Error Correcting Codes: Combinatorics, Algorithms and Applications (Fall 2007)

Lecture 21: Reed-Muller Codes

October 15, 2007

Lecturer: Atri Rudra *Scribe: Michel Kulhandjian*

Recall the question we raised earlier. Is there an explicit codes with $R > 0$ and $\delta > 0$ that have efficient unique decoding up to $\frac{\delta}{2}$ for $(q = 2)$? For RS codes we have optimal trade off between R and δ and efficient decoding algorithm exists. However, RS codes have the unfortunate property that q must be large and hence, we are interested in the above question for small alphabet.

1 Reed-Muller Codes

We now extend the definition of Reed-Solomon codes, to multivariate polynomials with v number of variables. These codes were termed Reed-Muller Codes after D. E. Muller [1] who discovered the codes and I. S. Reed who gave a decoding procedure [2]. The codes that are presented here are generalized to a range of parameters that were not covered in original work which mostly had focused on codes over $\mathbb{F}_2$. The way Reed-Muller codes are described here, they will be strict generalization of Reed-Solomon codes, although Reed-Solomon was discovered much later! Reed-Muller codes can have smaller q unlike Reed-Solomon, in fact there exist efficient decoding algorithms up to $\frac{\delta}{2}$ and $q = 2$. The trade off between R and δ however is not that satisfactory as Reed-Muller codes can not have $R > 0$, $\delta > 0$. Reed-Muller codes have found many uses in complexity theory and codeword testing.

Definition 1.1. *A code (family) is called asymptotically good if its rate satisfies $R > 0$ and relative distance satisfies $\delta > 0$.*

Recall that Reed-Solomon codes were defined using univariate polynomials. Now we define Reed-Muller Codes using multivariate polynomials.

Definition 1.2 (Reed-Muller code)**.** *Let $\mathbb{F}_q[x_1, x_2, ..., x_v]$ denote the $\mathbb{F}_q^v$-space of multivariate polynomials where all the coefficients are from $\mathbb{F}_q$. We define an encoding function for Reed-Muller code as $RM_q(t, v)$ as follows. A message symbol $\mathbf{m}$ with coefficients $\langle m_{i_1,i_2,...,i_v}\rangle \in \mathbb{F}_q$ is mapped to a v-variate polynomial $\mathbb{F}_q[x_1, x_2, ..., x_v]$ of degree $\leq t$ over $\mathbb{F}_q^v$ as follows.*

$$\mathbf{m} \mapsto p_{\mathbf{m}}(x_1, x_2, ..., x_v),$$

where

$$p_{\mathbf{m}}(x_1, x_2, ..., x_v) = \sum_{\substack{i_1,i_2,...,i_v \in \mathbb{Z}_{\geq 0} \\ i_1+i_2+...+i_v \leq t}} m_{i_1,i_2,...,i_v} \prod_{j=1}^{v} x_j^{i_j} \tag{1}$$

with $\mathbf{m} = \langle m_{i_1,i_2,...,i_v} \rangle_{\substack{(i_1,i_2,...,i_v) \\ i_1+i_2+...+i_v \leq t}}$, $0 \leq i_j \leq q-1$, *and* $t \leq v(q-1)$. *The encoding of* $\mathbf{m}$ *is the evaluation of* $p_{\mathbf{m}}(x)$ *at all the vectors in* $\mathbb{F}_q^v$*:*

$$RM(\mathbf{m}) = \langle p_{\mathbf{m}}(\alpha_1, \alpha_2, ..., \alpha_v) \rangle_{(\alpha_1,\alpha_2,...,\alpha_v) \in \mathbb{F}_q^v}$$

Remark 1.3. *It is easy to check that* $RM_q(t,1)$ *codes are equivalent to* $[q, t+1]_q$ *Reed-Solomon codes.*

It is obvious that the block length of the $RM_q(t,v)$ code is $n = q^v$. We will first look at the case where $t < q$ which is useful in applications of complexity theory. The message length k equals the number of v-long sequences of integers that sums to at most t and it turns out to be:

$$\begin{aligned} k &= |\{(i_1, i_2, ..., i_v) | i_1 + i_2 + ... + i_v \leq t\}| \\ &= \binom{v+t}{v} \end{aligned}$$

when $t \leq q-1$. Note that when $t \geq q$ it gets complicated because of the identity $x^q = x$, $\forall x \in \mathbb{F}_q$.

Remark 1.4. *The way the mapping is defined in (1),* $RM_q(t,v)$ *codes are linear codes. The proof is similar to the one we used to prove that* RS *codes are linear and is left as an exercise.*

Proposition 1.5. *The distance of* $RM_q(t,v)$ *code is* $\geq \left(1 - \frac{t}{q}\right) q^v$. *Hence for* $t < q$, $RM_q(t,v)$ *is an* $\left[q^v, \binom{v+t}{v}, \left(1 - \frac{t}{q}\right) q^v\right]_q$ *code.*

Let us look at some instantiations of the parameters.

1. When $t = 1$, $RM_q(t,v)$ is equivalent to RS for $(q = n)$.

2. When $v = t = \frac{q}{2}$ the code length becomes $n = q^v = \sqrt{q^q}$. We can show that $q = \Theta\left(\frac{\log n}{\log \log n}\right)$. For this set of parameters we show the asymptotic behavior of the rate of the code.

$$\begin{aligned} k &= \binom{v+t}{v} \\ &= \binom{q}{\frac{q}{2}} \\ &\leq (2e)^{\frac{q}{2}} \\ &= 2^{\Theta(q)} \end{aligned} \tag{2}$$

From (2) note that $R \to 0$ as $n \to \infty$ for $\delta = \frac{1}{2}$.

The proof of Proposition 1.5 follows immediately from the following lemma as $RM_q(t,v)$ is linear.

Lemma 1.6 (Schwartz-Zippel)**.** *Any non-zero v-variate polynomial in $\mathbb{F}_q[x_1, x_2, ..., x_v]$ of degree almost t has $\leq tq^{v-1}$ roots.*

Proof. We will prove the proposition by induction on the number of variables. For $v = 1$ it states the familiar result that a non-zero univariate polynomial has at most as many roots as its degree. Now assume that the induction hypothesis is true for the multivariate polynomial with up to $v - 1$ variables, for $v > 1$. Let $P(x_1, \ldots, x_v)$ be a degree t polynomial. Decompose the polynomial as follows:

$$P(x_1, x_2, ..., x_v) = \sum_{i=0}^{t_1} R_i(x_2, ..., x_v)x_1^i \tag{3}$$

W.l.o.g. we can assume that $t_1 \geq 1$ (as otherwise we have $v - 1$ variables and by induction we will have $\leq tq^{v-2} < tq^{v-1}$ roots) and t_1 may be strictly smaller than t. We consider the following two cases for a possible root $(\alpha_1, \ldots, \alpha_v)$:

- Case 1: $R_{t_1}(\alpha_2, ..., \alpha_v) = 0$ where $(\alpha_2, ..., \alpha_v) \in \mathbb{F}_q^{v-1}$. The number of such roots of R_{t_1} by the induction hypothesis is:

  $$\begin{aligned} |\{(\alpha_2, ..., \alpha_v) \mid R_t(\alpha_2, ..., \alpha_v) = 0\}| &\leq deg(R_{t_1})q^{v-2} \\ &\leq (t - t_1)q^{v-2}. \end{aligned}$$

  Thus, the number of roots of $P(x_1, \alpha_2..., \alpha_v)$ given $R_{t_1}(\alpha_2, ..., \alpha_v) = 0$ is

  $$\begin{aligned} |\{(\alpha_1, \alpha_2, \ldots, \alpha_v) \mid R_{t_1}(\alpha_2, ..., \alpha_v) = 0\}| &\leq (t - t_1)q^{v-2}q \\ &= (t - t_1)q^{v-1}, \end{aligned}$$

  where the inequality follows from the fact that any tuple $(\alpha_2, \ldots, \alpha_v)$ can be extended to at most q vectors in $\mathbb{F}_q^n$.

- Case 2: $R_{t_1}(\alpha_2, ..., \alpha_v) \neq 0$. Fix $(\alpha_2^\star, \ldots, \alpha_v^\star)$ such that $R_{t_1}(\alpha_2^\star, \ldots, \alpha_v^\star) \neq 0$. Since $P(x_1, \alpha_2^\star, ..., \alpha_v^\star)$ is a univariate polynomial of degree $\leq t_1$, by induction:

  $$|\{(\alpha_1, \alpha_2^\star, \ldots, \alpha_v^\star) \mid P(\alpha_1, \alpha_2^\star, \ldots, \alpha_v^\star) = 0\}| \leq t_1.$$

  Then the total number of roots $(\alpha_1, \alpha_2, ..., \alpha_v)$ such that $R_{t_1}(\alpha_2, ..., \alpha_v) \neq 0$ is

  $$|\{(\alpha_1, \alpha_2, \ldots, \alpha_v) \mid P(\alpha_1, \alpha_2, \ldots, \alpha_v) = 0 \wedge R_{t_1}(\alpha_2, ..., \alpha_v) \neq 0\}| \leq t_1 q^{v-1},$$

  where the inequality follows from the fact that there can be at most q^{v-1} distinct tuples $(\alpha_2^\star, \ldots, \alpha_v^\star)$.

Now combining two cases we get the total number of roots as follows:

$$
\begin{aligned}
|\{(\alpha_1, \alpha_2,\ & \dots, \alpha_v) \mid P(\alpha_1, \alpha_2, \dots, \alpha_v) = 0\}| \\
&= |\{(\alpha_1, \alpha_2, \dots, \alpha_v) \mid P(\alpha_1, \dots, \alpha_v) = 0 \wedge R_{t_1}(\alpha_2, ..., \alpha_v) = 0\}| \\
&\quad + |\{(\alpha_1, \alpha_2, \dots, \alpha_v) \mid P(\alpha_1, \alpha_2, \dots, \alpha_v) = 0 \wedge R_{t_1}(\alpha_2, ..., \alpha_v) \neq 0\}| \\
&\leq |\{(\alpha_1, \alpha_2, \dots, \alpha_v) \mid R_{t_1}(\alpha_2, ..., \alpha_v) = 0\}| \\
&\quad + |\{(\alpha_1, \alpha_2, \dots, \alpha_v) \mid P(\alpha_1, \alpha_2, \dots, \alpha_v) = 0 \wedge R_{t_1}(\alpha_2, ..., \alpha_v) \neq 0\}| \\
&\leq (t - t_1)q^{v-1} + t_1 q^{v-1} \\
&= tq^{v-1},
\end{aligned}
$$

as desired. □

From the Schwartz-Zippel lemma, the distance of $RM_q(t, v)$ is at least the total number of non-zero values of $P(x_1, x_2, ..., x_v)$ which is equal to the number of possible $(x_1, x_2, \dots, x_v)$ minus the total number of roots of $P(x_1, x_2, ..., x_v)$. Therefore,

$$d \geq \left(1 - \frac{t}{q}\right) q^v$$

and it is a tight bound. So $RM_q(t, v)$ is an

$$\left[q^v, \binom{v+t}{v}, \left(1 - \frac{t}{q}\right) q^v\right]_q$$

code. Hence, this proves the Proposition 1.5.

If we take Reed-Muller code with parameters $q = 2$ and $t > 2$ we get $RM_2(t, v)$ code. The polynomial P defined by (3) of degree $\leq t$ turns out to be,

$$P(x_1, x_2, \dots, x_v) = \sum_{\substack{S \subseteq [v] \\ |S| \leq t}} m_s \prod_{i \in S} x_i,$$

That is P is a multi linear polynomial where $m_s \in \mathbb{F}_2$ and $\mathbf{m} = \langle m_s \rangle_{\substack{S \subseteq [v] \\ |S| \leq t}}$. We have the following result:

Claim 1.7. *$RM_2(t, v)$ is a* $\left[2^v, \sum_{i=0}^{t} \binom{v}{i}, 2^{m-v}\right]_2$

This result will be proved in the next lecture.

References

[1] D. E. Muller. Application of boolean algebra to switching circuit design and to error detection. *IEEE Transactions on Computers*, 3:6–12, 1954.

[2] Irving S. Reed. A class of multiple-error-correcting codes and the decoding scheme. *IEEE Transactions on Information Theory*, 4:38–49, 1954.