Capacity per Unit Cost

4-25-06.

Based on Verdú 1990.

Consider a DMC W with inputs from $\mathcal{X}$ and outputs in $\mathcal{Y}$. For $x \in \mathcal{X}$, let $a(x) \geq 0$ denote the cost to transmit x over the channel. The cost to transmit a message $i \in \{1, \ldots, N\}$ using $\underline{x}_i \in \mathcal{X}^n$ is $\sum_{k=1}^{n} a(x_{ik})$.

__Def.__ A number $R \geq 0$ is achievable per unit cost if $\forall \lambda > 0$, $\Delta R > 0$, $\exists \alpha_0$ $\forall \alpha \geq \alpha_0$, for some block length n and number of messages N such that

$$\frac{\log N}{\alpha} > R - \Delta R,$$

$\exists$ codewords $\underline{x}_1, \ldots, \underline{x}_N$ in $\mathcal{X}^n$ with $\sum_{k=1}^{n} a(x_{ik}) \leq \alpha$ for all $i = 1, \ldots, N$, $\exists$ a decoder $\varphi : \mathcal{Y}^n \to \{1, \ldots N\}$ with

$$e_n(f, \varphi) := \frac{1}{N} \sum_{i=1}^{N} W^n(\{\underline{y} : \varphi(\underline{y}) \neq i\} | \underline{x}_i) < \lambda.$$

__Theorem.__ The set of rates $R \geq 0$ that are achievable per unit cost is $[0, \underline{C}]$, where

$$\underline{C} := \sup_{\alpha > 0} \frac{C(\alpha)}{\alpha},$$

where

$$C(\alpha) := \sup_{p : E[a(X)] \leq \alpha} I(p \times W).$$

WLOG, assume $n_0 \geq \frac{2C(\alpha)}{\Delta R \cdot \alpha} + 1$

Proof. We first show that for fixed $\alpha > 0$, $C(\alpha)/\alpha$ is an achievable rate per unit cost. Let $\lambda > 0$ and $\Delta R > 0$ be given. Since $C(\alpha)$ is achievable in the usual sense, $\exists n_0 \forall n \geq n_0$, $\exists N$ with $\frac{\log N}{n} > C(\alpha) - \frac{\Delta R \cdot \alpha}{2}$, $\exists \underline{x}_1, \ldots, \underline{x}_N$ with

$$\frac{1}{n} \sum_{k=1}^{n} a(x_{ik}) \leq \alpha \quad \text{and } \exists \phi \text{ with } e_n(f, \phi) < \lambda.$$

Put $\alpha_0 := \alpha n_0$. Let $\alpha' \geq \alpha_0$.

First suppose $\alpha' = n\alpha$ for some $n \geq 1$. Then $\alpha' \geq \alpha_0 \Rightarrow n \geq n_0$ and $\exists$ a code as above; in particular,

$$\sum_{k=1}^{n} a(x_{ik}) \leq n\alpha = \alpha' \quad \text{and} \quad \frac{\log N}{n} > C(\alpha) - \frac{\Delta R \cdot \alpha}{2}$$

$$\Rightarrow \quad \frac{\log N}{\alpha'} > \frac{C(\alpha)}{\alpha} - \frac{\Delta R}{2}.$$

Second, suppose $n\alpha < \alpha' < (n+1)\alpha$. Then using the same code,

$$\sum_{k=1}^{n} a(x_{ik}) \leq n\alpha < \alpha' \quad \text{and} \quad \frac{\log N}{\alpha'} = \frac{\log N}{n\alpha} \cdot \frac{n\alpha}{\alpha'}$$

$$(n+1)\alpha > \alpha' \geq \alpha_0 = \alpha n_0 \geq \alpha\left(\frac{2C(\alpha)}{\Delta R \cdot \alpha} + 1\right)$$

$$n > \frac{2C(\alpha)}{\Delta R \cdot \alpha}$$

$$\frac{1}{n} < \frac{\Delta R \cdot \alpha}{2C(\alpha)}$$

$$\Rightarrow \left[\frac{C(\alpha)}{\alpha} - \frac{\Delta R}{2}\right] \cdot \frac{n\alpha}{\alpha'}$$

$$\geq \left[\frac{C(\alpha)}{\alpha} - \frac{\Delta R}{2}\right] \cdot \frac{n\alpha}{(n+1)\alpha}$$

$$= \left[\frac{C(\alpha)}{\alpha} - \frac{\Delta R}{2}\right] \frac{1}{1 + 1/n}$$

$$\geq \left[\frac{C(\alpha)}{\alpha} - \frac{\Delta R}{2}\right] \frac{1}{1 + \frac{\Delta R \cdot \alpha}{2C(\alpha)}}$$

Then $\frac{\log N}{\alpha'} > \frac{c(\alpha)}{\alpha} - \Delta R$,

which follows from the fact that

$$(u-v)\cdot\frac{1}{1+v/u} = (u-v)\frac{u}{u+v}$$

$$= u\frac{[u+v-2v]}{u+v}$$

$$= u - \frac{2uv}{u+v}$$

$$\geq u - \frac{2uv}{u} = u - 2v.$$

For the converse, $\underbrace{\qquad}_{\text{let } \lambda>0,\ \Delta R>0 \text{ be given, } \exists \alpha_0 \forall \alpha' \geq \alpha_0,}$ we have a code with $\frac{\log N}{\alpha'} > R - \Delta R$ and $\sum_{k=1}^{n} a(x_{ik}) \leq \alpha'$ and $e_n(f,\varphi) < \lambda$. Let $P(M=i, \underline{X}=\underline{x}, \underline{Y}=\underline{y}) = \frac{1}{N}\delta_{\underline{x}_i}(\underline{x}) W^n(\underline{y}|\underline{x})$. By Fano's inequality,

$$H(M) \leq e_n(f,\varphi)\log N + \log 2 + I(M \wedge \varphi(\underline{Y}))$$

$\therefore \log N < \lambda \log N + \log 2 + I(\underline{X} \wedge \underline{Y})$, by the Data Proc. Lemma

$\therefore (1-\lambda)\log N < I(\underline{X} \wedge \underline{Y}) + \log 2$

Now

$$E\left[\sum_{k=1}^{n} a(X_k)\right] = \frac{1}{N}\sum_{i=1}^{N}\sum_{\underline{x}\in\mathcal{X}^n}\delta_{\underline{x}_i}(\underline{x})\sum_{k=1}^{n} a(x_k), \quad \underline{x}=(x_1,\dots,x_n)$$

$$= \frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{n} a(x_{ik}) \leq \alpha'$$

Thus, $E\left[\frac{1}{n}\sum_{k=1}^{n} a(X_k)\right] \leq \alpha'/n$.

So, 5-3-06

$$\frac{\log N}{\alpha'} < \frac{1}{1-\lambda}\left\{ \frac{n}{\alpha'} \cdot \sup_{E[\frac{1}{n}\sum_{k=1}^{n} a(X_k)] \le \alpha'/n} \frac{1}{n} I(\underline{X} \wedge \underline{Y}) + \frac{\log 2}{\alpha'} \right\} \quad (*)$$

Put $\alpha_k' := E[a(X_k)]$. Then $I(X_k \wedge Y_k) \le C(\alpha_k')$

As we show later C is concave. We can therefore write in (*)

$$\frac{1}{n} I(\underline{X} \wedge \underline{Y}) \le \frac{1}{n}\sum_{k=1}^{n} I(X_k \wedge Y_k) \le \frac{1}{n}\sum_{k=1}^{n} C(\alpha_k')$$

$$\le C\Big(\underbrace{\frac{1}{n}\sum_{k=1}^{n} \alpha_k'}_{= \frac{1}{n}\sum_{k=1}^{n} E[a(X_k)] \le \alpha'/n}\Big)$$

$$\le C(\alpha'/n), \text{ since } C \text{ is nondecreasing.}$$

Thus,

$$\frac{\log N}{\alpha'} < \frac{1}{1-\lambda}\left\{ \frac{n}{\alpha'} \cdot C\left(\frac{\alpha'}{n}\right) + \frac{\log 2}{\alpha'} \right\}$$

and

$$R < \Delta R + \frac{1}{1-\lambda}\left\{ \sup_{\alpha > 0} \frac{C(\alpha)}{\alpha} + \frac{\log 2}{\alpha'} \right\}$$

Since $\lambda > 0$, $\Delta R > 0$, & $\alpha' \ge \alpha_0$ are arbitrary,

$$R \le \sup_{\alpha > 0} \frac{C(\alpha)}{\alpha}.$$

∎

<u>Lemma</u> $C(\alpha) := \sup_{p:\, E[a(X)] \le \alpha} I(p \times W)$ is concave.

<u>Proof.</u> Let $\alpha_{min} := \min_x a(x)$. For $\alpha_1, \alpha_2 \ge \alpha_{min}$ and $0 \le \lambda \le 1$, we must show that

$$C(\lambda\alpha_1 + (1-\lambda)\alpha_2) \ge \lambda C(\alpha_1) + (1-\lambda) C(\alpha_2)$$

For $i = 1, 2$
let p_i be such that $C(\alpha_i) = I(p_i \times W)$ with

$\sum_x p_i(x) a(x) \le \alpha_i$. Then

$$\lambda C(\alpha_1) + (1-\lambda) C(\alpha_2) = \lambda I(p_1 \times W) + (1-\lambda) I(p_2 \times W)$$

$$\le I([\lambda p_1 + (1-\lambda) p_2] \times W)$$

since $I(p \times W)$ is a concave function of p. Now,

since

$$\sum_x [\lambda p_1(x) + (1-\lambda) p_2(x)] a(x)$$

$$= \lambda \sum_x p_1(x) a(x) + (1-\lambda) \sum_x p_2(x) a(x)$$

$$\le \lambda \alpha_1 + (1-\lambda) \alpha_2$$

$$I([\lambda p_1 + (1-\lambda) p_2] \times W) \le C(\lambda\alpha_1 + (1-\lambda)\alpha_2).$$ ▨