IERG 6154: Network Information Theory
Homework 5

(Degraded broadcast channels and superposition coding)

Due: March 25, 2019

1. Let $X \to (Y, Z)$ be a broadcast channel. Show that the following two regions are identical.

   Region I: The union over rate pairs (R_1, R_2) that satisfy

   $$\begin{aligned} R_2 &\leq I(V; Z) \\ R_1 &\leq I(X; Y|V) \\ R_1 + R_2 &\leq I(X; Y) \end{aligned}$$

   for any $V \to X \to (Y, Z)$ is Markov.

   Region II: The union over rate pairs (R_1, R_2) that satisfy

   $$\begin{aligned} R_2 &\leq I(V; Z) \\ R_1 + R_2 &\leq I(V; Z) + I(X; Y|V) \\ R_1 + R_2 &\leq I(X; Y) \end{aligned}$$

   for any $V \to X \to (Y, Z)$ is Markov.

   Hint: consider the corner points.

2. Let $X \to (Y, Z)$ be a broadcast channel. Complete the details of the following achievability scheme. Consider a triple (U, V, X) such that U and V are independent, and $X = x(u, v)$ (a function of (U, V)). Generate 2^{nR_2} codewords $V^n(m_2)$ independently according to $\prod_i p(v_i)$. Independently generate 2^{nR_1} codewords $U^n(m_1)$ independently according to $\prod_i p(u_i)$. The codeword $X^n(m_1, m_2)$ is formed as $x_i(m_1, m_2) = f(u_i(m_1), v_i(m_2))$. Show that there is a decoding scheme to achieve the following rate pairs (Note that Y only needs to deode M_1 correctly.)

   $$\begin{aligned} R_2 &\leq I(V; Z) \\ R_1 &\leq I(U; Y|V) = I(U; YV) \\ R_1 + R_2 &\leq I(UV; Y) \end{aligned}$$

3. Let $X \to (Y, Z)$ be a broadcast channel. Show that the region obtained by taking the union of rate pairs satisfying

   $$\begin{aligned} R_2 &\leq I(V; Z) \\ R_1 &\leq I(U; Y|V) = I(U; YV) \\ R_1 + R_2 &\leq I(UV; Y) \end{aligned}$$

   over all $p(u)p(v)x(u, v)$ (as in Question 2 above) yields the superposition coding region (Region I) for a given broadcast channel.

   Note that the codebook storage space is greatly simplified $\left(2^{nR_1} + 2^{nR_2} \text{ vs } 2^{n(R_1+R_2)}\right)$ when using the description in Question 2. This also generalizes the coding scheme for the two BSC's.

   Hint: You can use the following functional representation lemma. Given $p(v, x)$ we can always find a U independent of V such that X is a function of (U, V).

4. Let $X \to Y$ be a Z-channel with $\mathrm{P}(Y = 0|X = 0) = 1, \mathrm{P}(Y = 1|X = 1) = p$ and $X \to Z$ is another Z-channel with $\mathrm{P}(Z = 0|X = 0) = 1, \mathrm{P}(Z = 1|X = 1) = pq$ where $0 \leq p \leq 1, 0 \leq q \leq 1$.

   (a) Show that one can consider $X \to (Y, Z)$ to be a stochastically degraded broadcast channel.

(b) Show that when $\lambda q \leq 1$ then $I(X;Y) - \lambda I(X;Z)$ is concave in $p(x)$. Hence conclude that $\max_{(R_1,R_2)\in\mathcal{C}} \lambda R_2 + R_1 = I(X;Y)$.

(c) Show that when $\lambda \geq \frac{1-pq}{q-pq}$ then $I(X;Y) - \lambda I(X;Z)$ is convex in $p(x)$. Hence conclude that $\max_{(R_1,R_2)\in\mathcal{C}} \lambda R_2 + R_1 = \lambda I(X;Z)$.

(d) When $\frac{1}{q} < \lambda < \frac{1-pq}{q-pq}$ show that the function $f(x) = I(X;Y) - \lambda I(X;Z)$ (where $x = \mathrm{P}(X=1)$) is convex for $x \in [0, x_*]$ and concave when $x \in [x_*, 1]$ for some $x_* \in [0,1]$.

(e) Using above show that $\mathfrak{C}[f]$ has the following shape: (i) $\mathfrak{C}[f]$ is a line segment joining the point 0 at $x=0$ to the point $f(x^\sharp)$ at $x^\sharp$ for some $x^\sharp \in [x_*, 1]$, (ii) $\mathfrak{C}[f](x) = f(x)$ in the segment $[x^\sharp, 1]$.

(f) Hence conclude that to obtain $\max_{(R_1,R_2)\in\mathcal{C}} \lambda R_2 + R_1$ it suffices to consider $U \to X \sim Z(s)$ for some $s \in [0,1]$. (You can assume $\lambda \geq 1$.) Hence $Z(s)$ is a Z channel as above with $\mathrm{P}(X=0|U=0) = 1, \mathrm{P}(X=1|U=1) = s$. Note that I am not insisting that the distribution on U be uniform.

5. *Three receiver less noisy broadcast channel*
Consider a broadcast channel where the sender X has access to three independent messages (M_1, M_2, M_3) and he wishes to communicate these messages to receivers Y_1, Y_2, Y_3 respectively. Assume that Y_1 is less noisy than Y_2 and Y_2 is less noisy than Y_3. Consider the region, $\mathcal{R}$,defined by the set of all rate-pairs satisfying

$$\begin{aligned} R_3 &\leq I(U_3;Y_3) \\ R_2 &\leq I(U_2;Y_2|U_3) \\ R_1 &\leq I(X;Y_1|U_2,U_3). \end{aligned}$$

for some $(U_3, U_2) \to X \to (Y_1, Y_2, Y_3)$ forming a Markov chain.

Show the following:

a) If $(R_1, R_2, R_3) \in \mathcal{R}$ then both $(R_1, R_2 + R_3, 0)$ and $(R_1 + R_2 + R_3, 0, 0)$ belong to $\mathcal{R}$.

b) Show that using superposition coding the region $\mathcal{R}$ is achievable.

c) Show that for $\lambda_2, \lambda_3 \geq 0$ the function defined by

$$F_{\vec{\lambda}}(X) := \lambda_3 I(X;Y_3) + \mathfrak{C}[\lambda_2 I(X;Y_2) - \lambda_3 I(X;Y_3) + \mathfrak{C}[I(X;Y_1) - \lambda_2 I(X;Y_2)]]$$

is sub-additive. Here $\mathfrak{C}[]$ denotes the upper concave envelope w.r.t. $p(x)$.

d) Hence deduce that $\mathcal{R}$ is the capacity region for the 3-receiver less noisy broadcast channel.

6. *Characterization of degraded, less noisy, and more capable broadcast channel*
A binary input binary output channel can be characterized by two parameters $c, d \in [0,1] \times [0,1]$ where $c = \mathrm{P}(Y = 1|X=0), d = \mathrm{P}(Y=0|X=1)$. Further, w.l.o.g. we can assume that $c + d \leq 1$ (otherwise just swap the output labels). Indeed we can assume a strict inequality (otherwise Y is independent of X, which is not interesting). Define new parameters (α, β) as follows:

$$\alpha = \frac{c}{1-d}, \quad \beta = \frac{d}{1-c}.$$

Consider two channels $W_1(y_1|x)$ and $W_2(y_2|x)$ parameterized by (α_1, β_1) and (α_2, β_2) respectively. Show that

a) Y_2 is a degraded version of Y_1 *if and only if* $\alpha_1 \leq \alpha_2, \beta_1 \leq \beta_2$.

b) Y_2 is a less noisy version of Y_1 *if and only if*

$$\frac{\alpha_1}{(1-\alpha_1)(1-\beta_1)} \leq \frac{\alpha_2}{(1-\alpha_2)(1-\beta_2)}, \quad \frac{\beta_1}{(1-\alpha_1)(1-\beta_1)} \leq \frac{\beta_2}{(1-\alpha_2)(1-\beta_2)}.$$

c) Reverting back to the c, d notation, show that Y_1 is more-capable than Y_2 *if and only if*

$$D(c_1\|1-d_1) \geq D(c_2\|1-d_2), \quad D(1-d_1\|c_1) \geq D(1-d_2\|c_2),$$

where $D(x\|y)$ is the two-point divergence defined by $D(x\|y) = x\log_2(x/y) + (1-x)\log_2((1-x)/(1-y))$.