Homework Set #3
MAC and Compound Channel

1. **Converse for the Gaussian multiple access channel.** Prove the converse for the Gaussian multiple access channel by extending the converse in the discrete case to take into account the power constraint on the codewords.

   Converse for the Gaussian multiple access channel. The proof of the converse for the Gaussian case proceeds on very similar lines to the discrete case. However, for the Gaussian case, the two stages of proof that were required in the discrete case, namely, of finding a new expression for the capacity region and then proving a converse, can be combined into one single step.

   By the code construction, it is possible to estimate (W_1, W_2) from the received sequence Y^n with a low probability of error. Hence the conditional entropy of (W_1, W_2) given Y^n must be small. By Fano's inequality,

$$H(W_1, W_2|Y^n) \leq n(R_1 + R_2)P_e^{(n)} + H(P_e^{(n)}) \stackrel{\triangle}{=} n\epsilon_n. \quad (1)$$

   It is clear that $\epsilon_n \to 0$ as $P_e^{(n)} \to 0$.

   Then we have

$$\begin{aligned} H(W_1|Y^n) &\leq H(W_1, W_2|Y^n) \leq n\epsilon_n, && (2)\\ H(W_2|Y^n) &\leq H(W_1, W_2|Y^n) \leq n\epsilon_n. && (3) \end{aligned}$$

We can now bound the rate R_1 as

$$
\begin{aligned}
nR_1 &= H(W_1) && (4)\\
&= I(W_1;Y^n) + H(W_1|Y^n) && (5)\\
&\overset{(a)}{\leq} I(W_1;Y^n) + n\epsilon_n && (6)\\
&\overset{(b)}{\leq} I(X_1^n(W_1);Y^n) + n\epsilon_n && (7)\\
&= H(X_1^n(W_1)) - H(X_1^n(W_1)|Y^n) + n\epsilon_n && (8)\\
&\overset{(c)}{\leq} H(X_1^n(W_1)|X_2^n(W_2)) - H(X_1^n(W_1)|Y^n, X_2^n(W_2)) + n\epsilon_n && (9)\\
&= I(X_1^n(W_1);Y^n|X_2^n(W_2)) + n\epsilon_n && (10)\\
&= h(Y^n|X_2^n(W_2)) - h(Y^n|X_1^n(W_1), X_2^n(W_2)) + n\epsilon_n && (11)\\
&\overset{(d)}{=} h(Y^n|X_2^n(W_2)) - h(Z^n|X_1^n(W_1), X_2^n(W_2)) + n\epsilon_n && (12)\\
&\overset{(e)}{=} h(Y^n|X_2^n(W_2)) - h(Z^n) + n\epsilon_n && (13)\\
&\overset{(f)}{=} h(Y^n|X_2^n(W_2)) - \sum_{i=1}^{n} h(Z_i) + n\epsilon_n && (14)\\
&\overset{(g)}{\leq} \sum_{i=1}^{n} h(Y_i|X_2^n(W_2)) - \sum_{i=1}^{n} h(Z_i) + n\epsilon_n && (15)\\
&\overset{(h)}{\leq} \sum_{i=1}^{n} h(Y_i|X_{2i}) - \sum_{i=1}^{n} h(Z_i) + n\epsilon_n && (16)\\
&\overset{(i)}{=} \sum_{i=1}^{n} h(X_{1i} + Z_i|X_{2i}) - \sum_{i=1}^{n} h(Z_i) + n\epsilon_n && (17)\\
&\overset{(j)}{=} \sum_{i=1}^{n} h(X_{1i} + Z_i) - \sum_{i=1}^{n} h(Z_i) + n\epsilon_n && (18)\\
&\overset{(k)}{\leq} \sum_{i=1}^{n} \frac{1}{2}\log 2\pi e(P_{1i} + N) - \frac{1}{2}\log 2\pi eN + n\epsilon_n && (19)\\
&= \sum_{i=1}^{n} \frac{1}{2}\log\left(1 + \frac{P_{1i}}{N}\right) + n\epsilon_n && (20)
\end{aligned}
$$

where
(a) follows from Fano's inequality,
(b) from the data processing inequality,

(c) from the fact that since W_1 and W_2 are independent, so are $X_1^n(W_1)$ and $X_2^n(W_2)$, and hence it follows that $H(X_1^n(W_1)|X_2^n(W_2)) = H(X_1^n(W_1))$, and $H(X_1^n(W_1)|Y^n, X_2^n(W_2)) \leq H(X_1^n(W_1)|Y^n)$ by conditioning,
(d) from the fact that $Y^n = X_1^n + X_2^n + Z^n$,
(e) from the fact that Z^n is independent of X_1^n and X_2^n,
(f) from the fact that the noise is i.i.d.,
(g) from the chain rule and removing conditioning,
(h) from removing conditioning,
(i) from the fact that $Y_i = X_{1i} + X_{2i} + Z_i$,
(j) from the fact that X_{1i} and Z_i are independent of X_{2i}, and
(k) from the entropy maximizing property of the normal (Theorem 9.6.5), after defining $P_{1i} = EX_{1i}^2$.

Hence, we have

$$R_1 \leq \frac{1}{n} \sum_{i=1}^{n} \frac{1}{2} \log \left(1 + \frac{P_{1i}}{N}\right) + \epsilon_n. \tag{21}$$

Similarly, we have

$$R_2 \leq \frac{1}{n} \sum_{i=1}^{n} \frac{1}{2} \log \left(1 + \frac{P_{2i}}{N}\right) + \epsilon_n. \tag{22}$$

To bound the sum of the rates, we have

$$
\begin{aligned}
n(R_1+R_2) &= H(W_1,W_2) && (23)\\
&= I(W_1,W_2;Y^n)+H(W_1,W_2|Y^n) && (24)\\
&\stackrel{(a)}{\le} I(W_1,W_2;Y^n)+n\epsilon_n && (25)\\
&\stackrel{(b)}{\le} I(X_1^n(W_1),X_2^n(W_2);Y^n)+n\epsilon_n && (26)\\
&= h(Y^n)-h(Y^n|X_1^n(W_1),X_2^n(W_2))+n\epsilon_n && (27)\\
&\stackrel{(c)}{=} h(Y^n)-h(Z^n)+n\epsilon_n && (28)\\
&\stackrel{(d)}{=} h(Y^n)-\sum_{i=1}^n h(Z_i)+n\epsilon_n && (29)\\
&\stackrel{(e)}{\le} \sum_{i=1}^n h(Y_i)-\sum_{i=1}^n h(Z_i)+n\epsilon_n && (30)\\
&\stackrel{(f)}{\le} \sum_{i=1}^n \frac{1}{2}\log 2\pi e(P_{1i}+P_{2i}+N)-\frac{1}{2}\log 2\pi e N+n\epsilon_n && (31)\\
&= \frac{1}{2}\log\left(1+\frac{P_{1i}+P_{2i}}{N}\right)+n\epsilon_n && (32)
\end{aligned}
$$

where
(a) follows from Fano's inequality,
(b) from the data processing inequality,
(c) from the fact that $Y^n = X_1^n + X_2^n + Z^n$, and Z^n is independent of X_1^n and X_2^n,
(d) from the fact that Z_i are i.i.d., (e) follows from the chain rule and removing conditioning, and
(f) from the entropy maximizing property of the normal, and the definitions of P_{1i} and P_{2i}.

Hence we have

$$R_1+R_2 \le \frac{1}{n}\sum_{i=1}^n \frac{1}{2}\log\left(1+\frac{P_{1i}+P_{2i}}{N}\right)+\epsilon_n. \qquad (33)$$

The power constraint on the codewords imply that

$$\frac{1}{n}\sum_{i=1}^{n} P_{1i} \leq P_1, \tag{34}$$

and

$$\frac{1}{n}\sum_{i=1}^{n} P_{2i} \leq P_2. \tag{35}$$

Now since log is concave function, we can apply Jensens inequality to the expressions in (21), (22) and (33). Thus we obtain

$$R_1 \leq \frac{1}{2}\log\left(1 + \frac{\frac{1}{n}\sum_{i=1}^{n} P_{1i}}{N}\right) + \epsilon_n \tag{36}$$

$$R_2 \leq \frac{1}{2}\log\left(1 + \frac{\frac{1}{n}\sum_{i=1}^{n} P_{2i}}{N}\right) + \epsilon_n \tag{37}$$

$$R_1 + R_2 \leq \frac{1}{2}\log\left(1 + \frac{\frac{1}{n}\sum_{i=1}^{n} P_{1i} + P_{2i}}{N}\right) + \epsilon_n. \tag{38}$$

which when combined with the power constraints, and taking the limit at $n \to \infty$, we obtain the desired converse, i.e.,

$$R_1 < \frac{1}{2}\log(1 + \frac{P_1}{N}), \tag{39}$$

$$R_2 < \frac{1}{2}\log(1 + \frac{P_2}{N}), \tag{40}$$

$$R_1 + R_2 < \frac{1}{2}\log(1 + \frac{P_1 + P_2}{N}). \tag{41}$$

2. **A multiple access identity.**
Let $C(x) = \frac{1}{2}\log(1 + x)$ denote the channel capacity of a Gaussian channel with signal to noise ratio x. Show

$$C\left(\frac{P_1}{N}\right) + C\left(\frac{P_2}{P_1 + N}\right) = C\left(\frac{P_1 + P_2}{N}\right).$$

This suggests that 2 independent users can send information as well as if they had pooled their power.

Solutions.

$$C(\frac{P_1+P_2}{N}) = \frac{1}{2}\log(1+\frac{P_1+P_2}{N}) \quad (42)$$
$$= \frac{1}{2}\log(\frac{N+P_1+P_2}{N}) \quad (43)$$
$$= \frac{1}{2}\log(\frac{N+P_1+P_2}{N+P_1}\cdot\frac{N+P_1}{N}) \quad (44)$$
$$= \frac{1}{2}\log(\frac{N+P_1+P_2}{N+P_1}) + \frac{1}{2}\log(\frac{N+P_1}{N}) \quad (45)$$
$$= C(\frac{P_2}{P_1+N}) + C(\frac{P}{N_1}) \quad (46)$$

3. **Gaussian multiple access.**

A group of m users, each with power P, is using a Gaussian multiple access channel at capacity, so that

$$\sum_{i=1}^{m} R_i = C\left(\frac{mP}{N}\right), \quad (47)$$

where $C(x) = \frac{1}{2}\log(1+x)$ and N is the receiver noise power.

A new user of power P_0 wishes to join in.

(a) At what rate can he send without disturbing the other users?

(b) What should his power P_0 be so that the new users rate is equal to the combined communication rate $C(mP/N)$ of all the other users?

Solustions Gaussian multiple access.

(a) If the new user can be decoded while treating all the other senders as part of the noise, then his signal can be subtracted out before decoding the other senders, and hence will not disturb the rates of the other senders. Therefore if

$$R_0 < \frac{1}{2}\log\left(1+\frac{P_0}{mP+N}\right), \quad (48)$$

the new user will not disturb the other senders.

(b) The new user will have a rate equal to the sum of the existing senders if

$$\frac{1}{2}\log\left(1+\frac{P_0}{mP+N}\right)=\frac{1}{2}\log\left(1+\frac{mP}{N}\right) \tag{49}$$

or

$$P_0=(mP+N)\frac{mP}{N} \tag{50}$$

4. **Frequency Division Multiple Access (FDMA).** Maximize the throughput $R_1+R_2=W_1\log(1+\frac{P_1}{NW_1})+(W-W_1)\log(1+\frac{P_2}{N(W-W_1)})$ over W_1 to show that bandwidth should be proportional to transmitted power for FDMA.

Solustions Frequency Division Multiple Access (FDMA).

Allocating bandwidth W_1 and $W_2=W-W_1$ to the two senders, we can achieve the following rates

$$R_1=W_1\log\left(1+\frac{P_1}{NW_1}\right), \tag{51}$$

$$R_2=W_2\log\left(1+\frac{P_2}{NW_2}\right). \tag{52}$$

To maximimize the sum of the rates, we write

$$R=R_1+R_2=W_1\log\left(1+\frac{P_1}{NW_1}\right)+(W-W_1)\log\left(1+\frac{P_2}{N(W-W_1)}\right) \tag{53}$$

and differentiating with respect to W_1, we obtain

$$\log\left(1+\frac{P_1}{NW_1}\right)+\frac{W_1}{1+\frac{P_1}{NW_1}}\left(-\frac{P_1}{NW_1^2}\right)$$
$$-\log\left(1+\frac{P_2}{N(W-W_1)}\right)+\frac{W-W_1}{1+\frac{P_2}{N(W-W_1)}}\left(\frac{P_2}{N(W-W_1)^2}\right)=0 \tag{54}$$

Instead of solving this equation, we can verify that if we set

$$W_1=\frac{P_1}{P_1+P_2}W \tag{55}$$

so that

$$\frac{P_1}{NW_1} = \frac{P_2}{NW_2} = \frac{P_1 + P_2}{NW} \tag{56}$$

that (54) is satisfied, and that using bandwidth proportional to the power optimizes the total rate for Frequency Division Multiple Access.

5. **Compound channel with feedback**. In the class we introduced the memoryless compound channel $(\mathcal{X}; p(y|x,s); \mathcal{Y})$ where $s \in \mathcal{S}$ is the state of the channel. Thought this question we assume that the alphabets $\mathcal{X}, \mathcal{Y}, \mathcal{S}$ are all finite. A $(2^{nR}, n)$ code for the compound channel is defined in the same way as for the DMC (see lecture notes). The average probability of error is defined as

$$P_e^{(n)} = \sup_s P\left\{\hat{W} \neq W, s \text{ is the actual channel}\right\}$$

A rate R is achievable if there exists a sequence of $(2^{nR}, n)$ codes with $P_e^{(n)} \to 0$.

(a) What is the capacity of the discrete compound channel with feedback? Prove converse and achievability.

(b) Compute the capacity of the compound binary eraser channel with feedback where the probability of an eraser is one of the four values $(0, 0.1, 0.2, 0.25)$.

(c) Write an expression and then sort from lower to higher the

(i) Capacity of compound channel with feedback when the state is not known.

(ii) Capacity of compound channel with no feedback when the state is not known.

(iii) Capacity of compound channel with feedback when the state known only at the encoder.

(iv) Capacity of compound channel with no feedback when the state known at the encoder.

(d) If the probability distribution that achieves the capacity of each channel is the same, does it imply that the capacity with feedback and without feedback are equal? If it does, prove it and if it does not give a counter example.

(e) If the capacity of the compound channel without feedback is zero, does it imply that the capacity with feedback is also zero? If it does, prove it and if it does not give a counter example.

(f) Under what conditions the capacity of the compound channel with feedback and without feedback has the same capacity.

Solution:

(a) The capacity of the compound channel with feedback is given by

$$C = \min_{s} \max_{p(x)} I(X; Y_s),$$

where the notation Y_s means that we condition on the channel state $S = s$.

converse We need to show that if rate R is achievable, then it satisfies $R \leq C$. Achievabilty of rate R for a compound channel implies that exist a sequence of codes $(2^{nR}, n)$ such that for any state s the probability of error goes to zero as $n \to \infty$, hence by using Fano inequality we get

$$H(W|Y_s^n) \leq 1 + P_e^{(n)} nR = n\epsilon_n, \qquad \forall s \in \mathcal{S}.$$

The message W is distributed uniformly $[1, 2^{nR}]$ hence,

$$\begin{aligned} nR &= H(W) \\ &= I(W; Y_s^n) + H(W|Y_s^n) \\ &\leq I(W; Y_s^n) + n\epsilon_n, \qquad \forall s \in \mathcal{S}. \end{aligned}$$

The equation holds for all $s \in \mathcal{S}$ and in particular for

$$\begin{aligned}
nR &\leq \min_s I(W; Y_s^n) + n\epsilon_n \\
&= \min_s \sum_{i=1}^{n} H(Y_{si}|Y_s^{i-1}) - I(Y_{si}|W, Y_s^{i-1}) + n\epsilon_n \\
&= \min_s \sum_{i=1}^{n} H(Y_{si}|Y_s^{i-1}) - I(Y_{si}|W, Y_s^{i-1}, X^i) + n\epsilon_n \\
&= \min_s \sum_{i=1}^{n} H(Y_{si}|Y_s^{i-1}) - I(Y_{si}|X_i) + n\epsilon_n \\
&\leq \min_s \sum_{i=1}^{n} H(Y_{si}) - I(Y_{si}|X_i) + n\epsilon_n \\
&\leq \min_s \sum_{i=1}^{n} I(Y_{si}; X_i) + n\epsilon_n \\
&\leq \min_s \sum_{i=1}^{n} I(Y_{si}; X_i) + n\epsilon_n \\
&= nC + n\epsilon_n
\end{aligned}$$

Achievability The achievabilty proof for discrete memoryless compound channel with feedback is based on the achievability of a DMC. The decoder first send to the encoder the state of the channel and then the encoder uses the code that is design for this state.

Under this coding scheme a rate R that satisfies $R < \max_{p(x|s)} \min_s I(X; Y_s)$ is achievable for any DMC realization $s \in \mathcal{S}$ and therefore is achievable for the compound channel. Now note that,

$$\max_{p(x|s)} \min_s I(X; Y_s) = \min_s \max_{p(x|s)} I(X; Y_s) = \min_s \max_{p(x)} I(X; Y_s)$$

(b) The capacity of an binary eraser channel with eraser probability p is $1 - p$, hence the capacity is C=0.75.

(c) $C_1 = C_3 = C_4 > C_2$

(d) yes is the same since

$$\min_s \max_{p(x)} I(X;Y_s) = \max_{p(x)} \min_s I(X;Y_s)$$

holds.

(e) yes. If the capacity of compound channel without feedback is zero then

$$\max_{p(x)} \min_s I(X;Y_s) = 0,$$

and this implies that exist a state s^* that for any input distribution $I(X;Y_{s^*}) = 0$. Hence

$$C_{feedback} = \min_s \max_{p(x)} I(X;Y_s) \leq \max_{p(x)} I(X;Y_{s^*}) = 0$$