Homework Set #3
Methods of types

1. **Large deviations.**
Let $X_1, X_2, \ldots$ be i.i.d. random variables drawn according to the Bernoulli distribution
$$\Pr\{X_i = 1\} = \Pr\{X_i = -1\} = \frac{1}{2}.$$
Let S_n be the random walk defined by
$$S_n = \sum_{i=1}^{n} X_i.$$
Find the function $f(\alpha)$ such that, for all $\alpha > 0$,
$$\lim_{n\to\infty} -\frac{1}{n} \log \Pr\{S_n \geq n\alpha\} = f(\alpha).$$

Solution: Large deviations.
We wish to look at
$$Pr\left\{\sum_{i=1}^{n} X_i \geq n\alpha\right\}$$
where X_i is generated by a bernoulli distribution P. We define a set
$$E = \left\{P : \sum_{a\in\{1,-1\}} P(a)a \geq \alpha\right\}$$
Then,
$$\begin{aligned} Pr\left\{\sum_{i=1}^{n} X_i \geq n\alpha\right\} &= Pr(P_{X^n} \in E) \\ &\doteq 2^{-n \min_{P' \in E} D(P^{'} \| P)} \end{aligned}$$

Since $D(P^{'}\|P) = \sum_a P^{'}(a)\log(\frac{P^{'}(a)}{1/2}) = 1 - H(P^{'})$, the minimum of $D(P^{'}\|P)$ is achieved when $H(P^{'})$ is maximized. From the constraint, we have $\sum_a P^{'}(a)a = P^{'}(0) - (1 - P^{'}(0)) \geq \alpha$, and thus $P^{'}(0) \geq \frac{\alpha+1}{2}$. From $\alpha > 0$, we know that $\frac{\alpha+1}{2} > \frac{1}{2}$, and hence, $H(P^{'})$ is maximized when $P^{'}(0) = \frac{1+\alpha}{2}$. Therefore, we have

$$
\begin{aligned}
Pr\left\{\sum_{i=1}^{n} X_i \geq n\alpha\right\} &\doteq 2^{-n \min_{P^{'}\in E} D(P^{'}\|P)} \\
&= 2^{-n(1-H(\frac{1+\alpha}{2}))}
\end{aligned}
$$

Then,

$$
f(\alpha) = -\frac{1}{n}\log 2^{-n(1-H(\frac{1+\alpha}{2}))} = 1 - H(\frac{1+\alpha}{2})
$$

2. **Counting.**
Let $\mathcal{X} = \{1, 2, \ldots, m\}$. Show that the number of sequences $x^n \in \mathcal{X}^n$ satisfying $\frac{1}{n}\sum_{i=1}^{n} g(x_i) \geq \alpha$ is approximately equal to 2^{nH^*}, to first order in the exponent, where

$$
H^* = \max_{P:\sum P(i)g(i)\geq\alpha} H(P).
$$

Solution: Counting.
We wish to count the number of sequences satisfying a certain property. Instead of directly counting the sequences, we will calculate the probability of the set under a uniform distribution. Since the uniform distribution puts a probability of $\frac{1}{m^n}$ on every sequence of length n, we can count the sequences by multiplying the probability of the set by m^n.

The probability of the set can be calculated easily from Sanov's theorem. Let Q be the uniform distribution, and let E be the set of sequences of length n satisfying $\frac{1}{n}\sum g(x_i) \geq \alpha$. Then by Sanov's theorem, we have

$$
Q^n(E) \doteq 2^{-nD(P^*||Q)},
$$

where P^* is the type in E that is closest to Q. Since Q is the uniform distribution, $D(P||Q) = \log m - H(P)$, and therefore P^* is the type in E that has maximum entropy. Therefore, if we let

$$H^* = \max_{P:\sum_{i=1}^m P(i)g(i)\geq\alpha} H(P),$$

we have

$$Q^n(E) \doteq 2^{-n(\log m - H^*)}.$$

Multiplying this by m^n to find the number of sequences in this set, we obtain

$$|E| \doteq 2^{-n\log m}2^{nH^*}m^n = 2^{nH^*}.$$

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

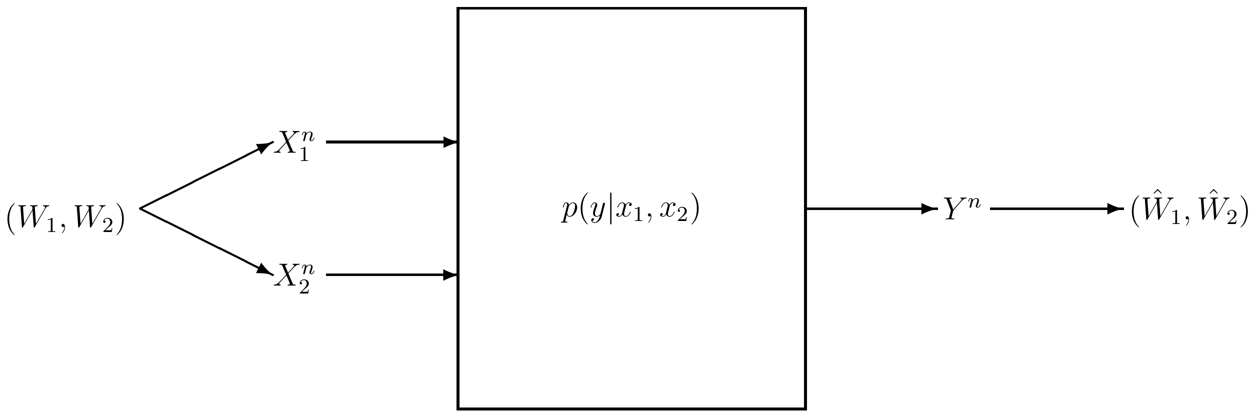


Figure 1: Multiple access channel with cooperating senders.

(a) Suppose X_1 and X_2 have access to *both* indices $W_1 \in \{1, 2^{nR_1}\}, W_2 \in \{1, 2^{nR_2}\}$. Thus the codewords $X_1^n(W_1, W_2), X_2^n(W_1, W_2)$ depend on both indices. Find the capacity region.

(b) Evaluate this region for the binary erasure multiple access channel $Y = X_1 + X_2$, $X_i \in \{0, 1\}$. Compare to the non-cooperative region.

Solution: The cooperative capacity of a multiple access channel

(a) When both senders have access to the pair of messages to be transmitted, they can act in concert. The channel is then equivalent to a single user channel with the input $X = (X_1, X_2) \in \mathcal{X}_1 \times \mathcal{X}_2$, and the message $W = (W_1, W_2)$. The capacity of this single user channel is $C = \max_{p(x)} I(X;Y) = \max_{p(x_1,x_2)} I(X_1, X_2; Y)$. The two senders can send at any combination of rates with the total rate

$$R_1 + R_2 \leq C.$$

(b) When the two senders cooperate to send a common message, the capacity is

$$C = \max_{p(x_1,x_2)} I(X_1, X_2; Y) = \max H(Y) = \log 3,$$

achieved by (for example) a uniform distribution on the pairs, (0,0), (0,1) and (1,1). The cooperative and non-cooperative regions are illustrated in Figure 2.

4. **Capacity of multiple access channels.**
Find the capacity region for each of the following multiple access channels:

(a) Additive modulo 2 multiple access access channel. $X_1 \in \{0,1\}, X_2 \in \{0,1\}, Y = X_1 \oplus X_2$.

(b) Multiplicative multiple access channel. $X_1 \in \{-1,1\}, X_2 \in \{-1,1\}, Y = X_1 \cdot X_2$.

Solution: Capacity of multiple access channels.

(a) Additive modulo 2 multiple access channel.

Quite clearly we cannot send at a total rate of more than 1 bit, since $H(Y) \leq 1$. We can achieve a rate of 1 bit from sender 1 by setting $X_2 = 0$, and similarly we can send 1 bit/transmission from sender 2. By simple time sharing we can achieve the entire capacity region which is shown in Figure 3.

(b) Multiplicative multiple access channel.
$X_1, X_2 \in \{-1, 1\}, Y = X_1 \cdot X_2$.
This channel is equivalent to the previous channel with the mapping $-1 \to 1$ and $1 \to 0$. Hence the capacity region is the same as the previous channel.

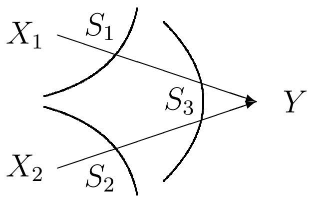
5. **Cut-set interpretation of capacity region of multiple access channel.**
For the multiple access channel we know that (R_1, R_2) is achievable if

$$R_1 < I(X_1; Y \mid X_2), \tag{1}$$
$$R_2 < I(X_2; Y \mid X_1), \tag{2}$$
$$R_1 + R_2 < I(X_1, X_2; Y), \tag{3}$$

for X_1, X_2 independent. Show, for X_1, X_2 independent, that

$$I(X_1; Y \mid X_2) = I(X_1; Y, X_2).$$

Thus R_1 is less than the mutual information between X_1 and everything else.



Interpret the information bounds as bounds on the rate of flow across cutsets S_1, S_2 and S_3.

Solution: Cut-set interpretation of capacity region of multiple access channel.
By the chain rule for mutual information and the independence of X_1 and X_2,

$$I(X_1; Y, X_2) = I(X_1; X_2) + I(X_1; Y|X_2) = I(X_1; Y|X_2).$$

We can interpret $I(X_1; Y, X_2)$ as the maximum amount of information that could flow across the cutset S_1. This is an upper bound on the rate R_1. Similarly, we can interpret the other bounds.

6. **Multiple-access channel.**

Let the output Y of a multiple-access channel be given by

$$Y = X_1 + \text{sgn}(X_2)$$

where X_1, X_2 are both real and power limited,

$$\begin{aligned} E[X_1^2] &\le P_1, \\ E[X_2^2] &\le P_2, \end{aligned}$$

and $\text{sgn}(x) = \begin{cases} 1, & x > 0 \\ -1, & x \le 0 \end{cases}$.

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

(a) Find the capacity region.

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

Solution: Multiple-access channel.

(a) It is easy to see that we can achieve any rate pair (R_1, R_2) satisfying

$$R_2 \le 1.$$

(In other words, R_1 can be arbitrarily large.) Although it can be calculated from the generic capacity formula

$$\begin{aligned} R_1 &\le I(X_1; Y|X_2) = \infty \\ R_2 &\le I(X_2; Y|X_1) \le 1 \\ R_1 + R_2 &\le I(X_1, X_2; Y) = \infty, \end{aligned}$$

we can find the capacity region from first principles.

Clearly, we can't get better than $R_1 = \infty$. For the bound on R_2, for each fixed X_1, the second user cannot transmit more than 1 bit.

(b) Achieving $R_1 = \infty, R_2 = 1$ is also trivial. The first user sends the fractional part of an integer (satisyfing the power constraint) and the second user sends 1 bit per transmission by sending $\pm\sqrt{P_2}$.

7. **Csiszar Sum Equality**. Let X^n and Y^n be two random vectors with arbitrary joint probability distribution. Show that:

$$\sum_{i=1}^{n} I(X_{i+1}^n; Y_i | Y^{i-1}) = \sum_{i=1}^{n} I(Y^{i-1}; X_i | X_{i+1}^n) \tag{4}$$

As we shall see this inequality is useful in proving converses to several multiple user channels. (Hint: You can prove this by induction or by expanding the terms on both sides using the chain rule.)

Solution: Csiszar Sum Equality.

$$\begin{aligned}
\sum_{i=1}^{n} I(X_{i+1}^n; Y_i | Y^{i-1}) &= \sum_{i=1}^{n} \sum_{j=i+1}^{n} I(X_j; Y_i | Y^{i-1}, X_{j+1}^n) \\
&= \sum_{j=2}^{n} \sum_{i=1}^{j-1} I(X_j; Y_i | Y^{i-1}, X_{j+1}^n) \\
&= \sum_{j=2}^{n} I(X_j; Y^{j-1} | X_{j+1}^n) \\
&= \sum_{j=1}^{n} I(X_j; Y^{j-1} | X_{j+1}^n) \\
&= \sum_{i=1}^{n} I(Y^{i-1}; X_i | X_{i+1}^n)
\end{aligned}$$

where the first and third equalities follow from chain rule, and the second equality follows from switching of the summations.

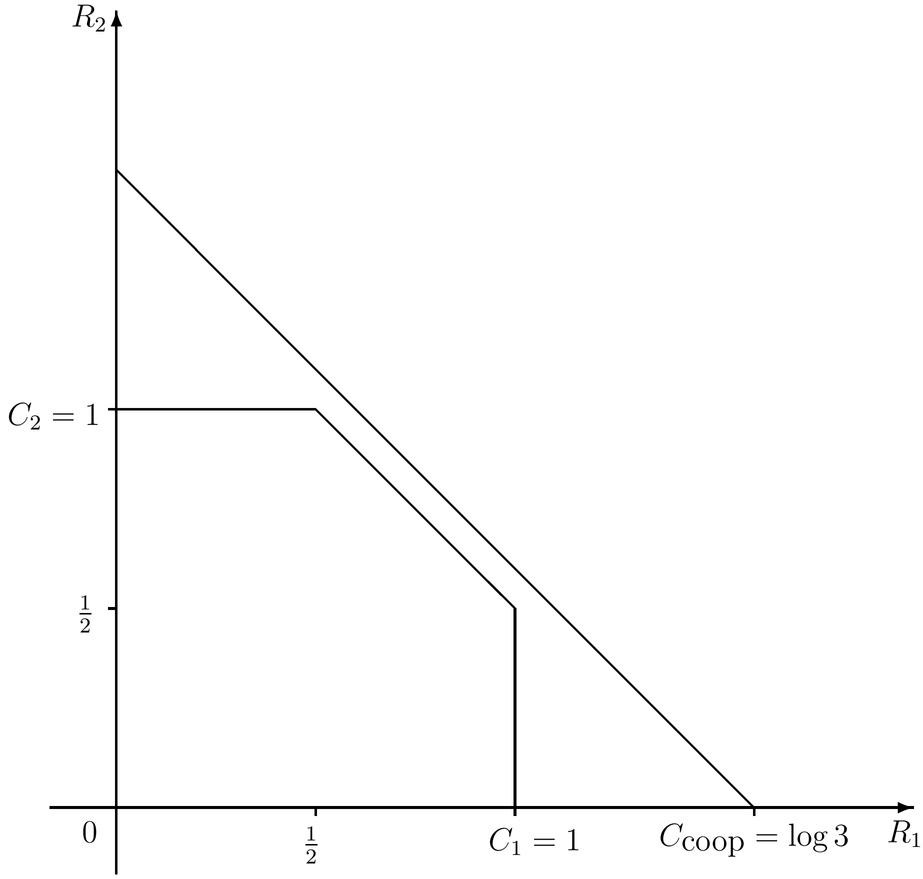


Figure 2: Cooperative and non-cooperative capacity for a binary erasure multiple access channel.

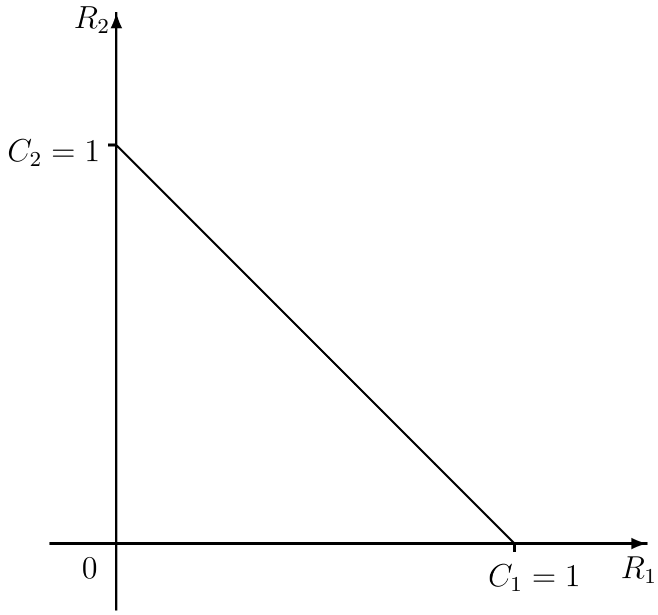


Figure 3: Capacity region of additive modulo 2 MAC.