Electrical Engineering 229A Lecture 22 Notes

Daniel Raban

November 9, 2021

1 Network Information Theory

In this lecture, we will be discussing **multiuser/network information theory**. There is a recent book on it by Abbas El Gamal and Young-Han Kim.

1.1 Shannon capacity region of a multiuser DMC

In the **multiple access channel** model, there are multiple transmitters and a single receiver. For example, we could think of a cell tower receiving multiple signals. The channel is modeled as in the DMC case and also as in the power constrained Gaussian channel model. We will study the Shannon capacity region in Shannon's block coding formulation. Good news: This is known, unlike many problems in network information theory.

Consider the 2 transmitter case:

- $\mathscr{X}_1$ is the input alphabet of transmitter 1.
- $\mathscr{X}_2$ is the input alphabet of transmitter 2.
- $\mathscr{Y}$ is the output alphabet.
- The channel model in a simple use is $(p(y \mid x_1, x_2) \geq 0, \sum_y p(y \mid x_1, x_2) = 1 \, \forall x_1, x_2)$.
- The encoding map of transmitter 1 and 2 are

$$e_n^{(1)} : [M_n^{(1)}] \mapsto \mathcal{X}_1^n, \qquad e_n^{(2)} : [M_n^{(2)}] \mapsto \mathcal{X}_2^n,$$

  and the decoding map is

$$d_n : \mathcal{Y}^n \mapsto [M_n^{(1)}] \times [M_n^{(2)}].$$

Like a Pavlovian dog, let's turn the Shannon crank.

Definition 1.1. If there exists $((e_n^{(1)}, e_n^{(2)}, d_n), n \geq 1)$ such that

$$\liminf_n \frac{1}{n} \log M_n^{(1)} \geq R_1,$$

$$\liminf_n \frac{1}{n} \log M_n^{(2)} \geq R_2,$$

$$\mathbb{P}(d_n(e_n^{(1)}(W_{1,n}), e_n^{(2)}(W_{2,n})) \neq (W_{1,n}, W_{2,n}) \to 0,$$

where $W_{i,n} \sim \text{Unif}[M_n^{(i)}]$, and $W_{1,n} \amalg W_{2,n}$, we say that the rate pair (R_1, R_2) is **achievable**.

Theorem 1.1. *The closure*[1] *of the set of achievable rate pairs is the closed convex hull of the union of all the sets of rate pairs of the type*

$$\begin{aligned} \mathcal{R}_{p(x_1)p(x_2)} = \{(R_1, R_2) : & R_1 < I(X_1; Y \mid X_2), \\ & R_2 < I(X_2; Y \mid X_1), \\ & R_1 + R_2 < I(X_1, X_2; Y)\} \end{aligned}$$

for some $p(x_1)p(x_2)$*, where* $p(y \mid x_1, x_2)$ *is given by the channel.*

In general, each of these regions looks like a polyhedron.

Remark 1.1. A more elegant way to write this region is as

$$\begin{aligned} \{(R_1, R_2) : & R_1 < I(X_1; Y \mid X_2, Q), \\ & R_2 < I(X_2; Y \mid X_1, Q), \\ & R_1 + R_2 < I(X_1, X_2; Y \mid Q)\}, \end{aligned}$$

where the joint distribution is

$$p(q)p(x_1 \mid q)p(x_2 \mid q)p(y \mid x_1, x_2),$$

and $Q \in \mathscr{Q}$, a finite set of size ≤ 4.

Proof. Achievability is via a random coding argument. Given $p(x_1)p(x_2)$ and $(R_1, R_2) in \mathcal{R}_{p(x_1)p(x_2)}$ and clock length n, transmitter 1 constructs the random codebook

$$\begin{bmatrix} X_{1,1}(1) & \cdots & X_{1,n}(1) \\ \vdots & & \vdots \\ X_{1,1}(m_1) & \cdots & X_{1,n}(m_1) \\ \vdots & & \vdots \\ X_{1,1}(\lceil 2^{n(R_1-\delta)} \rceil) & \cdots & X_{1,n}(\lceil 2^{n(R_1-\delta)} \rceil) \end{bmatrix},$$

[1]We take the closure because this is an engineering class, where we don't want to bother with the boundary.

and transmitter 2 constructs the random codebook

$$\begin{bmatrix} X_{2,1}(1) & \cdots & X_{2,n}(1) \\ \vdots & & \vdots \\ X_{2,1}(m_1) & \cdots & X_{2,n}(m_1) \\ \vdots & & \vdots \\ X_{2,1}(\lceil 2^{n(R_2-\delta)} \rceil) & \cdots & X_{2,n}(\lceil 2^{n(R_2-\delta)} \rceil) \end{bmatrix}.$$

Let $W_{1,n} = \text{Unif}([M_n^{(1)}])$ and $W_{2,n} = \text{Unif}([M_n^{(2)}])$, where $M_n^{(1)} = 2^{nR_1}$ and $M_n^{(2)} = 2^{nR_2}$. Then decode via

$$d_n(Y^n) = \begin{cases} (m_1, m_2) & \text{if there is a unique } (m_1, m_2) \text{ such that} \\ & \quad ((X_1)_1^n(m_1), (X_2)_1^n(m_2), Y_1^n) \text{ is } \varepsilon\text{-jointly weakly typical} \\ \text{arbitrary} & \text{if there is no such } (m_1, m_2) \text{ or more than 1 such } (m_1, m_2). \end{cases}$$

Then, by symmetry,

$$\begin{aligned} &\mathbb{P}(d_n(e_n^{(1)}(W_{1,n}), e_n^{(2)}(W_{2,n})) \neq (W_{1,n}, W_{2,n})) \\ &\qquad = \mathbb{P}(d_n(e_n^{(1)}(1), e_n^{(2)}(1)) \neq (1,1)) \\ &\qquad \leq \mathbb{P}(E_{1,1}^c) + \sum_{i \neq 1} \mathbb{P}(E_{i,1}) + \sum_{j \neq 1} \mathbb{P}(E_{1,j}) + \sum_{i \neq 1, j \neq 1} \mathbb{P}(E_{i,j}), \end{aligned}$$

where $E_{i,j}$ is the event that $((X_1)_1^n(i), (X_2)_1^n(j), Y_1^n)$ us ε-jointly weakly typical. Then $\mathbb{P}(E_{1,1}) \to 0$ by the weak law of large numbers, and

$$\begin{aligned} \mathbb{P}(E_{i,1}) &= \sum_{((x_1)_1^n, (x_2)_1^n, y_1^n) \in A_\varepsilon^{(n)}} p((x_1)_1^n) p((x_2)_1^n, y_1^n) \\ &\leq |A_\varepsilon^{(n)}| 2^{-nH(X_1)} 2^{n\varepsilon} \\ &\leq 2^{nH(X_1, X_2, Y)} 2^{n\varepsilon} 2^{-nH(X_2, Y)} 2^{n\varepsilon} 2^{-nH(X_1)} 2^{n\varepsilon} \\ &= 2^{-nI(X_1; X_2; Y)} 2^{3n\varepsilon} \\ &= 2^{-nI(X_1; Y | X_2)} 2^{3n\varepsilon} \end{aligned}$$

because $I(X_1; X_2) = 0$. Hence,

$$\sum_{i \neq 1} \mathbb{P}(E_{i,1}) \leq 2^{n(R_1 - \delta)} 2^{-nI(X_1; X_2 | Y)} 2^{3n\varepsilon},$$

so if $R_1 < I(X_1; X_2 \mid Y) - 3\varepsilon + \delta$, then this goes to 0 as $n \to \infty$. We can apply a similar argument to $\mathbb{P}(E_{1,j})$.

When $i \neq 1$ and $j \neq 1$,

$$\begin{aligned}\mathbb{P}(E_{i,j}) &\leq \sum_{((x_1)_1^n,(x_2)_1^n,y_1^n)\in A_\varepsilon^{(n)}} \underbrace{p((x_1)_1^n)}_{\prod_{t=1}^n p_{X_1}(x_{1,t})} \underbrace{p((x_2)_1^n)}_{\prod_{t=1}^n p_{X_2}(x_{2,t})} p(y_1^n) \\ &\leq |A_\varepsilon^{(n)}| 2^{-H(X_1)} 2^{-n\varepsilon} 2^{-nH(X_2)} 2^{n\varepsilon} 2^{-nH(Y)} 2^{n\varepsilon} \\ &\leq 2^{n(H(X_1,X_2,Y)-H(X_1)-H(X_2)-H(Y))} 2^{4n\varepsilon}.\end{aligned}$$

This tells us that if $R_1+R_2 \leq I(X_1,X_2;Y)-4\varepsilon+2\delta$, then $\sum_{i\neq 1, j\neq 1} p(E_{i,j}) \to 0$ as $n \to \infty$.

For the converse, we use Fano's inequality 3 times. For any $((e_n^{(1)}, e_n^{(2)}, d_n), n \geq 1$,

$$\begin{aligned}nR_1 + o(n) &= \log\lceil 2^{nR_1}\rceil \\ &= H(W_1) \\ &= I(_1; Y_1^n) + H(W_1 \mid Y_1^n) \\ &\leq I(W_1 \mid Y_1^n) + n\varepsilon_n,\end{aligned}$$

where $\varepsilon_n \to 0$ from Fano's inequality because $H(W_1 \mid Y_1^n) \leq H(W_1, W_2 \mid Y_1^n)$ and $H(W_1, W_2 \mid Y_1^n) \leq h(p_{\text{error}}^{(n)}) + n(R_1+R_2)p_{\text{error}}^{(n)}$.

$$\begin{aligned}&\leq I((X_1)1^n(W_1); Y_1^n) \\ &= H((X_1)_1^n(W_1)) - H((X_1)_1^n(W_1) \mid Y_1^n) + n\varepsilon_n \\ &= H((X_1)_1^n(W_1) \mid (X_1)_1^n(W_2)) - H((X_1)_1^n(W_1) \mid Y_1^n, (X_2)_1^n(W_2)) + n\varepsilon_n \\ &= I((X_1)_1^n(W_1); Y_1^n \mid (X_2)_1^n(W_2)) + n\varepsilon_n \\ &= H(Y_1^n \mid (X_2)_1^n(W_2)) - H(Y_1^n \mid (X_1)_1^n(W_1), (X_2)_1^n(W_2)) + n\varepsilon_n \\ &\leq \sum_{i=1}^n H(Y_i \mid (X_2)_1^n(W_2)) - \sum_{i=1}^n H(Y_i \mid X_{1,i}(W_1), X_{2,i}(W_2)) + \varepsilon_n \\ &= \sum_{i=1}^n I(Y_i; X_{1,i} \mid X_{2,i}) + n\varepsilon_n.\end{aligned}$$

We get $R_1 \leq \frac{1}{n}\sum_{i=1}^n I(Y_i; X_{1,n} \mid X_{2,i}) + \varepsilon_n$ and similar bounds for R_2 and R_1+R_2. □

1.2 Achievable rate pairs of a multiuser AWGN channel

In the case of Gaussian noise, we have input X_1 with power constraint P_1 and input X_2 with power constraint P_2. WIth $\mathcal{N}(0,\sigma^2)$ noise, the result is more explicit:

Theorem 1.2. *With Gaussian noise, the set of achievable rate pairs is*

$$\left\{(R_1, R_2) : R_1 \leq \frac{1}{2}\log\left(1+\frac{P_1}{\sigma^2}\right)\right.$$

$$
\left.\begin{aligned}
R_2 &\leq \frac{1}{2}\log\left(1+\frac{P_2}{\sigma^2}\right)\\
R_1+R_2 &\leq \frac{1}{2}\log\left(1+\frac{P_1+P_2}{\sigma^2}\right)
\end{aligned}\right\}.
$$