MPRI **Year 2024–25**
Module 2.13.2: Error-correcting codes and applications to cryptography

Mid-term exam, December 4, 2024

- *You have 1h30. You can write your answers either in French or in English.*
- *Exercises are independent.*
- *Questions marked with a* ($\star$) *are harder than the other ones.*

Exercise 1. We study cyclic codes of length 16 over $\mathbb{F}_3$.

1°) Give the list of corresponding minimal cyclotomic classes.

Answer : The minimal cyclotomic classes are $\{0\}, \{1,3,9,11\}, \{2,6\}, \{4,12\}, \{5,15,13,7\}, \{8\}, \{10,14\}$.

2°) What is the number of cyclic codes of length 16 over $\mathbb{F}_3$.

Answer : 128.

3°) What is the number of cyclic codes of length 16 and dimension 11 over $\mathbb{F}_3$?

Answer : We need to count the number of cyclotomic classes of size 5. From the minimal ones, one can have

- A union of a minimal one of size 1 and an minimal one of size 4. Since there are 2 minimal classes of size 1 and 2 of size 4, this yields 4 possibilities;
- A union one of size 1 and two of size 2. There are $2\binom{3}{2} = 6$ possibilities.

In summary there are 10 such codes.

4°) Recall that BCH bound asserts that if a cyclotomic class contains an arithmetic progression of s elements, *i.e.* a sequence of the form $a, a+b, a+2b, \ldots, a+(s-1)b$, then the corresponding code has minimum distance at least $s+1$.

Prove the existence of a cyclic code of parameters $[16, 8, d]_3$ with $d \geqslant 9$.

Answer : Consider the cyclotomic class $\{0,2,4,6,8,10,12,14\}$. Or $\{1,3,5,7,9,11,13,15\}$.

5°) Prove that actually $d = 9$.

Answer : Singleton bound asserts that $d \leqslant 9$. Thus the code is MDS.

Exercise 2. A linear code $\mathscr{C} \subseteq \mathbb{F}_q^n$ is said to be *Complementary Dual* (CD) if $\mathscr{C} \cap \mathscr{C}^\perp = \{0\}$.

1°) Prove that for such a CD code, $\mathbb{F}_q^n = \mathscr{C} \oplus \mathscr{C}^\perp$.

Answer : Since $\dim \mathscr{C} + \dim \mathscr{C}^\perp$, the very definition of CD codes entails that the direct sum $\mathscr{C} \oplus \mathscr{C}^\perp$ equals $\mathbb{F}_q^n$.

2°) Let $\mathscr{C} \subseteq \mathbb{F}_q^n$ be a code of dimension k and $\mathbf{G} \in \mathbb{F}_q^{k\times n}$. Prove that $\mathscr{C}$ is CD if and only if $\mathbf{G}\mathbf{G}^\top$ is nonsingular (*i.e.* invertible).

Answer : Suppose $\mathscr{C}$ is LCD and that there exists $u \in \mathbb{F}_q^n$ such that $\mathbf{G}\mathbf{G}^\top u^\top = 0$. This means that $u\mathbf{G} \in \mathscr{C}^\perp$ while it is in $\mathscr{C}$. Since $\mathscr{C}$ is CD, we deduce that $u\mathbf{G} = 0$ and hence $u = 0$ since $\mathbf{G}$ has full rank ($\mathscr{C}$ has dimension k, $\mathbf{G} \in \mathbb{F}_q^{k\times n}$, its rows form a basis of $\mathscr{C}$). Therefore $\mathbf{G}\mathbf{G}^\top$ is nonsingular.
Suppose that $\mathbf{G}\mathbf{G}^\top$ is nonsingular. Let $u\mathbf{G} \in \mathscr{C}\cap\mathscr{C}^\perp$, we have $\mathbf{G}(u\mathbf{G})^\top = 0$ and hence $u \in \ker \mathbf{G}\mathbf{G}^\top$ which entails $u = 0$. Thus $\mathscr{C} \cap \mathscr{C}^\perp = \{0\}$.

3°) Let $\mathbf{G} = (\mathbf{I}_k \mid \mathbf{A}) \in \mathbb{F}_2^{k\times n}$ be a systematic ($\mathbf{I}_k$ denotes the $k \times k$ identity matrix) generator matrix of a **binary** code $\mathscr{C}$.

(a) Prove that

$$\mathbf{G}_1 = (\mathbf{I}_k \mid \mathbf{A} \mid \mathbf{A})$$

is a generator matrix of a CD code.

Answer :

$$\mathbf{G}_1\mathbf{G}_1^\top = \mathbf{I}_k + 2\mathbf{A}\mathbf{A}^\top = \mathbf{I}_k$$

From Question 2, we get the result.

(b) If $\mathscr{C}$ is $[n, k, d]_2$, then prove that the code $\mathscr{C}_1$ with generator matrix $\mathbf{G}_1$ is $[2n-k, k, d_1]$ with $d_1 \geqslant d$.

Answer : Clearly $\mathbf{G}_1 \in \mathbb{F}_2^{k\times(2n-k)}$ and has full rank (due to the $\mathbf{I}_k$ block in $\mathbf{G}_1$). Only minimum distance should be checked. Let $u \in \mathbb{F}_2^k$ such that $u\mathbf{G}_1$ is a minimum-weight codeword of $\mathscr{C}_1$. Then $\mathrm{wt}(u\mathbf{G}_1) \geqslant \mathrm{wt}(u\mathbf{G})$. Moreover, if $u\mathbf{G} = 0$ then $u = 0$. Therefore $d_1 \geqslant d$

(c) Prove that $\mathscr{C}_1^{\perp}$ has minimum distance $\leqslant 2$.

Answer : There are pairs of repeated columns in $\mathbf{G}_1$ which yield codewords of weight 2 in $\mathscr{C}_1^{\perp}$.

4°) Prove the existence of sequences of binary CD codes which are asymptotically good (*i.e.* the sequences of rates and relative distances go to (R, δ) where both R and δ are positive).

Answer : Since random codes are asymptotically on Gilbert Varshamov bound, we know the existence of sequences of random codes with asymptotic parameters $R, \delta > 0$. After permuting their columns, which has no influence on the parameters, all these codes have a systematic generator matrix. Using the previous operation, we obtain sequences of codes with asymptotic parameters $(R_1 = \frac{R}{2-R},\ \delta_1 \geqslant \frac{\delta}{2-R})$

Turn the page please $\longrightarrow$

Exercise 3. For $\mathbf{y} \in \mathbb{F}_2^n$ and $0 \leqslant r \leqslant n$, $\mathcal{B}_{\mathrm{H}}(\mathbf{y}, r)$ denotes the Hamming ball with center $\mathbf{y}$ and radius r:

$$\mathcal{B}_{\mathrm{H}}(\mathbf{y}, r) \stackrel{\text{def}}{=} \{\mathbf{x} \in \mathbb{F}_2^n : \ d_{\mathrm{H}}(\mathbf{x}, \mathbf{y}) < r\} \quad \text{and} \quad \forall \rho \in [0,1], \quad \mathrm{Vol}(n, \rho) \stackrel{\text{def}}{=} \sharp \mathcal{B}_{\mathrm{H}}(\mathbf{0}, \rho n).$$

In what follows $\mathcal{C}$ denotes a binary code with minimum distance $\geqslant \delta n$ for some constant $\delta \in [0, 1/2]$.

1°) Explain where the Hamming bound, defined as follows, comes from:

$$\sharp \mathcal{C} \cdot \mathrm{Vol}(n, \delta/2) \leqslant 2^n.$$

The aim of this exercise is to prove the so-called *Elias-Bassalygo* bound stating that

$$\sharp \mathcal{C} \cdot \mathrm{Vol}(n, J(\delta)) \leqslant (2n) 2^n \quad \text{where} \ \ J(\delta) \stackrel{\text{def}}{=} \frac{1 - \sqrt{1 - 2\delta}}{2} \tag{1}$$

which **asymptotically** improves Hamming bound for any $\delta \in [0, 1/2]$. More precisely, we want to prove the following:

$$\exists \rho \in (\delta/2, 1/2) \text{ such that } \quad \sharp \mathcal{C} \cdot \mathrm{Vol}(n, \rho) \leqslant 2n \cdot 2^n \tag{2}$$

2°) Explain why (2) would improve upon the asymptotic Hamming bound.

Answer : We know that

$$\mathrm{Vol}(n, \rho) = 2^{n(h(\rho) + o(1))}$$

where h is the binary entropy which is an increasing function over $[0, 1/2]$. Therefore, as $\rho \in [\delta/2, 1/2]$,

$$\frac{\mathrm{Vol}(n, \delta/2)}{\mathrm{Vol}(n, \rho)} = 2^{-\Omega(n)} \tag{3}$$

But Equation (2) can be rewritten as:

$$\sharp \mathcal{C} \cdot \mathrm{Vol}(n, \delta/2) \leqslant 2^n \frac{\mathrm{Vol}(n, \delta/2) \ (2n)}{\mathrm{Vol}(n, \rho)}$$

Therefore using Equation (3) concludes the question.

3°) Given any $\rho \in [0, 1]$, show that there exists at least one $\mathbf{y}_0 \in \mathbb{F}_2^n$ such that,

$$\sharp \Big(\mathcal{B}_{\mathrm{H}}(\mathbf{y}_0, \rho n) \cap \mathcal{C} \Big) \geqslant \frac{\sharp \mathcal{C} \cdot \mathrm{Vol}(n, \rho)}{2^n}.$$

Hint: *pick* $\mathbf{y}$ *uniformly at random and compute the expected value of* $\sharp \Big(\mathcal{B}_{\mathrm{H}}(\mathbf{y}, \rho n) \cap \mathcal{C} \Big)$.

Answer : Let us consider the following random variable (where $\mathbf{y}$ is picked uniformly at random in $\mathbb{F}_2^n$),

$$X \stackrel{\text{def}}{=} \sharp\Big(\mathcal{B}_{\mathrm{H}}(\mathbf{y}, \rho n) \cap \mathcal{C}\Big)$$

Notice that,

$$X = \sum_{\substack{\mathbf{c}\in\mathscr{C} \\ \mathbf{x}\in\mathbb{F}_2^n:\ \mathrm{wt}(\mathbf{x})<\rho n}} 1_{\mathbf{y}=\mathbf{c}+\mathbf{x}}$$

Therefore, by linearity of the expectation:

$$\mathbb{E}(X) = \sum_{\substack{\mathbf{c}\in\mathscr{C} \\ \mathbf{x}\in\mathbb{F}_2^n:\ \mathrm{wt}(\mathbf{x})<\rho n}} \mathbb{P}_{\mathbf{y}}\left(\mathbf{y} = \mathbf{c}+\mathbf{x}\right) = \sum_{\substack{\mathbf{c}\in\mathscr{C} \\ \mathbf{x}\in\mathbb{F}_2^n:\ \mathrm{wt}(\mathbf{x})<\rho n}} \frac{1}{2^n} = \frac{\sharp\mathscr{C}\cdot \mathrm{Vol}(n,\rho)}{2^n}$$

where in the second equality we used that $\mathbf{y}$ is uniformly distributed. Since this is the expected value of the size, there must exist at least one $\mathbf{y}_0 \in \mathbb{F}_2^n$ such that

$$\sharp\Big(\mathcal{B}_{\mathrm{H}}(\mathbf{y}_0, \rho n) \cap \mathcal{C}\Big) \geqslant \mathbb{E}(X) = \frac{\sharp\mathscr{C}\cdot \mathrm{Vol}(n,\rho)}{2^n}$$

4°) Deduce how (2) is implied by the following property:

$$\exists \rho \in (\delta/2, 1/2) \text{ such that } \quad \forall \mathbf{y} \in \mathbb{F}_2^n,\ \sharp\Big(\mathcal{B}_{\mathrm{H}}(\mathbf{y}, \rho n) \cap \mathcal{C}\Big) \leqslant 2n.$$

Answer : We just need to use the result of the question above.

5°) Let us define

$$\forall \mathbf{x} \in \mathbb{F}_2^n, \quad v_{\mathbf{x}} \stackrel{\text{def}}{=} \left((-1)^{x_1}, \ldots, (-1)^{x_n}\right).$$

Show that,

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{F}_2^n, \quad \langle v_{\mathbf{x}}, v_{\mathbf{y}} \rangle = n - 2d_{\mathrm{H}}(\mathbf{x}, \mathbf{y}).$$

where $\langle \cdot, \cdot \rangle$ denotes the Euclidean scalar product on $\mathbb{R}^n$.

Answer : By definition of the Euclidean scalar product,

$$\begin{aligned}
\langle v_{\mathbf{x}}, v_{\mathbf{y}} \rangle &= \sum_{i=1}^{n} (-1)^{x_i + y_i} \\
&= \sharp\{i \in [1,n] :\ x_i = y_i = 1\} - \sharp\{i \in [1,n] :\ x_i \neq y_i\} \\
&= \Big(n - d_{\mathrm{H}}(\mathbf{x}, \mathbf{y}))\Big) - d_{\mathrm{H}}(\mathbf{x}, \mathbf{y})
\end{aligned}$$

concluding the question.

6°) ($\star$) Suppose that $\mathcal{C} \cap \mathcal{B}_{\mathrm{H}}(\mathbf{y}_0, \rho n)$ is a set of m codewords: $\{\mathbf{c}_1, \ldots, \mathbf{c}_m\}$. Show that for $\alpha > 0$,

$$4\rho \leqslant 2 - \alpha + \frac{2\delta - 1}{\alpha} \quad \Longrightarrow \quad \forall 1 \leqslant i < j \leqslant m, \quad \langle v_{\mathbf{c}_i} - \alpha v_{\mathbf{y}_0}, v_{\mathbf{c}_j} - \alpha v_{\mathbf{y}_0} \rangle \leqslant 0.$$

Answer : We have the following computation,

$$\begin{aligned} \langle v_{\mathbf{c}_i} - \alpha v_{\mathbf{y}_0}, v_{\mathbf{c}_j} - \alpha v_{\mathbf{y}_0} \rangle &= \langle v_{\mathbf{c}_i}, v_{\mathbf{c}_j} \rangle - \alpha \left(\langle v_{\mathbf{y}_0}, v_{\mathbf{c}_j} \rangle + \langle v_{\mathbf{c}_i}, v_{\mathbf{y}_0} \rangle \right) + \alpha^2 \langle v_{\mathbf{y}_0}, v_{\mathbf{y}_0} \rangle \\ &\leqslant n - 2\delta n - 2\alpha(1 - 2\rho)n + \alpha^2 n \\ &= n \left(1 - 2\delta - 2\alpha + 4\rho\alpha + \alpha^2 \right) \end{aligned} \tag{4}$$

where in the second inequality we used that $\mathrm{wt}(\mathbf{c}_i - \mathbf{c}_j) \geqslant \delta n$ and $\mathrm{wt}(\mathbf{c}_i - \mathbf{y}_0) < \rho n$ for any i, j. Notice now that for $\alpha > 0$,

$$1 - 2\delta - 2\alpha + 4\rho\alpha + \alpha^2 \leqslant 0 \iff 4\rho \leqslant \frac{-\alpha^2 + 2\alpha}{\alpha} + \frac{2\delta - 1}{\alpha} = -\alpha + 2 + \frac{2\delta - 1}{\alpha}$$

To conclude, we just need to plug this into (4).

7°) Using the previous question, prove that for $\rho \leqslant J(\delta)$ (defined in (1)), there exists one $\alpha > 0$ such that,

$$\forall 1 \leqslant i < j \leqslant m, \quad \langle v_{\mathbf{c}_i} - \alpha v_{\mathbf{y}_0}, v_{\mathbf{c}_j} - \alpha v_{\mathbf{y}_0} \rangle \leqslant 0.$$

Answer : Let us show that for all $\rho \leqslant J(\delta)$ we have $4\rho \leqslant 2 - \alpha + \frac{2\delta - 1}{\alpha}$. It will remain to use the result of the previous question to conclude. Let $f : \alpha \mapsto -\alpha + 2 + \frac{2\delta - 1}{\alpha}$ for $\alpha > 0$. Let us prove that for all $\rho \leqslant J(\delta)$ it exists some $\alpha > 0$ such that $f(\alpha) \geqslant 4\rho$. We have,

$$f'(\alpha) = -1 - \frac{2\delta - 1}{\alpha^2} = \frac{-\alpha^2 - 2\delta + 1}{\alpha^2}$$

Therefore f is maximum for $\alpha_0 \stackrel{\text{def}}{=} \sqrt{1 - 2\delta}$. But in that case,

$$f(\alpha_0) = 2 + \frac{2\delta - 1 - \alpha_0^2}{\alpha_0} = 2 + \frac{4\delta - 2}{\sqrt{1 - 2\delta}} = 2 \left(1 - \frac{1 - 2\delta}{\sqrt{1 - 2\delta}} \right) = 2 \left(1 - \sqrt{1 - 2\delta} \right) = 4J(\delta)$$

where in the second equality we used that $\delta \leqslant 1/2$. It concludes the question.

8°) ($\star$) Let $u_1, \ldots, u_m$ be m non-zero elements in $\mathbb{R}^n$ such that, for all $1 \leqslant i < j \leqslant m$, $\langle u_i, u_j \rangle \leqslant 0$. Show that $m \leqslant 2n$.

Answer : The result clearly holds for $n = 1$. Let us suppose that it is true for some $n \in \mathbb{N}\backslash\{0\}$. Let $u_1, \dots, u_m \in \mathbb{R}^{n+1}$ be non-zero vector and π be the orthogonal projection onto $\mathrm{Span}(u_1)^\perp$. Notice that $\pi(u_i) = 0$ for $i > 1$ implies that $u_i = \lambda_i u_1$ for $\lambda < 0$. It exists at most one such i, otherwise we would have $\langle u_i, u_j\rangle = \lambda_i \cdot \lambda_j > 0$. Therefore, there is at least $m-2$ non-zero vectors in the family $(\pi(u_i))_{1\leqslant i\leqslant m}$. Let us show now that $\langle \pi(u_i), \pi(u_j)\rangle \leqslant 0$. By induction it will show that $m - 2 \leqslant 2n$ which implies that $m \leqslant 2(n+1)$ concluding the recurrence.
We have,

$$\begin{aligned}\langle \pi(u_i), \pi(u_j)\rangle &= \left\langle u_i - \frac{\langle u_1, u_i\rangle}{\|u_1\|^2}u_1, u_j - \frac{\langle u_1, u_j\rangle}{\|u_1\|^2}u_1 \right\rangle \\ &= \langle u_i, u_j\rangle - \frac{\langle u_1, u_j\rangle \cdot \langle u_i, u_1\rangle}{\|u_1\|^2} - \frac{\langle u_1, u_i\rangle \cdot \langle u_1, u_j\rangle}{\|u_1\|^2} + \frac{\langle u_1, u_i\rangle \cdot \langle u_1, u_j\rangle}{\|u_1\|^2} \\ &= \langle u_i, u_j\rangle - \frac{\langle u_1, u_j\rangle \cdot \langle u_i, u_1\rangle}{\|u_1\|^2}\end{aligned}$$

which is $\leqslant 0$ as all the $\langle u_i, u_j\rangle$ are $\leqslant 0$ for $1 \leqslant i < j \leqslant m$.

Hint: *reason by induction over n and consider the orthogonal projection onto* $\mathrm{Span}(u_1)^\perp$. *Note that this projection can be made explicit as:* $x \mapsto x - \frac{\langle u_1, x\rangle}{\langle u_1, u_1\rangle}u_1$.

9°) Conclude.

Answer : Combining the previous question with question 7 shows that m (the number of codewords in $\mathscr{C} \cap \mathcal{B}_{\mathrm{H}}(\mathbf{y}_0, \rho n)$) is at most $2n$ for all $\rho \leqslant J(\delta)$ showing that the inequality from question 4 holds for the relative radius $\rho \leqslant J(\delta)$. It concludes the proof of Elias Bassalygo' bound.