Handout 27 Advanced Digital Communications
Solutions to Homework 11 Dec. 12, 2016

SOLUTION 1.

(a) That $\mathcal{C}^\perp$ is a linear subspace of $\{0,1\}^n$ is obvious: For all $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ in $\mathcal{C}^\perp$, $G(\boldsymbol{y}_1+\boldsymbol{y}_2)^T = G\boldsymbol{y}_1^T + G\boldsymbol{y}_2^T = \mathbf{0}$. Therefore, $\boldsymbol{y}_1 + \boldsymbol{y}_2$ is in $\mathcal{C}^\perp$.

Moreover since G has rank k (otherwise the dimension of the space it spans will be smaller than k), the rank–nullity theorem tells that its null space (which is exactly $\mathcal{C}^\perp$) has dimension $n-k$.

(b) Since $\mathcal{C}^\perp$ is a binary linear block code of dimension k it can be written as $\mathcal{C}^\perp = \{\boldsymbol{x} \in \{0,1\}^n : \boldsymbol{x} = \boldsymbol{u}H : \boldsymbol{u} \in \{0,1\}^{n-k}\}$

(c) Pick $\boldsymbol{x} \in \mathcal{C}$. Then $\forall \boldsymbol{y} \in \mathcal{C}^\perp$, $\boldsymbol{x}\boldsymbol{y}^T = \boldsymbol{u}G\boldsymbol{y}^T = \boldsymbol{u}\mathbf{0}^T = 0$. In particular, since every row of of H is a codeword of $\mathcal{C}^T$, $\boldsymbol{x}H^T = \mathbf{0}$ which implies $H\boldsymbol{x}^T = \mathbf{0}$. This shows

$$\mathcal{C} \subset \{\boldsymbol{x} \in \{0,1\}^n : H\boldsymbol{x}^T = \mathbf{0}\}.$$

The same reasoning as (a) shows that the right-hand-side of the above, which is the null space of H, has dimension k thus 2^k elements. $\mathcal{C}$ has also 2^k elements. Thus we conclude that $\mathcal{C} = \{\boldsymbol{x} \in \{0,1\}^n : H\boldsymbol{x}^T = \mathbf{0}\}$.

SOLUTION 2. Let S_0 be the set of codewords at Hamming distance d from $\boldsymbol{x}_0$ and S_1 be the set of codewords at Hamming distance d from $\boldsymbol{x}_1$. For each $\boldsymbol{y}$ in S_0, note that $\boldsymbol{x}_1 + \boldsymbol{y}$ is at distance d from $\boldsymbol{x}_1$, and thus $\{\boldsymbol{x}_1 + \boldsymbol{y} : \boldsymbol{y} \in S_0\} \subset S_1$. Similarly, since for every $\boldsymbol{y} \in S_1$, $\boldsymbol{x}_1 + \boldsymbol{y}$ has d ones, it is at distance d from $\boldsymbol{x}_0$ (the all-zero codeword) thus, $\{\boldsymbol{x}_1 + \boldsymbol{y} : \boldsymbol{y} \in S_1\} \subset S_0$. These two relationships yield $|S_0| \leq |S_1|$ and $|S_1| \leq |S_0|$, leading to the conclusion that $|S_0| = |S_1|$.

SOLUTION 3.

(a) Note first that the sum of two even-weight codewords is of even weight, the sum of two odd-weight codewords is of even weight and the sum of an odd-weight codeword with an even-weight codeword is of odd weight.

If the code contains no odd-weight codeword then we are done. Otherwise let $\boldsymbol{x}$ be an odd-weight codeword. Then the mapping $\boldsymbol{y} \mapsto \boldsymbol{x}+\boldsymbol{y}$ is a bijection between even-weight and odd-weight codewords, and we conclude that there must be an equal number of odd-weight and even-weight codewords.

(b) The same proof above applies: either all codewords have a zero at the i^{th} digit, or there is a codeword $\boldsymbol{x}$ with has a 1 in its i^{th} digit. The mapping $\boldsymbol{y} \mapsto \boldsymbol{x} + \boldsymbol{y}$ gives a bijection between codewords who have a zero at the i^{th} digit and codewords which have a 1 at the i^{th} digit.

In the first case, when all codewords have a zero at the i^{th} digit, one can improve the code by simply deleting the i^{th} digit from each codeword: no matter what the message is, the same symbol would have been transmitted, giving no additional information.

(c) To find the average number of 1's per codewords, one would find the total number of 1's in all codewords, and divide this sum by the number of codewords. Suppose there are M codewords. Arrange the codewords in rows, and count the total number of 1's by going over columns one by one. Since each column contains at most $M/2$ ones, and there are n columns, the total number of 1's is less than or equal to $Mn/2$. Dividing by M we see that the average number of 1's per codeword is at most $n/2$.

(d) Since the code is proper we know that there is no position $i \in \{1, 2 \ldots, n\}$ for which $x_{m,i} = 0$ for all m. (Recall that $x_{m,i}$ denotes the i^{th} digit of the m^{th} codeword.) Thus, by (b) we know that for every i, half of $x_{m,i}$ are zero and half are one. Consequently, if a codeword is chosen uniformly at random then $x_{m,i}$ will indeed be equally probable to be 0 or 1.

SOLUTION 4.

(a) Since $\mathrm{d_H}(\boldsymbol{x}, \boldsymbol{y}) = \mathrm{w_H}(\boldsymbol{x} - \boldsymbol{y})$ we always have

$$d_{\min}(\mathcal{C}) = \min_{\substack{\boldsymbol{x},\boldsymbol{y}\in\mathcal{C} \\ \boldsymbol{x}\neq\boldsymbol{y}}} \mathrm{w_H}(\boldsymbol{x} - \boldsymbol{y}).$$

For a linear code $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{C}$ imply $\boldsymbol{x} - \boldsymbol{y} \in \mathcal{C}$. Thus,

$$d_{\min}(\mathcal{C}) = \min_{\substack{\boldsymbol{x}\in\mathcal{C} \\ \boldsymbol{x}\neq\boldsymbol{0}}} \mathrm{w_H}(\boldsymbol{x}).$$

(b) Let us do the following assignment: To each codeword, we assign *all* binary sequences that are *no more* than $\lfloor (d_{\min} - 1)/2 \rfloor$ bit flips away. With this assignment, *each binary string* of length n will be assigned to *exactly* one of the codewords since the Hamming distance of any two codewords is at least $d_{\min}$. Note that some binary strings may not be assigned at all because they are too far away from any codeword. Therefore, to each codeword we assign in total

$$\sum_{i=0}^{\lfloor (d_{\min}-1)/2 \rfloor} \binom{n}{i}$$

different binary strings. Since there are M codewords in total, we have

$$M \sum_{i=0}^{\lfloor (d_{\min}-1)/2 \rfloor} \binom{n}{i}$$

different binary strings that are accounted for. Finally, since there are only 2^n different binary strings of length n, we must have

$$M \sum_{i=0}^{\lfloor (d_{\min}-1)/2 \rfloor} \binom{n}{i} \;\leq\; 2^n.$$

(c) We can weaken the lower bound of (a) as

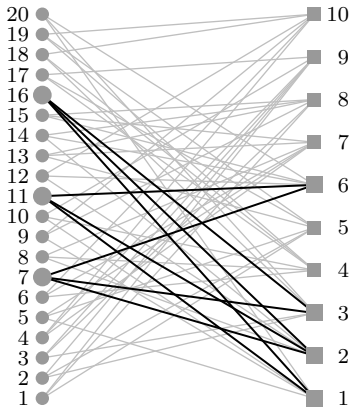
$$\binom{n}{\lfloor (d_{\min} - 1)/2} \leq \frac{2^n}{M}$$

Taking the logarithm of both sides of the above and dividing by n, we have

$$\frac{1}{n}\log_2\binom{n}{\lfloor (d_{\min}-1)/2\rfloor} \leq \frac{n-\log(M)}{n}$$

Using the Stirling approximation (given in the hint) and the fact that $\lim_{n\to\infty}\frac{\lfloor (d_{\min}-1)/2\rfloor}{n} = \delta/2$ the result follows.

SOLUTION 5. Consider the code described by the following Tanner graph:



The parity check-matrix of the code is

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

The set of variable nodes $\mathcal{S} = \{7, 11, 16\}$ forms a stopping set: all check nodes which are connected to $\mathcal{S}$ are connected to $\mathcal{S}$ *at least twice*.

Suppose the channel erases exactly the positions in $\mathcal{S}$, namely the seventh, the eleventh, and the sixteenth bits. Then BP decoder cannot recover those positions because each of the check nodes 1 2, and 6 receives an erasure message from at least two variable nodes.

On the other side, to recover the values of x_7, x_{11}, and x_{16}, the MAP decoder needs to solve the following system of linear equations:

$$H_{\{7,11,16\}}\begin{bmatrix}x_7\\x_{11}\\x_{16}\end{bmatrix} = H_{\sim\{7,11,16\}}\boldsymbol{y}^T_{\sim\{7,11,16\}}.$$

In the above $H_{\{7,11,16\}}$ denotes the 10×3 sub-matrix of H corresponding to its 7^{th}, 11^{th}, and 16^{th} columns, and $H_{\sim\{7,11,16\}}$ denotes its 10×17 sub-matrix obtained by *deleting* the 7^{th}, 11^{th}, and 16^{th} columns. Similarly $\boldsymbol{y}_{\sim\{7,11,16\}}$ denotes the row vector of received bits *excluding* y_7, y_{11}, and y_{16} (which are actually erasures).

Note that x_7, x_{11}, and x_{16} of the sent codeword *is* a solution to the above system (since $\boldsymbol{y}_{\sim\{7,11,16\}} = \boldsymbol{x}_{\sim\{7,11,16\}}$ and $H\boldsymbol{x}^T = \mathbf{0}$). Therefore, the only 'bad' thing that may happen is that the above would be *under-determined* system of equations having more than one solution. We can check that this is not the case. Indeed,

$$H_{\{7,11,16\}} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

is a full-rank matrix: its first three rows are linearly independent (in GF(2)).