# Class 14 Exercises

CS250/EE387, Winter 2022

In the lecture videos/notes, we saw *Locally Correctable Codes* (LCCs). Recall the definition of an LCC:

**Definition 1** (LCC)**.** *A code $\mathcal{C} \subseteq \mathbb{F}_q^n$ is a $(\delta, Q, \gamma)$-LCC if there is a randomized algorithm $A$ so that the following holds. For all $w \in \mathbb{F}_q^n$ so that $\delta(c, n) \leq \delta$, and for all $i \in [n]$, $A$ makes $Q$ queries to $w$ and outputs $A^w(i)$ so that*

$$\Pr[A^w(i) = c_i] \geq 1 - \gamma.$$

We saw several examples of LCCs. One was Reed-Muller codes:

**Definition 2** ($q$-ary Reed-Muller Code)**.** *The $m$-variate Reed-Muller code over $\mathbb{F}_q$ with degree $d$ is given by*

$$RM_q(m, d) = \left\{ \langle f(\vec{\alpha}) \rangle_{\vec{\alpha} \in \mathbb{F}_q^m} \,:\, f \in \mathbb{F}_q[X_1, \ldots, X_m], \deg(f) \leq d \right\}.$$

**For the rest of class, we will specialize to the case where $m = 2$ and $d = q - 2$.**

As we saw in the videos, the RM codes that were decent LCCs had low rate. Today, we'll see one way to modify them, through a process called *lifting*, to make the rate close to 1.

0. Consider the following property of a (possibly high-degree) polynomial $p(X, Y)$, which we'll call property $\mathcal{P}(d)$:

   **Definition 3.** *We say that $p(X, Y) \in \mathbb{F}_q[X, Y]$ satisfies* property $\mathcal{P}(d)$ *if for any line $\ell(T) = (aT + b, cT + d)$, the univariate polynomial $p(\ell(T))$ is equivalent to (that is, has all the same evaluations as) a polynomial of degree at most $d$.*

   (a) Explain why any polynomial $p(X, Y)$ with total degree at most $d$ has property $\mathcal{P}(d)$.

   (b) Let $\mathcal{F} \subseteq \mathbb{F}_q[X, Y]$, so that every polynomial $f \in \mathcal{F}$ has property $\mathcal{P}(d)$. In the mini-lecture, we saw a proof that $RM_q(2, d)$ was an LCC (for appropriate values of $q, d$). Explain why the same argument works for the code

   $$\mathcal{C} = \left\{ \langle f(\vec{\alpha}) \rangle_{\vec{\alpha} \in \mathbb{F}_q^2} \,:\, f \in \mathcal{F} \right\}.$$

   (c) What might be the advantage of considering a code like $\mathcal{C}$?

For the rest of today's class, we will see how to come up with a code $\mathcal{C}$ as in part (c) above with rate way better than the corresponding Reed-Muller code!

1. Let's start with an example. Consider the polynomial $p(X,Y) = X^2Y^2 \in \mathbb{F}_4[X,Y]$.
   (a) Is (the evaluation vector of) $p(X,Y)$ in $RM_{q=4}(m=2, d=2)$?
   (b) Show that $p(X,Y)$ has property $\mathcal{P}(2)$.
   Hint: You may use the **facts** that for any $x, y \in \mathbb{F}_4$, we have $(x+y)^2 = x^2 + y^2$, and $x^4 = x$.
   (c) Reflect on the fact that this is pretty weird. Can you come up with an example like this over the real numbers?

2. **From now until the end of class, let $q = 2^t$. As before, we have $m = 2$, and $d = q-2$.**
   (a) What is the rate of $RM_q(m = 2, d = q-2)$? (Or at least, what is its limit as $q$ gets large?)
   (b) Consider the following theorem:

   **Theorem 1.** *Let $q = 2^t$ for some $t$. The the number of $f(X,Y) \in \mathbb{F}_q[X,Y]$ so that $\mathcal{P}(q-2)$ holds for $f(X,Y)$ is at least $q^{4^t - 3^t - 1}$.*

   Assuming that theorem is true, explain why this implies the existence of a LCC $\mathcal{C}$ of length $N = q^2$ and rate that tends to 1 as $N \to \infty$, with parameters $\delta = \frac{1}{100\sqrt{N}}$, $Q = \sqrt{N}$, and $\gamma = 0.1$.
   **Note:** In case it is helpful, $\log_3(4) \approx 1.26$.

3. In this part, we will (mostly) prove Theorem 1.
   Say that a monomial $M_{ij}(X,Y) := X^iY^j$ is *good* if $\mathcal{P}(q-2)$ holds for $M_ij(X,Y)$.
   (a) Explain why, to prove Theorem 1, it is enough to show that the number of good monomials is at least $4^t - 3^t - 1$.
   (b) Let $\ell(T) = (T, aT + b)$. (We are going to restrict ourselves to lines that look like this for simplicity; the general case is basically the same). Suppose that $(i,j) \neq (q-1, q-1)$. Consider the univariate polynomial
   $$P_{ij}(T) = M_{ij}(\ell(T)).$$
   Show that the coefficient on $T^{q-1}$ in $P_{ij}(T)$ is
   $$\begin{cases} \binom{j}{q-i-1} a^{q-i-1} b^{j-(q-i-1)} & j \geq q-i-1 \\ 0 & j < q-i-1 \end{cases}$$
   where when we refer to an integer like $\binom{j}{q-i-1}$ as a element of $\mathbb{F}_{2^t}$, we mean $1+1+\ldots+1$ that many times.
   Conclude that if $\binom{j}{q-i-1} \equiv 0 \mod 2$, then $M_{ij}(X,Y)$ is good.[1] (For the "conclude" part, you can use the fact that $1+1=0$ in $\mathbb{F}_{2^t}$).

[1] Here, you can ignore the fact that we didn't consider general lines of the form $(aT+b, cT+d)$, only lines like $(T, aT+b)$. The argument for the more general case is exactly the same, just slightly more tedious. Notice that by restricting to these simpler lines, we are only leaving out the "horizontal" lines of the form $(c, aT+b)$ for some constant $c$.

(c) To finish the proof, we will use the following corollary of Lucas' theorem (which we will not prove):

**Fact 2.** *For an integer $m < 2^t$, let $b(m) \in \{0,1\}^t$ denote the binary expansion of $m$. For example if $t = 3$, we have $b(5) = 101$. For a vector $v \in \{0,1\}^t$, write $\overline{v} \in \{0,1\}^t$ to denote the coordinate-wise flip of $v$. For example, $\overline{b(5)} = 010$.*

*For two vectors $v, w \in \{0,1\}^t$, we say that $v$ "lies in the 2-shadow of $w$" if $v_i = 1$ implies that $w_i = 1$. For example, $v = 100$ lies in the 2-shadow of $w = 101$, since whenever $v$ has a 1, $w$ also has a 1. However, $v = 110$ does not lie in the 2-shadow of $w = 101$, since $v_2 = 1$ but $w_2 = 0$.*

*With this notation, the **fact** is that, for $q = 2^t$,*

$$\binom{j}{q-i-1} \neq 0 \Leftrightarrow \overline{b(i)} \leq_2 b(j).$$

This fact may seem weird, but it is true! Convince yourself of this by example by applying it to with $j = 5$ and $i = 3$, and for $j = 5$ and $i = 4$ (and with $t = 3$ for both).

(d) Show that the number of good monomials is at least $4^t - 3^t - 1$, proving the theorem.
Hint: It might be helpful that $\sum_{s=0}^{t} \binom{t}{s} 2^s = 3^t$. (There's also a way to do it where this is not helpful).

4. **(Bonus)** Try to use the same ideas for $d < q - 2$ and $m > 2$ to come up with an LCC with rate close to 1 and parameters $\delta = 0.01, Q = N^{0.01}, \gamma = 0.01$.