nonnegative.)*

For $z \in \{0,1\}^n$, define $f_z : \{0,1\}^n \to \mathbb{R}$ by $f_z(x) = f(z + x)$. Finally, define $F_B : \{0,1\}^n \to \mathbb{R}$ by:

$$F_B = \frac{1}{2^n} \sum_{z \in C^\perp} f_z$$

Again, for convenience, we set F to be F_B. Note that $\mathrm{Supp}(F) \subseteq \bigcup_{z \in C^\perp} (z + B)$.

We will sketch the high-level idea of the proof. In order to lower bound the $\bigcup_{z \in C^\perp}(z + B)$, we show that the F must have a large support. We establish this by showing that $\mathbb{E}[F]^2$ is comparable to $\mathbb{E}\left[F^2\right]$.

For this we will focus on the quantity $\langle AF, F\rangle$ and establish lower and upper bounds on it.

Lemma 10 (Lower bound on $\langle AF, F\rangle$)

$$\langle AF, F\rangle \geq \lambda \mathbb{E}\left[F^2\right]$$

Lemma 11 (Upper bound on $\langle AF, F\rangle$)

$$\langle AF, F\rangle \leq n\mathbb{E}[F]^2 + (n - 2d)\mathbb{E}\left[F^2\right]$$

Before proving the above lemmata, we will show how to obtain Theorem 6 using them.

Corollary 12 (F has a large support)

$$\left|\bigcup_{z\in C^\perp}(z+B)\right| \geq \frac{\lambda-(n-2d)}{n}2^n$$

PROOF: The lower and upper bounds (Lemmata 10 and 11) together imply

$$\lambda\mathbb{E}\left[F^2\right] \leq n\mathbb{E}\left[F\right]^2 + (n-2d)\mathbb{E}\left[F^2\right],$$

giving

$$\mathbb{E}\left[F\right]^2 \geq \frac{\lambda-(n-2d)}{n}\mathbb{E}\left[F^2\right] \tag{3}$$

On the other hand, note that F is supported on $S = \bigcup_{z\in C^\perp}(z+B)$. Therefore,

$$\mathbb{E}\left[F\right]^2 = \left(\frac{1}{2^n}\sum_{x\in S}F(x)\right)^2 = \langle 1_S, F\rangle^2 \leq \mathbb{E}\left[1_S^2\right]\mathbb{E}\left[F^2\right] = \frac{|S|}{2^n}\mathbb{E}\left[F^2\right], \tag{4}$$

using the Cauchy-Schwartz inequality.

Finally, Equations (3) and (4) together give the required bound:

$$|S| \geq \frac{\lambda-(n-2d)}{n}2^n$$

□

For $\lambda \geq n-2d+1$ (as assumed in Theorem 6), this bound simplifies to

$$\left|\bigcup_{z\in C^\perp}(z+B)\right| \geq \frac{2^n}{n}$$

2.1 Lower bound on $\langle AF, F\rangle$

In this section, we prove Lemma 10. The proof is based on the spectral property of B. Fix an $x \in \{0,1\}^n$. Then,

$$\begin{aligned} AF(x) &= \sum_{i=1}^{n}F(x+e_i) = \frac{1}{2^n}\sum_i\sum_{z\in C^\perp}f_z(x+e_i) = \frac{1}{2^n}\sum_{z\in C^\perp}\sum_i f(x+z+e_i) \\ &= \frac{1}{2^n}\sum_{z\in C^\perp}Af(x+z) \geq \frac{1}{2^n}\sum_{z\in C^\perp}\lambda f(x+z) = \frac{\lambda}{2^n}\sum_{z\in C^\perp}f_z(x) = \lambda F(x), \end{aligned}$$

Therefore,

$$\langle AF, F\rangle = \mathbb{E}_x\left[AF(x)F(x)\right] \geq \lambda\mathbb{E}_x\left[(F(x))^2\right] = \lambda\mathbb{E}\left[F^2\right]$$

which gives the claim.

2.2 Upper bound on $\langle AF, F\rangle$

In this section, we prove Lemma 11. The proof is based on the properties of the Fourier transform of F.

The following simple result is a crucial ingredient in calculating $\hat{F}$.

Lemma 13 *For $g : \{0,1\}^n \to \mathbb{R}$ and $z \in \{0,1\}^n$, define $g_z : \{0,1\}^n \to \mathbb{R}$ by $g_z(x) = g(x+z)$. Then,*

$$\widehat{g_z}(\alpha) = (-1)^{\alpha\cdot z}\hat{g}(\alpha)$$

PROOF: We have

$$\widehat{g_z}(\alpha) = \frac{1}{2^n}\sum_x g_z(x)(-1)^{\alpha\cdot x} = \frac{1}{2^n}\sum_x g(x+z)(-1)^{\alpha\cdot x} = \frac{1}{2^n}\sum_y g(y)(-1)^{\alpha\cdot(y+z)}$$

where we make the substitution $y = x + z$. Therefore,

$$\widehat{g_z}(\alpha) = (-1)^{\alpha\cdot z}\frac{1}{2^n}\sum_y g(y)(-1)^{\alpha\cdot y} = (-1)^{\alpha\cdot z}\hat{g}(\alpha)$$

□

Lemma 14 (Fourier transform of F)

$$\hat{F}(\alpha) = \widehat{1_{C^\perp}}(\alpha)\hat{f}(\alpha) = \frac{|C^\perp|}{2^n}1_C(\alpha)\hat{f}(\alpha)\ .$$

PROOF: From Lemma 13, we know that $\widehat{f_z}(\alpha) = (-1)^{\alpha\cdot z}\hat{f}(\alpha)$. Therefore,

$$\begin{aligned}\hat{F}(\alpha) &= \frac{1}{2^n}\sum_{z\in C^\perp}\widehat{f_z}(\alpha) = \frac{1}{2^n}\sum_{z\in C^\perp}(-1)^{\alpha\cdot z}\hat{f}(\alpha) = \frac{1}{2^n}\hat{f}(\alpha)\sum_{z\in C^\perp}(-1)^{\alpha\cdot z}\\ &= \hat{f}(\alpha)\mathbb{E}_z[1_{C^\perp}(z)(-1)^{\alpha\cdot z}] = \hat{f}(\alpha)\widehat{1_{C^\perp}}(\alpha)\ .\end{aligned}$$

The full claim follows since $\widehat{1_{C^\perp}} = \frac{|C^\perp|}{2^n}1_C$. □

Remark 15 *Lemma 14 can be directly obtained as follows. Note that*

$$F(x) = \sum_{z\in C^\perp} f(x+z) = \sum_{z\in\{0,1\}^n} 1_{C^\perp}(z)f(x+z) = \left(1_{C^\perp} * f\right)(x)$$

where $*$ represents the 'convolution' operation. It is a well known property that the Fourier transform of a convolution is the product of the Fourier transforms. Formally, for $f, g : \{0,1\}^n \to \mathbb{R}$, we have $\widehat{f * g} = \hat{f}\cdot\hat{g}$. This establishes the claim.

Corollary 16 *Let C be a binary linear code of distance at least d. Suppose $\alpha \in \{0,1\}^n$ is such that $\alpha \neq \mathbf{0}$ and $\mathrm{wt}(\alpha) < d$. Then, $\hat{F}(\alpha) = 0$.*

PROOF: Since $\mathrm{wt}(\alpha) < d$, it follows that $\alpha \notin C$. Therefore, from Lemma 14, $\hat{F}(\alpha) = 0$. □

Lemma 17 (Fourier transform of AF) *For $g : \{0,1\}^n \to \mathbb{R}$,*

$$\widehat{Ag}(\alpha) = \hat{g}(\alpha)(n - 2\mathrm{wt}(\alpha))$$

PROOF: We know that

$$Ag = \sum_i g_{e_i}$$

Therefore,

$$\widehat{Ag}(\alpha) = \sum_i \widehat{g_{e_i}}(\alpha) = \hat{g}(\alpha) \sum_i (-1)^{\alpha \cdot e_i} = \hat{g}(\alpha) \sum_i (-1)^{\alpha_i}$$

where $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)^T$. It is easy to check that for $\alpha_i \in \{0,1\}$, $(-1)^{\alpha_i} = 1 - 2\alpha_i$. Plugging this in the previous equation, we get

$$\widehat{Ag}(\alpha) = \hat{g}(\alpha) \sum_i (1 - 2\alpha_i) = \hat{g}(\alpha)(n - 2\mathrm{wt}(\alpha)).$$

□

With these claims in place, we can establish an upper bound on $\langle AF, F \rangle$.

$$\begin{aligned}
\langle AF, F \rangle &= \sum_\alpha \widehat{AF}(\alpha)\hat{F}(\alpha) \\
&= \sum_\alpha \hat{F}(\alpha)^2(n - 2\mathrm{wt}(\alpha)) \\
&= n\hat{F}(0)^2 + \sum_{\alpha : \mathrm{wt}(\alpha) \geq d} \hat{F}(\alpha)^2(n - 2\mathrm{wt}(\alpha))
\end{aligned}$$

using Corollary 16. We now complete the upper bound.

$$\begin{aligned}
\langle AF, F \rangle &\leq n\hat{F}(0)^2 + (n - 2d) \sum_{\alpha : \mathrm{wt}(\alpha) \geq d} \hat{F}(\alpha)^2 \\
&\leq n\hat{F}(0)^2 + (n - 2d) \sum_\alpha \hat{F}(\alpha)^2 \\
&= n\mathbb{E}[F]^2 + (n - 2d)\mathbb{E}\left[F^2\right],
\end{aligned}$$

using $\hat{F}(0) = \mathbb{E}[F]$.

3 Lower bound on the maximum eigenvalue of Hamming balls

We are interested in lower bounding the maximum eigenvalue of the Hamming ball $B(0,r)$, where $r = \gamma n$ for $\gamma < 1/2$. We will restrict ourselves to an $f : \{0,1\}^n \to \mathbb{R}$, such that $f(x)$ depends on *only* on the weight of x, and has support $\{x : r - M \leq \mathrm{wt}(x) \leq r\} \subseteq B(0,r)$, for some $M = o(n)$. (For instance, we could choose $M = n^{3/4}$.) Define f as follows:

$$f(x) = \begin{cases} \frac{1}{\sqrt{\binom{n}{i}}}, & \text{if } \mathrm{wt}(x) = i \in [r-M, r], \\ 0, & \text{otherwise.} \end{cases}$$

For convenience, we will denote by $f(i)$ the evaluation of f at any x with weight i. Let us now compute $\langle f, f\rangle$ and $\langle Af, f\rangle$ respectively.

$$2^n\langle f, f\rangle = \sum_{i=r-M}^{r} \sum_{x:\mathrm{wt}(x)=i} f(x)^2 = \sum_{i=r-M}^{r} \binom{n}{i} f(i)^2 = M + 1 \leq M(1 + o(1)) \tag{5}$$

On the other hand,

$$2^n\langle Af, f\rangle = \sum_x Af(x)f(x) = \sum_i \sum_{x:\mathrm{wt}(x)=i} Af(x)f(x)$$

Fix an $x \in \{0,1\}^n$ of weight i. Therefore, of the n neighbors of x, i have a weight $i-1$, and the remaining $n-i$ have a weight $i+1$. Therefore,

$$Af(x)f(x) = f(x)\sum_{j=1}^{n} f(x + e_j) = f(i)\left(if(i-1) + (n-i)f(i+1)\right)$$

Therefore,

$$\begin{aligned} 2^n\langle Af, f\rangle &= \sum_{i=1}^{n} \left(i\binom{n}{i} f(i)f(i-1) + (n-i)\binom{n}{i} f(i)f(i+1) \right) \\ &= \sum_{i=r-M+1}^{r} i\sqrt{\frac{\binom{n}{i}}{\binom{n}{i-1}}} + \sum_{i=r-M}^{r-1} (n-i)\sqrt{\frac{\binom{n}{i}}{\binom{n}{i+1}}} \\ &= \sum_{i=r-M+1}^{r} \sqrt{i(n-i+1)} + \sum_{i=r-M}^{r-1} \sqrt{(n-i)(i+1)} \\ &= 2\sum_{i=r-M+1}^{r} \sqrt{i(n-i+1)} \end{aligned}$$

For the values of r and i we are interested in, it is easy to see that

$$i(n-i+1) \geq (r-M+1)(n-r+M) \geq r(n-r) - o(n^2).$$

Hence,

$$2^n\langle Af, f\rangle \geq M\left(2\sqrt{r(n-r)} - o(n)\right) \tag{6}$$

Combining Equations (6) and (5), we get

$$\lambda_{B(0,r)} \geq \frac{\langle Af, f\rangle}{\langle f, f\rangle} \geq \frac{2\sqrt{r(n-r)} - o(n)}{1 + o(1)} = 2\sqrt{r(n-r)} - o(n)$$

For $r = \gamma n$, this gives the bound

$$\lambda_{B(0,r)} \geq 2n\sqrt{\gamma(1-\gamma)} - o(n)$$

4 Some remarks and the second MRRW bound

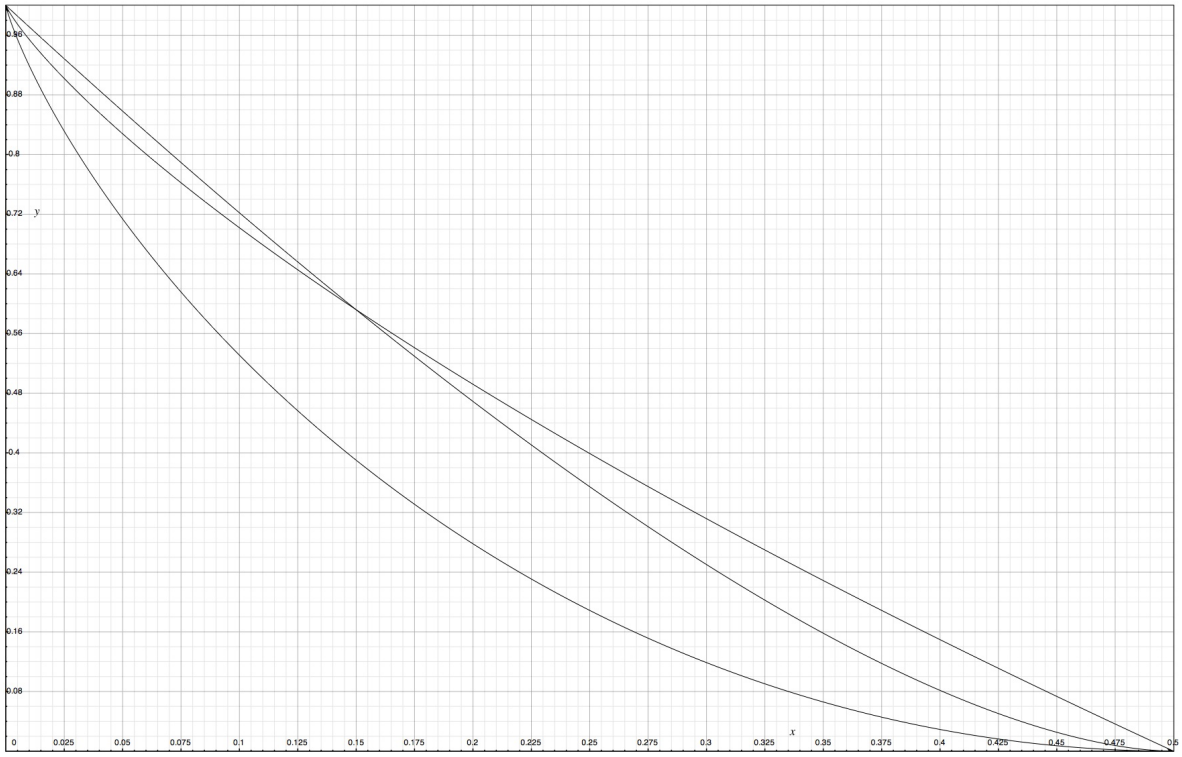
Though we proved the bound only for linear codes, the bound holds also for general codes, and in fact can be proved within the same Fourier analytic framework; see the final section of the paper by Navon and Samorodnitsky for the details.

The first MRRW bound can also be generalized to larger alphabets, giving the following statement.

Theorem 18 (First MRRW bound for larger alphabets) *The rate of a q-ary code of relative distance δ, $0 < \delta < 1 - 1/q$, is at most*

$$h_q\Big(\frac{1}{q}\big(q - 1 - (q-2)\delta - 2\sqrt{(q-1)\delta(1-\delta)}\big)\Big) + o(1) \ .$$

Below is a plot of the best bounds on the best possible rate $R(\delta)$ as a function of relative distance δ we have seen so far for binary codes: the Gilbert-Varshamov lower bound $R(\delta) \geq 1 - h(\delta)$, the Elias-Bassalygo bound $R(\delta) \leq 1 - h(\frac{1-\sqrt{1-2\delta}}{2})$, and the first MRRW bound $R(\delta) \leq h(\frac{1}{2} - \sqrt{\delta(1-\delta)})$.



Note that the first MRRW bound is weaker than the Elias-Bassalygo bound for δ smaller than about 0.15 (in fact it is even weaker than the Hamming bound for $\delta \lesssim 0.11$). There is a strenghtening of the bound, called the second MRRW bound, which uses the inequality

$$A(n,d) \leq \frac{2^n}{\binom{n}{w}} A(n,d,w)$$

together with an upper bound on the size $A(n,d,w)$ of constant-weight codes via Delsarte's linear programming approach applied to the "Johnson" association scheme. We state the bound here without proof. This bound beats the Elias-Bassalygo bound for the entire range $\delta \in [0,1/2]$. The bound coincides with the first MRRW bound for $\delta > 0.273$. The second MRRW bound gives the best upper bound on $R(\delta)$ for the entire range of δ and has not been improved upon in over three decades!

Theorem 19 (Second MRRW bound for binary codes) *Let $0 < \delta < 1/2$. The largest rate of a binary code of relative distance δ is at most* $\mathrm{MRRW}^{(2)}(\delta) + o(1)$ *where*

$$\mathrm{MRRW}^{(2)}(\delta) = \min_{\delta/2 \leq \xi \leq 1/2} \left\{ 1 - h(\xi) + R_{\mathrm{cw}}(\xi, \delta) \right\} \tag{7}$$

with

$$R_{\mathrm{cw}}(\xi,\delta) = \begin{cases} h\left(\frac{1}{2}\left(1 - \sqrt{1 - \left(\sqrt{4\xi(1-\xi) - 2\delta + \delta^2} - \delta\right)^2}\right)\right) & \text{if} \quad \delta \leq 2\xi(1-\xi) \\ 0 & \text{otherwise.} \end{cases}$$

The above bound encompasses the Hamming, Elias-Bassalygo, and first MRRW bounds.

Exercise 4 *1. Verify that the choice $\xi = 1/2$ in the the minimization in (7) yields the first MRRW bound.*

2. Verify that the choice $\xi = \delta/2$ in the minimization in (7) yields the Hamming bound.

3. Verify that picking ξ so that $2\xi(1-\xi) = \delta$ in the minimization in (7) yields the Elias-Bassalygo bound.

As far as I am aware the second MRRW bound only applies for binary codes and has not been extended to larger alphabets.