Handout 29 Information Theory and Coding
Solutions to Homework 11 Dec. 13, 2021

PROBLEM 1. Recall that the minimum distance is also given by the weight of the minimum weight codeword. Now observe that there exists a codeword x of weight w iff $xH = 0$ where H is the parity-check matrix with n rows. This is equivalent to saying that some w rows of H are linearly dependent. We then know that there exist d rows that are linearly dependent. However, no combination of $d-1$ rows or less are dependent since this case would give rise to a codeword of weight less or equal to $d-1$. This concludes the proof.

PROBLEM 2.

(a) At the first step, we can choose any non-zero column vector with r coordinates. This will be the first row of our $n \times r$ parity-check matrix. Now suppose we have chosen i rows so that no $d-1$ are linearly dependent. They are all non-zero rows. There are at most
$$\binom{i}{1} + \cdots + \binom{i}{d-2}$$
distinct linear combinations of these i rows taken $d-2$ or fewer at a time.

(b) The total number of r-tuples (include the all-zero one) is 2^r. We can then choose a new row different from the previous ones, linearly independent from the previous ones, and keep the property that every $d-1$ rows are independent.

(c) We can iterate the procedure and we keep doing so as long as
$$1 + \binom{i}{1} + \cdots \binom{i}{d-2} < 2^r$$
where the first term counts the all-zero vector. At the last step, we can do so iff
$$1 + \binom{n-1}{1} + \cdots \binom{n-1}{d-2} < 2^r.$$

(d) Multiply both sides of the previous inequality by $M = 2^k$ gives the result since $r = n - k$.

PROBLEM 3. Let S_0 be the set of codewords at Hamming distance n from $\mathbf{x}_0$ and S_1 be the set of codewords at Hamming distance n from $\mathbf{x}_1$. For each $\mathbf{y}$ in S_0, note that $\mathbf{x}_1 + \mathbf{y}$ is at distance n from $\mathbf{x}_1$, and thus $\{\mathbf{x}_1 + \mathbf{y} : \mathbf{y} \in S_0\} \subset S_1$. Similarly, $\{\mathbf{x}_1 + \mathbf{y} : \mathbf{y} \in S_1\} \subset S_0$. These two relationships yield $|S_0| \leq |S_1|$ and $|S_1| \leq |S_0|$, leading to the conclusion that $|S_0| = |S_1|$.

PROBLEM 4.

(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 x be an odd-weight codeword. Then the mapping $y \mapsto x + y$ 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 nth digit, or there is a codeword x with has a 1 in its nth digit. The mapping $y \mapsto x + y$ gives a bijection between codewords who have a zero at the nth digit and codewords which have a 1 at the nth digit. In the first case, when all codewords have a zero at the nth digit, one can improve the code by simply deleting the nth 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 $N/2$.

Problem 5.

(a) Any codeword of $\mathcal{C}$ is of the from $\langle \mathbf{a}, \mathbf{a} \oplus \mathbf{b} \rangle$ with $\mathbf{a} \in \mathcal{C}_1$ and $\mathbf{b} \in \mathcal{C}_2$. Given two codewords $\langle \mathbf{u}', \mathbf{u}' \oplus \mathbf{v}' \rangle$ and $\langle \mathbf{u}'', \mathbf{u}'' \oplus \mathbf{v}'' \rangle$ of $\mathcal{C}$, their sum is $\langle \mathbf{u}, \mathbf{u} \oplus \mathbf{v} \rangle$ with $\mathbf{u} = \mathbf{u}' \oplus \mathbf{u}''$ and $\mathbf{v} = \mathbf{v}' \oplus \mathbf{v}''$. Since $\mathcal{C}_1$ and $\mathcal{C}_2$ are linear codes $\mathbf{u} \in \mathcal{C}_1$ and $\mathbf{v} \in \mathcal{C}_2$. Thus the sum of any two codewords of $\mathcal{C}$ is a codeword of $\mathcal{C}$ and we conclude that $\mathcal{C}$ is linear.

(b) If $(\mathbf{u}, \mathbf{v}) \neq (\mathbf{u}', \mathbf{v}')$, then either $\mathbf{u} \neq \mathbf{u}'$, or, $\mathbf{u} = \mathbf{u}'$ and $\mathbf{v} \neq \mathbf{v}'$. In either case $\langle \mathbf{u} | \mathbf{u} \oplus \mathbf{v} \rangle \neq \langle \mathbf{u}' | \mathbf{u}' \oplus \mathbf{v}' \rangle$: in the first case the first halves differ, in the second case the second halves differ. Thus no two of the $(\mathbf{u}, \mathbf{v})$ pairs are mapped to the same element of $\mathcal{C}$, and the code has exactly $M_1 M_2$ elements. Its rate is $\frac{1}{2n} \log(M_1 M_2) = \frac{1}{2} R_1 + \frac{1}{2} R_2$.

(c) As $\mathbf{v} = \mathbf{u} \oplus \mathbf{u} \oplus \mathbf{v}$,

$$w_H(\mathbf{v}) = w_H(\mathbf{u} \oplus \mathbf{u} \oplus \mathbf{v}) \leq w_H(\mathbf{u}) + w_H(\mathbf{u} \oplus \mathbf{v})$$

by the triangle inequality. Noting that the right hand side is $w_H(\langle \mathbf{u} | \mathbf{u} \oplus \mathbf{v} \rangle)$ completes the proof.

(d) If $\mathbf{v} = \mathbf{0}$ we have $\langle \mathbf{u} | \mathbf{u} \oplus \mathbf{v} \rangle = \langle \mathbf{u} | \mathbf{u} \rangle$ which has twice the Hamming weight of $\mathbf{u}$. Otherwise (c) gives $w_H(\langle \mathbf{u} | \mathbf{u} \oplus \mathbf{v} \rangle) \geq w_H(\mathbf{v})$.

(e) Since $\mathcal{C}$ is linear its minimum distance equals the minimum weight of its non-zero codewords. If $\langle \mathbf{u} | \mathbf{u} \oplus \mathbf{v} \rangle$ is non-zero either $\mathbf{v} \neq \mathbf{0}$, or, $\mathbf{v} = \mathbf{0}$ and $\mathbf{u} \neq \mathbf{0}$. By (d), in the first case $w_H(\langle \mathbf{u} | \mathbf{u} \oplus \mathbf{v} \rangle) \geq w_H(\mathbf{v}) \geq d_1$, in the second case $w_H(\langle \mathbf{u} | \mathbf{u} \oplus \mathbf{v} \rangle) \geq 2 w_H(\mathbf{u}) \geq 2 d_2$. Thus $d \geq \min\{2d_1, d_2\}$.

(f) Let $\mathbf{u}_0$ be the minimum weight non-zero codeword of $\mathcal{C}_1$ and let $\mathbf{v}_0$ be the minimum weight non-zero codeword of $\mathcal{C}_2$. Note that $\langle \mathbf{u}_0 | \mathbf{u}_0 \rangle$ is a non-zero codeword of $\mathcal{C}$ (corresponding to the choice $\mathbf{u} = \mathbf{u}_0$, $\mathbf{v} = \mathbf{0}$). It has weight $2d_1$. Similarly, $\langle \mathbf{0} | \mathbf{v}_0 \rangle$ is also a non-zero codeword of $\mathcal{C}$ (corresponding to the choice $\mathbf{u} = \mathbf{0}$, $\mathbf{v} = \mathbf{v}_0$). It has weight d_2. Consequently $d \leq \min\{2d_1, d_2\}$. In light of (e) we find $d = \min\{2d_1, d_2\}$.

This method of constructing a longer code from two shorter ones is known under several names: 'Plotkin construction', 'bar product', '$(u|u+v)$ construction' appear regularly in the literature. Compare this method to the 'obvious' method of letting the codewords to be $\langle \mathbf{u} | \mathbf{v} \rangle$. The simple method has the same block-length and rate as we have here, but

its minimum distance is only $\min\{d_1, d_2\}$. The factor two gained in d_1 by the bar product is significant, and many practical code families can be built from very simple base codes by a recursive application of the bar product. Notable among them are the family of Reed–Muller codes.