Lecture 14: Codes Better than Random

Error-Correcting Codes (Spring 2016)
Rutgers University
Swastik Kopparty
Scribes: Shahab Raji

Over $\{0,1\}$ alphabet with constant R and δ best known codes are Greedy and Random codes (i.e., the GV bound)[1]. Until the 80's, this was true for every alphabet with size $q = O(1)$. In the 80's, the algebraic geometric codes with a tradeoff on R and δ was introduced which work better than the random codes.

When q is in the form p^{2k} for some prime number p, and $q \geq 49$ the tradeoff is

$$R = 1 - \delta - \frac{1}{\sqrt{q} - 1} - o(1)$$

For other q's, there are different tradeoffs.

1 Code Construction

We will use Reed-Solomon construction sketch; We take polynomials in $\mathbb{F}_q$ of degree $\leq d$ and evaluate it on all the points of $\mathbb{F}_q$ (as a subset of the whole plane).

Problem: Fixing q will give you a code with fixed length.

Idea: Instead of just taking polynomials, we take something longer. One solution is Reed-Muller codes. But it will have negative effect on R and δ.

1.1 Construction Steps

- Take an algebraic curve in $\mathbb{F}_q^2$, e.g $x^2 + y^2 = 1$
- Mod out equivalent polynomials.
- Evaluate low degree algebraic functions on it.

Based on *Goppa's method*[1] , we need to understand:

1. Dimension of low degree polynomials on the curve, (k).
2. Number of zero's that a low degree polynomial can have on the curve, (a).
3. Number of points on the curve, (n).

[1]For constant distance we already saw BCH codes that do better, for large alphabets we saw Reed-Solomon codes that can do better, and for huge $\delta = 1/2 - \Theta(1/\sqrt{n})$, there are dual-BCH codes that we briefly saw.

Question from the class: why curves and not higher dimensional varieties? In general, curves give you good R,δ-tradeoff, and two or more dimensional varieties gives you bad R,δ-tradeoff - similar to Reed-Solomon vs Reed-Muller.

If we can find curves over $\mathbb{F}_q^m$ where $n \to \infty$, then we get longer codes over a fixed alphabet. The following theorem states that such curves can be found.

Theorem 1 (Tsfasman-Vladut-Zink '82[2])**.** *We can find curves over $\mathbb{F}_q^m$ with $n \to \infty$ and $\forall k$ $a \leq k + \frac{n}{\sqrt{q}-1}$.*

2 An Example : Hermitian code

A very simple Algebraic Geometric code on $\mathbb{F}_q$; with $q = p^2$ for some prime p. Let Curve $V = \{(x, y) \in \mathbb{F}_q^2 \text{ s.t. } x^p + x = y^{p+1}\}$.

Number of points on $V = \sum_{a \in \mathbb{F}_p} |Tr^{-1}(a)| \cdot |\mathcal{N}^{-1}(a)| \approx p^3$. (Tight for Schwartz–Zippel)

$x^p + x$ is the Field Trace of $\mathbb{F}_q$ to $\mathbb{F}_p$. $(Tr_{\mathbb{F}_q/\mathbb{F}_p}(x))$. And y^{p+1} is the Norm.$(\mathcal{N}_{\mathbb{F}_q/\mathbb{F}_q})$

Claim. $x^p + x \in \mathbb{F}_p$.
Proof. $(x^p + x)^p = (x^p)^p + x^p = x^{p^2} + x^p = x^q + x^p = x + x^p$ □

Claim. $y^{p+1} \in \mathbb{F}_p$.
Proof. $(y^{p+1})^p = y^{p^2+p} = y^q \cdot y^p = y \cdot y^p = y^{p+1}$ □

Fact 2. *$x^p + x - y^{p+1}$ is an irreducible polynomial in $\mathbb{F}_q[x, y]$. ($Trace - Norm$ is always an irreducible polynomial.)*

Two polynomials which their difference is divisible by $x^p + x - y^{p+1}$, both evaluate the same on the solution set of V.

Pick a parameter D. A code of degree D is

$$C_D = \{R(x, y) \in \mathbb{F}_q[x, y],\ deg(R) = D,\ deg_y(R) \leq p\}$$

where $R(x, y)$ is the set of all polynomials and $deg_y(R) \leq p$ is to ensure that there are no two polynomials that evaluate the same on V.

$$dim(C_D) = \begin{cases} \frac{D^2}{2} & D \leq p \\ (D - \frac{p}{2})p & D > p \end{cases}$$

We construct codewords of R by evaluating R at all the points of V. Then $n = p^3$ and $k = |C_D|$.

Claim 3. *Two elements of C_D, say R_1 and R_2, differ on many points.*

Proof. This follows from the fact that $R_1 - R_2$ is nonzero on many points of V. $R_1 - R_2$ can vanish on the intersection points of $R_1 - R_2$ and $x^p + x - y^{p+1}$. Since $x^p + x - y^{p+1}$ is irreducible and $deg_y(R_1 - R_2) < p+1$, therefore $R_1 - R_2$ and $x^p + x - y^{p+1}$ are relatively prime. Then by Bézout's Theorem[3],

$$\text{number of common zeros} \leq D(p+1) \approx Dp$$

□

2.1 Properties of the code

- $n = p^3$
- $k = \begin{cases} \frac{D^2}{2} & D \leq p \\ (D - \frac{p}{2})p & D > p \end{cases}$
- $R = \begin{cases} \frac{D^2}{2p^3} & D \leq p \\ \frac{D - p/2}{p^2} & D > p \end{cases}$
- $\delta \geq 1 - \frac{D}{p^2}$
- Alphabet Size=p^2

Note that for $D = \Omega(p^2)$, $R \approx 1 - \delta$. This looks like Reed-Solomon parameters, but over smaller alphabets.

3 A code that slightly beats the GV-Bound over {0,1}

We have different methods of creating codes like the greedy method and random linear method. We will now show that a random *quasicyclic* code meets (and beats) the GV bound. This beautiful argument is due to Gaborit-Zemor.

A *circulant matrix* is a matrix where each row vector is created by rotating the preceding row one element to the right.

$$\begin{pmatrix} c_1 & c_2 & c_3 & \cdots & c_n \\ c_n & c_1 & c_2 & \cdots & c_{n-1} \\ c_{n-1} & c_n & c_1 & \cdots & c_{n-2} \\ & & \vdots & & \end{pmatrix}$$

A linear code is called **quasicyclic** if it has a generator matrix of the form: $G = \left[I \middle| A\right]_{k \times 2k}$. This gives us a code with dimension k and length $2k$. Equivalently, a linear code is quasicyclic if for every codeword $(x, y) \in \mathbb{F}_2^{2k}$ (with $x, y \in \mathbb{F}_2^k$), then for any i, the rotation $(Rot^i(x), Rot^i(y))$ is also in the code.

3.1 Meeting the GV bound

We first show that such a We want to show that this code has a good distance, i.e. $\forall$message $x \neq 0$, $wt(xG) \geq d$. We do this by estimating $Pr[wt(xG) \geq d]$ and then use the Union bound over all x's. (Fix $x \neq 0$)

3.1.1 Estimating $Pr[wt(xG) < d]$

Since $xG = (x|xA)$, then $wt(xG) = wt(x) + wt(xA)$. Note that the $wt(x)$ is not random and $wt(xA)$ is the random part.

$$wt(xG) < d \Leftrightarrow wt(xA) < \underbrace{d - wt(x)}_{u}$$

We need to understand what the probability that xA has low Hamming weight is. More generally, what is the distribution of xA for fixed x and a random circulant matrix A?

We will see that when the integer k is good, and for all x (except some trivial cases), xA is *almost uniformly distributed*!

Let X be a circulant matrix with first column x^T. Let a be a row vector equal to the first column of A. Then we have the

$$xA = aX$$

Thus we now have a random vector a rather than a random matrix A. To understand the distribution of xA, we thus need to understand the *rank* of X.

Theorem 4. *If k is s.t. 1) k is prime. and 2) 2 is a generator of $\mathbb{F}_k^*$ then unless x is all-one or all-zero,*

$$Rank(X) \geq k - 1.$$

Proof. We will study the ring of $k \times k$ circulant matrices over $\mathbb{F}_2$.

Let M be the "rotate by 1" $k \times k$ matrix.

The main observation is that the ring generated by M is the ring of all circulant matrices.

From here, it is not hard to show that the ring of $k \times k$ circulant matrices is isomorphic to the ring $R = \mathbb{F}_2[Y]/\langle Y^k - 1\rangle$ (map Y to M).

Represent elements of R by polynomials of degree at most $k-1$. Note that if you have a polynomial $P(Y) = a_0 + a_1 Y + \cdots + a_{k-1} Y^{k-1}$, then $Y \cdot P(Y)$ rotates the coefficients.

What is the structure of R? It depends on the factorization of $Y^k - 1$. When k satisfies the hypotheses, we have $Y^k - 1 = (Y-1)(Y^{k-1} + \cdots + 1)$, where both these factors are irreducible over $\mathbb{F}_2$ (The proof involves taking a root α of $Y^k - 1$, and asking: for which value of ℓ does $\alpha^{2^\ell} = \alpha$?).

Thus

$$R \simeq \mathbb{F}_2[Y]/(Y^k - 1) \simeq \mathbb{F}_2[Y]/\langle Y - 1\rangle \oplus \mathbb{F}_2[Y]/\langle Y^{k-1} + \cdots + 1\rangle \simeq \mathbb{F}_2 \oplus \mathbb{F}_{2^{k-1}}$$

Each element λ of R can thus be represented as $(b, \beta) \in \mathbb{F}_2 \times \mathbb{F}_{2^{k-1}}$. Given an element $\lambda = (b, \beta)$, the rank of the corresponding circulant matrix equals the dimension of the image of R under multiplicaiton by λ, i.e., λR. What is λR? If b, β are both nonzero, then it equals R. If b zero and β is nonzero then it equals $\{0\} \times \mathbb{F}_{2^{k-1}}$. If b is nonzero and β is zero, then it equals $\mathbb{F}_2 \times \{0\}$. Otherwise it equals 0.

In the first two cases, the dimension of λR is at least $k - 1$, as desired. Note that this covers all but 2 values for λ: $(1, 0)$ and $(0, 0)$. One can check that these cases correspond to the all 1's matrix and the all 0's matrix. □

If x is not the all 1 or all 0 vector, then since every $y \in \mathbb{F}_2^k$ has at most two $a \in \mathbb{F}_2^k$ such that $Xa = y$,

$$Pr[wt(Xa) < u] \leq \frac{2|B_k(u)|}{2^k}$$

3.1.2 Applying the Union Bound

Let k satisfy the criteria of the previous lemma. It is not known that infinitely many such k exist, but it is widely believed (and we will assume it for this lecture). Gaborit-Zemor find a way around this unconditionally.

Now we use the Union Bound on all x's.

$$\begin{aligned}
Pr[\exists x \neq 0 \text{ s.t. } wt(xG) < d] &\leq \sum_{x \neq 0} 2 \cdot \frac{|B_k(d - wt(x))|}{2^k} \\
&= \sum_{x \text{ s.t. } wt(x) < d} 2 \cdot \frac{|B_k(d - wt(x))|}{2^k} \\
&= \sum_{0 < i < d} 2 \cdot \frac{\overbrace{|B_k(d-i)|\binom{k}{i}}^{\text{strings with weight } d \text{ over } \{0,1\}^{2k}}}{2^k} \\
&= 2 \cdot \frac{B_{2k}(d)}{2^k}
\end{aligned}$$

If $\frac{B_{2k}(d)}{2^k} < \frac{1}{100}$, then with high probability the code is good. Since $|C| = 2^k$, this condition is the same as $\frac{|C| B_{2k}(d)}{2^{2k}} < \frac{1}{100}$. So $|C| < \frac{1}{100} \frac{2^{2k}}{B_{2k}(d)}$ suffices.

This meets the GV bound (upto a constant), but we still did not beat the GV-bound.

3.2 Improved Proof

The key observation: *We only need to worry about one x per rotation class.* This improved proof helps us to beat the GV-bound.

Let $S \subseteq \mathbb{F}_2^k$ be a set of strings, such that there is one string per rotation class in S. $|S| \approx \frac{2^k}{k}$.

$$\begin{aligned} Pr[\exists x \neq 0 \text{ s.t. } wt(xG) < d] &= Pr[\exists x \neq 0 \wedge x \in S \text{ s.t. } wt(xG) < d] \\ &\leq \sum_{x \neq 0, x \in S} 2 \cdot \frac{|B_k(d - wt(x))|}{2^k} \\ &= \sum_{x \in S \text{ s.t. } wt(x) < d} 2 \cdot \frac{|B_k(d - wt(x))|}{2^k} \\ &= \sum_{0 < i < d} 2 \cdot \frac{B_k(d-i)}{2^k} \cdot \frac{\binom{k}{i}}{k} \\ &= 2\frac{B_{2k}(d)}{2^k} \cdot \frac{1}{k} \end{aligned}$$

Thus we may take d such that $B_{2k}(d) < \frac{1}{100} \cdot 2^k$ and get that C has distance d with high probability. Thus we get a code C with $|C| \geq \frac{1}{100} \frac{k \cdot 2^{2k}}{B_{2k}(d)}$ and distance d

Thus the number of codewords in the code beats the GV bound by a factor k! This doesn't show up on the R vs δ graph unfortunately. This is the best asymptotic existence result in the constant relative distance regime.

References

[1] V. D. Goppa, *"Codes on algebraic curves"* (in Russian), Dokl. Akad. Nauk SSSR , vol. 259, pp. 12891290, 1981.

[2] M. A. Tsfasman, S. G. Vladut, and T. Zink, *"Modular curves, Shimura curves, and Goppa codes, better than Varshamov–Gilbert bound"* Math. Nachr., vol. 109, pp. 2128, 1982.

[3] Weisstein, Eric W. *"Bézout's Theorem."* From MathWorld–A Wolfram Web Resource. `http://mathworld.wolfram.com/BezoutsTheorem.html`