Lecture 5: BCH Codes

Error-Correcting Codes (Spring 2016)
Rutgers University
Swastik Kopparty
Scribes: Brandon Butch and Sijian Tang

1 Introduction

Last time we saw codes with constant $R, \delta > 0$ as $n \to \infty$. Today's focus will be codes with constant distance d (not relative distance) that meet the volume packing bound. Such codes are called **BCH Codes**.

1.1 Essentials of finite fields

BCH codes take advantage of certain properties of finite fields (that may not hold true for fields in general).

Fact 1. *Let $\mathbb{F}_{2^m}$ denote the finite field of 2^m elements. The following hold*

1. *$\mathbb{F}_{2^m}$ is a vector space of dimension m over $\mathbb{F}_2$*
2. *$\mathbb{F}_{2^m}$ has characteristic 2 ($\forall x \in \mathbb{F}_{2^m} : 2x = x + x = 0$)*
3. *$\forall x, y \in \mathbb{F}_{2^m} : (x+y)^2 = x^2 + 2xy + y^2 = x^2 + y^2$ since $2xy = 0$*

Example 2. *For $m = 2$ we have $\mathbb{F}_{2^2} = \{0, 1, \alpha, \alpha + 1\}$ with field operations summarized in the following tables*

$+$	0	1	α	$\alpha+1$
0	0	1	α	$\alpha+1$
1	1	0	$\alpha+1$	α
α	α	$\alpha+1$	0	1
$\alpha+1$	$\alpha+1$	α	1	0

Addition for $\mathbb{F}_{2^2}$

$\times$	0	1	α	$\alpha+1$
0	0	0	0	0
1	0	1	α	$\alpha+1$
α	0	α	$\alpha+1$	1
$\alpha+1$	0	$\alpha+1$	1	α

Multiplication for $\mathbb{F}_{2^2}$

Let $\alpha_1, \dots, \alpha_m$ be a basis for $\mathbb{F}_{2^m}/\mathbb{F}_2$. Then every element of $\mathbb{F}_{2^m}$ can be written as

$$\sum_{i=1}^{m} c_i \alpha_i$$

where all $c_i \in \mathbb{F}_2$. We will represent elements using the following function

$$\phi : \mathbb{F}_{2^m} \to \mathbb{F}_{2^m}$$
$$\phi(\beta) = (c_1 c_2 \cdots c_m)$$

such that $\sum c_i \alpha_i = \beta$. $\phi(\beta)$ can be viewed as the concatenation of the c_i's. This is nothing more than a representation of the elements of $\mathbb{F}_{2^m}$ as m bit strings. (Note that there are 2^m elements of $\mathbb{F}_{2^m}$, the same as the number of binary strings of length m.) Each $\phi(\beta)$ can be thought of as a bit string that corresponds uniquely with an element of $\mathbb{F}_{2^m}$, but the operations (particularly multiplication) would need to be redefined more carefully if one wanted to perform them directly on the bit string representations.

Example 3. *Using $\{1, \alpha\}$ as a basis for $\mathbb{F}_{2^2}$, we can express each element as $c_1 1 + c_2 \alpha$ where $c_1, c_2 \in \{0, 1\}$. This gives us the following representations*

$$\phi(0) = 00$$
$$\phi(1) = 10$$
$$\phi(\alpha) = 01$$
$$\phi(\alpha + 1) = 11.$$

1.2 Essentials of Vandermonde matrices

Define a $k \times k$ matrix

$$V = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \gamma_1 & \gamma_2 & \cdots & \gamma_k \\ \gamma_1^2 & \gamma_2^2 & \cdots & \gamma_k^2 \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_1^{k-1} & \gamma_2^{k-1} & \dots & \gamma_k^{k-1} \end{pmatrix}.$$

In general, element γ_j^{i-1} occupies coordinate (i, j). This is called a **Vandermonde matrix**.

Fact 4. *If $\gamma_1, \dots, \gamma_k$ are all distinct, then this matrix is nonsingular.*

Proof. Suppose otherwise. Let u be a nonzero vector such that $Vu^T = 0$. Explicitly, we have $\forall i$

$$u_0 + u_1 \gamma_i + u_2 \gamma_i^2 + \cdots + u_{k-1} \gamma_i^{k-1} = 0$$

which would imply that the polynomial $\sum_{j=0}^{k-1} u_j x^j$ vanishes at k distinct points (namely $\gamma_1, \dots, \gamma_k$). This is a contradiction, since the polynomial has degree $k - 1$, therefore it can have at most $k - 1$ distinct roots. □

2 BCH Codes

2.1 Warm Up

As an introduction to BCH Codes, let's examine the case for $d = 5$. For $\mathbb{F}_{2^m}$ with a fixed basis and ϕ as above, define the party check matrix

$$H = \begin{pmatrix} | & | & & | \\ \phi(\alpha_1) & \phi(\alpha_2) & \cdots & \phi(\alpha_{2^m-1}) \\ | & | & & | \\ \phi(\alpha_1^3) & \phi(\alpha_2^3) & \cdots & \phi(\alpha_{2^m-1}^3) \\ | & | & & | \end{pmatrix}$$

where $\{\alpha_1, \alpha_2, \ldots, \alpha_{2^m-1}\} = \mathbb{F}_{2^m} \setminus \{0\}$. This is a $2m \times 2^m - 1$ matrix. Each column in the upper half (i.e. the upper m rows) consists of the image under ϕ of a *nonzero* element of $\mathbb{F}_{2^m}$; the lower half of the column consists of the image of that element cubed. Note that because of the way we defined ϕ, the upper half will simply be all nonzero bit strings of length m.

Claim 5. *The code described by H has distance at least* 5*.*

Proof. It suffices to show that any 4 columns of H are linearly independent over $\mathbb{F}_2$. Consider 4 arbitrary columns and let $\beta_1, \beta_2, \beta_3, \beta_4$ be the corresponding representations in $\mathbb{F}_{2^m} \setminus \{0\}$. Suppose the columns were dependent with coefficients $e_1, e_2, e_3, e_4 \in \mathbb{F}_2$ such that

$$\sum_{i=1}^{4} e_i \phi(\beta_i) = 0$$
$$\sum_{i=1}^{4} e_i \phi(\beta_i^3) = 0.$$

By the linearity of ϕ, this gives us

$$\phi\left(\sum_{i=1}^{4} e_i \beta_i\right) = 0$$
$$\phi\left(\sum_{i=1}^{4} e_i \beta_i^3\right) = 0$$

which implies

$$\sum_{i=1}^{4} e_i \beta_i = 0 \tag{1}$$
$$\sum_{i=1}^{4} e_i \beta_i^3 = 0. \tag{2}$$

We square (1), using the fact that the field has characteristic 2, to obtain

$$\left(\sum_{i=1}^{4} e_i\beta_i\right)^2 = \sum_{i=1}^{4} e_i\beta_i^2 = 0. \tag{3}$$

Repeat this step, squaring again to obtain

$$\left(\sum_{i=1}^{4} e_i\beta_i^2\right)^2 = \sum_{i=1}^{4} e_i\beta_i^4 = 0. \tag{4}$$

Now we express $(1), (2), (3), (4)$ as a combined system

$$\begin{pmatrix} \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ \beta_1^2 & \beta_2^2 & \beta_3^2 & \beta_4^2 \\ \beta_1^3 & \beta_2^3 & \beta_3^3 & \beta_4^3 \\ \beta_1^4 & \beta_2^4 & \beta_3^4 & \beta_4^4 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

If the matrix of β's is nonsingular (i.e. its determinant is nonzero), then the only way this equation can be satisfied is if

$$\begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

The matrix

$$M = \begin{pmatrix} \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ \beta_1^2 & \beta_2^2 & \beta_3^2 & \beta_4^2 \\ \beta_1^3 & \beta_2^3 & \beta_3^3 & \beta_4^3 \\ \beta_1^4 & \beta_2^4 & \beta_3^4 & \beta_4^4 \end{pmatrix}$$

is *almost* Vandermonde. It is easy to work out that, since all β_i are nonzero and distinct, M is nonsingular. Therefore, it must be the case that $e_1 = e_2 = e_3 = e_4 = 0$, which implies that any 4 columns of H are indeed linearly independent. □

How large is this code? Recall H has dimensions $2m \times 2^m - 1$. Let $n = 2^m - 1$, then

$$|C| \geq \frac{2^n}{2^{2m}} = \frac{2^n}{(n+1)^2}.$$

Recall the volume packing bound for distance 5:

$$|C| \leq \frac{2^n}{B(2)} = \frac{2^n}{(n+1)\binom{n}{2}} = \Theta\left(\frac{2^n}{n^2}\right).$$

2.2 BCH Codes in general

How can we generalize the above construction to larger distances? For some $t \in \mathbb{N}$, we define the parity check matrix

$$H = \begin{pmatrix} | & | & & | \\ \phi(\alpha_1) & \phi(\alpha_2) & \cdots & \phi(\alpha_{2^m-1}) \\ | & | & & | \\ \phi(\alpha_1^3) & \phi(\alpha_2^3) & \cdots & \phi(\alpha_{2^m-1}^3) \\ | & | & & | \\ \phi(\alpha_1^5) & \phi(\alpha_2^5) & \cdots & \phi(\alpha_{2^m-1}^5) \\ | & | & & | \\ \vdots & \vdots & \ddots & \vdots \\ | & | & & | \\ \phi(\alpha_1^{2t-1}) & \phi(\alpha_2^{2t-1}) & \cdots & \phi(\alpha_{2^m-1}^{2t-1}) \end{pmatrix}.$$

This is an extension the earlier definition, where row i corresponds with exponent $2i-1$. As above, the α's run over all elements of $\mathbb{F}_{2^m} \setminus \{0\}$. H has dimensions $tm \times 2^m - 1$.

Claim 6. *The code described by H has distance at least $2t+1$.*

Proof. The proof mimics the one given earlier. We need to show that any $2t$ columns are linearly independent. Suppose columns corresponding to $\beta_1, \ldots, \beta_{2t}$ were dependent with coefficients $e_1, \ldots, e_{2t} \in \mathbb{F}_2$. Then, as before, using linearity of ϕ, we know that $\forall j \leq t$

$$\sum_{i=1}^{2t} e_i \beta_i^{2j-1} = 0.$$

We need to show $\forall l \leq 2t$

$$\sum_{i=1}^{2t} e_i \beta_i^{l} = 0.$$

We know that this holds for odd $l \leq 2t$. But, if we know that it holds for some l', then we know it holds for $2l'$ because

$$\sum_{i=1}^{2t} e_i \beta_i^{2l'} = \left(\sum_{i=1}^{2t} e_i \beta_i^{l'} \right)^2$$

since our field has characteristic 2. Now construct the matrix M as in the earlier proof, and realize that M is a column-scaling of a Vandermonde matrix with distinct, nonzero β's. Therefore, M is nonsingular and so the columns of H are linearly independent. $\square$

What is the size of this code? Let $n = 2^m - 1$, then

$$|C| \geq \frac{2^n}{2^{nt}} = \frac{2^n}{(n+1)^t} = \Theta\left(\frac{2^n}{n^t}\right).$$

This matches the volume packing bound up to a constant factor.

3 Relation between BCH code and RS code

Consider the dual of RS code:
Take RS code over $\mathbb{F}_q$ with evaluation set $\mathbb{F}_q$ and code word be all polynomials of degree $< k$. What is $C^{\perp}$?

$$C^{\perp} = \{f : \mathbb{F}_q \to \mathbb{F}_q | \forall p(x) \text{ of degree } < k, \sum_{x \in \mathbb{F}_q} f(x)p(x) = 0\}$$

Claim 7. $\forall m < q-1$,

$$\sum_{x \in \mathbb{F}_q} x^m = 0$$

Take $y \in \mathbb{F}_q \setminus \{0\}$, s.t. $y^m \neq 1$. (exist since $deg(x^m - 1) < q - 1$.)
Then:

$$\sum_{x \in \mathbb{F}_q} (xy)^m = \sum_{x \in \mathbb{F}_q} x^m$$
$$(y^m - 1) \cdot \sum_{x \in \mathbb{F}_q} x^m = 0$$
$$\implies \sum_{x \in \mathbb{F}_q} x^m = 0$$

This directly implies that $x^i \perp x^j$ if $i + j < q - 1$
So $x^i \in C^{\perp}$ for each $i \leq q - 1 - k$. Let $S = span\{1, x, ..., x^{q-1-k}\}$. Then $dim(S) = x^{q-k}$ That means: $S \subset C^{\perp}$ and $dim(S) = dim(C^{\perp})$. So $S = C^{\perp}$.
So this is the parity check matrix for C.

$$\begin{pmatrix} 1 & \dots & 1 & \dots & 1 \\ & & x & & \\ & & x^2 & & \\ & & \vdots & & \\ & & x^{q-1-k} & & \end{pmatrix}$$

This gives a quick proof that C has distance $\geq q - k - 1$: Any $q - k$ colums form a Vandermonde Matrix and so are linear independent.

Let $q = 2^m$, $\mathbb{F}_2 \subset \mathbb{F}_q$.

Claim 8. *Let* $\widetilde{C}$ *=BCH code with parameter* $t = \lfloor \frac{q-k}{2} \rfloor$, *then*

$$\widetilde{C} = \widehat{C} := \{p \in C \text{ s.t. } p(x) \in \mathbb{F}_2 \text{ for all } x \in \mathbb{F}_q\}$$

Proof. For any $v \in ((F)_2)^n$, we want to show that:

$$Hv = 0 \text{ if and only if } \widetilde{H}v = 0$$

Where $\widetilde{H}$ is the parity check matrix for $\widehat{C}$, we have:

$$\widetilde{H} = \begin{pmatrix} \phi(x) \\ \phi(x^3) \\ \vdots \\ \phi(x^{q-1-k}) \end{pmatrix}$$

If $\widetilde{H}v = 0$, then $\widetilde{\widetilde{H}}v = 0$, where $\widetilde{\widetilde{H}} = \widetilde{H}$+even rows.
We can see that $\widetilde{\widetilde{H}} = \phi(H)$, which finishes the proof. □

So BCH code with parameter t are contained in RS code with distance $2t+1$.
So using BerlekampWelch algorithm one can decode BCH code of parameter t r from t errors in time $ploy(n)$. (n^t is trivial).

Dual of BCH codes are called Dual-BCH codes. Dual-BCH codes with parameter t is a code with $C \subset \mathbb{F}_2^n$, $|C| = O(n^t)$.
Turns out that C has distance $\frac{1}{2} - \frac{t}{\sqrt{n}}$. (Follows from Weil Bound)

Remarkable because:

1. Greedy/Random code with n^t codeword has distance $\frac{1}{2} - \sqrt{\frac{\log(n)}{n}}$
2. Optimal tradeoff for distance vs. size in the region.

How do codeword of C looks like?

$$f : \mathbb{F}_{2^m} \to \mathbb{F}_2$$
$$f(x) = \text{first bit of } \phi(\sum_{i=0}^{2t-1} a_i x^i)$$

Equivalently, pick $\mathbb{F}_2$ linear function $\ell : \mathbb{F}_{2^m} \to \mathbb{F}_2$.

$$f(x) = \ell(\sum_{i=0}^{2t-1} a_i x^i)$$