Solutions of assignment 1

Response 1
We consider the $(5, 2)$ linear code over $\mathbb{F}_2$ with parity-check matrix

$$H = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

- All codewords $C = \{c \in \mathbb{F}_2^5 \ ; \ Hc^T = 0\}$ with $c = (c_1, c_2, \cdots, c_5)$

$$\begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{pmatrix} = \begin{pmatrix} c_2 + c_3 \\ c_1 + c_4 \\ c_1 + c_2 + c_5 \end{pmatrix} = (000)^T$$

$$\begin{cases} c_2 + c_3 = 0 \\ c_1 + c_4 = 0 \\ c_1 + c_2 + c_5 = 0 \end{cases} \iff \begin{cases} c_1 = c_2 + c_5 = c_4 \\ c_2 = c_3 \end{cases}$$

Hence

$$C = \{(c_1, c_2, c_2, c_1, c_1 + c_2); \ c_1, c_2 \in \mathbb{F}_2\}$$

i.e.

$$\begin{array}{rccc} f: & \mathbb{F}_2^2 & \longrightarrow & \mathbb{F}_2^5 \\ & (c_1, c_2) & \longmapsto & (c_1, c_2, c_2, c_1, c_1 + c_2) \end{array}$$

$$\begin{aligned} f(0,0) &= (0,0,0,0,0) \\ f(1,0) &= (1,0,0,1,1) \\ f(0,1) &= (0,1,1,0,1) \\ f(1,1) &= (1,1,1,1,0) \end{aligned}$$

So $C = \{(0,0,0,0,0), (1,0,0,1,1), (0,1,1,0,1), (1,1,1,1,0)\}$

- By definition the minimum distance $d_C = min_{c \in C}(w(c)) = min\{3, 4\} = 3$.

- Standard array

message row	00	10	01	11
codewords	0000	10011	01101	11110
remaining cosets	10000	00011	11101	01110
	01000	11011	00101	10110
	00100	10111	01001	11010
	00010	10001	01111	11100
	00001	10010	01100	11111
	11000	01011	10101	00110
	$\underbrace{10100}_{\text{column of coset leaders}}$	00111	11001	01010

- Decoding : We use the standard array to decode the following 3 vectors
  1. $(11111) - (00001) = 11110$ which corresponds to the message 11.
  2. (01101) is a codeword that corresponds to the message 01.
  3. $(01100) - (00001) = 01101$ which corresponds to the message 01.

Response 2: Let f be the corresponding function

$$f : (a_1, a_2, a_3) \longmapsto (a_1, a_2, a_3, a_1 + a_2, a_1 + a_3, a_2 + a_3)$$

So the dimension is $k = 3$ and the length of the code is $n = 6$.

- Let's compute the generator matrix

$$\begin{aligned} f(1,0,0) &= (1,0,0,1,1,0) \\ f(0,1,0) &= (0,1,0,1,0,1) \\ f(0,0,1) &= (0,0,1,0,1,1) \end{aligned}$$

  So the generator matrix $G = (I_3, A^T)$ is

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$

  where the matrix A is given by

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

- The parity check matrix $H = (A, I_3)$ is

$$\begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

- To compute the minimum distance let's determine all the codewords C. All $c \in C$ satisfies $c = aG$ for some vector a.

$$\text{Codewords} \quad \begin{cases} (0,0,0)G = (0,0,0,0,0,0) \\ (1,0,0)G = (1,0,0,1,1,0) \\ (0,1,0)G = (0,1,0,1,0,1) \\ (0,0,1)G = (0,0,1,0,1,1) \\ (1,1,0)G = (1,1,0,0,1,1) \\ (1,0,1)G = (1,0,1,1,0,1) \\ (0,1,1)G = (0,1,1,1,1,0) \\ (1,1,1)G = (1,1,1,0,0,0) \end{cases}$$

  By definition the minimum distance $d_C = min_{c \in C}(w(c)) = 3$.

- Standard array

message row	000	100	010	001	110	101	011	111
codewords	000000	100110	010101	001011	110011	101101	011110	111000
remaining cosets	100000	000110	110101	101011	010011	001101	111110	011000
	010000	110110	000101	011011	100011	111101	001110	101000
	001000	101110	011101	000011	111011	100101	010110	110000
	000100	100010	010001	001111	110111	101001	011100	111010
	000010	100100	010111	001001	110001	101111	011100	111010
	000001	100111	010100	001010	110010	101100	011111	111001
	$\underbrace{100001}_{\text{column of coset leaders}}$	000111	110100	101010	010010	001100	111111	011001

- Let's decode the following vectors

  1. $y = 100111$
     $y - e = 100111 - 000001 = 100110$ which corresponds to 100
  2. $y = 110101$
     $y - e = 110101 - 100000 = 010101$ which corresponds to 010
  3. $y = 001011$. For this case we have $e = 000000$ and y is a codeword that corresponds to message 001.

Response 3:

a) For each set $S_i, i = 1, 2$ let's list the elements $< S_i >$:

$$S_1 = \{010, 011, 111\}, \quad S_2 = \{1010, 0101, 1111\}.$$

For S_1 we can consider the linear combinations of its elements.

$$\begin{aligned}
0(010) + 0(011) + 0(111) &= 000 \\
1(010) + 0(011) + 0(111) &= 010 \\
0(010) + 1(011) + 0(111) &= 011 \\
0(010) + 0(011) + 1(111) &= 111 \\
0(010) + 1(011) + 1(111) &= 100 \\
1(010) + 0(011) + 1(111) &= 101 \\
1(010) + 1(011) + 0(111) &= 001 \\
1(010) + 1(011) + 1(111) &= 110
\end{aligned}$$

So $< S_1 >$ is equal to the whole space $\mathbb{F}_2^3$. For S_2 we note that the elements are linearly dependent. Indeed

$$1010 + 0101 + 1111 = 0000$$

So we consider the two elements 1010 and 0101 of S_2 that are linearly independent

$$\begin{aligned}
0(1010) + 0(0101) &= 0000 \\
0(1010) + 1(0101) &= 0101 \\
1(1010) + 0(0101) &= 1010 \\
1(1010) + 1(0101) &= 1111
\end{aligned}$$

So $< S_2 >= \{0000, 1010, 0101, 1111\}$.

b) Dual code for each of the code C_i. We have $C_i^{\perp} = \{v \in \mathbb{F}_2^3 \ ; \ v \cdot c = 0 \text{ for all } c \in C_i\}$.

- For C_1 since $< S_1 >= \mathbb{F}_2^3$ then $C_1^{\perp} = (0000)$.
- For C_2 we have

$$\begin{cases} (c_1c_2c_3c_4) \cdot (1010) = 0 \\ (c_1c_2c_3c_4) \cdot (0101) = 0 \end{cases} \iff \begin{cases} c_1 + c_3 = 0 \\ c_2 + c_4 = 0 \end{cases}$$

So $C_2^{\perp} = \{c_1c_2c_3c_4 \ ; \ c_1 + c_3 = 0 \text{ and } c_2 + c_4 = 0\}$. Since C_2 is of dimension 2 then $dim(C_2^{\perp}) = 2$ and

$$C_2^{\perp} = \{0000, 1010, 0101, 1111\} = C_2$$

We have a self dual code.

c) Let's prove $C^{\perp}$ is a linear code (independent of whether or not C is linear). Let $a, b \in C^{\perp}, k \in \mathbb{F}_2$ and H the matrix given by

$$H = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

We have $H(a+b)^T = Ha^T + Hb^T = 0$ and $H(ka^T) = kHa^T = 0$. So $C^{\perp}$ is a linear code.

Response 4:

a) For a linear code (n, k) over $\mathbb{F}_q$ we know that the number of codewords M is q^k. Since

$$k + d - 1 \leq n \Longleftrightarrow k \leq n - d + 1$$

we have $M = q^k \leq q^{n-d+1}$ which is a bound in term of n, d and q.

b) Comparison with Hamming bound:
For C a t-error correcting code over $\mathbb{F}_q$ of length n, with M codewords the Hamming bound is

$$M\left(1 + \binom{n}{1}(q-1) + \binom{n}{2}(q-1)^2 + \cdots + \binom{n}{t}(q-1)^t\right) \leq q^n.$$

So for the binary code we have

$$M\left(1 + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{t}\right) \leq q^n \Longleftrightarrow M \leq \frac{2^n}{1 + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{t}}$$

Hence if any bound is better than Hamming bound we have

$$2^{n-d+1} < \frac{2^n}{1 + \sum_{i=0}^{t} \binom{n}{i}} \Longrightarrow \frac{1}{2^{d-1}} < \frac{1}{1 + \sum_{i=1}^{t} \binom{n}{i}}$$
$$\Longrightarrow 2^{d-1} > 1 + \sum_{i=1}^{t} \binom{n}{i}$$

Let's show some pairs of d and n such that our bound is better than Hamming's bound

- Let's take $d = 3$, we know that $t \leq \frac{d-1}{2} \Rightarrow t = 1$. So

$$2^2 > 1 + \binom{n}{1} \Rightarrow 3 > n$$

Thus for $d = 3$ if $n < 3$ our bound is better and if $n > 3$ Hamming's bound is better.

- Let's take $d = 5$. Then $t = 2$ and

$$2^4 > 1+\binom{n}{1}+\binom{n}{2} \Rightarrow 15 > n+\frac{n(n-1)}{2} \Rightarrow 30 > n^2+n \Rightarrow n^2+n-30 < 0 \Rightarrow (n+6)(n-5) < 0$$

Since n is positive we have $n < 5$. For $d = 5$ if $n < 5$ our bound is better and if $n > 5$ Hamming's bound is better.

Response 5:

- We prove the first implication (C can detect s or fewer errors) by contradiction. Suppose that $d_C \leq s$. By hypothesis there is an error with weight $\leq s$. But this error cannot be detected because there are two codewords with distance $\leq s$. This is a contradiction.
- If $d_C \geq s + 1$ and s' errors have occurred with $s' \leq s$ then

$$d(y, y + e) = s' \leq s < s + 1.$$

Since $d_C \geq s + 1$, the syndrome will detect the errors.

Response 6:
To prove the two questions we are going to use the Hamming bound. In a binary case a $t-$error correcting code of length n is perfect if

$$M\left(1 + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{t}\right) = 2^n$$

a) In a repetition code, there are two words $(000\cdots0,\ 111\cdots1)$ so $M = 2$. Moreover, $t = \frac{n-1}{2}$ which is an integer, as n is odd by hypothesis. Hence

$$2\left(1 + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{\frac{n-1}{2}}\right) =$$

$$\left(\binom{n}{n} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{\frac{n-1}{2}}\right)+\left(\binom{n}{\frac{n-1}{2}+1} + \binom{n}{\frac{n-1}{2}+2} + \cdots + \binom{n}{n}\right) = (1+1)^n = 2^n$$

b) In a binary hamming code C_m we have the following properties:

- C_m can correct 1 error $\Rightarrow t = 1$.
- Length $n = 2^m - 1$
- Dimension $k = n - m = 2^m - 1 - m$
- $M = 2^k = 2^{2^m-1-m}$

So by the Hamming bound we get

$$\begin{aligned} 2^{2^m-1-m}\left(1+\binom{n}{1}\right) &= 2^{2^m-1-m}(1+2^m-1) \\ &= 2^{2^m-1-m}\cdot 2^m = 2^{2^m-1} \\ &= 2^n \end{aligned}$$

This shows that the two codes in a) and b) are perfect.

Response 7:

a) To prove this equivalence we will use the definition of an *e-erasure decodable* code C of length n.

- Let's suppose that C is not an e-erasure decodable and show that the minimum distance $d_C < e+1$. If C is not e-erasure decodable we may assure that there is more than one codeword that agrees with received word y in $n-s$ positions (where s is the number of erasures and $s \leq e$). We note the two codewords by c_1 and c_2.

  $$d(c_1, c_2) \leq n-(n-s) = s \Longrightarrow d(c_1, c_2) \leq s \leq e$$

  So $d(c_1, c_2) < e+1$ and the minimum distance of the code $d_C \leq d(c_1, c_2) < e+1$. Thus $d_C < e+1$.
- Now suppose that $d_C < e+1$ and show that C is not $e-$erasure decodable. If $d_C < e+1$ then there exist two codeword c_1 and c_2 such that the distance $d(c_1, c_2) \leq e$. Suppose these codewords differ in exactly $s \leq e$ positions. This means that they must agree in $n-s$positions. To show that this code is not e-erasure decodable, we note that should a codeword y agree with c_1, in $n-s$ positions, then it must also agree with c_2 in $n-s$ positions. But this means that there is not a unique codeword agreeing with y in $n-s$ positions and so the code is not e-erasure decodable.

Response 9:

1. Let's show that the sum S of all elements of a finite field $\mathbb{F}_q$ is zero, except for $\mathbb{F}_2$.

   - For $\mathbb{F}_2$ we have two elements $0, 1$ and $S = 0+1 = 1 \neq 0$.
   - For $\mathbb{F}_q = \mathbb{F}_p$ with p a prime number. The elements are $\{0, 1, \cdots, p-1\}$

   $$S = 0+1+2+3+\cdots p-1 = \frac{p(p-1)}{2} \equiv 0 \pmod{p}$$

- For $\mathbb{F}_q = \mathbb{F}_{p^n}$ with p a prime number and $n \geq 2$ the space $\mathbb{F}_{p^n}$ is given by

$$\mathbb{F}_{p^n} = \{\sum_{i=0}^{n-1} a_i \alpha^i \ ; \ a_i \in \mathbb{F}_p\}$$

where the a_i are called scalars. When we add all coefficients of α^i in $\mathbb{F}_{p^n}$ we get the sum of the elements of $\mathbb{F}_p$ which is equal to 0.

2. Let's show (by induction) that for $a \in \mathbb{F}_q$ and $n \in \mathbb{N}$ the polynomial $x^{q^n} - x + na$ is divisible by $x^q - x + a$ over $\mathbb{F}_q$.

- For $n = 1$ it is trivial
- Let's suppose $x^q - x + a | x^{q^n} - x + na$. Let r be a root of $x^q - x + a$. Then

$$r^q = r - a \Rightarrow (r^q)^{q^n} = (r-a)^{q^n} = r^{q^n} - a^{q^n}$$

Since a is an element of $\mathbb{F}_q$, $a^q = a = a^{q^n}$. By hypothesis we have $r^{q^n} = r - na$. Therefore

$$(r^q)^{q^n} = r - na - a = r - (n+1)a$$

Hence r is a root of $x^{q^{n+1}} - x + (n+1)a$ i.e $x^q - x + a$ divides $x^{q^{n+1}} - x + (n+1)a$

3. Let F be a finite extension of $K = \mathbb{F}_q$ and $\alpha = \beta^q - \beta$ for some $\beta \in F$. let's prove that $\alpha = \gamma^q - \gamma$ with $\gamma \in F$ if and only if $\beta - \gamma \in K$.

- For the first implication we have

$$\begin{aligned} \alpha = \beta^q - \beta = \gamma^q - \gamma &\Longrightarrow \beta^q - \gamma^q = \beta - \gamma \\ &\Longrightarrow (\beta - \gamma)^q = \beta - \gamma \\ &\Longrightarrow \beta - \gamma \in \mathbb{F}_q \end{aligned}$$

- For the second implication we have

$$\begin{aligned} \beta - \gamma \in K &\Longrightarrow (\beta - \gamma)^q = \beta - \gamma \\ &\Longrightarrow \beta^q - \beta = \gamma^q - \gamma \\ &\Longrightarrow \alpha = \gamma^q - \gamma \end{aligned}$$

This γ is in F. Indeed $\gamma = \beta - (\beta - \gamma)$. Since $\beta \in F$ then $\gamma \in F$.