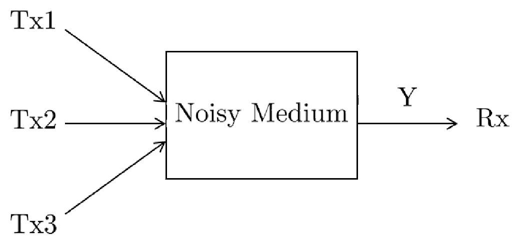
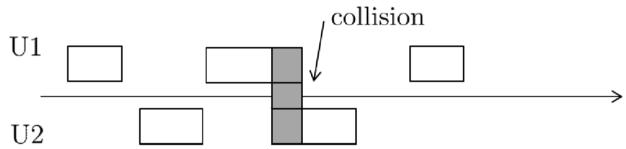
26.1 Problem motivation and main results



Note: In network community, people are mostly interested in channel access control mechanisms that help to detect or avoid data packet collisions so that the channel is shared among multiple users.



The famous ALOHA protocal achieves

$$\sum_i R_i \approx 0.37C$$

where C is the (single-user) capacity of the channel.[1]

In information theory community, the goal is to achieve

$$\sum_i R_i > C$$

The key to achieve this is to use coding so that collisions are resolvable.

In the following discussion we shall focus on the case with two users. This is without loss of much generality, as all the results can easily be extended to N users.

Definition 26.1.

- Multiple-access channel: $\{P_{Y^n|A^n,B^n} : \mathcal{A}^n \times \mathcal{B}^n \to \mathcal{Y}^n, n = 1, 2, \dots\}$.
- a (n, M_1, M_2, ϵ) code is specified by

$$\begin{aligned} &f_1 : [M_1] \to \mathcal{A}^n, \quad f_2 : [M_2] \to \mathcal{B}^n \\ &g : \mathcal{Y}^n \to [M_1] \times [M_2] \end{aligned}$$

[1] Note that there is a lot of research about how to achieve even these 37%. Indeed, ALOHA in a nutshell simply postulates that everytime a user has a packet to transmit, he should attempt transmission in each time slot with probability p, independently. The optimal setting of p is the inverse of the number of actively trying users. Thus, it is non-trivial how to learn the dynamically changing number of active users without requiring a central authority. This is how ideas such as exponential backoff etc arise.

$$\begin{tikzcd}
W_1 \arrow[r, "f_1"] & A^n \arrow[dr] & & \hat{W}_1 \\
& & Y^n \arrow[ur, "g"] \arrow[dr] & \\
W_2 \arrow[r, "f_2"] & B^n \arrow[ur] & & \hat{W}_2
\end{tikzcd}$$

$$P$$

$W_1, W_2 \sim$ uniform, and the codes achieves

$$\mathbb{P}[\{W_1 \neq \hat{W}_1\} \bigcup \{W_2 \neq \hat{W}_2\}] \leq \epsilon$$

- Fundamental limit of capacity region

$$\mathcal{R}^*(n, \epsilon) = \{(R_1, R_2) : \exists \text{ a } (n, 2^{nR_1}, 2^{nR_2}, \epsilon) \text{ code}\}$$

- Asymptotics:

$$\mathcal{C}_\epsilon = \left[\liminf_{n \to \infty} \mathcal{R}^*(n, \epsilon)\right]$$

where $[\cdot]$ denotes the closure of a set.

Note: $\liminf$ and $\limsup$ of a sequence of sets $\{A_n\}$:

$$\liminf_n A_n = \{\omega : \omega \in A_n, \forall n \geq n_0\}$$
$$\limsup_n A_n = \{\omega : \omega \text{ infinitely occur}\}$$

- $$\mathcal{C} = \lim \mathcal{C}_\epsilon = \bigcap_{\epsilon > 0} \mathcal{C}_\epsilon$$

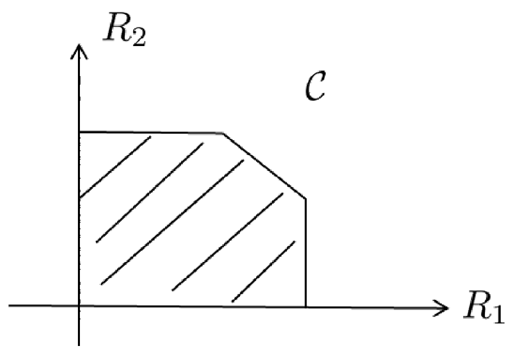
Theorem 26.1 (Capacity region)**.**

$$\mathcal{C}_\epsilon = \overline{co} \bigcup_{P_A, P_B} Penta(P_A, P_B) \tag{26.1}$$
$$= \Big[\bigcup_{P_{U,A,B} = P_U P_{A|U} P_{B|U}} Penta(P_{A|U}, P_{B|U} | P_U) \Big] \tag{26.2}$$

where $\overline{co}$ is the set operator of constructing the convex hull followed by taking the closure, and $Penta(\cdot, \cdot)$ is defined as follows:

$$Penta(P_A, P_B) = \left\{ (R_1, R_2) : \begin{array}{r} 0 \leq R_1 \leq I(A; Y | B) \\ 0 \leq R_2 \leq I(B; Y | A) \\ R_1 + R_2 \leq I(A, B; Y) \end{array} \right\}$$
$$Penta(P_{A|U}, P_{B|U} | P_U) = \left\{ (R_1, R_2) : \begin{array}{r} 0 \leq R_1 \leq I(A; Y | B, U) \\ 0 \leq R_2 \leq I(B; Y | A, U) \\ R_1 + R_2 \leq I(A, B; Y | U) \end{array} \right\}$$

Note: the two forms in (26.1) and (26.2) are equivalent without cost constraints. In the case when constraints such as $\mathbb{E}\mathsf{c}_1(A) \leq P_1$ and $\mathbb{E}\mathsf{c}_2(B) \leq P_2$ are present, only the second expression yields the true capacity region.



Penta

26.2 MAC achievability bound

First, we introduce a lemma which will be used in the proof of Theorem 26.1.

Lemma 26.1. *$\forall P_A, P_B, P_{Y|A,B}$ such that $P_{A,B,Y} = P_A P_B P_{Y|A,B}$, and $\forall \gamma_1, \gamma_2, \gamma_{12} > 0$, $\forall M_1, M_2$, there exists a (M_1, M_2, ϵ) MAC code such that:*

$$\begin{aligned}\epsilon \leq &\mathbb{P}\Big[\{i_{12}(A,B;Y) \leq \log\gamma_{12}\} \bigcup \{i_1(A;Y|B) \leq \log\gamma_1\} \bigcup \{i_2(B;Y|A) \leq \log\gamma_2\}\Big] \\ &+ (M_1 - 1)(M_2 - 1)e^{-\gamma_{12}} + (M_1 - 1)e^{-\gamma_1} + (M_2 - 1)e^{-\gamma_2}\end{aligned} \tag{26.3}$$

Proof. We again use the idea of random coding.

Generate the codebooks

$$c_1, \ldots, c_{M_1} \in \mathcal{A}, \quad d_1, \ldots, d_{M_2} \in \mathcal{B}$$

where the codes are drawn i.i.d from distributions: $c_1, \ldots, c_{M_1} \sim$ *i.i.d.* P_A, $d_1, \ldots, d_{M_2} \sim$ *i.i.d.* P_B.

The decoder operates in the following way: report (m, m') if it is the unique pair that satisfies:

$$\begin{aligned}(P_{12}) \quad & i_{12}(c_m, d_{m'}; y) > \log\gamma_{12} \\ (P_1) \quad & i_1(c_m; y|d_{m'}) > \log\gamma_1 \\ (P_2) \quad & i_2(d_{m'}; y|c_m) > \log\gamma_2\end{aligned}$$

Evaluate the expected error probability:

$$\begin{aligned}\mathbb{E}P_e(c_1^{M_1}, d_1^{M_2}) = \mathbb{P}\Big[&\{(W_1, W_2) \text{ violate } (P_{12}) \text{ or } (P_1) \text{ or } (P_2)\} \\ &\bigcup \{\exists \text{ impostor } (W_1', W_2') \text{ that satisfy } (P_{12}) \text{ and } (P_1) \text{ and } (P_2)\}\Big]\end{aligned}$$

by symmetry of random codes, we have

$$\begin{aligned}P_e = \mathbb{E}[P_e|W_1 = m, W_2 = m'] = \mathbb{P}\Big[&\{(m, m') \text{ violate } (P_{12}) \text{ or } (P_1) \text{ or } (P_2)\} \\ &\bigcup \{\exists \text{ impostor } (i \neq m, j \neq m') \text{ that satisfy } (P_{12}) \text{ and } (P_1) \text{ and } (P_2)\}\Big]\end{aligned}$$

$$\Rightarrow P_e \leq \mathbb{P}\Big[\{i_{12}(A,B;Y) \leq \log\gamma_{12}\} \bigcup \{i_1(A;Y|B) \leq \log\gamma_1\} \bigcup \{i_2(B;Y|A) \leq \log\gamma_2\}\Big] + \mathbb{P}[E_{12}] + \mathbb{P}[E_1] + \mathbb{P}[E_2]$$

where

$$
\begin{aligned}
\mathbb{P}[E_{12}] &= \mathbb{P}[\{\exists (i \neq m, j \neq m') \text{ s.t. } i_{12}(c_m, d_{m'}; y) > \log \gamma_{12}\}] \\
&\leq (M_1 - 1)(M_2 - 1)\mathbb{P}[i_{12}(\overline{A}, \overline{B}; Y) > \log \gamma_{12}] \\
&= \mathbb{E}[e^{-i_{12}(A,B;Y)} \mathbf{1}\{i_{12}(A, B; Y) > \log \gamma_{12}\}] \\
&\leq e^{-\gamma_{12}} \\
\mathbb{P}[E_2] &= \mathbb{P}[\{\exists (j \neq m') \text{ s.t. } i_2(d_j; y | c_i) > \log \gamma_2\}] \\
&\leq (M_2 - 1)\mathbb{P}[i_2(\overline{B}; Y | A) > \log \gamma_2] \\
&= \mathbb{E}_A[e^{-i_2(B;Y|A)} \mathbf{1}\{i_2(B; Y|A) > \log \gamma_2\} | A] \\
&\leq \mathbb{E}_A[e^{-\gamma_2} | A] = e^{-\gamma_2} \\
\text{similarly} \quad \mathbb{P}[E_1] &\leq e^{-\gamma_1}
\end{aligned}
$$

□

Note: [Intuition] Consider the decoding step when a random codebook is used. We observe Y and need to solve an M-ary hypothesis testing problem: Which of $\{P_{Y|A=c_m, B=d_{m'}}\}_{m,m' \in [M_1] \times [M_2]}$ produced the sample Y?

Recall that in P2P channel coding, we had a similar problem and the M-ary hypothesis testing problem was converted to M binary testing problems:

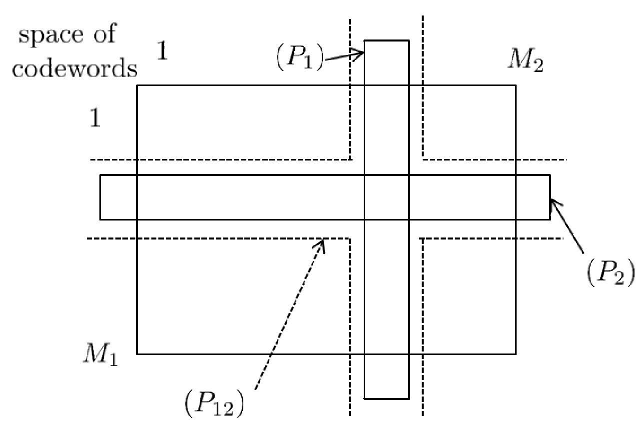
$$
P_{Y|X=c_j} \quad \text{vs} \quad P_{Y_{-j}} \triangleq \sum_{i \neq j} \frac{1}{M-1} P_{Y|X=c_i} \approx P_Y
$$

I.e. distinguish c_j (hypothesis H_0) against the average distribution induced by all other codewords (hypothesis H_1), which for a random coding ensemble $c_j \sim P_X$ is very well approximated by $P_Y = P_{Y|X} \circ P_X$. The optimal test for this problem is roughly

$$
\frac{P_{Y|X=c}}{P_Y} \gtrsim \log(M-1) \quad \Longrightarrow \quad \text{decide } P_{Y|X=c_j} \tag{26.4}
$$

since the prior for H_0 is $\frac{1}{M}$, while the prior for H_1 is $\frac{M-1}{M}$.

The proof above followed the same idea except that this time because of the two-dimensional grid structure:



there are in fact binary HT of three kinds

$$(P12) \quad \sim \quad \text{test } P_{Y|A=c_m,B=d_{m'}} \text{ vs } \frac{1}{(M_1-1)(M_2-1)}\sum_{i\neq m}\sum_{j\neq m'} P_{Y|A=c_i,B=d_j} \approx P_Y$$

$$(P1) \quad \sim \quad \text{test } P_{Y|A=c_m,B=d_{m'}} \text{ vs } \frac{1}{M_1-1}\sum_{i\neq m} P_{Y|A=c_i,B=d_{m'}} \approx P_{Y|B=d_{m'}}$$

$$(P2) \quad \sim \quad \text{test } P_{Y|A=c_m,B=d_{m'}} \text{ vs } \frac{1}{M_2-1}\sum_{j\neq m'} P_{Y|A=c_m,B=d_j} \approx P_{Y|A=c_m}$$

And analogously to (26.4) the optimal tests are given by comparing the respective information densities with $\log M_1M_2$, $\log M_1$ and $\log M_2$.

Another observation following from the proof is that the following decoder would also achieve exactly the same performance:

- Step 1: rule out all cells (i,j) with $i_{12}(c_i,d_j;Y) \lesssim \log M_1M_2$.
- Step 2: If the messages remaining are NOT all in one row or one column, then FAIL.
- Step 3a: If the messages remaining are all in one column m' then declare $\hat{W}_2 = m'$. Rule out all entries in that column with $i_1(c_i;Y|d_{m'}) \lesssim \log M_1$. If more than one entry remains, FAIL. Otherwise declare the unique remaining entry m as $\hat{W}_1 = m$.
- Step 3b: Similarly with column replaced by row, i_1 with i_2 and $\log M_1$ with $\log M_2$.

The importance of this observation is that in the regime when RHS of (26.3) is small, the decoder always finds it possible to basically decode one message, "subtract" its influence and then decode the other message. Which of the possibilities 3a/3b appears more often depends on the operating point (R_1,R_2) inside $\mathcal{C}$.

26.3 MAC capacity region proof

Proof. 1. Show $\mathcal{C}$ is convex.

Take $(R_1,R_2) \in \mathcal{C}_{\epsilon/2}$, and take $(R_1',R_2') \in \mathcal{C}_{\epsilon/2}$.

We merge the $(n, 2^{nR_1}, 2^{nR_2}, \epsilon/2)$ code and the $(n, 2^{nR_1}, 2^{nR_2}, \epsilon/2)$ code in the following time-sharing way: in the first n channels, use the first set of codes, and in the last n channels, use the second set of codes.

Thus we formed a new $(2n, 2^{R_1+R_1'}, 2^{R_2+R_2'}, \epsilon)$ code, we know that

$$\frac{1}{2}\mathcal{C}_{\epsilon/2} + \frac{1}{2}\mathcal{C}_{\epsilon/2} \subset \mathcal{C}_\epsilon$$

take limit at both sides

$$\frac{1}{2}\mathcal{C} + \frac{1}{2}\mathcal{C} \subset \mathcal{C}$$

also we know that $\mathcal{C} \subset \frac{1}{2}\mathcal{C} + \frac{1}{2}\mathcal{C}$, therefore $\mathcal{C} = \frac{1}{2}\mathcal{C} + \frac{1}{2}\mathcal{C}$ is convex.

Note: the set addition is defined in the following way:

$$\mathcal{A} + \mathcal{B} \triangleq \{(a+b) : a \in \mathcal{A}, b \in \mathcal{B}\}$$

2. Achievability

   STP: $\forall P_A, P_B, \forall (R_1, R_2) \in Penta(P_A, P_B)$, $\exists (n, 2^{nR_1}, 2^{nR_2}, \epsilon) code.$

   Apply Lemma 26.1 with:

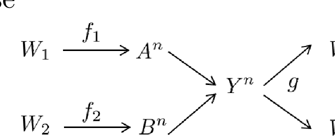
$$
\begin{aligned}
&P_A \to P_A^n, \quad P_B \to P_B^n, \quad P_{Y|A,B} \to P_{Y|A,B}^n \\
&M_1 = 2^{nR_1}, \quad M_2 = 2^{nR_2}, \\
&\log \gamma_{12} = n(I(A,B;Y) - \delta), \quad \log \gamma_1 = n(I(A;Y|B) - \delta), \quad \log \gamma_2 = n(I(B;Y|A) - \delta).
\end{aligned}
$$

   we have that there exists a (M_1, M_2, ϵ) code with

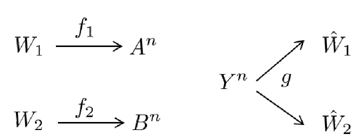
$$
\begin{aligned}
\epsilon \le & \underbrace{\mathbb{P}\Big[\{\frac{1}{n}\sum_{k=1}^{n} i_{12}(A_k, B_k; Y_k) \le \log \gamma_{12} - \delta\} \bigcup \{\frac{1}{n}\sum_{k=1}^{n} i_1(A_k; Y_k|B_k) \le \log \gamma_1 - \delta\} \bigcup \{\frac{1}{n}\sum_{k=1}^{n} i_2(B_k; Y_k|A_k) \le \log \gamma_2 - \delta\}\Big]}_{①} \\
& + \underbrace{(2^{nR_1} - 1)(2^{nR_2} - 1)e^{-\gamma_{12}} + (2^{nR_1} - 1)e^{-\gamma_1} + (2^{nR_2} - 1)e^{-\gamma_2}}_{②}
\end{aligned}
$$

   by WLLN, the first part goes to zero, and for any (R_1, R_2) such that $R_1 < I(A;Y|B) - \delta$ and $R_2 < I(B;Y|A) - \delta$ and $R_1 + R_2 < I(A,B;Y) - \delta$, the second part goes to zero as well. Therefore, if $(R_1, R_2) \in$ interior of the Penta, there exists a $(M_1, M_2, \epsilon = o(1))$ code.

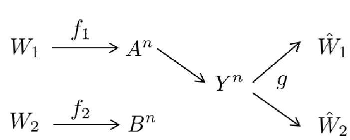
3. Weak converse



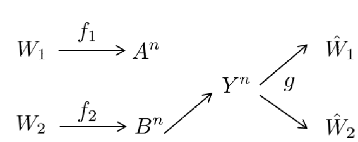
P



$Q \in (*)$



$Q_1 \in (*1)$



$Q_2 \in (*2)$

$$
\mathbb{Q}[W_1 = \hat{W}_1, W_2 = \hat{W}_2] = \frac{1}{M_1 M_2}, \quad \mathbb{P}[W_1 = \hat{W}_1, W_2 = \hat{W}_2] \ge 1 - \epsilon
$$

   d-proc:

$$
\begin{aligned}
&d(1 - \epsilon \| \frac{1}{M_1 M_2}) \le \inf_{\mathbb{Q} \in (*)} D(P\|Q) = I(A^n, B^n; Y^n) \\
\Rightarrow & R_1 + R_2 \le \frac{1}{n} I(A^n, B^n; Y^n) + o(1)
\end{aligned}
$$

To get separate bounds, we apply the same trick to evaluate the information flow from the link between $A \to Y$ and $B \to Y$ separately:

$$\mathbb{Q}_1[W_2 = \hat{W}_2] = \frac{1}{M_2}, \quad \mathbb{P}[W_2 = \hat{W}_2] \ge 1 - \epsilon$$

d-proc:

$$\begin{aligned} d(1-\epsilon \| \frac{1}{M_2}) &\le \inf_{\mathbb{Q}_1 \in (*1)} D(P \| Q_1) = I(B^n; Y^n | A^n) \\ \Rightarrow R_2 &\le \frac{1}{n} I(B^n; Y^n | A^n) + o(1) \end{aligned}$$

similarly we can show that

$$R_2 \le \frac{1}{n} I(A^n; Y^n | B^n) + o(1)$$

For memoryless channels, we know that $\frac{1}{n} I(A^n, B^n; Y^n) \le \frac{1}{n} \sum_k I(A_k, B_k; Y_k)$. Similarly, since given B^n the channel $A^n \to Y^n$ is still memoryless we have

$$I(A^n; Y^n | B^n) \le \sum_{k=1}^n I(A_k; Y_k | B^n) = \sum_{k=1}^n I(A_k; Y_k | B_k)$$

Notice that each (A_i, B_i) pair corresponds to (P_{A_k}, P_{B_k}), $\forall k$ define

$$Penta_k(P_{A_k}, P_{B_k}) = \left\{ (R_{1,k}, R_{2,k}) : \begin{array}{r} 0 \le R_{1,k} \le I(A_k; Y_k | B_k) \\ 0 \le R_{2,k} \le I(B_k; Y_k | A_k) \\ R_{1,k} + R_{2,k} \le I(A_k, B_k; Y_k) \end{array} \right\}$$

therefore

$$\begin{aligned} (R_1, R_2) &\in \left[\frac{1}{n} \sum_k Penta_k \right] \\ \Rightarrow C &\in \overline{co} \bigcup_{P_A, P_B} Penta \end{aligned}$$

□