Mathematical methods in communication — November 9th, 2009

Lecture No. 4

Lecturer: Haim Permuter — *Scribe: Alon Kipnis*

I. Multiple Access Channel[1]

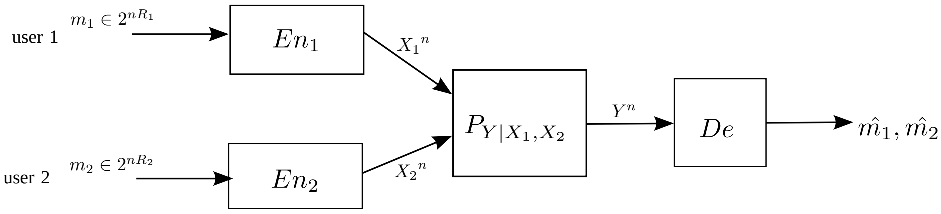


Fig. 1. A scheme of a multiple access channel

In the previous lecture we have defined:

Definition 1 A pair rate (R_1, R_2) is called *achievable* if there exists a sequence of $(2^{nR_1}, 2^{nR_2}, n)$ codes such that $P_e^{(n)} \to 0$.

Definition 2 *capacity region* $\mathcal{R}$ is the closure of all achievable rates.

Theorem 1 The capacity region $\mathcal{R}$ of a memoryless MAC is the convex closure of all (R_1, R_2) satisfying,

$$R_1 \leq I(X_1; Y|X_2), \tag{1}$$

$$R_2 \leq I(X_2; Y|X_1), \tag{2}$$

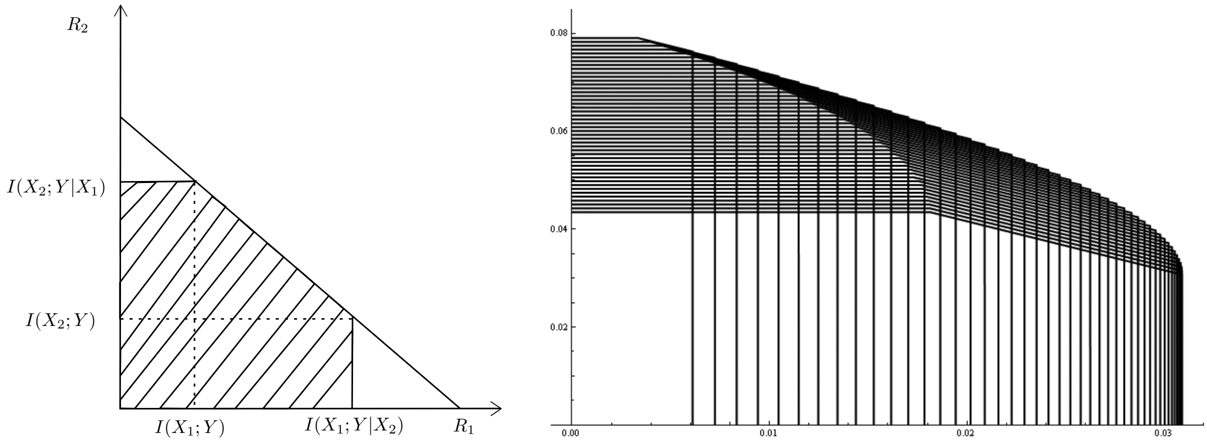
$$R_1 + R_2 \leq I(X_1, X_2; Y). \tag{3}$$

for some product distribution $p(x_1)p(x_2)$ on $\mathcal{X}_1 \times \mathcal{X}_2$.

Equivalently, $\mathcal{R}$ is the closure of the set:

$$\bigcup_{p(q)p(x_1|q)p(x_2|q)} \left\{ \begin{array}{r} R_1 \leq I(X_1; Y|X_2, Q), \\ R_2 \leq I(X_2; Y|X_1, Q), \\ R_1 + R_2 \leq I(X_1, X_2; Y, Q). \end{array} \right. \tag{4}$$

[1]The multiple-access channel capacity region was found by Ahlswede [2] and Liao [3] and was extended to the case of the multiple-access channel with common information by Slepian and Wolf [4]. Gaarder and Wolf [5] were the first to show that feedback increases the capacity of a discrete memoryless multiple-access channel.



(a) (b)

Fig. 2. (a) The region defined by eq (1)-(3) for some fixed $p_1(x_1)p_2(x_2)$. (b) The region defined by (1)-(3) for various $p_1(x_1)p_2(x_2)$ and the binary channel $Y \sim (p, 1-p)$ where $p = f(X_1, X_2)$ defined by: $f(0,0) = \frac{1}{4}$, $f(0,1) = \frac{1}{3}$, $f(1,0) = \frac{1}{4}$, $f(1,1) = \frac{1}{3}$.

Note that since X_1 and X_2 are independent,

$$I(X_1;Y|X_2) = I(X_1;Y,X_2) \geq I(X_1;Y). \tag{5}$$

Example 1 (Binary Additive Noise MAC) Let the inputs be $\mathcal{X}_1 = \mathcal{X}_2 = \{0,1\}$, and $Z \sim Bernuli(p)$ be an additive noise. The output is given by $Y = X_1 \oplus X_2 \oplus Z$. What is the capacity region of this MAC?
Solution: Consider,

$$\tag{6}$$

$$R_1 \leq I(X_1;Y|X_2,Q) \tag{7}$$

$$= H(Y|X_2,Q) - H(Y|X_1,X_2,Q) \tag{8}$$

$$\leq 1 - H(Z), \tag{9}$$

Similarly,

$$R_2 \leq 1 - H(Z),$$

$$R_1 + R_2 \leq I(X_1,X_2;Y,Q),$$

$$\leq 1 - H(Z).$$

Note that if $X_1 \sim Bernuli(\frac{1}{2})$ we have equality in (6). This is because, $X_1 \sim Bernuli(\frac{1}{2})$ implies $X_1 \oplus Z \sim Bernuli(\frac{1}{2})$. The same if $X_2 \sim Bernuli(\frac{1}{2})$. Hence the capacity region of this MAC is given by Fig. 4.

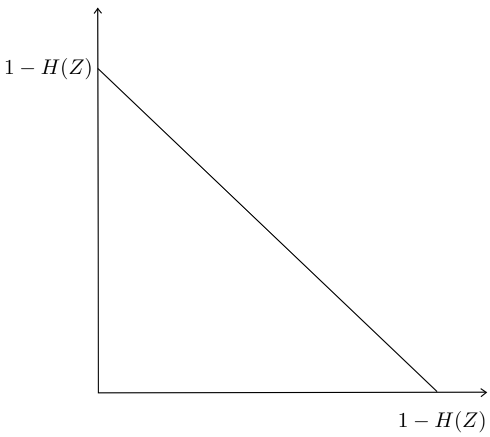


Fig. 3. The capacity region of Example (1).

Gaussian MAC

Two senders, X_1 and X_2, communicate to the single receiver, Y. The received signal at time i is

$$Y_i = X_{1,i} + X_{2,i} + Z_i.$$

Where $\{Z_i\} \sim Norm(0, \sigma^2)$ and i.i.d. each. We also assume the power constraint P_j on sender j; that is, for each sender, for all messages, we must have

$$\frac{1}{n}\sum_{i=1}^{n} {x_{ij}}^2(w_j) \leq P_j,$$

$$w_j \in \{1, 2, ..., 2^{nR_j}\}, j = 1, 2.$$

We can extend the proof for the discrete multiple-access channel to the Gaussian multiple-access channel. The converse can also be extended similarly, so the capacity region of the Gaussian multiple-access channel is the convex closure of all (R_1, R_2) satisfying,

$$R_1 \leq I(X_1; Y|X_2), \tag{10}$$

$$R_2 \leq I(X_2; Y|X_1), \tag{11}$$

$$R_1 + R_2 \leq I(X_1, X_2; Y), \tag{12}$$

for some input distribution $f(x_1)f(x_2)$ satisfying ${EX_1}^2 \leq P_1$ and ${EX_2}^2 \leq P_2$.

Now, we can expand the mutual information in terms of differential entropy, and thus

$$R_1 \leq I(X_1; Y|X_2, Q)$$

$$
\begin{aligned}
&= h(Y|X_2,Q) - h(Y|X_1,X_2,Q) \\
&\overset{(a)}{=} h(X_1+Z|X_2,Q) - h(Z) \\
&\leq h(X_1+Z|X_2) - h(Z) \\
&\leq h(X_1+Z) - h(Z) \\
&\overset{(b)}{\leq} \frac{1}{2}\log 2\pi e(P_1+\sigma^2) - \frac{1}{2}\log 2\pi e\sigma^2 \\
&= \frac{1}{2}\log(1+SNR_1).
\end{aligned}
$$

where

(a) follows from the fact that $h(Y|X_1,X_2,Q) = h(Z)$.

(b) follows from the fact that the maximum differential entropy for X_1+Z is $\frac{1}{2}\log 2\pi e\left(P_1+\sigma^2\right)$.

and we denoted $SNR_1 = \frac{P_1}{\sigma^2}$.

Similarly,

$$R_2 \leq \frac{1}{2}\log(1+SNR_2),$$

and

$$
\begin{aligned}
R_1+R_2 &\leq I(Y;X_1,X_2|Q) \qquad (13)\\
&= h(Y|Q) - h(Y|X_1,X_2,Q) \\
&= h(Y|Q) - h(Z) \\
&\leq \frac{1}{2}\log 2\pi e(P_1+P_2+\sigma^2) - \frac{1}{2}\log 2\pi e\sigma^2 \\
&= \frac{1}{2}\log(1+SNR_1+SNR_2).
\end{aligned}
$$

Exercise 1 Show that if $X_1 \sim Norm(0,\sigma_1)$ and $X_2 \sim Norm(0,\sigma_1)$ then we have equality in (13).

Now we shall prove the converse of theorem 1:

Given a sequence of $(2^{nR_1}, 2^{nR_2}, n)$ codes s.t. $P_e^{(n)} \to 0$, we will show that there exist a joint distribution $p(q)p(x_1|q)p(x_2|q)$ s.t.

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

Proof: Given a sequence of codes $\left(2^{nR_1}, 2^{nR_2}, n\right)$ and a probability of error such that $P_e^{(n)} \longrightarrow 0$ as $n \longrightarrow \infty$. Fix a code with rate (R_1,R_2) and a probability of error $P_e^{(n)}$. Fix n. Consider the given code of block length n. The joint distribution on $\mathcal{M}_1 \times \mathcal{M}_2 \times {\mathcal{X}_1}^n \times {\mathcal{X}_2}^n \times \mathcal{Y}^n$ is

$$p\left(m_1, m_2, {x_1}^n, {x_2}^n, y^n\right) = \frac{1}{2^{nR_1}} \frac{1}{2^{nR_2}} p\left({x_1}^n|m_1\right) p\left({x_2}^n|m_2\right) \prod_{i=1}^{n} p\left(y_i|x_{1,i}, x_{2,i}\right), \tag{14}$$

where $p\left({x_1}^n|m_1\right)$ is either 1 or 0, depending on whether ${x_1}^n = {x_1}^n\left(m_1\right)$, the codeword corresponding to m_1, or not, and similarly for $p\left({x_2}^n|m_2\right)$. The follow are calculated with respect to this distribution.

$$\begin{aligned}
nR_1 &= H(M_1) = H(M_1|{X_2}^n) \\
&= H(M_1|{X_2}^n) - H(M_1|{X_2}^n, Y^n) + H(M_1|{X_2}^n, Y^n) \\
&= I(Y^n; M_1|{X_2}^n) + H(M_1|{X_2}^n, Y^n) \\
&\overset{(a)}{\leq} H(Y^n|{X_2}^n) - H(Y^n|{X_2}^n, M_1) + n\epsilon_n \\
&= H(Y^n|{X_2}^n) - H(Y^n|{X_2}^n, {X_1}^n, M_1) + n\epsilon_n \\
&\overset{(b)}{=} H(Y^n|{X_2}^n) - H(Y^n|{X_2}^n, {X_1}^n) + n\epsilon_n \\
&\overset{(c)}{=} H(Y^n|{X_2}^n) - \sum_{i=1}^{n} H(Y_i|X_{2,i}, X_{1,i}) + n\epsilon_n \\
&= \sum_{i=1}^{n} H(Y_i|Y^{i-1}, {X_2}^n) - \sum_{i=1}^{n} H(Y_i|X_{2,i}, X_{1,i}) + n\epsilon_n \\
&\leq \sum_{i=1}^{n} H(Y_i|X_{2,i}) - \sum_{i=1}^{n} H(Y_i|X_{2,i}, X_{1,i}) + n\epsilon_n \\
&= \sum_{i=1}^{n} I(Y_i; X_{1,i}|X_{2,i}) + n\epsilon_n,
\end{aligned} \tag{15}$$

where

(a) follows from Fano's inequality and we denoted $\epsilon_n = \frac{1}{n} + R_1 P_e^{(n)}$.

(b) follows from the Markov chain $M_1 \rightarrow (X_{1,i}, X_{2,i}) \rightarrow Y_i$.

(c) follows from the memoryless and no feedback property of the channel.

Similar calculation leads us to

$$nR_2 = \sum_{i=1}^{n} I(Y_i; X_{2,i}|X_{1,i}) + n\epsilon_n, \tag{16}$$

and

$$nR_1 + nR_2 = \sum_{i=1}^{n} I(X_{1,i}, X_{2,i}; Y_i) + n\epsilon_n, \tag{17}$$

Let us define Q to be uniform over $(1, 2, ..., n)$. Let $X_{1,q}$ be the q^{th} element of $(X_{1,1}, ..., X_{1,n})$, then

$X_{1,Q}$ is uniform over $(X_{1,1},...,X_{1,n})$. RHS of (15) becomes,

$$nR_1 \leq n\sum_{i=1}^{n}\frac{1}{n}I(Y_Q;X_{1,Q}|X_{2,Q},Q=i)+n\epsilon_n \tag{18}$$

$$= nI(Y_Q;X_{1,Q}|X_{2,Q})+n\epsilon_n, \tag{19}$$

and similarly,

$$nR_2 \leq nI(Y_Q;X_{2,Q}|X_{1,Q})+n\epsilon_n, \tag{20}$$

$$nR_1+nR_2 \leq nI(X_{1,Q},X_{2,Q};Y_Q)+n\epsilon_n, \tag{21}$$

Therefore, by taking $X_1 = X_{1,Q}$, $X_2 = X_{2,Q}$ and $Y = Y_Q$ we get a new random variables whose distributions depends on Q in the same way as the distributions of $X_{1,i}$, $X_{2,i}$ depend on i. Moreover, $X_{1,i}\,(M_1)$ and $X_{1,i}\,(M_1)$ are independent since M_1 and M_2 are independent, so given Q, $X_{1,Q}$ and $X_{2,Q}$ are independent as well. Hence, by taking the limit $\epsilon_n = \frac{1}{n}+R_1P_e^{(n)} \longrightarrow 0$ as $n \longrightarrow \infty$ we get

$$R_1 \leq I(Y_Q;X_{1,Q}|X_{2,Q}), \tag{22}$$

$$nR_2 \leq nI(Y_Q;X_{2,Q}|X_{1,Q}), \tag{23}$$

$$nR_1+nR_2 \leq nI(X_{1,Q},X_{2,Q};Y_Q). \tag{24}$$

for some choice of joint distribution $p\,(q)\,p\,(x_1|q)\,p\,(x_2|q)\,p\,(y|x_1,x_2)$. ■

II. Method of Types (Large Deviation)

Assume that n Bernuli experiments are being done with probability $p=(\frac{1}{2},\frac{1}{2})$. What is the probability that for large n the result will be distributed $q=(0.2,0.8)$?
We will see that the answer to that is approximately $2^{-nD(p||q)}$.
For a sequence X^n over $\mathcal{X}$ we define:

Definition 3 The *type* P_{x^n} is the relative proportion of occurrences of each symbol of $\mathcal{X}$ (i.e. $P_{x^n} = N(a|X^n)/n$ for all $a \in \mathcal{X}$, where $N(a|x^n)$ is the number of times the symbol a occurs in the sequence $x^n \in \mathcal{X}^n$).

We will also use the notation: $P_{x^n}(a)=\frac{N(a|x^n)}{n}$. Thus, if $x^n = 00110$ then $P_{x^n}(0)=\frac{3}{5}$ and $P_{x^n}=\left(\frac{3}{5},\frac{2}{5}\right)$.

Definition 4 Let $\mathcal{P}_n$ denote the *set of types with denominator n*.

For example, if $\mathcal{X}=\{0,1\}$, the set of possible types with denominator n is

$$\mathcal{P}=\left\{(P(0),P(1)):\left(\frac{0}{n},\frac{n}{n}\right),\left(\frac{1}{n},\frac{n-1}{n}\right),...,\left(\frac{n}{n},\frac{0}{n}\right)\right\}. \tag{25}$$

Lemma 1 An upper bound for $|\mathcal{P}_n|$:

$$|\mathcal{P}_n| \leq (n+1)^{|\mathcal{X}|}. \tag{26}$$

Proof: There are $|\mathcal{X}|$ components in the vector that specifies P_{x^n}. The numerator in each component can take on only $n+1$ values. So there are at most $(n+1)^{|\mathcal{X}|}$ choices for the type vector. ■

Definition 5 let $P \in \mathcal{P}_n$. The *type class* of P, denoted by $T(P)$, is the set of sequences of length n with type P. I.e,

$$T(P) = (x^n \in \mathcal{X}^n : P_{x^n} = P). \tag{27}$$

Lemma 2 Let $\{X_i\}_{i\geq 1}$ be an i.i.d sequence distributed according to a distribution $Q(x)$. Let x^n be a specific sequence of type P, then $Q^n(x^n) = 2^{-nH(P)+D(P||Q)}$.

Proof:

Since $\{X_i\}_{i\geq 1}$ are i.i.d,

$$Q^n(x^n) = \prod_{i=1}^{n} Q(x_i). \tag{28}$$

Now consider

$$\begin{aligned}
\log Q^n(x^n) &= \sum_{i=1}^{n} \log Q(x_i) && (29)\\
&\overset{(a)}{=} \sum_{a\in\mathcal{X}} N(a|x^n) \log Q(a) && (30)\\
&\overset{(b)}{=} n \sum_{a\in\mathcal{X}} P_{x^n}(a) \log Q(a) && (31)\\
&= n \sum_{a\in\mathcal{X}} P_{x^n}(a) \log \frac{Q(a)}{P_{x^n}(a)} \cdot P_{x^n}(a) && (32)\\
&= n(-H(P) - D(P||Q)), && (33)
\end{aligned}$$

where

(a) follows because each $a \in \mathcal{X}$ contributes exactly $\log Q(a)$ times it's number of occurences in x^n to the sum in (29).

(b) follows from the definition of $P_{x^n}(a)$.

Hence we obtained

$$Q^n(x^n) = 2^{-nH(P)+D(P||Q)}. \tag{34}$$

■

REFERENCES

[1] T. M. Cover and J.A. Thomas *Elements of Information Theory*. Jhon Wiley & Sons, Hoboken, Ney Jersy 2006

[2] R. Ahlswede. *Multi-way communication channels.* In Proc. 2nd Int. Symp. Inf. Theory (Tsahkadsor, Armenian S.S.R.), pages 2352. Hungarian Academy of Sciences, Budapest, 1971.

[3] S. Kullback. *Information Theory and Statistics*. Wiley, New York, 1959.

[4] D. Slepian and J. K. Wolf. *A coding theorem for multiple access channels with correlated sources*. Bell Syst. Tech. J., 52:10371076, 1973.

[5] T. Gaarder and J. K. Wolf. *The capacity region of a multiple-access discrete memoryless channel can increase with feedback*. IEEE Trans. Inf. Theory, IT-21:100102, 1975.

[6] T. M. Cover and C. S. K. Leung. *An achievable rate region for the multiple access channel with feedback*. IEEE Trans. Inf. Theory, IT-27:292298, 1981.

[7] F. M. J. Willems. *The feedback capacity of a class of discrete memoryless multiple access channels*. IEEE Trans. Inf. Theory, IT-28:9395, 1982.

[8] L. H. Ozarow. *The capacity of the white Gaussian multiple access channel with feedback.* IEEE Trans. Inf. Theory, IT-30:623629, 1984.