EE/Ma 126b Information Theory - Homework Set #6

Ling Li, ling@cs.caltech.edu

February 28, 2001

6.1 *The cooperative capacity of a multiple access channel.*

(a) Since X_1 and X_2 know both indices, we can regard the pair of (X_1, X_2) as one single codeword X. Thus for some distribution $p(x_1, x_2)$, or, distribution $p(x)$, we can design code $X(1), X(2), \dots, X(2^{R_1+R_2})$, i.i.d. $\sim p(x)$ and thus we will get the error $P_e^{(n)} \to 0$ when $n \to \infty$ if

$$R_1 + R_2 < I(X_1, X_2; Y).$$

(This is just the situation of one user channel.) And we also can get the converse that all code with error $P_e^{(n)} \to 0$ when $n \to \infty$ should have

$$R_1 + R_2 \le I(X_1, X_2; Y).$$

Thus we know that as taking X_1 and X_2 as a single codeword, the rate sum of R_1 and R_2 is bounded by

$$\boxed{R_1 + R_2 \le \max_{p(x_1,x_2)} I(X_1, X_2; Y) = C}. \tag{1}$$

And since we can achieve $(C, 0)$ by setting $X_1 = X$ and X_2 is null, or $(0, C)$ by setting X_1 null and $X_2 = X$, we know the capacity region is just (1).

(b) For $Y = X_1 + X_2$, $X_i \in \{0, 1\}$, we get

$$I(X_1, X_2; Y) = H(Y) - H(Y|X_1, X_2) = H(Y) \le \log 3,$$

since $Y \in \{0, 1, 2\}$. And we can achieve $\log 3$ by setting $p(0, 0) = p(1, 1) = \frac{1}{3}$ and $p(1, 0) + p(0, 1) = \frac{1}{3}$. Thus $C = \log 3$, and the capacity region is

$$\boxed{R_1 + R_2 \le \log 3}.$$

This region is larger than the one without cooperation between X_1 and X_2, since $\log 3 > 1.5$.

6.2 *Capacity of multiple access channels.*

(a) We can achieve the rate pair $(R_1, R_2) = (1, 0)$ by setting $X_2 = 0$. Similarly, setting $X_1 = 0$, we can achieve rate pair $(0, 1)$. For any distribution $p_1(x_1)p_2(x_2)$, we always have

$$I(X_1, X_2; Y) \le H(Y) \le 1,$$

since Y is binary. Thus the capacity region is just $\boxed{R_1 + R_2 \le 1}$.

(b) If we use the 1-1 mapping $\{-1,1\} \to \{0,1\} : -1 \mapsto 1, 1 \mapsto 0$ for both X_i and Y, the function $Y = X_1 \cdot X_2$ over $\{-1,1\}$ is exactly the function $Y = X_1 \oplus X_2$ over $\{0,1\}$. Since for discrete channels, we do not care what the symbol is, the capacity region is exactly the same as that in (a). That is, $\boxed{R_1 + R_2 \le 1}$.

6.3 *Gaussian multiple access channel capacity.* Let $f_1(x_1)$ be some normal distribution with variance $P_1 - \epsilon$ and $f_2(x_2)$ be some normal distribution with variance $P_2 - \epsilon$. The bold alphabets denote vectors of length n.

Codebook. Generate the codewords $\mathbf{X}_1(1), \mathbf{X}_1(2), \ldots, \mathbf{X}_1(2^{nR_1})$ i.i.d. $\sim \prod_{i=1}^n f_1(x_{1,i})$, and $\mathbf{X}_2(1), \mathbf{X}_2(2), \ldots, \mathbf{X}_2(2^{nR_2})$ i.i.d. $\sim \prod_{i=1}^n f_2(x_{2,i})$. These codewords form the codebook, which is revealed to the senders and the receiver.

Encoding. To send index i, sender 1 sends $\mathbf{X}_1(i)$. Similarly, to send j, sender 2 sends $\mathbf{X}_2(j)$.

Decoding. Let $A_\epsilon^{(n)}$ denote the set of typical $(\mathbf{X}_1, \mathbf{X}_2, \mathbf{Y})$ sequences. Then the receiver chooses the pair (i,j) such that

$$(\mathbf{X}_1(i), \mathbf{X}_2(j), \mathbf{Y}) \in A_\epsilon^{(n)}$$

if such a pair (i,j) exists and is unique; otherwise, an error is declared.

Error probability. By symmetry, assume that pair $(1,1)$ is sent. Define the events

$$E_1 = \left\{ \frac{1}{n} \sum_{i=1}^n X_{1,i}^2 > P_1 \right\}, E_2 = \left\{ \frac{1}{n} \sum_{i=1}^n X_{2,i}^2 > P_2 \right\},$$

and

$$E_{i,j} = \left\{ (\mathbf{X}_1(i), \mathbf{X}_2(j), \mathbf{Y}) \in A_\epsilon^{(n)} \right\}.$$

Then an error occurs if E_1 or E_2 occurs (the power constraints are violated) or $E_{1,1}^c$ occurs (the transmitted codewords and the received sequence are not jointly typical) or $E_{i,j}$ with $(i,j) \neq (1,1)$ occurs (some wrong codewords is jointly typical with the received sequence). Let P denote *the expectation over all codebooks of the conditional probability given* $(1,1)$ *sent.* That is, for some event $\mathcal{E}$,

$$P(\mathcal{E}) = E \Pr \{ \mathcal{E} | (1,1) \text{ sent} \}.$$

Thus

$$\begin{aligned} EP_e^{(n)} &= P\left(E_1 \cup E_2 \cup E_{1,1}^c \cup \bigcup_{(i,j)\neq(1,1)} E_{i,j} \right) \\ &\le P(E_1) + P(E_2) + P(E_{1,1}^c) + \sum_{i\neq 1} P(E_{i,1}) + \sum_{j \neq 1} P(E_{1,j}) + \sum_{i\neq 1}\sum_{j\neq 1} P(E_{i,j}). \end{aligned}$$

By the WLLN, $P(E_1) \to 0$ and $P(E_2) \to 0$ as $n \to \infty$. By the joint AEP, $P(E_{1,1}^c) \le \epsilon$ for n sufficiently large. Since by the code generation process, $\mathbf{X}_1(1)$ and $\mathbf{X}_1(i)$ are independent $(i \neq 1)$, so are $\mathbf{Y}$ and $\mathbf{X}_1(i)$ given $\mathbf{X}_2(1)$. Hence by the joint AEP, $P(E_{i,1}) \le$

$2^{-n(I(X_1;Y|X_2)-\epsilon)}$. Similarly, $P(E_{1,j}) \leq 2^{-n(I(X_2;Y|X_1)-\epsilon)}$. And for $i \neq 1$ and $j \neq 1$, $\mathbf{X}_1(i), \mathbf{X}_2(j)$ and $\mathbf{Y}$ are independent, so by joint AEP $P(E_{i,j}) \leq 2^{-n(I(X_1,X_2;Y)-\epsilon)}$. Thus

$$\begin{aligned} EP_e^{(n)} &\leq 3\epsilon + 2^{-n(I(X_1;Y|X_2)-R_1-\epsilon)} + 2^{-n(I(X_2;Y|X_1)-R_2-\epsilon)} \\ &\quad + 2^{-n(I(X_1,X_2;Y)-R_1-R_2-\epsilon)} \\ &\leq 6\epsilon \end{aligned}$$

for n sufficiently large and

$$R_1 < I(X_1;Y|X_2), \tag{2}$$
$$R_2 < I(X_2;Y|X_1), \tag{3}$$
$$R_1 + R_2 < I(X_1, X_2; Y). \tag{4}$$

This proves the existence of good $((2^{nR_1}, 2^{nR_2}), n)$ code. And by choosing a good codebook and deleting the worst half of both sets of codewords, we obtain a code with low maximal probability of error. In particular, the power constraints are satisfied by each of the remaining codewords (similar to to proof for the single user Gaussian channel).

The bounds (2), (3) and (4) give the achievable region for a specific distribution. The full achievable region is the closure of the convex hull of those regions for any distribution. Since X_1 and X_2 are independent (by the code generation), and the noise Z is also independent of X_1 and X_2, we have

$$I(X_1;Y|X_2) = I(X_1; X_1 + X_2 + Z|X_2) = h(X_1 + Z) - h(X_1) \leq \frac{1}{2}\log\left(1 + \frac{P_1}{N}\right), \tag{5}$$

and we can achieve the equality by setting $X_1 \sim \mathcal{N}(0, P_1)$. Similarly, $I(X_2;Y|X_1) \leq \frac{1}{2}\log\left(1 + \frac{P_2}{N}\right)$ with equality when $X_2 \sim \mathcal{N}(0, P_2)$. And

$$I(X_1, X_2; Y) = I(X_1, X_2; X_1 + X_2 + Z) = h(X_1 + X_2 + Z) - h(Z) \leq \frac{1}{2}\log\left(1 + \frac{P_1 + P_2}{N}\right) \tag{6}$$

with equality when $X_1 \sim \mathcal{N}(0, P_1)$ and $X_2 \sim \mathcal{N}(0, P_2)$. So finally, we get the achievable region is

$$\begin{aligned} R_1 &< \frac{1}{2}\log\left(1 + \frac{P_1}{N}\right), \\ R_2 &< \frac{1}{2}\log\left(1 + \frac{P_2}{N}\right), \\ R_1 + R_2 &< \frac{1}{2}\log\left(1 + \frac{P_1 + P_2}{N}\right). \end{aligned}$$

6.4 *Converse for the Gaussian multiple access channel.* Consider any $((2^{nR_1}, 2^{nR_2}), n)$ code that satisfies the power constraints and with error probability $P_e^{(n)} \to 0$ when $n \to \infty$. Then

$$
\begin{aligned}
nR_1 &= H(W_1) = I(W_1;\mathbf{Y}) + H(W_1|\mathbf{Y}) \\
&\le I(W_1;\mathbf{Y}) + n\epsilon_n \\
&\le I(\mathbf{X}_1(W_1);\mathbf{Y}) + n\epsilon_n \\
&= h(\mathbf{X}_1(W_1)) - h(\mathbf{X}_1(W_1)|\mathbf{Y}) + n\epsilon_n \\
&\le h(\mathbf{X}_1(W_1)|\mathbf{X}_2(W_2)) - h(\mathbf{X}_1(W_1)|\mathbf{Y},\mathbf{X}_2(W_2)) + n\epsilon_n \\
&= I(\mathbf{X}_1(W_1);\mathbf{Y}|\mathbf{X}_2(W_2)) + n\epsilon_n \\
&= h(\mathbf{Y}|\mathbf{X}_2(W_2)) - h(\mathbf{Y}|\mathbf{X}_1(W_1),\mathbf{X}_2(W_2)) + n\epsilon_n \\
&= h(\mathbf{Y}|\mathbf{X}_2(W_2)) - \sum_{i=1}^{n} h(Y_i|Y_1^{i-1},\mathbf{X}_1(W_1),\mathbf{X}_2(W_2)) + n\epsilon_n \\
&= h(\mathbf{Y}|\mathbf{X}_2(W_2)) - \sum_{i=1}^{n} h(Y_i|X_{1,i},X_{2,i}) + n\epsilon_n \\
&\le \sum_{i=1}^{n} h(Y_i|\mathbf{X}_2(W_2)) - \sum_{i=1}^{n} h(Y_i|X_{1,i},X_{2,i}) + n\epsilon_n \\
&\le \sum_{i=1}^{n} h(Y_i|X_{2,i}) - \sum_{i=1}^{n} h(Y_i|X_{1,i},X_{2,i}) + n\epsilon_n \\
&= \sum_{i=1}^{n} I(X_{1,i};Y_i|X_{2,i}) + n\epsilon_n.
\end{aligned}
$$

This is very similar to the proof in the discrete case and all the reasons are the same (though in continuous version). Now let $P_{1,i}$ be the average power of the ith codeword $X_{1,i}$, i.e.,

$$P_{1,i} = \frac{1}{2^{nR_1}} \sum_w X_{1,i}^2(w).$$

Then since $Y_i = X_{1,i} + X_{2,i} + Z_i$, and $X_{1,i}$, $X_{2,i}$ and Z_i are independent (since W_1 and W_2 are independent and Z are independent of (X_1, X_2)), we have (similar to (5))

$$I(X_{1,i};Y_i|X_{2,i}) = h(X_{1,i}+Z_i) - h(X_{1,i}) \le \frac{1}{2}\log\left(1+\frac{P_{1,i}}{N}\right).$$

Continuing with the inequalities of the converse, we get

$$
\begin{aligned}
R_1 &\le \frac{1}{n}\sum_{i=1}^{n}\frac{1}{2}\log\left(1+\frac{P_{1,i}}{N}\right) + \epsilon_n \\
&\le \frac{1}{2}\log\left(1+\frac{1}{n}\sum_{i=1}^{n}\frac{P_{1,i}}{N}\right) + \epsilon_n \\
&\le \frac{1}{2}\log\left(1+\frac{P_1}{N}\right) + \epsilon_n,
\end{aligned}
$$

since $\frac{1}{2}\log(1+x)$ is a concave function of x and we also have the power constraint $\frac{1}{n}\sum_i P_{1,i} \le P_1$. Together with $\epsilon_n \to 0$ when $n \to \infty$, we have

$$\boxed{R_1 \le \frac{1}{2}\log\left(1+\frac{P_1}{N}\right)}.$$

Similarly, we have

$$\boxed{R_2 \le \frac{1}{2}\log\left(1+\frac{P_2}{N}\right)}.$$

To bound the sum of the rates, we have (oh, I won't write all the details. They are just an analog to the proof in the discrete case.)

$$n(R_1+R_2) \le \sum_{i=1}^{n} I(X_{1,i}, X_{2,i}; Y_i) + n\epsilon_n.$$

With (6), we obtain $I(X_{1,i}, X_{2,i}; Y_i) \le \frac{1}{2}\log\left(1+\frac{P_{1,i}+P_{2,i}}{N}\right)$. Thus

$$\begin{aligned} R_1+R_2 &\le \frac{1}{n}\sum_{i=1}^{n}\frac{1}{2}\log\left(1+\frac{P_{1,i}+P_{2,i}}{N}\right)+\epsilon_n \\ &\le \frac{1}{2}\log\left(1+\frac{1}{n}\sum_{i=1}^{n}\frac{P_{1,i}+P_{2,i}}{N}\right)+\epsilon_n \\ &\le \frac{1}{2}\log\left(1+\frac{P_1+P_2}{N}\right)+\epsilon_n. \end{aligned}$$

Since $\epsilon_n \to 0$ when $n \to \infty$, we get

$$\boxed{R_1+R_2 \le \frac{1}{2}\log\left(1+\frac{P_1+P_2}{N}\right)}.$$

This finally finishes the proof of the converse.

6.5 *Unusual multiple access channel.*

(a) If we set $X_1 = 0$, the channel becomes $Y = X_2$ and we can send X_2 at a rate $R_2 = 1$. Thus the rate pair $(0,1)$ is achievable. Symmetrically, the pair $(1,0)$ is also achievable.

(b) For any non-degenerate distribution $p_1(x_1)p_2(x_2)$, we have

$$\begin{aligned} I(X_1,X_2;Y) &= H(Y) - H(Y|X_1,X_2) \\ &= H(Y) - \sum_{x_1,x_2} H(Y|X_1=x_1, X_2=x_2)p_1(x_1)p_2(x_2) \\ &= H(Y) - H(Y|X_1=1,X_2=1)p_1(1)p_2(1) \\ &= H(Y) - p_1(1)p_2(1) \\ &< 1. \end{aligned}$$

The reason for the inequality is that Y is binary so $H(Y) \le 1$, and $p_1(1)p_2(1) > 0$.

(c) Consider the rate pair $(\frac{1}{2},\frac{1}{2})$. It can be achieved by timesharing of those two rate pairs $(1,0)$ and $(0,1)$. However, for any non-degenerate distribution $p_1(x_1)p_2(x_2)$, we have $I(X_1,X_2;Y) < 1$, and for any degenerate distribution, we have $I(X_1;Y|X_2) = 0$ and/or $I(X_2;Y|X_1) = 0$. Thus the pair $(\frac{1}{2},\frac{1}{2})$ can not be within the region defined by

$$\begin{aligned} R_1 &\le I(X_1;Y|X_2), \\ R_2 &\le I(X_2;Y|X_1), \\ R_1+R_2 &\le I(X_1,X_2;Y), \end{aligned}$$

for any $p_1(x_1)p_2(x_2)$.