1. **Problem 15.7. Convexity of capacity region of broadcast channel.** Let $C \subseteq \mathbb{R}^2$ be the capacity region of all achievable rate pairs $R = (R_1, R_2)$ for the broadcast channel. Show that $C$ is a convex set by using a time-sharing argument. Specifically, show that if $R^{(1)}$ and $R^{(2)}$ are achievable, $\lambda R^{(1)} + (1 - \lambda)R^{(2)}$ is achievable for $0 \leq \lambda \leq 1$.

**Solution:**

We have two rate pairs that are achievable: $R^{(1)} = (R_1^{(1)}, R_2^{(1)})$ and $R^{(2)} = (R_1^{(2)}, R_2^{(2)})$ for which, we have two sequences of codes: $((2^nR_1^{(1)}, 2^nR_2^{(1)}), n)$ and $((2^nR_1^{(2)}, 2^nR_2^{(2)}), n)$. Like it is done in the proof of Theorem 15.3.2 for the case of the multiple-access channel in the textbook, here we can apply a similar argument and construct a third codebook of length $n$ at a rate $\lambda R^{(1)} + (1 - \lambda)R^{(2)}$ using the first codebook for the first $\lambda n$ symbols, and the second codebook for the last $(1 - \lambda)n$.

We have that the number of $X_1$ codewords for the new code is given by:

$$2^n\lambda R_1^{(1)} 2^n(1-\lambda)R_1^{(2)} = 2^n(\lambda R_1^{(1)} + (1-\lambda)R_1^{(2)})$$

And the number of $X_2$ codewords for the new code is given by:

$$2^n\lambda R_2^{(1)} 2^n(1-\lambda)R_2^{(2)} = 2^n(\lambda R_2^{(1)} + (1-\lambda)R_2^{(2)})$$

So, we have obtained a rate: $\lambda R^{(1)} + (1 - \lambda)R^{(2)}$. Recalling that the overall probability of error is less that the sum of the probabilities of error for each of the segments:

$$P_e^{(n)} \leq P_e^{(\lambda n)}(1) + P_e^{((1-\lambda)n)}(2)$$

We see that the probability of error goes to 0 as $n \to \infty$, hence the rate is achievable.
2. **Problem 15.11.** Converse for the degraded broadcast channel. The following chain of inequalities proves the converse for the degraded discrete memoryless broadcast channel. Provide reasons for each of the labeled inequalities.

*Setup for converse for degraded broadcast channel capacity:*

\[(W_1, W_2)_{\text{indep}} \rightarrow X^n(W_1, W_2) \rightarrow Y_1^n \rightarrow Y_2^n\]

- **Encoding:** \(f_n : 2^{nR_1} \times 2^{nR_2} \rightarrow X^n\)
- **Decoding:** \(g_n : Y_1^n \rightarrow 2^{nR_1}, h_n : Y_2^n \rightarrow 2^{nR_2}\). Let \(U_i = (W_2, Y_i^{i-1})\).

\[
nR_2 \leq_{\text{Fano}} I(W_2; Y_2^n) \leq \sum_{i=1}^{n} I(W_2; Y_i^{i-1}) \leq \sum_{i} (H(Y_i^{i-1}) - H(Y_i^{i-1} | W_2)) \leq \sum_{i} (H(Y_i^{i-1}) - H(Y_i^{i-1} | W_2, Y_1^{i-1}, Y_2^{i-1})) = \sum_{i} I(U_i; Y_2)
\]

**Solution:**

Reasons for each of the labeled inequalities:

(a) is given by the Chain rule.

(b) corresponds to the definition of conditional mutual information.

(c) conditioning reduces entropy.

(d) since \(Y_2^i\) is conditionally independent of \(Y_2^{i-1}\) given \(Y_1^{i-1}\).

(e) since \(U_i = (W_2, Y_i^{i-1})\), we can apply the identity \(I(X; Y) = H(Y) - H(Y|X)\).

*Continuation of converse: Give reasons for the labeled inequalities:*

\[
nR_1 \leq_{\text{Fano}} I(W_1; Y_1^n) \leq \sum_{i=1}^{n} I(W_1; Y_1^n | W_2) \leq \sum_{i} I(W_1; Y_1^n | Y_1^{i-1}, Y_2) = \sum_{i=1}^{n} I(X_i; Y_1 | U_i)
\]
Solution:
Reasons for each of the labeled inequalities:

(f) can be obtained by using the chain rule for mutual information, and because of the non-negative
tivity of the mutual information:

\[ I(W_1; W_2 | Y^n_1) \geq 0 \]

(12)

\[ I(W_1; Y^n_1, W_2) = I(W_1; Y^n_1) + I(W_1; W_2 | Y^n_1) \]

(13)

(g) can be obtained by using the chain rule for mutual information, and since

\[ I(W_1; W_2) = 0 \]

because of the independence between

\[ W_1 \text{ and } W_2. \]

(14)

\[ I(W_1; Y^n_1, W_2) = I(W_1; W_2) + I(W_1; Y^n_1 | W_2) \]

(15)

(h) corresponds to the chain rule of mutual information.

(i) is the data processing inequality and the definition of \( U_i = (W_2, Y^{i-1}) \).

Now let \( Q \) be a time-sharing random variable with \( Pr(Q = i) = 1/n, i = 1, 2, \ldots, n \).

Justify the following:

\[ R_1 \leq I(X_Q; Y_{1Q} | U_Q, Q), \]

\[ R_2 \leq I(U_Q; Y_{2Q} | Q) \]

(16)

(17)

for some distribution \( p(q)p(u|q)p(x|u,q)p(y_1, y_2| x) \).

Solution:

For this part we can use a similar analysis like the one given in the textbook (pp.540-542) for

the proof of the converse of the Multiple-Access channel. Then:

\[ nR_1 = (i) \sum_{i=1}^{n} I(X_i; Y_{1i} | U_i) \]

(18)

We can write:

\[ R_1 \leq \frac{1}{n} \sum_{i=1}^{n} I(X_i; Y_{1i} | U_i) \]

(19)

\[ \leq \frac{1}{n} \sum_{i=1}^{n} I(X_Q; Y_{1Q} | U_Q, Q = i) \]

(20)

\[ \leq I(X_Q; Y_{1Q} | U_Q, Q) \]

(21)

For \( R_2 \) we have:

\[ nR_2 = (e) \sum_{i} I(U_i; Y_{2i}) \]

(22)
We can write:

\[ R_2 \leq \frac{1}{n} \sum_i I(U_i; Y_{2i}) \tag{23} \]
\[ \leq \frac{1}{n} \sum_i I(U_q; Y_{2q}|Q = i) \tag{24} \]
\[ \leq \sum_i I(U_Q; Y_{2Q}|Q) \tag{25} \]

By appropriately redefining \( U \), argue that this region is equal to the convex closure of regions of the form

\[ R_1 \leq I(X; Y_1|U), \tag{26} \]
\[ R_2 \leq I(U; Y_2) \tag{27} \]

for some joint distribution \( p(u)p(x|u)p(y_1, y_2|x) \).

**Solution:**
From the previous part we have:

\[ R_1 \leq I(X_Q; Y_{1Q}|U_Q, Q) \tag{28} \]
\[ R_2 \leq I(U_Q; Y_{2Q}|Q) \tag{29} \]

So we can define a new set of random variables whose distributions depend on \( Q \): \( X = X_Q, \ Y_1 = Y_{1Q}, \ Y_2 = Y_{2Q} \) and \( U = (U_Q, Q) \), such that the previous inequalities can be re-written as:

\[ R_1 \leq I(X_Q; Y_{1Q}|U_Q, Q) \tag{30} \]
\[ \leq I(X; Y_1|U) \tag{31} \]
\[ R_2 \leq I(U_Q; Y_{2Q}|Q) \tag{32} \]
\[ \leq I(U_Q, Q; Y_{2Q}) - I(Q; Y_{2Q}) \tag{33} \]
\[ \leq I(U; Y_2) \tag{34} \]
3. **Problem 15.16. Multiple-access channel.** Let the output $Y$ of a multiple-access channel be given by:

$$Y = X_1 + sgn(X_2)$$

where $X_1, X_2$ are both real and power limited,

$$E(X_1^2) \leq P_1$$

$$E(X_2^2) \leq P_2$$

and

$$sgn(x) = \begin{cases} 1, & x > 0, \\ -1, & x \leq 0 \end{cases}$$

Note that there is interference but no noise in this channel.

(a) **Find the capacity region.**

For this problem we can use an auxiliary random variable defined as $\theta = sgn(X_2)$ such that channel becomes: $Y = X_1 + \theta$. Here we have two random variables. In the particular case of $\theta$, it that can take on only two values $-1$ and $+1$ depending on the value of $X_2$. So, initially we can set $X_1 = 0$ such that the channel becomes $Y = \theta$, and given the binary nature of $\theta$, we see that the maximum rate $R_2$ is limited by 1. We can also use the well known expression for the capacity region for the rate $R_2$ in the multiple-access channel:

$$R_2 \leq I(\theta; Y|X_1)$$

$$\leq H(Y|X_1) - H(Y|\theta, X_1)$$

$$\leq H(X_1 + \theta|X_1)$$

$$\leq H(\theta)$$

$$\leq 1$$

On the other hand, by fixing $X_2$ for example to any positive value (under the power constraint limit), for example $X_2 = \sqrt{P_2}$ we obtain $\theta = 1$, and the channel becomes $Y = X_1 + 1$. We can see that subtracting 1 from $Y$ is enough to recover $X_1$ which allows us to transmit as much information as we want yielding an infinite bound for $R_1 \leq \infty$. Hence $R_1 + R_2 \leq \infty$.

(b) **Describe a coding scheme that achieves the capacity region.**

Since we can transmit an infinite rate $R_1$, we can use one of those bits to encode the information of $X_2$ since $R_2 \leq 1$. This scheme presents an achievable rate pair $(R_1 = \infty, R_2 = 1)$ recalling that we have to respect the power constraints for both random variables, $E[X_1] \leq \sqrt{P_1}$ and $E[X_2] \leq \sqrt{P_2}$.  

5
4. **Problem 15.19. SlepianWolf.** Two senders know random variables $U_1$ and $U_2$, respectively. Let the random variables $(U_1, U_2)$ have the following joint distribution:

<table>
<thead>
<tr>
<th>$U_1 \backslash U_2$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>$\cdots$</th>
<th>$m-1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$\alpha$</td>
<td>$\frac{\beta}{m-1}$</td>
<td>$\frac{\beta}{m-1}$</td>
<td>$\cdots$</td>
<td>$\frac{\beta}{m-1}$</td>
</tr>
<tr>
<td>1</td>
<td>$\frac{\gamma}{m-1}$</td>
<td>0</td>
<td>0</td>
<td>$\cdots$</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>$\frac{\gamma}{m-1}$</td>
<td>0</td>
<td>0</td>
<td>$\cdots$</td>
<td>0</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\cdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$m-1$</td>
<td>$\frac{\gamma}{m-1}$</td>
<td>0</td>
<td>0</td>
<td>$\cdots$</td>
<td>0</td>
</tr>
</tbody>
</table>

where $\alpha + \beta + \gamma = 1$. Find the region of rates $(R_1, R_2)$ that would allow a common receiver to decode both random variables reliably.

**Solution:**

From **Theorem 15.4.1** in the textbook, we know that the achievable region for the Slepian-Wolf will be given by:

$$R_1 \geq H(U_1|U_2)$$  \hspace{1cm} (40)

$$R_2 \geq H(U_2|U_1)$$  \hspace{1cm} (41)

$$R_1 + R_2 \geq H(U_1, U_2)$$ \hspace{1cm} (42)

First we compute the marginal distributions for $U_1$ and $U_2$ as follows:

$$p(u_1) = \{\alpha + \beta, \frac{\gamma}{m-1}, \frac{\gamma}{m-1}, \cdots, \frac{\gamma}{m-1}(\text{times})\}$$  \hspace{1cm} (43)

$$p(u_2) = \{\alpha + \gamma, \frac{\beta}{m-1}, \frac{\beta}{m-1}, \cdots, \frac{\beta}{m-1}(\text{times})\}$$  \hspace{1cm} (44)
Now the regions are given by:

\[ R_1 \geq H(U_1|U_2) \]  
\[ \geq \sum_{i=0}^{m-1} P(U_2 = i) H(U_1|U_2 = i) \]  
\[ \geq P(U_2 = 0) H(U_1|U_2 = 0) + \sum_{i=1}^{m-1} P(U_2 = i) H(U_1|U_2 = i) \]  
\[ \geq (\alpha + \gamma) H \left( \frac{\alpha}{\alpha + \gamma}, \frac{\gamma}{\alpha + \gamma}, \frac{\gamma}{\alpha + \gamma}, \ldots, \frac{\gamma}{\alpha + \gamma} \right) \]  
\[ \sum_{i=1}^{m-1} P(U_2 = i) H(U_1|U_2 = i) = 0. \]

Since, \( \sum_{i=1}^{m-1} P(U_2 = i) H(U_1|U_2 = i) = 0. \)

Similarly we can obtain \( R_2 \) as follows:

\[ R_2 \geq H(U_2|U_1) \]  
\[ \geq \sum_{i=0}^{m-1} P(U_1 = i) H(U_2|U_1 = i) \]  
\[ \geq P(U_1 = 0) H(U_2|U_1 = 0) + \sum_{i=1}^{m-1} P(U_1 = i) H(U_2|U_1 = i) \]  
\[ \geq (\alpha + \beta) H \left( \frac{\alpha}{\alpha + \beta}, \frac{\beta}{\alpha + \beta}, \frac{\beta}{\alpha + \beta}, \ldots, \frac{\beta}{\alpha + \beta} \right) \]  
\[ \sum_{i=1}^{m-1} P(U_1 = i) H(U_2|U_1 = i) = 0. \]

Since, \( \sum_{i=1}^{m-1} P(U_1 = i) H(U_2|U_1 = i) = 0. \)

\[ R_1 + R_2 \geq H(U_1, U_2) \]  
\[ \geq H \left( \frac{\alpha}{m-1}, \frac{\gamma}{m-1}, \ldots, \frac{\gamma}{m-1}, \frac{\beta}{m-1}, \frac{\beta}{m-1}, \ldots, \frac{\beta}{m-1} \right) \]  
\[ \geq -\alpha \log_2 (\alpha) - \gamma \log_2 \left( \frac{\gamma}{m-1} \right) - \beta \log_2 \left( \frac{\beta}{m-1} \right) \]  
\[ \geq -\alpha \log_2 (\alpha) - \gamma \log_2 (\gamma) - \beta \log_2 (\beta) + (\gamma + \beta) \log_2 (m-1) \]  
\[ \geq H(\alpha, \beta, \gamma) + (\gamma + \beta) \log_2 (m-1) \]
5. **Problem 15.30. Parallel Gaussian channels from a mobile telephone.** Assume that a sender $X$ is sending to two fixed base stations. Assume that the sender sends a signal $X$ that is constrained to have average power $P$. Assume that the two base stations receive signals $Y_1$ and $Y_2$, where

$$Y_1 = \alpha_1 X + Z_1$$
$$Y_2 = \alpha_2 X + Z_2,$$

where $Z_i \sim \mathcal{N}(0, N_i)$, $Z_2 \sim \mathcal{N}(0, N_2)$, and $Z_1$ and $Z_2$ are independent. We will assume the $\alpha$’s are constant over a transmitted block.

(a) Assuming that both signals $Y_1$ and $Y_2$ are available at a common decoder $Y = (Y_1, Y_2)$, what is the capacity of the channel from the sender to the common receiver?

We can consider this problem as having two parallel Gaussian channels with inputs $\alpha_1 X, \alpha_2 X$ and corresponding outputs given by $Y_1, Y_2$. Since we have the power constraint $P$ in the input $X$.

So we have: $Y = \begin{bmatrix} Y_1 = \alpha_1 X + Z_1 \\ Y_2 = \alpha_2 X + Z_2 \end{bmatrix}$

We know the capacity is given by: $C = \max_{p(X)} I(X; Y)$

Since we know that the output is Gaussian the mutual information is maximized by a Gaussian input for $p(X)$ so that $X \sim \mathcal{N}(0, P)$. Hence we have:

$$X = \begin{bmatrix} \alpha_1 X \sim \mathcal{N}(0, \alpha_1^2 P) \\ \alpha_2 X \sim \mathcal{N}(0, \alpha_2^2 P) \end{bmatrix}$$

In order to compute the capacity we need the determinant $|K_Y| = |K_X + K_Z|$. Since we have independent noises we have:

$$K_Z = \begin{bmatrix} N_1 & 0 \\ 0 & N_2 \end{bmatrix}$$

$$K_X = \begin{bmatrix} \alpha_1^2 P & COV[\alpha_1 X \alpha_2 X] \\ COV[\alpha_1 X \alpha_2 X] & \alpha_2^2 P \end{bmatrix}$$

$$COV[\alpha_1 X \alpha_2 X] = E[\alpha_1 \alpha_2 X^2] = \alpha_1 \alpha_2 P$$

(58)

Then,

$$K_X = \begin{bmatrix} \alpha_1^2 P & \alpha_1 \alpha_2 P \\ \alpha_1 \alpha_2 P & \alpha_2^2 P \end{bmatrix}$$

(59)
\(|K_Y| = |K_X + K_Z|\)  
\[\begin{align*}
\text{(60)} & \quad = \left[ \begin{array}{cc} \alpha_1^2 P & \alpha_1 \alpha_2 P \\ \alpha_1 \alpha_2 P & \alpha_2^2 P \end{array} \right] + \left[ \begin{array}{cc} N_1 & 0 \\ 0 & N_2 \end{array} \right] \\
\text{(61)} & \quad = \left[ \begin{array}{cc} \alpha_1^2 P & \alpha_1 \alpha_2 P + N_1 \\ \alpha_1 \alpha_2 P & \alpha_2^2 P + N_2 \end{array} \right] \\
\text{(62)} & \quad = (\alpha_1^2 P + N_1)(\alpha_2^2 P + N_2) - \alpha_1^2 \alpha_2^2 P^2 \\
\text{(63)} & \quad = \alpha_1^2 \alpha_2^2 P^2 + \alpha_1^2 PN_2 + \alpha_2^2 PN_1 + N_1 N_2 - \alpha_1^2 \alpha_2^2 P^2 \\
\text{(64)} & \quad = (\alpha_1^2 PN_2 + \alpha_2^2 PN_1 + N_1 N_2) \\
\text{(65)} & \quad = (\alpha_1^2 PN_2 + \alpha_2^2 PN_1 + N_1 N_2)
\end{align*}\]

To find the capacity of a set of two parallel Gaussian channels we have:
\[C = \frac{1}{2} \log_2 \left( \frac{|K_Y|}{|K_Z|} \right)\]  
\[\text{(66)}  \]
\[= \frac{1}{2} \log_2 \left( \frac{\alpha_1^2 PN_2 + \alpha_2^2 PN_1 + N_1 N_2}{N_1 N_2} \right)\]  
\[\text{(67)}\]
\[= \frac{1}{2} \log_2 \left( 1 + \frac{\alpha_1^2 PN_2 + \alpha_2^2 PN_1}{N_1 N_2} \right)\]  
\[\text{(68)}\]

(b) If, instead, the two receivers \(Y_1\) and \(Y_2\) each decode their signals independently, this becomes a broadcast channel. Let \(R_1\) be the rate to base station 1 and \(R_2\) be the rate to base station 2. Find the capacity region of this channel.

Here we have a sender with power \(P\) and two receivers, with corresponding noises \(N_1\) and \(N_2\). From the textbook, we have that the channel is modeled as: \(Y_1 = X + Z_1\) and \(Y_2 = X + Z_2\) for two arbitrary correlated Gaussian random variables \(Z_1 \sim \mathcal{N}(0,N_1)\) and \(Z_2 \sim \mathcal{N}(0,N_2)\). Then for a Gaussian broadcast channel we can use the expression given in the textbook for the region in: Eq.(15.11) and Eq.(15.12) for a given \(0 \leq \lambda \leq 1\). Assuming without loss of generality that \(N_1 < N_2\):

\[R_1 \leq C \left( \frac{\lambda P}{N_1} \right)\]  
\[\text{(69)}\]
\[R_2 \leq C \left( \frac{(1 - \lambda) P}{\lambda P + N_2} \right)\]  
\[\text{(70)}\]

So, for this case we have:

\[R_1 \leq \frac{1}{2} \log_2 \left( 1 + \frac{\lambda \alpha_1^2 P}{N_1} \right)\]  
\[\text{(71)}\]
\[R_2 \leq \frac{1}{2} \log_2 \left( 1 + \frac{(1 - \lambda) \alpha_2^2 P}{\lambda \alpha_1^2 P + N_2} \right)\]  
\[\text{(72)}\]