**Handout 29**

**Homework 12 Dec. 12, 2017**

**Information Theory and Coding**

---

**Problem 1.** Suppose the alphabet \( \mathcal{X} \) has \( q \) elements and it forms a finite field when equipped with the operations + and \( \cdot \). Let \( \alpha_0, \ldots, \alpha_{m-1} \) be \( m \) distinct elements of \( \mathcal{X} \). We will describe the codewords of a block code \( C \) of length \( n \) \((n \geq m)\) as follows: a sequence \( x = (x_0, \ldots, x_{n-1}) \in \mathcal{X}^n \) is a codeword if and only if

\[
x(\alpha_i) = 0 \quad \text{for every } i = 0, \ldots, m-1
\]

where \( x(D) = x_0 + x_1 D + \cdots + x_{n-1} D^{n-1} \).

(a) Show that the code \( C \) is linear.

(b) Let \( g(D) = \prod_{i=0}^{m-1}(D - \alpha_i) \). Show that \((x_0, \ldots, x_{n-1})\) is a codeword if and only if

\[
x(D) = g(D)h(D), \quad \text{for some } h(D),
\]

and conclude that the code has \( q^{n-m} \) codewords.

Suppose now that the \( \alpha_i \) are have the form \( \alpha_i = \beta^i \), i.e., \( \alpha_0 = 1, \alpha_1 = \beta, \ldots, \alpha_{m-1} = \beta^{m-1} \).

(c) Let \( A \) be the \( n \times m \) matrix

\[
A = \begin{bmatrix}
1 & 1 & 1 & \cdots & 1 \\
1 & \beta & \beta^2 & \cdots & \beta^{m-1} \\
1 & \beta^2 & \beta^4 & \cdots & \beta^{2(m-1)} \\
1 & \beta^3 & \beta^6 & \cdots & \beta^{3(m-1)} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & \beta^{n-1} & \beta^{2(n-1)} & \cdots & \beta^{(n-1)(m-1)}
\end{bmatrix}
\]

Show that the columns of \( A \) are linearly independent.

**Hint:** Suppose they were dependent so that there is a column vector \( u = [u_0 \ u_1 \ \ldots \ u_{m-1}]^T \) such that \( Au = 0 \). How many roots does \( u(D) \) have?

(d) Show that the code has minimum distance \( d = m+1 \).

**Hint:** Part (c) says that the rank of the matrix \( A \) is \( m \).

**Problem 2.** Let \( h_2(p) = -p \log p - (1-p) \log (1-p) \) denote the binary entropy function defined on the interval \( [0, \frac{1}{2}] \). Note that on this interval \( h_2 \) is a bijection, so its inverse \( h_2^{-1} : [0, 1] \to [0, \frac{1}{2}] \) is well defined. Define \( p + q = p(1-q) + q(1-p) \) and let \( \oplus \) be the XOR operation. Suppose \( X_1 \) and \( X_2 \) are two binary independent random variables with \( H(X_1) = h_2(p_1), \ H(X_2) = h_2(p_2) \), where \( 0 \leq p_1, p_2 \leq \frac{1}{2} \).

(a) Show that \( H(X_1 \oplus X_2) = h_2(p_1 \cdot p_2) \).

(b) Suppose that \( (X_1, Y) \) is independent of \( X_2 \), where \( Y \) is a random variable in \( \mathcal{Y} \). For every \( y \in \mathcal{Y} \), let \( 0 \leq p_1(y) \leq \frac{1}{2} \) be such that \( H(X_1|Y = y) = h_2(p_1(y)) \). We again assume that \( H(X_2) = h_2(p_2) \) and \( 0 \leq p_2 \leq \frac{1}{2} \).

Show that \( H(X_1|Y) = \sum_y h_2(p_1(y))q(y), \ H(X_1 \oplus X_2|Y) = \sum_y h_2(p_2 \cdot p_1(y))q(y), \) where \( q(y) = \mathbb{P}_Y(y) \) for every \( y \in \mathcal{Y} \).
(c) Show that for every $0 \leq p_2 \leq \frac{1}{2}$, the mapping $f : [0, 1] \rightarrow \mathbb{R}$ defined as $f(h) = h_2(p_2 * h_2^{-1}(h))$ is convex.

*Hint:* The graph of $f(h)$ can be drawn by the parametric curve $p \rightarrow (h_2(p), h_2(p_2 * p))$ so it is enough to show that $p \rightarrow \frac{\partial}{\partial p} h_2(p_2 * p)$ is increasing in $0 \leq p \leq \frac{1}{2}$.

(d) Suppose $H(X_1|Y) = h_2(p_1)$, $H(X_2) = h_2(p_2)$. Show that $H(X_1 \oplus X_2|Y) \geq h(p_1 * p_2)$.

(e) Suppose $(X_1, Y_1)$ is independent of $(X_2, Y_2)$ and $H(X_1|Y_1) = h_2(p_1)$, $H(X_2|Y_2) = h_2(p_2)$. Show that $H(X_1 \oplus X_2|Y_1, Y_2) \geq h(p_1 * p_2)$.

**Problem 3.** Suppose $C_1$ and $C_2$ are binary linear codes of block-length $n$. Denote the number of codewords of $C_i$ by $M_i$ and the minimum distance of $C_i$ by $d_i$. For $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$ let $\langle u|v \rangle$ denote the concatenation of the two sequences, i.e.,

$$\langle u|v \rangle = (u_1, \ldots, u_n, v_1, \ldots, v_n).$$

Let $C$ denote the binary code of block-length $2n$ obtained from $C_1$ and $C_2$ as follows:

$$C = \{ \langle u|u \oplus v \rangle : u \in C_1, v \in C_2 \}.$$

(a) Is $C$ a linear code?

(b) How many codewords does $C$ have? Carefully justify your answer. What is the rate $R$ of $C$ in terms of the rates $R_1$ and $R_2$ of the codes $C_1$ and $C_2$?

(c) Show that the Hamming weight of $\langle u|u \oplus v \rangle$ satisfies

$$w_H(\langle u|u \oplus v \rangle) \geq w_H(v).$$

(d) Show that the Hamming weight of $\langle u|u \oplus v \rangle$ satisfies

$$w_H(\langle u|u \oplus v \rangle) \geq \begin{cases} w_H(v) & \text{if } v \neq 0 \\ 2w_H(u) & \text{else.} \end{cases}$$

(e) Show that the minimum distance $d$ of $C$ satisfies

$$d \geq \min\{2d_1, d_2\}.$$  

(f) Show that $d = \min\{2d_1, d_2\}$.

**Problem 4.** Let $W : \{0, 1\} \rightarrow \mathcal{Y}$ be a channel where the input is binary and where the output alphabet is $\mathcal{Y}$. The Bhattacharyya parameter of the channel $W$ is defined as

$$Z(W) = \sum_{y \in \mathcal{Y}} \sqrt{W(y|0)W(y|1)}.$$ 

Let $X_1, X_2$ be two independent random variables uniformly distributed in $\{0, 1\}$ and let $Y_1$ and $Y_2$ be the output of the channel $W$ when the input is $X_1$ and $X_2$ respectively, i.e.,

$${\mathbb{P}}_{Y_1,Y_2|X_1,X_2}(y_1,y_2|x_1,x_2) = W(y_1|x_1)W(y_2|x_2).$$

Define the channels $W^- : \{0, 1\} \rightarrow \mathcal{Y}^2$ and $W^+ : \{0, 1\} \rightarrow \mathcal{Y}^2 \times \{0, 1\}$ as follows:

- $W^-(y_1, y_2|u_1) = \mathbb{P}[Y_1 = y_1, Y_2 = y_2|X_1 \oplus X_2 = u_1]$ for every $u_1 \in \{0, 1\}$ and every $y_1, y_2 \in \mathcal{Y}$, where $\oplus$ is the XOR operation.
• \( W^+(y_1, y_2, u_1|u_2) = \mathbb{P}[Y_1 = y_1, Y_2 = y_2, X_1 \oplus X_2 = u_1|X_2 = u_2] \) for every \( u_1, u_2 \in \{0, 1\} \) and every \( y_1, y_2 \in \mathcal{Y} \).

(a) Show that \( W^-(y_1, y_2|u_1) = \frac{1}{2} \sum_{u_2 \in \{0, 1\}} W(y_1|u_1 \oplus u_2) W(y_2|u_2) \).

(b) Show that \( W^+(y_1, y_2, u_1|u_2) = \frac{1}{2} W(y_1|u_1 \oplus u_2) W(y_2|u_2) \).

(c) Show that \( Z(W^+) = Z(W)^2 \).

For every \( y \in \mathcal{Y} \) define \( \alpha(y) = W(y|0), \beta(y) = W(y|1) \) and \( \gamma(y) = \sqrt{\alpha(y)\beta(y)} \).

(d) Show that
\[
Z(W^-) = \sum_{y_1, y_2 \in \mathcal{Y}} \frac{1}{2} \sqrt{(\alpha(y_1)\alpha(y_2) + \beta(y_1)\beta(y_2)) (\alpha(y_1)\beta(y_2) + \beta(y_1)\alpha(y_2))}.
\]

(e) Show that for every \( x, y, z, t \geq 0 \) we have \( \sqrt{x + y + z + t} \leq \sqrt{x} + \sqrt{y} + \sqrt{z} + \sqrt{t} \).

Deduce that
\[
Z(W^-) \leq \frac{1}{2} \left( \sum_{y_1, y_2 \in \mathcal{Y}} \alpha(y_1)\gamma(y_2) \right) + \frac{1}{2} \left( \sum_{y_1, y_2 \in \mathcal{Y}} \alpha(y_2)\gamma(y_1) \right) + \frac{1}{2} \left( \sum_{y_1, y_2 \in \mathcal{Y}} \beta(y_1)\gamma(y_2) \right) + \frac{1}{2} \left( \sum_{y_1, y_2 \in \mathcal{Y}} \beta(y_2)\gamma(y_1) \right). \tag{1}
\]

(f) Show that every sum in (1) is equal to \( Z(W) \). Deduce that \( Z(W^-) \leq 2Z(W) \).

**Problem 5.** For a given value \( 0 \leq z_0 \leq 1 \), define the following random process:

\[
Z_0 = z_0, \quad Z_{i+1} = \begin{cases} Z_i^2 & \text{with probability } 1/2 \\ 2Z_i - Z_i^2 & \text{with probability } 1/2 \end{cases} \quad i \geq 0,
\]

with the sequence of random choices made independently. Observe that the \( Z \) process keeps track of the polarization of a Binary Erasure Channel with erasure probability \( z_0 \) as it is transformed by the polar transform: \( \mathbb{P}(Z_i = z) \) is exactly the fraction of Binary Erasure Channels having an erasure probability \( z \) among the \( 2^i \) BEC channels which are synthesized by the polar transform at the \( i \)th level. The aim of this problem is to prove that for any \( \delta > 0 \), \( \mathbb{P}[Z_i \in (\delta, 1-\delta)] \to 0 \) as \( i \) gets large.

(a) Define \( Q_i = \sqrt{Z_i(1-Z_i)} \). Find \( f_1(z) \) and \( f_2(z) \) so that

\[
Q_{i+1} = Q_i \times \begin{cases} f_1(Z_i) & \text{with probability } 1/2 \\ f_2(Z_i) & \text{with probability } 1/2 \end{cases},
\]

(b) Show that \( f_1(z) + f_2(z) \leq \sqrt{3} \). Based on this, find a \( \rho < 1 \) so that

\[
\mathbb{E}[Q_{i+1} \mid Z_0, \ldots, Z_i] \leq \rho Q_i.
\]

(c) Show that, for the \( \rho \) you found in (b), \( \mathbb{E}[Q_i] \leq \frac{1}{2} \rho^i \).

(d) Show that

\[
\mathbb{P}[Z_i \in (\delta, 1-\delta)] = \mathbb{P}[Q_i > \sqrt{\delta(1-\delta)}] \leq \frac{\rho^i}{2\sqrt{\delta(1-\delta)}}.
\]

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