PROBLEM 1. Consider a two-way communication system where two parties communicate via a common output they both can observe and influence. Denote the common output by \( Y \), and the signals emitted by the two parties by \( x_1 \) and \( x_2 \) respectively. Let \( p(y|x_1, x_2) \) model the memoryless channel through which the two parties influence the output.

We will consider feedback-free block codes, i.e., we will use encoding and decoding functions of the form

\[
\begin{align*}
\text{enc}_1 : \{1, \ldots, 2^{nR_1}\} &\to \mathcal{X}_1^n \\
\text{dec}_1 : \mathcal{Y}^n \times \{1, \ldots, 2^{nR_1}\} &\to \{1, \ldots, 2^{nR_2}\} \\
\text{enc}_2 : \{1, \ldots, 2^{nR_2}\} &\to \mathcal{X}_2^n \\
\text{dec}_2 : \mathcal{Y}^n \times \{1, \ldots, 2^{nR_2}\} &\to \{1, \ldots, 2^{nR_1}\}
\end{align*}
\]

with which the parties encode their own message and decode the other party’s messages. (Note that when a party is decoding the other party’s message, it can make use of the knowledge of its own message).

We will say that the rate pair \((R_1, R_2)\) is achievable, if for any \( \epsilon > 0 \), there exist encoders and decoders with the above form for which the average error probability is less than \( \epsilon \).

Consider the following ‘random coding’ method to construct the encoders:

(i) Choose probability distributions \( p_j \) on \( \mathcal{X}_j \), \( j = 1, 2 \).

(ii) Choose \( \{\text{enc}_i(m_i) : m_i = 1, \ldots, 2^{nR_i}, i = 1, \ldots, n\} \) i.i.d., each having distribution as \( p_1 \). Similarly, choose \( \{\text{enc}_2(m_2) : m_2 = 1, \ldots, 2^{nR_2}, i = 1, \ldots, n\} \) i.i.d., each having distribution as \( p_2 \), independently of the choices for \( \text{enc}_1 \).

For the decoders we will use typicality decoders:

(i) Set \( p(x_1, x_2, y) = p_1(x_1)p_2(x_2)p(y|x_1, x_2) \). Choose a small \( \epsilon > 0 \) and consider the set \( T \) of \( \epsilon \)-typical \( (x_1^n, x_2^n, y^n) \)'s with respect to \( p \).

(ii) For decoder 1: given \( y^n \) and the correct \( m_1 \), \( \text{dec}_1 \) will declare \( \hat{m}_2 \) if it is the unique \( m_2 \) for which \( (\text{enc}_1(m_1), \text{enc}_2(m_2), y^n) \in T \). If there is no such \( m_2 \), \( \text{dec}_1 \) outputs 0. (Similar description applies to Decoder 2.)

(a) Given that \( m_1 \) and \( m_2 \) are the transmitted messages, show that \( (\text{enc}_1(m_1), \text{enc}_2(m_2), Y^n) \in T \) with high probability.

(b) Given that \( m_1 \) and \( m_2 \) are the transmitted messages, and \( \hat{m}_1 \neq m_1 \) what is the probability distribution of \( (\text{enc}_1(\hat{m}_1), \text{enc}_2(m_2), Y^n) \)?

(c) Under the assumptions in (b) show that the \( \Pr\{(\text{enc}_1(\hat{m}_1), \text{enc}_2(m_2), Y^n) \in T\} \leq 2^{-nI(X_1; X_2 Y)} \).

(d) Show that all rate pairs satisfying

\[
R_1 \leq I(X_1; YX_2), \quad R_2 \leq I(X_2; YX_1)
\]

for some \( p(x_1, x_2) = p(x_1)p(x_2) \) are achievable.
(e) For the case when $X_1, X_2, Y$ are all binary and $Y$ is the product of $X_1$ and $X_2$, show that the achievable region is strictly larger than what we can obtain by ‘half duplex communication’ (i.e., the set of rates that satisfy $R_1 + R_2 \leq 1$.)

**Problem 2.** Suppose we are told that for any $n$ and $M$, for any binary code with blocklength $n$, with $M$ codewords, the minimum distance $d_{\text{min}}$ satisfies $d_{\text{min}} \leq d_0(M, n)$ where $d_0$ is a specified upper bound on minimum distance.

(a) Show that any upper bound $d_0$ can be improved to the following upper bound: for any $n, M$, for any binary code with blocklength $n$ with $M$ codewords

$$d_{\text{min}} \leq d_1(M, n)$$

where $d_1(M, n) = \min_{k \leq n} d_0([M/2^k], n - k)$.

(b) Consider the trivial bound

$$d_0(M, n) = \begin{cases} n, & M \geq 2 \\ \infty, & M \leq 1 \end{cases}$$

What is the bound $d_1$ constructed via (a) for this $d_0$?

(c) Suppose we are given a binary code with $M$ words of blocklength $n$. Fix $1 \leq i \leq n$ and let $a_1, \ldots, a_M$ be the $i$th bits if the $M$ codewords. Suppose $M_1$ of the $a_m$’s are ’1’ and $M_0$ of them are ’0’. Show that

$$\sum_{m=1}^{M} \sum_{m' = 1 \atop m' \neq m}^{M} d_H(a_m, a'_m) = 2M_0M_1 \leq M^2/2.$$ 

(d) Show that for any binary code with $M \geq 2$ codewords $x_1, \ldots, x_M$ of blocklength $n$

$$M(M - 1)d_{\text{min}} \leq \sum_{m=1}^{M} \sum_{m' = 1 \atop m' \neq m}^{M} d_H(x_m, x_{m'}) \leq nM^2/2;$$

consequently, $d_{\text{min}} \leq \lfloor \frac{1}{2} n \frac{M}{M - 1} \rfloor$.

**Problem 3.** 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., $P_{Y_1, Y_2|X_1, X_2}(y_1, y_2|x_1, x_2) = W(y_1|x_1)W(y_2|x_2)$. Define the channels $W^- : \{0, 1\} \rightarrow \mathcal{Y}^2$ and $W^+ : \{0, 1\} \rightarrow \mathcal{Y}^2 \times \{0, 1\}$ as follows:

- $W^-(y_1, y_2|u_1) = P[Y_1 = y_1, Y_2 = y_2|X_1 \oplus X_2 = u_1]$ for every $u_1 \in \{0, 1\}$ and every $y_1, y_2 \in \mathcal{Y}$, where $\oplus$ is the XOR operation.
\( W^+(y_1, y_2, u_1|u_2) = P[Y_1 = y_1, Y_2 = y_2, X_1 \oplus X_2 = u_1|X_2 = u_2] \) 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
\[
Z(W^-) = \sum_{y_1, y_2 \in \mathcal{Y}} \frac{1}{2} \sqrt{\left(\alpha(y_1)\alpha(y_2) + \beta(y_1)\beta(y_2)\right)\left(\alpha(y_1)\beta(y_2) + \beta(y_1)\alpha(y_2)\right)}.
\]

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

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

**Problem 4.** 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: \( P(Z_i = z) \) 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 \( \delta > 0 \), \( P[Z_i \in (\delta, 1-\delta)] \to 0 \) 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

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

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

(d) Show that

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

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