PROBLEM 1 (RANDOM CODING). Recall that the Binary Symmetric Channel with crossover probability $p \leq 1/2$, in other words BSC($p$), is a discrete memoryless channel with input alphabets $\mathcal{X} = \mathcal{Y} = \{0, 1\}$. The transition probabilities of a BSC($p$) are given by
\[
W(y = 0|x = 0) = W(y = 1|x = 1) = 1 - p, \\
W(y = 1|x = 0) = W(y = 0|x = 1) = p.
\]

In this problem we consider a random codebook as in the lectures. Let the set of messages be $\mathcal{U}$, $|\mathcal{U}| = 2^{nR}$. The codebook is constructed by assigning to each message $u \in \mathcal{U}$, a binary string $C_n(u)$ sampled uniformly at random from $\{0, 1\}^n$. To send message $u$, we transmit $X^n = C_n(u)$ through BSC($p$) and the receiver observes a binary string $Y^n$. The decoder $\text{dec}_n : \{0, 1\}^n \rightarrow \mathcal{U}$ operates as follows:

1. It constructs a set of candidate messages $\hat{\mathcal{U}}(Y^n) = \arg\min_{u \in \mathcal{U}} d_h(Y^n, C_n(u))$, where $d_h(., .)$ is the Hamming distance, i.e., $\hat{U}(Y^n)$ consists of codewords at minimal distance to $Y^n$.

2. Depending on the size of $\hat{\mathcal{U}}(Y^n)$:
   - If $|\hat{\mathcal{U}}(Y^n)| = 1$, return the only element of $\hat{\mathcal{U}}(Y^n)$.
   - If $|\hat{\mathcal{U}}(Y^n)| > 1$, return an element uniformly at random from $\hat{\mathcal{U}}(Y^n)$.

We will now derive an upper bound to the average error probability for this setting. For this purpose, let $U$ be a uniform random variable in $\mathcal{U}$.

a) Show that the following holds:
\[
\Pr\left(u \notin \hat{\mathcal{U}}(Y^n) \mid d_h(Y^n, C_n(u)) = t, \ U = u\right) \leq \sum_{i=0}^{t-1} \binom{n}{i} 2^{-n(1-R)}
\]

b) Show that the following holds:
\[
\Pr\left(u \in \hat{\mathcal{U}}(Y^n), \ \text{dec}_n(Y^n) \neq u \mid d_h(Y^n, C_n(u)) = t, \ U = u\right) \leq \frac{1}{2} \binom{n}{t} 2^{-n(1-R)}
\]

c) Show that there exists a code $C_n$ such that
\[
\Pr(\text{dec}_n(Y^n) \neq U) \leq \sum_{t=0}^{n} \binom{n}{t} p^t(1-p)^{n-t} \min \left\{ 1, \sum_{i=0}^{t-1} \binom{n}{i} 2^{-n(1-R)} + \frac{1}{2} \binom{n}{t} 2^{-n(1-R)} \right\}.
\]
d) Prove that for \( t \leq n/2 \)
\[
\sum_{i=0}^{t} \binom{n}{i} \leq 2^{n h_2(t/n)}
\]
where \( h_2(p) \triangleq p \log \left( \frac{1}{p} \right) + (1 - p) \log \left( \frac{1}{1-p} \right) \) is the binary entropy function. [Hint: Consider the binomial expansion of \((\rho + \bar{\rho})^n\), where \( \rho = t/n \) and \( \bar{\rho} \triangleq 1 - \rho \).]

e) Show that
\begin{enumerate}[label=(i),ref=(i)]
\item for any \( q < p \),
\[
\sum_{i=0}^{[nq]} \binom{n}{i} p^i (1-p)^{n-i} \leq 2^{-n D_2(q\|p)}
\]
\item for any \( q > p \),
\[
\sum_{i=[nq]+1}^{n} \binom{n}{i} p^i (1-p)^{n-i} \leq 2^{-n D_2(q\|p)}
\]
\end{enumerate}
where \( D_2(q\|p) \triangleq q \log \left( \frac{q}{p} \right) + (1 - q) \log \left( \frac{1-q}{1-p} \right) \).

Hint: The following expansion might be useful:
\[
p^i (1-p)^{n-i} = \left( \frac{p}{q} \right)^i \left( \frac{1-p}{1-q} \right)^{n-i} q^i (1-q)^{n-i}
\]

f) Recall that the capacity of a BSC\((p)\) is given by \( C = 1 - h_2(p) \). Using the bound in part (c), show that for a sequence of codes \( \{C_n\} \) such that \( R < C \),
\[
\Pr(\text{dec}_n(Y^n) \neq U) \xrightarrow{n \to \infty} 0.
\]

[Hint: For \( R < C \), show that there is a \( q > p \) such that \( 1 - h_2(q) > R \). Then split the outer sum in part (c) into two parts as \( \sum_{t=0}^{[nq]} \) and \( \sum_{t=[nq]+1}^{n} \). Upper bound the first sum by using part (d) and the second sum by using part (e).]
**Problem 2 (Feinstein’s Theorem).** In homework 7 problem 4, we have shown for a discrete memoryless channel with input and output alphabets $X, Y$, and channel transition probabilities $W(y|x)$, there exists a code with length $n$, rate $R < I(X;Y)$ and maximal probability of error $\epsilon$ such that for any $\delta > 0$

$$\epsilon \leq \Pr(i(X^n_1;Y^n_1) < n(R + \delta)) + 2^{-n\delta}$$

where the input $X$ has distribution $P, i(x; y) = \log_2 \frac{W(y|x)}{W(y)}$ and $W(y) = \sum_{x \in X} W(y|x) P(x)$.

We will specialize this bound for the case of BSC($p$) described in Problem 1, with $P_X(0) = P_X(1) = \frac{1}{2}$.

a) We will define the Bhattacharya parameter of the channel as $Z(W) = \sum_y \sqrt{W(y|0)W(y|1)}$ and the dispersion of the channel as $V(W) = \text{Var}(i(X;Y))$. For a BSC($p$), show that

$$Z(W) = 2\sqrt{p(1-p)}, \quad V(W) = p(1-p) \log_2 \frac{1-p}{p}.$$  

We will refer to $Z(W)$ as $Z_p$ and $V(W)$ as $V_p$.

b) Show that for a BSC($p$) the following holds:

$$\Pr(i(X^n_1;Y^n_1) < n(R + \delta)) \leq \exp \left(-n \frac{(I(X;Y) - R - \delta)^2 Z_p^2}{2V_p} \right)$$

[Hint : Use Hoeffding’s inequality.]

c) Prove the following inequality:

$$\epsilon \leq \min_{0 < \gamma < 1} \exp \left(-n \frac{(\gamma Z_p(I(X;Y) - R))^2}{2V_p} \right) + \exp \left(-n \frac{(1 - \gamma)(I(X;Y) - R)}{\log_2 \epsilon} \right)$$
Problem 3 (Strong Converse). We still consider the problem of reliable transmission on BSC\((p)\) with \(p < 1/2\). Let us consider any pair of encoder \(\text{enc} : \mathcal{U} \rightarrow \{0, 1\}^n\) and decoder \(\text{dec} : \{0, 1\}^n \rightarrow \mathcal{U}\) where \(\mathcal{U}\) is the alphabet of the message, with \(|\mathcal{U}| = 2^{nR}\). We assume that the transmitted message \(U\) is chosen uniformly at random from \(\mathcal{U}\).

Define the decoding regions for each message \(u \in \mathcal{U}\) as \(\mathcal{Y}(u) = \{y^n \in \{0, 1\}^n : \text{dec}(y^n) = u\}\). Note that these regions are disjoint and form a partition of \(\{0, 1\}^n\), i.e., \(\mathcal{Y}(u) \cap \mathcal{Y}(u') = \emptyset\) for all \(u, u' \in \mathcal{U}, u \neq u'\), and \(\bigcup_{u \in \mathcal{U}} \mathcal{Y}(u) = \{0, 1\}^n\).

We want to show that if \(R > C\), with \(C = 1 - h_2(p)\), then the average probability of \text{correct} decoding decays to 0 exponentially fast as \(n \rightarrow \infty\).

a) Let us define the \(q\)-hamming ball of message \(u\) as
\[
B_{q,n}(u) = \{y^n \in \{0, 1\}^n : d_h(y^n, \text{enc}(u)) \leq nq\}.
\]
Show that if \(q < p\) then for all \(u \in \mathcal{U}\)
\[
W^n(\mathcal{Y}(u) \cap B_{q,n}(u) | U = u) \leq 2^{-nD_2(q||p)}.
\]
[Hint : Use the result of Problem 1, part (e).]

b) Show that for every \(y^n \notin B_{q,n}(u)\)
\[
W^n(y^n | U = u) \leq p^{nq}(1 - p)^{n(1-q)}.
\]

c) Show that
\[
P(\text{dec}(Y^n) = u | U = u) \leq 2^{-nD_2(q||p)} + \frac{|\mathcal{Y}(u)|}{2^n} 2^n(1 - h_2(q) - D_2(q||p)).
\]
[Hint : Split the decoding region \(\mathcal{Y}(u)\) into \(\mathcal{Y}(u) \cap B_{q,n}(u)\) and \(\mathcal{Y}(u) \cap B_{q,n}^c(u)\).]

d) Show that
\[
P(\text{dec}(Y^n) = U) \leq 2^{-nD_2(q||p)} + 2^{-n(R - 1 + h_2(q) + D_2(q||p))}.
\]

e) Justify that for any \(R > C\) there exists a \(q < p\) such that
\[
C < 1 + q \log_2 p + (1 - q) \log_2(1 - p) < R.
\]

f) Show that if \(R > C\) then
\[
\lim_{n \rightarrow \infty} P(\text{dec}(Y^n) \neq U) = 1.
\]
[Hint: Combine the results of (d) and (e).]