1. (30 points) Erasure broadcast channel. Let \( p(y_1, y_2|x) \) be a discrete memoryless broadcast channel with capacity region \( C \). Suppose that this channel is followed immediately by an erasure channel \( p(y'_j|y_j) \) at each receiver \( j = 1, 2 \) that erases the symbols with probability \( p \), that is, 
\[
Y'_j = \begin{cases} 
Y_j & \text{with probability } 1 - p, \\
e & \text{with probability } p,
\end{cases}
\]
which results in a new erasure broadcast channel
\[
p(y'_1, y'_2|x) = p(y_1, y_2|x)p(y'_1|y_1)p(y'_2|y_2).
\]
Note that the erasures at the receivers are independent of each other. Suppose that the underlying broadcast channel \( p(y_1, y_2|x) \) is degraded.

(a) Is the erasure broadcast channel degraded, less noisy, or more capable?

(b) Find the capacity region of the erasure broadcast channel.

(c) Now suppose that the underlying broadcast channel is not necessarily degraded and has the capacity region \( C \). Comment on the capacity region of the erasure broadcast channel.

Solution.

(a) The erasure broadcast channel is (stochastically) degraded (therefore it is also less noisy and more capable). To see this, define \( \tilde{Y}'_1 \) as
\[
\tilde{Y}'_1 = \begin{cases} 
Y_1 & \text{if } Y'_2 = Y_2, \\
e & \text{if } Y'_2 = e.
\end{cases}
\]
Then, \( \tilde{Y}'_1|\{X = x\} \sim p_{Y'_1|X}(\tilde{y}'_1|x) \), i.e., \( \tilde{Y}'_1 \) has the same conditional pmf as \( Y'_1 \) given \( X \) and \( X \to \tilde{Y}'_1 \to Y'_2 \) form a Markov chain. Therefore, by definition, the erasure broadcast channel \( p(y'_1, y'_2|x) \) is degraded (and hence less noisy as well as more capable).

(b) Since it is degraded, by Theorem 5.2 in the text, the capacity region of the erasure broadcast channel \( p(y'_1, y'_2|x) \) is the set of rate pairs \( (R_1, R_2) \) such that
\[
R_1 \leq I(X; Y'_1|U), \\
R_2 \leq I(U; Y'_2)
\]
for some pmf \( p(u,x) \). To express this region in terms of \( C \), define binary random variables \( E_j \) for \( j = 1, 2 \) as
\[
E_j = \begin{cases} 
1 & \text{if } Y'_j = e, \\
0 & \text{if } Y'_j = Y_j.
\end{cases}
\]
For any \( p(u, x) \), we have

\[
I(U; Y_1') \overset{(a)}{=} I(U; Y_2', E_2),
\]
\[
= I(U; E_2) + I(U; Y_2'|E_2),
\]
\[
\overset{(b)}{=} I(U; Y_2'|E_2 = 0) P(E_2 = 0),
\]
\[
\overset{(c)}{=} (1 - p) I(U; Y_2),
\]

where (a) follows since \( E_2 \) is a function of \( Y'_2 \), (b) follows by the independence of \( E_2 \) and \( U \), and (c) follows since \( E_2 \) is independent from \((U, Y_2)\). Following similar steps, it is easy to verify that \( I(X; Y'_1|U) = (1 - p)I(X; Y_1|U) \). Thus, the capacity region of the erasure broadcast channel \( p(y'_1, y'_2|x) \) is the set of rate pairs \((R_1, R_2)\) such that

\[
R_1 \leq (1 - p)I(X; Y_1|U),
\]
\[
R_2 \leq (1 - p)I(U; Y_2)
\]

for some pmf \( p(u, x) \), which is equivalent to

\[
(1 - p)\mathcal{C} \triangleq \{(1 - p)(R_1, R_2) : (R_1, R_2) \in \mathcal{C}\}.
\]

(c) We can assume without loss of generality that erasures occur simultaneously at both receivers (why?). Therefore, if a genie provides the locations of erasures to the sender, then the capacity region (with side information) is \((1 - p)\mathcal{C}'\) and hence the capacity region of the erasure broadcast channel (without genie), denoted by \(\mathcal{C}'\), satisfies

\[
\mathcal{C}' \subseteq (1 - p)\mathcal{C}.
\]

A more rigorous proof is too large to fit in the margin.

2. (40 points) **Common-message broadcasting with state information.** Consider the DM-BC with DM state \( p(y_1, y_2|x, s)p(s) \). Establish the common-message capacity \( C_0 \) (that is, the maximum achievable common rate \( R_0 \) when \( R_1 = R_2 = 0 \)) for the following settings. You are expected to prove achievability and the converse.

(a) The state information is available only at decoder 1.

(b) The state information is causally available at both the encoder and decoder 1.

(c) The state information is noncausally available at the encoder and decoder 1.

(d) The state information is causally available only at the encoder.

**Solution.**

(a) The common-message capacity is

\[
C_0 = \max_{p(x)} \min \{I(X; Y_1|S), I(X; Y_2)\}.
\]

This follows immediately by considering the augmented output \( \hat{Y}_1 = (Y_1, S) \) at receiver 1 and using Problem 5.9(c). Alternatively, this is proved step-by-step as follows.

**Achievability:** Fix a pmf \( p(x) \) that attains \( C_0 \) and randomly and independently generate \( 2^{nR_0} \) codewords \( x^n(m) \), \( m \in [1 : 2^{nR_0}] \), each according to \( \prod_{i=1}^{n} p_X(x_i) \). To send \( m \), the encoder transmits \( x^n(m) \). Decoder 1 finds the unique message \( \hat{m}_1 \) such that \((x^n(\hat{m}_1), s^n, y^n_1) \in T^{(n)}_e(X, S, Y_1)\) and decoder 2 finds the unique message \( \hat{m}_2 \) such that \((x^n(\hat{m}_2), y^n_2) \in T^{(n)}_e(X, Y_2)\). By the LLN
and the packing lemma, the average probability of error for decoder 1 tends to 0 as \( n \to \infty \) if \( R_0 < I(X; Y_1 | S) - \delta(\epsilon) \), and the average probability of error for decoder 2 tends to 0 as \( n \to \infty \) if \( R_0 < I(X; Y_2) - \delta(\epsilon) \). The proof of achievability follows by letting \( \epsilon \to 0 \).

**Proof of the converse:** Consider

\[
\begin{align*}
nR_0 & \leq H(M) \\
& = I(M; Y_1^n, S^n) + H(M|Y_1^n, S^n) \\
& \overset{(a)}{=} I(M; Y_1^n, S^n) + n\epsilon_n \\
& = \sum_{i=1}^{n} I(M; Y_{1i}, S_i|Y_{1i}^{i-1}, S_{i-1}) + n\epsilon_n \\
& \overset{(b)}{=} \sum_{i=1}^{n} I(X_i; Y_{1i}, S_i) + n\epsilon_n,
\end{align*}
\]

where (a) follows by Fano’s inequality, and (b) follows since \((M, Y_{1i}^{i-1}, S_{i-1}) \to (X_i, S_i) \to Y_{1i}\) form a Markov chain and \(S_i\) is independent of \((M, Y_{1i}^{i-1}, S_{i-1}, X_i)\). Similarly,

\[
\begin{align*}
nR_0 & \leq H(M) \\
& = I(M; Y_2^n) + H(M|Y_2^n) \\
& \leq I(M; Y_2^n) + n\epsilon_n \\
& = \sum_{i=1}^{n} I(M; Y_{2i}|Y_{2i}^{i-1}) + n\epsilon_n \\
& \leq \sum_{i=1}^{n} I(M, Y_{2i}^{i-1}, S_{i-1}, X_i; Y_{2i}) + n\epsilon_n \\
& = \sum_{i=1}^{n} I(X_i; Y_{2i}) + n\epsilon_n.
\end{align*}
\]

The rest of the proof follows by introducing a time sharing random variable \( Q \sim \text{Unif}[1 : n] \) independent of \((M, X^n, Y_1^n, Y_2^n, S^n)\), defining \( X = X_Q, Y_1 = Y_{1Q}, Y_2 = Y_{2Q} \), and \( S = S_Q \), and letting \( n \to \infty \).

(b) The common-message capacity is

\[
C_0 = \max_{p(u), x(u,s)} \min\{I(U; Y_1 | S), I(U; Y_2)\} \tag{1}
\]

where \( I(U; Y_1 | S) = I(X; Y_1 | S) \) since \( X \) is a function of \((U, S)\) and \( U \to (X, S) \to Y_1 \) form a Markov chain. To see this, first consider the case with causal state information available only at the encoder (see part (d)). Then, by the Shannon strategy, the common message capacity is

\[
\max_{p(u), x(u,s)} \min\{I(U; Y_1), I(U; Y_2)\}.
\]

(The converse proof follows by adapting the case with a single receiver.) With state information available at receiver 1 as well, we use the output augmentation trick in part (a) to establish (1). Alternatively, this is proved step-by-step as follows.
Achievability: Fix $p(u)$ and $x(u, s)$ that attain $C_0$ and randomly and independently generate $2^{nR_0}$ sequences $u^n(m)$, $m \in [1 : 2^{nR_0}]$, each according to $\prod_{i=1}^{n} p_{U_i}(u_i)$. To send $m$ given the state $s_i$, the encoder transmits $x(u_i(m), s_i)$ at time $i \in [1 : n]$. Decoder 1 finds the unique message $\hat{m}_1$ such that $(u^n(\hat{m}_1), s^n, y^n_1) \in T^{(a)}(U, S, Y_1)$ and decoder 2 finds the unique message $\hat{m}_2$ such that $(u^n(\hat{m}_2), y^n_2) \in T^{(a)}(U, Y_2)$. By the LLN and the packing lemma, the average probability of error for decoder 1 tends to 0 as $n \to \infty$ if $R_0 < I(U; Y_1|S) - \delta(\epsilon)$, and the average probability of error for decoder 2 tends to 0 as $n \to \infty$ if $R_0 < I(U; Y_2) - \delta(\epsilon)$. The result follows by letting $\epsilon \to 0$.

Proof of the converse: Define $U_i = (M, S^{i-1})$. Then, we have

$$nR_0 \leq H(M)$$

$$= I(M; Y_1^n, S^n) + H(M|Y_1^n, S^n)$$

$$\leq I(M; Y_1^n, S^n) + n\epsilon_n$$

$$= \sum_{i=1}^{n} I(M; Y_{1i}, S_i|Y_1^{i-1}, S^{i-1}) + n\epsilon_n$$

$$\leq \sum_{i=1}^{n} I(M, S^{i-1}, Y_1^{i-1}, X^{i-1}; Y_{1i}, S_i) + n\epsilon_n$$

$$= (a) \sum_{i=1}^{n} I(M, S^{i-1}, X^{i-1}; Y_{1i}, S_i) + n\epsilon_n$$

$$= \sum_{i=1}^{n} I(M, S^{i-1}; Y_{1i}, S_i) + n\epsilon_n$$

$$= \sum_{i=1}^{n} I(U_i; Y_{1i}, S_i) + n\epsilon_n,$$

where $(a)$ follows since $(Y_{1i}, S_i) \to (X^{i-1}, S^{i-1}) \to Y_1^{i-1}$ form a Markov chain. Similarly,

$$nR_0 \leq H(M)$$

$$= I(M; Y_2^n) + H(M|Y_2^n)$$

$$\leq I(M; Y_2^n) + n\epsilon_n$$

$$= \sum_{i=1}^{n} I(M; Y_{2i}|Y_2^{i-1}) + n\epsilon_n$$

$$\leq \sum_{i=1}^{n} I(M, S^{i-1}, Y_2^{i-1}, X^{i-1}; Y_{2i}) + n\epsilon_n$$

$$= \sum_{i=1}^{n} I(M, S^{i-1}, X^{i-1}; Y_{2i}) + n\epsilon_n$$

$$= \sum_{i=1}^{n} I(M, S^{i-1}; Y_{2i}) + n\epsilon_n$$

$$= \sum_{i=1}^{n} I(U_i; Y_{2i}) + n\epsilon_n.$$

The rest of the proof follows by introducing a time sharing random variable $Q \sim \text{Unif}[1 : n]$ independent of $(M, X^n, Y_1^n, Y_2^n, S^n)$, defining $U = (U_Q, Q), Y_1 = Y_{1Q}, Y_2 = Y_{2Q}$, and $S = S_Q$ (note that $U$ and $S$ are independent), and letting $n \to \infty$. 

4
(c) The common-message capacity is
\[
C_0 = \max_{p(u,s), x(u,s)} \min\{I(U; Y_1 | S), I(U; Y_2) - I(U; S)\}
\]
\[
= \max_{p(u,s), x(u,s)} \min\{I(X; Y_1 | S), I(U; Y_2) - I(U; S)\}.
\]

Achievability: Apply Gelfand–Pinsker coding at the encoder. Fix a conditional pmf \(p(u,s)\) and function \(x(u,s)\) that attain \(C_0\), and let \(\tilde{R} > R_0\). For each message \(m \in [1 : 2^nR_0]\), generate a subcodebook \(C(m)\) consisting of \(2^n(\tilde{R} - R_0)\) randomly and independently generated sequences \(u^n(l), l \in [(m - 1)2^n(\tilde{R} - R_0) + 1 : m2^n(\tilde{R} - R_0)]\), each according to \(\prod_{i=1}^n p_U(u_i)\). To send \(m\) given \(s^n\), the encoder chooses a sequence \(u^n(l) \in C(m)\) such that \((u^n(l), s^n) \in T^n_{C(m)}(U, S)\). If no such sequence exists, it chooses \(l = 1\). Then, the encoder transmits \(x(u_i(l), s_i)\) at time \(i \in [1 : n]\). Let \(\epsilon > \epsilon'\). Decoder 1 finds the unique message \(\hat{m}_1\) such that \((u^n(l), s^n, y^n_1) \in T^n_{C(m)}(U, S, Y_1)\) for some \(u^n(l) \in C(\hat{m}_1)\), and decoder 2 finds the unique message \(\hat{m}_2\) such that \((u^n(l), y^n_2) \in T^n_{C(m)}(U, Y_2)\) for some \(u^n(l) \in C(\hat{m}_2)\). By the covering lemma, the conditional typicality lemma, and the packing lemma, the average probability of error tends to 0 as \(n \to \infty\) if \(\tilde{R} - R_0 > I(U; S) + \delta(\epsilon')\), \(\tilde{R} < I(U; Y_1, S) - \delta(\epsilon)\), and \(\tilde{R} < I(U; Y_2) - \delta(\epsilon)\). By Fourier–Motzkin elimination method, the average probability of error tends to 0 as \(n \to \infty\) if \(R_0 < \min\{I(U; Y_1) - \delta(\epsilon) - \delta(\epsilon'), I(U; Y_2) - I(U; S) - \delta(\epsilon) - \delta(\epsilon')\}\). The proof of achievability follows by letting \(\epsilon \to 0\).

Proof of the converse: Define \(U_i = (M, S^n_{i+1}, Y_{2i}^{i-1})\). Then, we have
\[
nR_0 \leq H(M)
\]
\[
= I(M; Y_2^n) + H(M | Y_2)
\]
\[
\leq I(M; Y_2^n) + n\epsilon_n
\]
\[
= \sum_{i=1}^n I(M; Y_{2i} | Y_2^{i-1}) + n\epsilon_n
\]
\[
\leq \sum_{i=1}^n I(M, S^n_{i+1}, Y_2^{i-1}; Y_{2i}) - \sum_{i=1}^n I(S^n_{i+1}; Y_{2i} | M, Y_2^{i-1}) + n\epsilon_n
\]
\[
= \sum_{i=1}^n I(M, S^n_{i+1}, Y_2^{i-1}; Y_{2i}) - \sum_{i=1}^n I(Y_2^{i-1}; S_i | M, S^n_{i+1}) + n\epsilon_n
\]
\[
\leq \sum_{i=1}^n I(U_i; Y_{2i}) - I(U_i; S_i) + n\epsilon_n,
\]
where \((a)\) follows by the Csizsár sum identity, and \((b)\) follows by the independence of \(S_i\) and \((M, S^n_{i+1})\). Similarly,\[
\]
\[
nR_0 \leq H(M)
\]
\[
= I(M; Y_1^n | S^n) + H(M | Y_1^n, S^n)
\]
\[
\leq I(M; Y_1^n | S^n) + n\epsilon_n
\]
\[
= \sum_{i=1}^n I(M; Y_1 | Y_1^{i-1}, S^n) + n\epsilon_n
\]
\[
\leq \sum_{i=1}^n I(M, Y_1^{i-1}, S_1^{i-1}, S^n_{i+1}, X_i; Y_1; S_i) + n\epsilon_n
\]
\[
\leq \sum_{i=1}^n I(M, Y_1^{i-1}, Y_1^{i-1}, S^n_{i+1}, X_i; Y_1; S_i) + n\epsilon_n
\]
\[ = \sum_{i=1}^{n} I(X_i; Y_i | S_i) + n\epsilon_n. \]

By introducing a time sharing random variable \( Q \sim \text{Unif}[1 : n] \) independent of \((M, U^n, Y'^n_1, Y'^n_2, S^n)\) and defining \( U = (U_Q, Q), Y_1 = Y_{1Q}, Y_2 = Y_{2Q}, S = S_Q \), we have

\[ C_0 \leq \max_{p(u|s), p(x|u, s)} \min \{ I(X; Y_1 | S), I(U; Y_2) - I(U; S) \}, \]

as \( n \to \infty \) since \( X_i \) is not a function of \((U_i, S_i)\). Now, we show that

\[ \max_{p(u|s), p(x|u, s)} \min \{ I(X; Y_1 | S), I(U; Y_2) - I(U; S) \} = \max_{p(x|u, s)} \min \{ I(X; Y_1 | S), I(U; Y_2) - I(U; S) \}. \]

One direction of inequality \((\geq)\) is trivial. For the other direction, note that by the functional representation lemma, any \( p(x|u, s) \) can be expressed as \( x(u, s, v) \), where \( V \) is a random variable independent of \((U, S)\) with finite cardinality. Thus, by defining \( \tilde{U} = (U, V) \), we have

\[
\max_{p(x|u, s)} \min \{ I(X; Y_1 | S), I(U; Y_2) - I(U; S) \} = \max_{p(x|u, s)} \min \{ I(X; Y_1 | S), I(U; Y_2) - I(U; S) \} \\
\leq \max_{p(x|u, s)} \min \{ I(X; Y_1 | S), I(U; V, Y_2) - I(U, V; S) \} \\
= \max_{p(x|u, s)} \min \{ I(X; Y_1 | S), I(\tilde{U}; Y_2) - I(\tilde{U}; S) \}.
\]

(d) The common-message capacity is

\[ C_0 = \max_{p(u), x(u, s)} \min \{ I(U; Y_1), I(U; Y_2) \}, \]

as argued before in part (b). We still provide a formal proof.

**Achievability:** Fix a pmf \( p(u) \) and \( x(u, s) \) that attain \( C_0 \) and randomly and independently generate \( 2^{nR_0} \) sequences \( u^n(m) \), \( m \in [1 : 2^{nR_0}] \), each according to \( \prod_{i=1}^{n} p_U(u_i) \). To send \( m \) given the state \( s_i \), the encoder transmits \( x(u(m), s_i) \) at time \( i \in [1 : n] \). Decoder 1 finds the unique message \( \hat{m}_1 \) such that \((u^n(\hat{m}_1), y^n_1) \in T_{\epsilon}(U, Y_1)\) and decoder 2 finds the unique message \( \hat{m}_2 \) such that \((u^n(\hat{m}_2), y^n_2) \in T_{\epsilon}(U, Y_2)\). By the LLN and the packing lemma, the average probability of error for decoder \( j = 1, 2 \) tends to 0 as \( n \to \infty \) if \( R_0 < I(U; Y_j) - \delta(\epsilon) \). The proof of achievability follows by letting \( \epsilon \to 0 \).

**Proof of the converse:** Define \( U_i = (M, S^i-1) \). Then, we have

\[
nR_0 \leq H(M) \\
= I(M; Y^n_j) + H(M | Y^n_j) \\
\leq I(M; Y^n_j) + n\epsilon_n \\
= \sum_{i=1}^{n} I(M; Y_{ji} | Y_{ji}^{i-1}) + n\epsilon_n \\
\leq \sum_{i=1}^{n} I(M, S^{i-1}, Y_{ji}^{i-1}, X_{ji}^{i-1}; Y_{ji}) + n\epsilon_n \\
= \sum_{i=1}^{n} I(M, S^{i-1}, X_{ji}^{i-1}; Y_{ji}) + n\epsilon_n \\
= \sum_{i=1}^{n} I(M, S^{i-1}; Y_{ji}) + n\epsilon_n
\]

6
\[ I(U_i; Y_{ji}) + n\epsilon_n \]

for \( j = 1, 2 \). The rest of the proof follows by introducing the standard time sharing random variable and letting \( n \to \infty \).

3. (30 points) Broadcasting with side information.
Consider the 3-receiver deterministic BC \( p(y_1, y_2, y_3|x) \) such that \( Y_j = y_j(X), j = 1, 2, 3 \), with message triple \((M_1, M_2, M_3)\) uniformly distributed over \([1 : 2^{nR_1}] \times [1 : 2^{nR_2}] \times [1 : 2^{nR_3}]\) as depicted below. The sender wishes to communicate \( M_j \) to receiver \( j \in \{1, 2, 3\} \). We assume that each receiver has side information of some other messages, denoted as \( M(A_j) \) for receiver \( j \). Find the capacity region for the following cases. You are expected to prove achievability and the converse.

(a) \( A_1 = \emptyset, A_2 = \emptyset, A_3 = \emptyset \), that is, the receivers have no side information.
(b) \( A_1 = \{2, 3\}, A_2 = \{1, 3\}, A_3 = \{1, 2\} \), that is, the receivers already know all the messages other than they are receiving.
(c) \( A_1 = \emptyset, A_2 = \{1\}, A_3 = \{1, 2\} \).

Solution.

(a) The capacity region consists of the rate triples \((R_1, R_2, R_3)\) such that

\[
\begin{align*}
R_1 &< H(Y_1), \\
R_2 &< H(Y_2), \\
R_3 &< H(Y_3), \\
R_1 + R_2 &< H(Y_1, Y_2), \\
R_1 + R_3 &< H(Y_1, Y_3), \\
R_2 + R_3 &< H(Y_2, Y_3), \\
R_1 + R_2 + R_3 &< H(Y_1, Y_2, Y_3)
\end{align*}
\]

for some pmf \( p(x) \). Achievability follows by Marton’s inner bound with \( U_j = Y_j = y_j(X) \).
For the converse, consider

\[ nR_j \leq H(M_j) \]
\[ I(M; Y^n_j) + H(M|Y^n) \]
\[ \leq I(M; Y^n_j) + n\epsilon_n \]
\[ \leq H(Y^n_j) + n\epsilon_n \]
\[ \leq \sum_{i=1}^n H(Y_{ji}) + n\epsilon_n \]

for \( j = 1, 2, 3 \). Similarly, for \( j \neq k \) and \( j, k \in \{1, 2, 3\} \), we have

\[ n(R_j + R_k) \leq H(M_j) + H(M_k|M_j) \]
\[ = I(M_j; Y^n_j) + H(M_j|Y^n_j) + I(M_k; Y^n_k|M_j) + H(M_k|Y^n, M_j) \]
\[ \leq I(M_j; Y^n_j) + H(M_j|Y^n_j) + I(M_k; Y^n_k|M_j) + H(M_k|Y^n_k) \]
\[ \leq I(M_j; Y^n_j) + I(M_k; Y^n_k|M_j) + 2n\epsilon_n \]
\[ \leq I(M_j, M_k; Y^n_j, Y^n_k) + 2n\epsilon_n \]
\[ \leq \sum_{i=1}^n H(Y_{ji}, Y_{ki}) + n\epsilon_n. \]

For the last inequality in the capacity region, consider

\[ n(R_1 + R_2 + R_3) \leq H(M_1) + H(M_2|M_1) + H(M_3|M_1, M_2) \]
\[ \leq I(M_1; Y^n_1) + H(M_1|Y^n_1) + I(M_2; Y^n_2|M_1) + H(M_2|Y^n_2) + \]
\[ + I(M_3; Y^n_3|M_1, M_2) + H(M_3|Y^n_3) \]
\[ \leq I(M_1, M_2, M_3; Y^n_1, Y^n_2, Y^n_3) + 3n\epsilon_n \]
\[ \leq \sum_{i=1}^n H(Y_{i1}, Y_{i2}, Y_{i3}) + n\epsilon_n. \]

The rest of the proof follows by introducing the standard time sharing random variable and letting \( n \to \infty \).

(b) The capacity region consists of the rate triples \((R_1, R_2, R_3)\) such that

\[ R_1 < H(Y_1), \]
\[ R_2 < H(Y_2), \]
\[ R_3 < H(Y_3) \]

for some pmf \( p(x) \).

\textbf{Achievability:} Fix a pmf \( p(x) \). Randomly and independently generate \( 2^{n(R_1+R_2+R_3)} \) sequences \( x^n(m_1, m_2, m_3), m_j \in [1:2^{nR_j}] \), for \( j = 1, 2, 3 \), each according to \( \prod_{i=1}^n p_X(x_i) \). To send the message triple \((m_1, m_2, m_3)\), the encoder transmits \( x^n(m_1, m_2, m_3) \). Decoder 1 finds the unique message \( \hat{m}_1 \) such that \((x^n(\hat{m}_1, m_2, m_3), y^n_1) \in T^{(n)}_x(X, Y_1) \), decoder 2 finds the unique message \( \hat{m}_2 \) such that \((x^n(m_1, \hat{m}_2, m_3), y^n_2) \in T^{(n)}_x(X, Y_2) \), and decoder 3 finds the unique message \( \hat{m}_3 \) such that \((x^n(m_1, m_2, \hat{m}_3), y^n_3) \in T^{(n)}_x(X, Y_3) \). By the LLN and the packing lemma, the average probability of error tends to 0 as \( n \to \infty \) if \( R_j < I(X; Y_j) - \delta(\epsilon) \). The proof of achievability follows by letting \( \epsilon \to 0 \).

\textbf{Proof of converse:} Consider

\[ nR_j \leq H(M_j|M(A_j)) \]
\[ = I(M_j; Y^n_j|M(A_j)) + H(M_j|Y^n_j, M(A_j)) \]
\[ I(M_j; Y_j^n | M(A_j)) + n\epsilon_n \leq H(Y_j^n) + n\epsilon_n \]
\[ \leq \sum_{i=1}^{n} H(Y_{ji}) + n\epsilon_n \]

for \( j = 1, 2, 3 \). The rest of the proof follows by standard arguments.

(c) The capacity region consists of the rate triples \((R_1, R_2, R_3)\) such that
\[
R_1 < H(Y_1), \\
R_2 < H(Y_2), \\
R_3 < H(Y_3), \\
R_1 + R_2 < H(Y_1, Y_2), \\
R_1 + R_3 < H(Y_1, Y_3), \\
R_2 + R_3 < H(Y_2, Y_3), \\
R_1 + R_2 + R_3 < H(Y_1, Y_2, Y_3)
\]
for some pmf \( p(x) \). Notice that the capacity is the same as part a) and does not increase with this set of receiver side information. Achievability follows from part (a). For the converse, consider
\[
nR_j \leq H(M_j | M(A_j)) \\
= I(M; Y_j^n | M(A_j)) + H(M | M_j, M(A_j)) \\
\leq I(M; Y_j^n | M(A_j)) + n\epsilon_n \\
\leq H(Y_j^n) + n\epsilon_n \\
\leq \sum_{i=1}^{n} H(Y_{ji}) + n\epsilon_n
\]
for \( j = 1, 2, 3 \). Now, consider
\[
n(R_1 + R_2) \leq H(M_1) + H(M_2 | M_1) \\
= I(M_1; Y_1^n) + H(M_1 | Y_1^n) + I(M_2; Y_2^n | M_1) + H(M_2 | Y_2^n, M_1) \\
\leq I(M_1; Y_1^n) + I(M_2; Y_2^n | M_1) + 2n\epsilon_n \\
\leq I(M_1, M_2; Y_1^n, Y_2^n) + 2n\epsilon_n \\
\leq \sum_{i=1}^{n} H(Y_{1i}, Y_{2i}) + 2n\epsilon_n.
\]
Similarly, we have
\[
n(R_2 + R_3) \leq H(M_2 | M_1) + H(M_3 | M_1, M_2) \\
= I(M_2; Y_2^n | M_1) + H(M_2 | Y_2^n, M_1) + I(M_3; Y_3^n | M_1, M_2) + H(M_3 | Y_3^n, M_1, M_2) \\
\leq I(M_2, M_3; Y_2^n, Y_3^n | M_1) + 2n\epsilon_n \\
\leq \sum_{i=1}^{n} H(Y_{2i}, Y_{3i}) + 2n\epsilon_n
\]
and
\[
n(R_1 + R_3) \leq H(M_1) + H(M_3 | M_1, M_2)
\]
\[
n(R_1 + R_2 + R_3) \leq H(M_1) + H(M_2 | M_1) + H(M_3 | M_1, M_2)
\]
\[
\begin{align*}
I(M_1; Y^n_1) + H(M_1 | Y^n_1) + I(M_3; Y^n_3 | M_1, M_2) + H(M_3 | Y^n_3, M_1, M_2) \\
\leq I(M_1, M_2, M_3; Y^n_1, Y^n_3) + 2n\epsilon_n \\
\leq \sum_{i=1}^n H(Y_{1i}, Y_{3i}) + 2n\epsilon_n.
\end{align*}
\]

Finally, for the last term, consider

\[
\begin{align*}
n(R_1 + R_2 + R_3) &\leq H(M_1) + H(M_2 | M_1) + H(M_3 | M_1, M_2) \\
&= I(M_1; Y^n_1) + H(M_1 | Y^n_1) + I(M_2; Y^n_2 | M_1) + H(M_2 | Y^n_2, M_2) + \\
&\quad + I(M_3; Y^n_3 | M_1, M_2) + H(M_3 | Y^n_3, M_1, M_2), \\
&\leq I(M_1, M_2, M_3; Y^n_1, Y^n_2, Y^n_3) + 3n\epsilon_n \\
&\leq \sum_{i=1}^n H(Y_{1i}, Y_{2i}, Y_{3i}) + 3n\epsilon_n.
\end{align*}
\]

The rest of the proof follows by standard arguments.