Handout 33 Information Theory and Coding
Solutions to Homework 13 Dec. 17, 2024

PROBLEM 1. (a) Given a code $\mathcal{C}$ with M codewords and blocklength n, and $0 \le k \le n$, partition the codewords into 2^k groups according to their first k bits. The group with the largest number of codewords will contain at least $M' = \lceil M/2^k \rceil$ codewords. The minimum distance within that group is upper bounded by $d_0(M', n-k)$ since all codewords in the group agree in their first k bits. Thus the minimum distance of the code $\mathcal{C}$ is upper bounded by $d_0(\lceil M/2^k \rceil, n-k)$. Since this is true for each $k \in \{0, \dots, n\}$ we conclude that $d_{\min} \le d_1(M, n)$.

(b) With $d_0(M,n) = \begin{cases} n & M \le 2 \\ \infty & M \le 1 \end{cases}$ the minimum over k is obtained by choosing k as large as possible while keeping $M/2^k > 1$. Thus the bound d_1 says "$d_{\min} \le n-k$ when $M > 2^{kn}$" which is the Singleton bound we derived in class.

(c) Each pair (m, m') contributes 1 to the sum when $a_m = 0$ and $a_{m'} = 1$ or when $a_m = 1$ and $a_{m'} = 0$. There are $M_0 M_1$ pairs of the first type and $M_1 M_0$ pairs of the second type. Thus the sum equals $2M_0M_1$. As $M_0 + M_1 = M$, we have $M_0 M_1 \le M^2/4$, from which the final inequality follows.

(d) As $d_H(\mathbf{x}_m, \mathbf{x}_{m'}) \ge d_{\min}$ for every $m \ne m'$, the first inequality follows by summing both sides. For the second write $d_H(\mathbf{x}_m, \mathbf{x}_{m'}) = \sum_{i=1}^n d_H(x_{mi}, x_{m'i})$ to obtain

$$\sum_{m=1}^{M} \sum_{\substack{m'=1 \\ m' \ne m}}^{M} d_H(\mathbf{x}_m, \mathbf{x}_{m'}) = \sum_{i=1}^{n} \sum_{m=1}^{M} \sum_{\substack{m'=1 \\ m' \ne m}}^{M} d_H(x_{mi}, x_{m'i}).$$

By (c) for each i the inner double-sum is upper bounded by $M^2/2$ and the conclusion follows.

PROBLEM 2.

(a) We have

$$\begin{aligned}
W^-(y_1, y_2|u_1) &= \mathbb{P}_{Y_1,Y_2|X_1 \oplus X_2}(y_1, y_2|u_1) = \frac{\mathbb{P}_{Y_1,Y_2,X_1 \oplus X_2}(y_1, y_2, u_1)}{\mathbb{P}_{X_1 \oplus X_2}(u_1)} \\
&\overset{(*)}{=} 2\mathbb{P}_{Y_1,Y_2,X_1 \oplus X_2}(y_1, y_2, u_1) \\
&= 2 \sum_{u_2 \in \{0,1\}} \mathbb{P}_{Y_1,Y_2,X_1 \oplus X_2, X_2}(y_1, y_2, u_1, u_2) \\
&\overset{(**)}{=} 2 \sum_{u_2 \in \{0,1\}} \mathbb{P}_{Y_1,Y_2,X_1,X_2}(y_1, y_2, u_1 \oplus u_2, u_2) \\
&= 2 \sum_{u_2 \in \{0,1\}} \mathbb{P}_{Y_1,Y_2|X_1,X_2}(y_1, y_2|u_1 \oplus u_2, u_2) \mathbb{P}_{X_1,X_2}(u_1 \oplus u_2, u_2) \\
&= 2 \sum_{u_2 \in \{0,1\}} W(y_1|u_1 \oplus u_2) W(y_2|u_2) \frac{1}{2^2} \\
&= \frac{1}{2} \sum_{u_2 \in \{0,1\}} W(y_1|u_1 \oplus u_2) W(y_2|u_2),
\end{aligned}$$

where $(*)$ follows from the fact that if X_1, X_2 are independent and uniform then $X_1 \oplus X_2$ is also uniform. $(**)$ follows from the fact that

$$(X_1 \oplus X_2 = u_1 \text{ and } X_2 = u_2) \Leftrightarrow (X_1 = u_1 \oplus u_2 \text{ and } X_2 = u_2).$$

(b) We have

$$\begin{aligned}
W^+(y_1, y_2, u_1|u_2) = \mathbb{P}_{Y_1,Y_2,X_1\oplus X_2|X_2}(y_1, y_2, u_1|u_2) &= \frac{\mathbb{P}_{Y_1,Y_2,X_1\oplus X_2,X_2}(y_1, y_2, u_1, u_2)}{\mathbb{P}_{X_2}(u_2)} \\
&= 2\mathbb{P}_{Y_1,Y_2,X_1\oplus X_2,X_2}(y_1, y_2, u_1, u_2) \\
&\stackrel{(*)}{=} 2\mathbb{P}_{Y_1,Y_2,X_1,X_2}(y_1, y_2, u_1 \oplus u_2, u_2) \\
&= 2\mathbb{P}_{Y_1,Y_2|X_1,X_2}(y_1, y_2|u_1 \oplus u_2, u_2)\mathbb{P}_{X_1,X_2}(u_1 \oplus u_2, u_2) \\
&= 2W(y_1|u_1 \oplus u_2)W(y_2|u_2)\frac{1}{2^2} \\
&= \frac{1}{2}W(y_1|u_1 \oplus u_2)W(y_2|u_2),
\end{aligned}$$

where $(*)$ follows from the fact that

$$(X_1 \oplus X_2 = u_1 \text{ and } X_2 = u_2) \Leftrightarrow (X_1 = u_1 \oplus u_2 \text{ and } X_2 = u_2).$$

(c) We have

$$\begin{aligned}
Z(W^+) &= \sum_{\substack{y_1,y_2\in\mathcal{Y},\\ u_1\in\{0,1\}}} \sqrt{W^+(y_1, y_2, u_1|0)W^+(y_1, y_2, u_1|1)} \\
&= \frac{1}{2} \sum_{\substack{y_1,y_2\in\mathcal{Y},\\ u_1\in\{0,1\}}} \sqrt{W(y_1|u_1 \oplus 0)W(y_2|0)W(y_1|u_1 \oplus 1)W(y_2|1)} \\
&= \frac{1}{2}\left(\sum_{y_1,y_2\in\mathcal{Y}} \sqrt{W(y_1|0 \oplus 0)W(y_2|0)W(y_1|0 \oplus 1)W(y_2|1)}\right) \\
&\quad + \frac{1}{2}\left(\sum_{y_1,y_2\in\mathcal{Y}} \sqrt{W(y_1|1 \oplus 0)W(y_2|0)W(y_1|1 \oplus 1)W(y_2|1)}\right) \\
&= \frac{1}{2}\left(\sum_{y_1,y_2\in\mathcal{Y}} \sqrt{W(y_1|0)W(y_2|0)W(y_1|1)W(y_2|1)}\right) \\
&\quad + \frac{1}{2}\left(\sum_{y_1,y_2\in\mathcal{Y}} \sqrt{W(y_1|1)W(y_2|0)W(y_1|0)W(y_2|1)}\right) \\
&= \frac{1}{2}\left(\sum_{y_1\in\mathcal{Y}} \sqrt{W(y_1|0)W(y_1|1)}\right)\left(\sum_{y_2\in\mathcal{Y}} \sqrt{W(y_2|0)W(y_2|1)}\right) \\
&\quad + \frac{1}{2}\left(\sum_{y_1\in\mathcal{Y}} \sqrt{W(y_1|0)W(y_1|1)}\right)\left(\sum_{y_2\in\mathcal{Y}} \sqrt{W(y_2|0)W(y_2|1)}\right) \\
&= \frac{1}{2}Z(W) \cdot Z(W) + \frac{1}{2}Z(W) \cdot Z(W) = Z(W)^2.
\end{aligned}$$

(d) For every $y_1, y_2 \in \mathcal{Y}$, we have:

$$\begin{aligned}
W^-(y_1, y_2|0) &= \frac{1}{2} \sum_{u_2 \in \{0,1\}} W(y_1|0 \oplus u_2) W(y_2|u_2) = \frac{1}{2} \sum_{u_2 \in \{0,1\}} W(y_1|u_2) W(y_2|u_2) \\
&= \frac{1}{2} W(y_1|0) W(y_2|0) + \frac{1}{2} W(y_1|1) W(y_2|1) = \frac{1}{2}\alpha(y_1)\alpha(y_2) + \frac{1}{2}\beta(y_1)\beta(y_2) \\
&= \frac{1}{2}(\alpha(y_1)\alpha(y_2) + \beta(y_1)\beta(y_2)),
\end{aligned}$$

and

$$\begin{aligned}
W^-(y_1, y_2|1) &= \frac{1}{2} \sum_{u_2 \in \{0,1\}} W(y_1|1 \oplus u_2) W(y_2|u_2) \\
&= \frac{1}{2} W(y_1|1 \oplus 0) W(y_2|0) + \frac{1}{2} W(y_1|1 \oplus 1) W(y_2|1) \\
&= \frac{1}{2} W(y_1|1) W(y_2|0) + \frac{1}{2} W(y_1|0) W(y_2|1) = \frac{1}{2}\beta(y_1)\alpha(y_2) + \frac{1}{2}\alpha(y_1)\beta(y_2) \\
&= \frac{1}{2}(\alpha(y_1)\beta(y_2) + \beta(y_1)\alpha(y_2)).
\end{aligned}$$

We have

$$\begin{aligned}
Z(W^-) &= \sum_{y_1, y_2 \in \mathcal{Y}} \sqrt{W^-(y_1, y_2|0) W^-(y_1, y_2|1)} \\
&= \frac{1}{2} \sum_{y_1, y_2 \in \mathcal{Y}} \sqrt{\Big(\alpha(y_1)\alpha(y_2) + \beta(y_1)\beta(y_2)\Big)\Big(\alpha(y_1)\beta(y_2) + \beta(y_1)\alpha(y_2)\Big)}.
\end{aligned}$$

(e) For every $x, y \geq 0$, we have $x + y \leq x + y + 2\sqrt{xy} = (\sqrt{x} + \sqrt{y})^2$ which implies that $\sqrt{x+y} \leq \sqrt{x} + \sqrt{y}$. Therefore, for every $x, y, z, t \geq 0$ we have:

$$\sqrt{x+y+z+t} \leq \sqrt{x+y} + \sqrt{z+t} \leq \sqrt{x} + \sqrt{y} + \sqrt{z} + \sqrt{t}.$$

Therefore,

$$\begin{aligned}
&Z(W^-) \\
&\quad = \frac{1}{2} \sum_{y_1, y_2 \in \mathcal{Y}} \sqrt{\Big(\alpha(y_1)\alpha(y_2) + \beta(y_1)\beta(y_2)\Big)\Big(\alpha(y_1)\beta(y_2) + \beta(y_1)\alpha(y_2)\Big)} \\
&\quad = \frac{1}{2} \sum_{y_1, y_2 \in \mathcal{Y}} \sqrt{\alpha(y_1)^2\gamma(y_2)^2 + \alpha(y_2)^2\gamma(y_1)^2 + \beta(y_2)^2\gamma(y_1)^2 + \beta(y_1)^2\gamma(y_2)^2} \\
&\quad \overset{(*)}{\leq} \frac{1}{2} \sum_{y_1, y_2 \in \mathcal{Y}} \left(\sqrt{\alpha(y_1)^2\gamma(y_2)^2} + \sqrt{\alpha(y_2)^2\gamma(y_1)^2} + \sqrt{\beta(y_2)^2\gamma(y_1)^2} + \sqrt{\beta(y_1)^2\gamma(y_2)^2}\right) \\
&\quad = \frac{1}{2}\left(\sum_{y_1, y_2 \in \mathcal{Y}} \alpha(y_1)\gamma(y_2)\right) + \frac{1}{2}\left(\sum_{y_1, y_2 \in \mathcal{Y}} \alpha(y_2)\gamma(y_1)\right) \\
&\qquad + \frac{1}{2}\left(\sum_{y_1, y_2 \in \mathcal{Y}} \beta(y_2)\gamma(y_1)\right) + \frac{1}{2}\left(\sum_{y_1, y_2 \in \mathcal{Y}} \beta(y_1)\gamma(y_2)\right),
\end{aligned}$$

where $(*)$ follows from the inequality $\sqrt{x+y+z+t} \leq \sqrt{x} + \sqrt{y} + \sqrt{z} + \sqrt{t}$.

(f) Note that $\sum_{y\in\mathcal{Y}} \alpha(y) = \sum_{y\in\mathcal{Y}} \beta(y) = 1$ and $\sum_{y\in\mathcal{Y}} \gamma(y) = Z(W)$. Therefore,

$$\begin{aligned}
Z(W^-) &\le \frac{1}{2}\left(\sum_{y_1,y_2\in\mathcal{Y}} \alpha(y_1)\gamma(y_2)\right) + \frac{1}{2}\left(\sum_{y_1,y_2\in\mathcal{Y}} \alpha(y_2)\gamma(y_1)\right) \\
&\quad + \frac{1}{2}\left(\sum_{y_1,y_2\in\mathcal{Y}} \beta(y_2)\gamma(y_1)\right) + \frac{1}{2}\left(\sum_{y_1,y_2\in\mathcal{Y}} \beta(y_1)\gamma(y_2)\right) \\
&= \frac{1}{2}\left(\sum_{y_1\in\mathcal{Y}} \alpha(y_1)\right)\left(\sum_{y_2\in\mathcal{Y}} \gamma(y_2)\right) + \frac{1}{2}\left(\sum_{y_2\in\mathcal{Y}} \alpha(y_2)\right)\left(\sum_{y_1\in\mathcal{Y}} \gamma(y_1)\right) \\
&\quad + \frac{1}{2}\left(\sum_{y_2\in\mathcal{Y}} \beta(y_2)\right)\left(\sum_{y_1\in\mathcal{Y}} \gamma(y_1)\right) + \frac{1}{2}\left(\sum_{y_1\in\mathcal{Y}} \beta(y_1)\right)\left(\sum_{y_2\in\mathcal{Y}} \gamma(y_2)\right) \\
&= \frac{1}{2}1\cdot Z(W) + \frac{1}{2}1\cdot Z(W) + \frac{1}{2}1\cdot Z(W) + \frac{1}{2}1\cdot Z(W) = 2Z(W).
\end{aligned}$$

PROBLEM 3.

(a) We have

$$\begin{aligned}
Q_{i+1} = \sqrt{Z_{i+1}(1-Z_{i+1})} &= \begin{cases} \sqrt{Z_i^2(1-Z_i^2)} & \text{w.p. } 1/2 \\ \sqrt{(2Z_i - Z_i^2)(1-2Z_i+Z_i^2)} & \text{w.p. } 1/2 \end{cases} \\
&= \begin{cases} \sqrt{Z_i^2(1-Z_i)(1+Z_i)} & \text{w.p. } 1/2 \\ \sqrt{(2-Z_i)Z_i(1-Z_i)^2} & \text{w.p. } 1/2 \end{cases} \\
&= \begin{cases} \sqrt{Z_i(1-Z_i)}\sqrt{Z_i(1+Z_i)} & \text{w.p. } 1/2 \\ \sqrt{Z_i(1-Z_i)}\sqrt{(2-Z_i)(1-Z_i)} & \text{w.p. } 1/2 \end{cases} \\
&= \sqrt{Z_i(1-Z_i)} \begin{cases} \sqrt{Z_i(1+Z_i)} & \text{w.p. } 1/2 \\ \sqrt{(2-Z_i)(1-Z_i)} & \text{w.p. } 1/2 \end{cases} \\
&= Q_i \begin{cases} f_1(Z_i) & \text{w.p. } 1/2 \\ f_2(Z_i) & \text{w.p. } 1/2 \end{cases},
\end{aligned}$$

where $f_1(z) = \sqrt{z(z+1)}$ and $f_2(z) = \sqrt{(2-z)(1-z)}$.

(b) We have

$$f_1'(z) = \frac{2z+1}{2\sqrt{z(z+1)}}$$

so

$$\begin{aligned}
f_1''(z) &= \frac{4\sqrt{z(z+1)} - (2z+1)\dfrac{2(2z+1)}{2\sqrt{z(z+1)}}}{\left(2\sqrt{z(z+1)}\right)^2} \\
&= \frac{4z(z+1) - (2z+1)^2}{4\big(z(z+1)\big)^{\frac{3}{2}}} = \frac{-1}{4\big(z(z+1)\big)^{\frac{3}{2}}} \le 0.
\end{aligned}$$

Therefore, f_1 is concave. By noticing that $f_2(z) = f_1(1-z)$, we obtain:

$$\begin{aligned} f_1(z) + f_2(z) &= f_1(z) + f_1(1-z) = 2\left(\frac{1}{2}f_1(z) + \frac{1}{2}f_1(1-z)\right) \\ &\overset{(*)}{\leq} 2f_1\left(\frac{1}{2}z + \frac{1}{2}(1-z)\right) = 2f_1\left(\frac{1}{2}\right) = 2\sqrt{\frac{1}{2}\left(\frac{1}{2}+1\right)} \\ &= 2\sqrt{\frac{1}{2}\cdot\frac{3}{2}} = 2\frac{\sqrt{3}}{2} = \sqrt{3}, \end{aligned}$$

where $(*)$ follows from the concavity of f_1. We have

$$\mathbb{E}\big[Q_{i+1} \bigm| Z_0, \ldots, Z_i\big] = \frac{1}{2}f_1(Z_i)Q_i + \frac{1}{2}f_2(Z_i)Q_i = \frac{1}{2}(f_1(Z_i) + f_2(Z_i))Q_i \leq \rho Q_i,$$

where $\rho = \frac{\sqrt{3}}{2} < 1$.

(c) We will show the claim by induction on $i \geq 0$. For $i = 0$, we have $Z_0 = z_0$ with probability 1. Therefore, $\mathbb{E}Q_0 = \sqrt{z_0(1-z_0)}$.

It is easy to that the function $[0,1] \to \mathbb{R}$ defined by $z \to \sqrt{z(1-z)}$ achieves its maximum at $z = \frac{1}{2}$, and so $\mathbb{E}Q_0 = \sqrt{z_0(1-z_0)} \leq \sqrt{\frac{1}{2}\left(1-\frac{1}{2}\right)} = \frac{1}{2}$. Therefore, the claim is true for $i = 0$.

Now suppose that the claim is true for $i \geq 0$, i.e., $\mathbb{E}Q_i \leq \frac{1}{2}\rho^i$. We have

$$\mathbb{E}Q_{i+1} = \mathbb{E}\left[\mathbb{E}\big[Q_{i+1} \bigm| Z_0, \ldots, Z_i\big]\right] \overset{(*)}{\leq} \mathbb{E}[\rho Q_i] = \rho\mathbb{E}[Q_i] \overset{(**)}{\leq} \rho \cdot \frac{1}{2}\rho^i = \frac{1}{2}\rho^{i+1},$$

where $(*)$ follows from Part (b) and $(**)$ follows from the induction hypothesis. We conclude that $\mathbb{E}Q_i \leq \frac{1}{2}\rho^i$ for every $i \geq 0$.

(d) By noticing that $\delta < z < 1-\delta$ if and only if $z(1-z) > \delta(1-\delta)$, we get:

$$\begin{aligned} \mathbb{P}\big[Z_i \in (\delta, 1-\delta)\big] &= \mathbb{P}\big[Z_i(1-Z_i) > \delta(1-\delta)\big] = \mathbb{P}\big[\sqrt{Z_i(1-Z_i)} > \sqrt{\delta(1-\delta)}\big] \\ &= \mathbb{P}\big[Q_i > \sqrt{\delta(1-\delta)}\big] \overset{(*)}{\leq} \frac{\mathbb{E}Q_i}{\sqrt{\delta(1-\delta)}} \overset{(**)}{\leq} \frac{\rho^i}{2\sqrt{\delta(1-\delta)}}, \end{aligned}$$

where $(*)$ follows from the Markov inequality and $(**)$ follows from Part (c). Now since $\rho < 1$, we have $\dfrac{\rho^i}{2\sqrt{\delta(1-\delta)}} \to 0$ as $i \to \infty$. We conclude that

$$\mathbb{P}\big[Z_i \in (\delta, 1-\delta)\big] \to 0 \text{ as } i \text{ gets large.}$$