# EE512: Error Control Coding

Solution for Assignment on Linear Block Codes

February 14, 2007

1. Code 1: $n = 4$, $n - k = 2$
   Parity Check Equations: $x_1 + x_3 = 0$, $x_1 + x_2 + x_4 = 0$
   Parity Bits: $x_3 = x_1$, $x_4 = x_1 + x_2$
   $C_1 = \{0000, 0101, 1011, 1110\}$

   Code 2: $n = 4$, $n - k = 2$
   Parity Check Equations: $x_2 + x_3 + x_4 = 0$, $x_1 + x_2 + x_4 = 0$
   Parity Bits: $x_3 = x_1$,$x_4 = x_1 + x_2$,
   Hence $C_2 = C_1$

   Systematic form:
   Generator Matrix:
   $$\mathbf{G} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$
   Parity Check Matrix:
   $$\mathbf{H} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

2. After Gaussian Elimination:
   $$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$
   Blocklength, $n = 8$, Dimension $k = 4$, Minimum Distance $d_{min} = 4$ (found by enumerating all codewords).

3. (a)
   $$\begin{aligned} d_H(\mathbf{u}, \mathbf{v}) &= \{1 \le i \le n : u_i \neq v_i\} \\ &= \{1 \le i \le n : u_i + v_i \neq 0\} \\ &= \mathrm{wt}(\mathbf{u} + \mathbf{v}). \end{aligned}$$

   (b) $\mathrm{wt}(\mathbf{u}) + \mathrm{wt}(\mathbf{v})$ gives the number of ones in $\mathbf{u}$ and $\mathbf{v}$ with the common positions counted twice. Hence, subtract $2\mathrm{wt}(\mathbf{u} * \mathbf{v})$, which is the number of nonzero positions common to both $\mathbf{u}$ and $\mathbf{v}$, from $\mathrm{wt}(\mathbf{u}) + \mathrm{wt}(\mathbf{v})$ to get $d_H(\mathbf{u}, \mathbf{v})$.

   (c) Follows from (3b).

   (d) Since $d_H(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^{n} d(u_i, v_i)$, it is enough to prove
   $$d(u_i, v_i) \le d(u_i, w_i) + d(w_i, v_i). \tag{1}$$
   If $u_i = v_i = w_i$, then (1) evaluates to $0 \le 0$. If $u_i \neq v_i$ and $u_i = w_i$, $v_i \neq w_i$, $1 = 1$. If $u_i \neq v_i$ and $u_i \neq w_i$, $v_i = w_i$, (1) evaluates to $1 \le 1$. If $u_i \neq v_i$ and $u_i \neq w_i$, $v_i \neq w_i$, (1) evaluates to $1 \le 2$. In all cases, inequality is satisified. Hence, $d_H(\mathbf{u}, \mathbf{v}) \le d_H(\mathbf{u}, \mathbf{w}) + d_H(\mathbf{w}, \mathbf{v})$.

(e) From (3d), we know that $\text{wt}(\mathbf{u}+\mathbf{w}) \leq \text{wt}(\mathbf{u}+\mathbf{v}) + \text{wt}(\mathbf{v}+\mathbf{w})$. Let $\mathbf{w} = \mathbf{0}$. It follows $\text{wt}(\mathbf{u}) \leq \text{wt}(\mathbf{u}+\mathbf{v}) + \text{wt}(\mathbf{v})$

4. (a) $(n, 8, 4)$ Code
By Hamming bound, $n \geq 12$. By Singleton bound, $n \geq 11$. GV Bound evaluates to $n \geq 15$. For $n = 11$ or $n = 12$, we cannot construct a $(n-1, 8, 3)$ code with minimum distance 3, since we do not have $n-1$ distinct columns. A $(13, 8, 4)$ code is constructed by extending a $(12, 8, 3)$ code. One sample parity-check matrix is shown below.

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

(b) $(n, 16, 4)$ Code
By Hamming bound, $n \geq 21$. By Singleton bound, $n \geq 19$. GV Bound evaluates to $n \geq 25$. For $n = 19$ or , $n = 20$, we cannot construct a $(n-1, 16, 3)$ code with minimum distance 3, since we do not have $n-1$ distinct columns. A $(22, 16, 4)$ code is constructed by extending a $(21, 16, 3)$ code.

(c) $(n, 32, 4)$ Code
By Hamming bound, $n \geq 38$. By Singleton bound, $n \geq 33$. GV bound evaluates to $n \geq 42$. A $(39, 32, 4)$ code is constructed by extending a $(38, 32, 3)$ code.

5. (a) Dual of the $(n, 1, n)$ repetition code is the $(n, n-1, 2)$ even weight codeword.

(b) Dual of the $(n, n-1, 2)$ even weight codeword is the $(n, 1, n)$ repetition code.

6. $\mathbf{G} = \mathbf{I}_k$. Hence $C = \{0, 1\}^k$.

7. (a) Block Length $n = n_1 + n_2$, Dimension $k$. Any codeword $\mathbf{x}$ can be written as $\mathbf{x} = [\mathbf{x}_1 \mathbf{x}_2]$, where $\mathbf{x}_1 = uG_1$ and $\mathbf{x}_2 = uG_2$. If $u \neq \mathbf{0}$, $\text{wt}(\mathbf{x}) \geq \text{wt}(\mathbf{x}_1) + \text{wt}(\mathbf{x}_2)$. Therefore, minimum distance $d \geq d_1 + d_2$.

(b) Block Length $n = n_1 + n_2$, Dimension $k = k_1 + k_2$, Minimum distance $d = \min(d_1, d_2)$. Let $\mathbf{x}_1$, $\mathbf{x}_2$ be codewords generated by $\mathbf{G}_1$, $\mathbf{G}_2$. The vectors $[\mathbf{x_1 0}]$, and $[\mathbf{0 x_2}]$ are codewords of the new code. Hence, minimum distance $= \min(d_1, d_2)$.

8. Let $C = C_e \cup C_o$, where $C_e$ and $C_o$ are the set of even and odd weight codewords in $C$ respectively. Since $\mathbf{0} \in C_e$ and sum of even weight codewords is an even weight codeword, $C_e$ is a linear code.

Let $\mathbf{c}_1 \in C_o$. Consider the coset of $C_e$ generated by $\mathbf{c}_1$ denoted as $\mathbf{c}_1 + C_e$. For $\mathbf{c}_2 \in C_o$, we see that $\mathbf{c}_1 + \mathbf{c}_2 \in C_e$, which implies that $\mathbf{c}_2 \in \mathbf{c}_1 + C_e$ and $C_o \subseteq \mathbf{c}_1 + C_e$. Also, for $\mathbf{c}_3 \in C_e$, $\mathbf{c}_1 + \mathbf{c}_3 \in C_o$, which implies that $\mathbf{c}_1 + C_e \subseteq C_o$. Thus, $C_e + \mathbf{c}_1 = C_o$. Since, $\mid C_e \mid = \mid C_e + \mathbf{c}_1 \mid$, $\mid C_e \mid = \mid C_o \mid$.

9. Let $C_0$ and $C_1$ be the set of codewords with 0 and 1 in any particular column. $C_0$ is a subspace, since it includes $\mathbf{0}$ and is closed. Let $\mathbf{c}_1, \mathbf{c}_2 \in C_1$. Consider the coset of $C_0$, generated by $\mathbf{c}_1$. We know that $\mathbf{c}_1 + \mathbf{c}_2 \in C_0$. Hence any vector $\mathbf{c}_2 \in C_1$, also belongs to $C_0 + \mathbf{c}_1$. Thus, $C_0 + \mathbf{c}_1 = C_1$. Since, $\mid C_0 \mid = \mid C_0 + \mathbf{c}_1 \mid$, $\mid C_0 \mid + \mid C_1 \mid = 2^k$. Therefore, $\mid C_0 \mid = \mid C_1 \mid = 2^(k-1)$.

10. The parity-check matrix for the code is

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 & \mathbf{0} \\ \mathbf{1} & 1 \end{bmatrix},$$

where $\mathbf{H}_1$ is the parity-check matrix of the (7,4) Hamming code. Note that $\mathbf{H}$ is a parity-check matrix for the (8,4) extended Hamming code. The syndrome is given by $\mathbf{s} = \mathbf{r}\mathbf{H}^T$ for a received

| **s** Syndrome | **e** Error Pattern |
|---|---|
| 0000 | 00000000 |
| 0001 | 00000001 |
| 1111 | 00000010 |
| 0111 | 00000100 |
| 1011 | 00001000 |
| 1101 | 00010000 |
| 1001 | 00100000 |
| 0101 | 01000000 |
| 0011 | 10000000 |
| 1110 | 00000011 |
| 1000 | 00000110 |
| 1100 | 00001100 |
| 0110 | 00011000 |
| 0100 | 00110000 |
| 0010 | 10000001 |
| 1010 | 00010100 |

Table 1: Syndrome table.

vector **r**. The code has a total of 16 cosets represented by the 16 syndromes. Table 1 presents a syndrome table for the maximum likelihood decoder over a BSC.

The probability of codeword error is given by

$$\Pr(\text{Codeword Error}) = 1-\Pr(\mathbf{e} = \text{Correctable Error Vector}) = 1-8p(1-p)^7-7p^2(1-p)^6-(1-p)^8.$$

11. Message Bit Error: In general, finding accurate probability of bit error involves more drudgery than probability of block error. Typically, simple bounds such as

$$\frac{\Pr(\text{block error})}{\text{message length})} \leq \Pr(\text{bit error}) \leq \Pr(\text{block error})$$

are sufficient in practice. We will illustrate the entire calculation for this problem.

The list of codewords is seen to be $C = \{000000, 101001, 111010, 010011, 110100, 011101, 001110, 100111\}$. The sydrome table is given in Table 2.

| **s** Syndrome | **e** Error Pattern |
|---|---|
| 000 | 000000 |
| 001 | 001000 |
| 010 | 010000 |
| 011 | 100010 |
| 100 | 100000 |
| 101 | 000001 |
| 110 | 000100 |
| 111 | 000010 |

Table 2: Syndrome Table.

The syndrome table is not sufficient for our calculation. Each coset needs to be explicitly determined in the form of a standard array as shown in Table 3. Notice that the first row of the standard array is the code itself, and the first column is the list of coset leaders.

Let us assume that the all-zero codeword [000000] is transmitted and the last three bits of a codeword are set to be the message bits. We will compute the average number of message-bit

| Syndrome | Words in coset |
|---|---|
| 000 | $\{000000, 101001, 111010, 010011, 110100, 011101, 001110, 100111\}$ |
| 001 | $\{001000, 100001, 110010, 011011, 111100, 010101, 000110, 101111\}$ |
| 010 | $\{010000, 111001, 101010, 000011, 100100, 001101, 011110, 110111\}$ |
| 011 | $\{100010, 001011, 011000, 110001, 010110, 111111, 101100, 000101\}$ |
| 100 | $\{100000, 001001, 011010, 110011, 010100, 111101, 101110, 000111\}$ |
| 101 | $\{000001, 101000, 111011, 010010, 110101, 011100, 001111, 100110\}$ |
| 110 | $\{000100, 101101, 111110, 010111, 110000, 011001, 001010, 100011\}$ |
| 111 | $\{000010, 101011, 111000, 010001, 110110, 011111, 001100, 100101\}$ |

Table 3: Standard array.

errors in one decoded codeword. More precisely, if $X$ denotes the number of message-bit errors in a decoded codeword, we will compute the PMF of $X$ given that [000000] is the transmitted codeword. Table 4 shows the value of $X$ corresponding to each possiblity for the decoded codeword.

| Decoded codeword | Decoded message vector | $X$ |
|---|---|---|
| 000000 | 000 | 0 |
| 101001 | 001 | 1 |
| 111010 | 010 | 1 |
| 010011 | 011 | 2 |
| 110100 | 100 | 1 |
| 011101 | 101 | 2 |
| 001110 | 110 | 2 |
| 100111 | 111 | 3 |

Table 4: Message-bit error table

From Table 4 and Table 3, we see the following:

$$\begin{aligned}
\Pr(X=0) &= \Pr(\mathbf{e} \text{ is in 1st column of array}),\\
\Pr(X=1) &= \Pr(\mathbf{e} \text{ is in 2nd, 3rd, or 5th column of array}),\\
\Pr(X=2) &= \Pr(\mathbf{e} \text{ is in 4th, 6th or 7th column of array}),\\
\Pr(X=3) &= \Pr(\mathbf{e} \text{ is in 8th column of array}).
\end{aligned}$$

Now,

$$\begin{aligned}
E[X] &= \Pr(X=1) + 2\Pr(X=2) + 3\Pr(X=3),\\
&= (7p^2(1-p)^4 + 8p^3(1-p)^3 + 7p^4(1-p)^2 + 2p^5(1-p))\\
&\quad +2(6p^2(1-p)^4 + 8p^3(1-p)^3 + 7p^4(1-p)^2 + 2p^5(1-p) + p^6)\\
&\quad +3(p^2(1-p)^4 + 4p^3(1-p)^3 + p^4(1-p)^2 + 2p^5(1-p)),\\
&= 22p^2(1-p)^4 + 36p^3(1-p)^3 + 24p^4(1-p)^2 + 12p^5(1-p) + 2p^6.
\end{aligned}$$

Assuming the transmission of each codeword is independent, we can see that the bit-error rate (or expected fraction of message bits in error) will evaluate to $E[X]/k = E[X]/3$.

We could also compute other probabilities using the standard array and the syndrome table. For

instance,

$$\begin{aligned}
\Pr(\text{first message bit is in error}) &= \Pr(\mathbf{e} \text{ is in 5th, 6th, 7th or 8th column of array}), \\
&= 7p^2(1-p)^4 + 12p^3(1-p)^3 + 8p^4(1-p)^2 + 4p^5(1-p) + p^6. \\
\Pr(\text{second message bit is in error}) &= \Pr(\mathbf{e} \text{ is in 3rd, 4th, 7th or 8th column of array}), \\
&= 8p^2(1-p)^4 + 12p^3(1-p)^3 + 8p^4(1-p)^2 + 4p^5(1-p). \\
\Pr(\text{third message bit is in error}) &= \Pr(\mathbf{e} \text{ is in 2nd, 4th, 6th or 8th column of array}), \\
&= 7p^2(1-p)^4 + 12p^3(1-p)^3 + 8p^4(1-p)^2 + 4p^5(1-p) + p^6.
\end{aligned}$$

Notice that the average of the above three probabilities also works out to be equal to $E[X]/3$. Are these two averages always equal? How about for linear codes? What about the conditioning on the all-zero codeword?

12. (a) Parity Check Matrix

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

(b) The parity bits are found using the following equations so that the codeword satisfies $\mathbf{H}[\mathbf{p}\ \mathbf{m}]^T = \mathbf{0}$.

$$\begin{aligned}
x_{12} &= x_5 + x_6 + x_7 + x_8 + x_9 + x_{10} + x_{11} \\
x_{13} &= x_2 + x_3 + x_4 + x_8 + x_9 + x_{10} + x_{11} \\
x_{14} &= x_1 + x_3 + x_4 + x_6 + x_7 + x_{10} + x_{11} \\
x_{15} &= x_1 + x_2 + x_4 + x_5 + x_7 + x_9 + x_{11}
\end{aligned}$$

Using the given message, we get the codeword to be [111111000000100].

(c) Syndrome
$\mathbf{s} = \mathbf{H}\mathbf{r}^T = [1010]^T$. The syndrome equals the tenth column of $\mathbf{H}$ implying that the tenth bit is in error. Therefore, decoded message = [1110001111].

13. To show that $C \cup (\mathbf{u} + C)$ is closed, let $\mathbf{a}, \mathbf{b} \in C \cup (\mathbf{u} + C)$. We consider the following cases:

case 1: $\mathbf{a}, \mathbf{b} \in C$ implies $\mathbf{a} + \mathbf{b} \in C$.
case 2: $\mathbf{a}, \mathbf{b} \in \mathbf{u} + C$ implies $\mathbf{a} + \mathbf{b} \in C$.
case 3: $\mathbf{a} \in C, \mathbf{b} \in \mathbf{u} + C$ implies $\mathbf{a} + \mathbf{b} \in \mathbf{u} + C$.

Hence, $C \cup (\mathbf{u} + C)$ is a linear code.

14. (a) Note that a vector is in $C$ iff it is orthogonal to all vectors in $C^\perp$. Similarly, a vector is in $C^\perp$ iff it is orthogonal to all vectors in $C$.
We need to show that $(C^\perp)^\perp = C$. Consider $x \in C$. By definition $x \cdot y = 0$ for all $y \in C^\perp$; hence, we conclude that $x \in (C^\perp)^\perp$. So, $C \subseteq (C^\perp)^\perp$.
Now consider $u \in (C^\perp)^\perp$. By definition $u \cdot v = 0$ for all $v \in C^\perp$; hence, we conclude that $u \in C$. So, $(C^\perp)^\perp \subseteq C$.
Hence proved.

(b) Similar to above part.

15. (a) Converting $\mathbf{G}$ into systematic form, we get the following:

$$\mathbf{G}_s = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Since $\mathbf{G}_s = \mathbf{G}'$, the two codes are identical.

(b) In systematic form, the matrix $\mathbf{G}'$ becomes

$$\mathbf{G}'_s = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

After a column permutation, we get

$$\mathbf{G}'_{eq} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

Since $\mathbf{G} = \mathbf{G}'_{eq}$, the two codes are equivalent.

16. Let the parity-check matrix be

$$\mathbf{H} = \begin{bmatrix} 1 & a & b & 1 & 0 & 0 & 0 \\ 0 & c & d & 0 & 1 & 0 & 0 \\ 1 & e & f & 0 & 0 & 1 & 0 \\ 1 & g & h & 0 & 0 & 0 & 1 \end{bmatrix},$$

where the variables $a$ through $g$ are bits that need to be determined. Since $\mathbf{H}[0110011]^T = \mathbf{0}$, we get the following equations:

$$\begin{aligned} a + b = 0 &\implies a = b \\ c + d = 0 &\implies c = d \\ e + f = 1 &\implies e \neq f \\ g + h = 1 &\implies g \neq h \end{aligned}$$

For a minimum distance of 4, we need to have atleast 3 linearly independent columns. This implies atleast 3 ones must be present in each of the columns. Therefore the parity check matrix is

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Note that we still need to check that the minimum distance is actually 4!

17. (a) Notice that the generator matrix can be written as $G = [G_1\ G_2]$, where

$$G_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

is a generator matrix of the (5,4,2) even-weight code (why? find its parity-check matrix in systematic form) and $G_2$ is a generator matrix of the (7,4,3) Hamming code (why? find parity-check matrix). By Problem (7a), we see that $d \geq 2 + 3 = 5$. Since the first row of $G$ is a codeword of weight 5, we see that $d = 5$.

Alternatively, by enumerating all 16 codewords, we find $d_{min} = 5$.

(b) By careful inspection, we can quickly see that the codeword closest to [111111111111] will be [111101111111]. Alternatively, we need to calculate the syndrome and find a suitable coset leader!

18. (a) Linear: Impossible
For $\mathbf{a} \in C$, $\mathbf{a} + \mathbf{a} = 0$ has even weight and does not belong to $C$. Hence, the code is nonlinear.

(b) Minimum Distance is 5: Impossible
Let $\mathbf{u}$, $\mathbf{v}$ belong to $C$. $d_H(\mathbf{u}, \mathbf{v}) = \text{wt}(\mathbf{u}) + \text{wt}(\mathbf{v}) - 2\text{wt}(\mathbf{uv})$ = some even number. Hence minimum distance cannot be an odd number.

(c) $C$ is self-orthogonal: Impossible
For a self-orthogonal code, any codeword is orthogonal to itself. For $\mathbf{a} \in C$, $\mathbf{a} \cdot \mathbf{a} = \text{wt}(\mathbf{a}) \mod 2 = 1$. Hence, $C$ cannot be self-orthogonal.