Problem 1

Channel Coding for a “Double”-Channel

Consider two memoryless channels $W^{(1)} (y^{(1)}|x)$ and $W^{(2)} (y^{(2)}|x)$ defined over a common input alphabet $\mathcal{X}$ and over possibly different output alphabets $\mathcal{Y}^{(1)}, \mathcal{Y}^{(2)}$. Let $Q(\cdot)$ be some input distribution on $\mathcal{X}$ and let

$$ R < \min \{ I(Q, W^{(1)}), I(Q, W^{(2)}) \}, $$

so that $R$ is smaller than the mutual information corresponding to the input distribution $Q(\cdot)$ on each channel. Given any $\epsilon > 0$, prove that for every sufficiently large blocklength $n$ there exists a rate-$R$ codebook $\mathcal{C}$ and two decoders

$$ \psi^{(1)} : (\mathcal{Y}^{(1)})^n \rightarrow \{1, \ldots, 2^{nR}\}, $$
$$ \psi^{(2)} : (\mathcal{Y}^{(2)})^n \rightarrow \{1, \ldots, 2^{nR}\} $$

such that when the codebook is used over channel $\nu$ with decoder $\psi^{(\nu)}$, the maximal block error probability is smaller than $\epsilon$ for $\nu = 1, 2$.

Note: Each channel has a different decoder, but you must show that there exists one codebook that is good for both channels.

Problem 2

Zero-Error Capacity

A channel with alphabet $\{0, 1, 2, 3, 4\}$ has transition probabilities of the form

$$ W(y|x) = \begin{cases} 
\frac{1}{2} & \text{if } y = x \pm 1 \text{ mod 5}, \\
0 & \text{otherwise}.
\end{cases} $$

a) Compute the capacity of this channel in bits.

b) The zero-error capacity of a channel is the number of bits per channel use that can be transmitted with zero probability of error. Clearly, the zero-error capacity of this pentagonal channel is at least 1 bit (transmit 0 or 1 with probability $\frac{1}{2}$). Find a block code that shows that the zero-error capacity is greater than 1 bit. Can you estimate the exact value of the zero-error capacity?

Hint: Consider codes of length 2 for this channel.

The zero-error capacity of this channel was found by Lovász.
Problem 3  

Symbol Error Rate vs. Message Error Rate

In proving the converse to the channel coding theorem we demonstrated that $R > C$ precludes the possibility that as the blocklength $n$ tends to infinity

$$\Pr[(X_1, \ldots, X_n) \neq (\hat{X}_1, \ldots, \hat{X}_n)] \to 0. \quad (1)$$

Here $(X_1, \ldots, X_n) = X(M) = f(M)$ is the transmitted $n$-tuple and $(\hat{X}_1, \ldots, \hat{X}_n) = f(\hat{M})$, where $\hat{M}$ is the decoder’s output, and $f(\cdot)$ is assumed to be one-to-one.

In this problem you are asked to prove a stronger statement, namely that $R > C$ precludes the possibility that

$$\frac{1}{n} \sum_{k=1}^{n} \Pr[X_k \neq \hat{X}_k] \to 0.$$ 

a) Show that you are indeed asked to prove a stronger statement by showing that the converse proved in class follows from the converse that you are about to prove in this problem.

b) Justify the following proof of a stronger version of Fano’s inequality: Let $U_1, \ldots, U_n$ and $Z_1, \ldots, Z_n$ be two sequences of random variables taking values in a finite set of cardinality $K$. Let

$$P_{e,i} = \Pr[U_i \neq Z_i], \quad i = 1, \ldots, n$$

and

$$P_b = \frac{1}{n} \sum_{i=1}^{n} P_{e,i}.$$

Then

$$\frac{1}{n} H(U_1, \ldots, U_n | Z_1, \ldots, Z_n) \overset{(i)}{=} \frac{1}{n} \sum_{i=1}^{n} H(U_i | U_1, \ldots, U_{i-1}, Z_1, \ldots, Z_n) \leq \frac{1}{n} \sum_{i=1}^{n} H(U_i | Z_1, \ldots, Z_n) \overset{(ii)}{=} \frac{1}{n} \sum_{i=1}^{n} H_b(P_{e,i}) + \frac{1}{n} \sum_{i=1}^{n} P_{e,i} \log(K - 1) \overset{(iii)}{=} H_b \left( \frac{1}{n} \sum_{i=1}^{n} P_{e,i} \right) + \log(K - 1) \frac{1}{n} \sum_{i=1}^{n} P_{e,i} \overset{(iv)}{=} H_b(P_b) + P_b \log(K - 1).$$

c) Prove the stronger version of the converse to the channel coding theorem using this new form of Fano’s inequality.

Problem 4  

An Elementary Converse for the Binary Erasure Channel

Consider a binary erasure channel with erasure probability $\rho \in [0, 1]$. Let the positive integer $n$ be the blocklength and $\mathcal{M}$ the message set of cardinality $|\mathcal{M}| = 2^R$, where $R$ is the rate of the code. Let $f: \mathcal{M} \to \mathcal{X}^n$ be the encoding function, and let $\phi: \mathcal{Y}^n \to \mathcal{M}$ be the decoding function. For every $i \in \{1, \ldots, n\}$ define the binary random variable

$$S_i = \begin{cases} 0 & Y_i \in \{0,1\}, \\ 1 & Y_i = \?, \end{cases}$$
and for \( l \in \{0, \ldots, n\} \) denote the set of binary length-\( n \) sequences with \( l \) ones by \( \mathcal{F}_l \)

\[
\mathcal{F}_l = \left\{ s \in \{0,1\}^n : \sum_{i=1}^n s_i = l \right\}.
\]

Observe that the random variables \( \{S_i\} \) are IID and independent of the channel input sequence \( X_1^n \), i.e., irrespective of the binary sequences \( x, s \in \{0,1\}^n \),

\[
\Pr[S_1^n = s | X_1^n = x] = \prod_{i=1}^n \Pr[S_i = s_i] = \rho^{\sum_{i=1}^n s_i} (1 - \rho)^{n - \sum_{i=1}^n s_i}.
\]

a) Show that for any number \( \kappa \in [0,1] \) such that \( n\kappa \) is a nonnegative integer the following statement holds: Conditional on the event that the number of erasures after \( n \) channel uses is \( n\kappa \), i.e.,

\[
\sum_{i=1}^n S_i = n\kappa,
\]

the average probability of a decoding error

\[
P_e(\mathcal{F}_{n\kappa}) = \frac{1}{|\mathcal{M}|} \sum_{m \in \mathcal{M}} \Pr[\phi(Y_1^n) \neq m | X_1^n = g(m), S_1^n \in \mathcal{F}_{n\kappa}]
\]

satisfies the lower bound

\[
P_e(\mathcal{F}_{n\kappa}) \geq \frac{2^n R - 2^n (1 - \kappa)}{2^n R}.
\]

b) Conclude from Part a) and the Law of Large Numbers that for \( R > 1 - \rho \) and for every number \( \epsilon > 0 \) there is a positive integer \( N \) such that for every block length \( n \geq N \) the average probability of a decoding error

\[
P_e^{(n)} = \frac{1}{|\mathcal{M}|} \sum_{m \in \mathcal{M}} \Pr[\phi(Y_1^n) \neq m | X_1^n = g(m)]
\]

satisfies the lower bound

\[
P_e^{(n)} > 1 - \epsilon.
\]

**Problem 5**  
**Source-Channel Separation and Average Bit-Error Probability**

Consider a combined source-channel coding scheme comprising a DMC of capacity \( C \), a source that emits IID \( \text{Ber}(1/2) \) bits \( U_1, \ldots, U_k \), an encoder that maps the \( k \) bits to \( n \) channel input symbols \( X_1, \ldots, X_n \), and a decoder that maps the resulting \( n \) channel output symbols \( Y_1, \ldots, Y_n \) to an estimate of the source bits \( \hat{U}_1, \ldots, \hat{U}_k \). Suppose that the rate (in bits per channel use) exceeds the capacity of the channel, i.e., \( k/n > C \). The task of this problem is to show that

\[
\frac{1}{k} \sum_{i=1}^k \Pr[U_i \neq \hat{U}_i] \geq H_b^{-1} \left( 1 - \frac{n}{k} C \right) > 0,
\]

where \( H_b^{-1}(\cdot) \) denotes the inverse of the binary entropy function restricted to \([0,1/2]\). In other words, if \( k/n > C \), then the average bit-error probability cannot be made arbitrarily small.
a) Show that 
\[ I(U^k; \hat{U}^k) \leq nC. \]

b) Justify the steps of the following computation

\[
I(U^k; \hat{U}^k) \overset{(a)}{=} \sum_{i=1}^{k} I(U_i; \hat{U}^k | U_{i-1}) \\
\overset{(b)}{=} \sum_{i=1}^{k} I(U_i; \hat{U}^k, U_{i-1}) \\
\overset{(c)}{\geq} k - \sum_{i=1}^{k} H_b \left( \Pr[U_i \neq \hat{U}_i] \right) \\
\overset{(d)}{\geq} k \left( 1 - H_b \left( \frac{1}{k} \sum_{i=1}^{k} \Pr[U_i \neq \hat{U}_i] \right) \right).
\]

c) Combine Parts a) and b) to obtain (2).