EE7330: Network Information Theory **2021**

Lecture Notes 11: Tools for proving converse results

Instructor: Shashank Vatedka *Scribe: Ritesh Kumar*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. Please email the course instructor in case of any errors.*

11.1 Fano's inequality

Consider M and $\hat{M}$ are jointly distributed and probability of error $P_e = Pr[M \neq \hat{M}]$. And M & $\hat{M} \in \mathbb{M}$. Then

$$H(M/\hat{M}) \leqslant H_2(P_e) + P_e \log|\mathbb{M}| \tag{11.1}$$

Proof :
Consider a indicator function E as :

$$E = \begin{cases} 1 & \text{if } \hat{M} \neq M \\ 0 & \text{if } \hat{M} = M \end{cases} \tag{11.2}$$

So,we have $P_E(1) = P_e, P_E(0) = 1 - P_e$, and we can write $H(E) = H_2(P_e)$. Now

$$H\left(M, E|\hat{M}\right) = H\left(M|\hat{M}\right) + H\left(E|M, \hat{M}\right) \text{ (chain rule of entropy)} \tag{11.3}$$

$$H\left(M|\hat{M}\right) = H\left(E|\hat{M}\right) + H\left(E|M, \hat{M}\right) - H\left(E|M\hat{M}\right) \tag{11.4}$$

Since E is independent from M and $\hat{M}$, we can write,

$$H\left(M|\hat{M}\right) = H\left(E|\hat{M}\right) + H\left(\frac{M}{E}, \hat{M}\right) \tag{11.5}$$

$$\leqslant H(E) + H\left(M|E, \hat{M}\right) \tag{11.6}$$

$$= H_2(P_e) + H\left(M|\hat{M}, E = 0\right) P_E(0) + H\left(M|\hat{M}, E = 1\right) P_E(1) \tag{11.7}$$

Here, $H\left(M|\hat{M}, E = 0\right) = 0$ because for E = 0 , M = $\hat{M}$, and hence entropy = 0.

$$= H_2(P_e) + H\left(M|\hat{M}, E = 1\right) P_E(1) \tag{11.8}$$

$$= H_2(P_e) + H\left(M|\hat{M}, E = 1\right) P_e \tag{11.9}$$

$$\leqslant H_2(P_e) + H(M) P_e \tag{11.10}$$

$$\leqslant H_2(P_e) + P_e \log_2|\mathbb{M}| \text{ Proved.} \tag{11.11}$$

11.2 Proof of converse channel coding theorem

Theorem : Consider any sequence with (ENC_n, DEC_n) for DMC with transition probability $\text{P}_{Y|X}$ such that ,

$$\liminf_{\text{n}\to\infty} \frac{K_n}{n} \geqslant C + \epsilon$$

then,

$$P_e = \limsup_{n\,\to\,\infty}\ = Pr\left[\hat{M}_i \neq M_i\right] \geqslant \frac{\epsilon}{R}$$

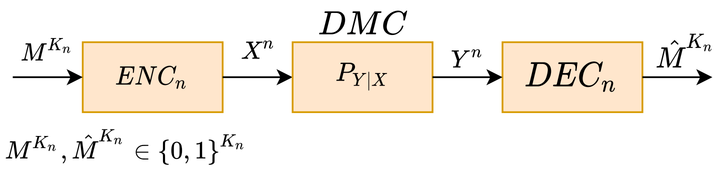

Figure 11.1: Single user over DMC

Proof:
We need some well-known inequalities to prove this. 1) Fano's inequality and 2) bound on mutual information. Fano's inequality we have discussed just above and now let us discuss bound on mutual information.
Lemma : For any P_{X^n}, if Y^n is obtained by passing X^n through Discrete memoryless channel having transition probability $\text{P}_{Y|X}$ then,

$$I\left(X^n;Y^n\right) \leqslant \sum_{i=1}^{n} I\left(X_i;Y_i\right) \leqslant nC \tag{11.12}$$

Proof of lemma :

Consider the expression of mutual information,

$$I\left(X^n;Y^n\right) = H\left(Y^n\right) - H\left(Y^n|X^n\right) \tag{11.13}$$

$$= \sum_{i=1}^{n}\left[H\left(Y_i|Y_1,Y_2\ldots Y_{i-1}\right) - H\left(Y_i|Y_1,Y_2\ldots Y_{i-1},X^n\right)\right] \tag{11.14}$$

(Chain rule of entropy)

$$\leqslant \sum_{i=1}^{n}\left[H\left(Y_i\right) - H\left(Y_i|Y_1,Y_2\ldots Y_{i-1},X^n\right)\right] \tag{11.15}$$

(Since conditioning decreases entropy)

Now,

$$H\left(Y_i|Y_1,Y_2\ldots Y_{i-1},X^n\right) = \sum_{x^n,y_1,\ldots,y_n} P_{Y|X}\left(y_i|y_1,\ldots y_{i-1},x^n\right) \times \log\left(\frac{1}{P_{Y|X}\left(y_i|y_1,\ldots y_{i-1},x^n\right)}\right) \tag{11.16}$$

Since for discrete memoryless channel present output depends only on the present input, hence we can write.

$$P_{Y|X}\left(y_i|y_1,\dots y_{i-1},x^n\right)=P_{Y|X}\left(y_i|x_i,y_1,x_i\dots y_{i-1},x_{i-1}\right)=P_{Y|X}\left(y_i|x_i\right) \tag{11.17}$$

Using eq (11.17) in eq (11.16), we can write,

$$H\left(Y_i|Y_1,Y_2\dots Y_{i-1},X^n\right)=\sum_{x^n,y_1,\dots,y_n}P_{Y|X}\left(y_i|x_i\right)\times\log_2\left(\frac{1}{P_{Y|X}\left(y_i|x_i\right)}\right)=H\left(Y_i|X_i\right) \tag{11.18}$$

By combining all i.e using eq(11.18) in eq11.15 we can write,

$$=\sum_{i=1}^{n}\left[H\left(Y_i\right)-H\left(Y_i|X_i\right)\right] \tag{11.19}$$

And we have,

$$\sum_{i=1}^{n}\left[H\left(Y_i\right)-H\left(Y_i|X_i\right)\right]=\sum_{i=1}^{n}I\left(X_i;Y_i\right) \tag{11.20}$$

From eq(11.13), eq(11.15) and eq(11.20) we have,

$$I\left(X^n;Y^n\right)\leqslant\sum_{i=1}^{n}I\left(X_i;Y_i\right)\leqslant nC \tag{11.21}$$

This conclude the proof of eq (11.12).

Proof of converse:

We have message, which is i.i.d and uniform so we can write,

$$K_n=H\left(M^{K_n}\right) \tag{11.22}$$

$$=H\left(M^{K_n}|\hat{M}^{K_n}\right)+I\left(M^{K_n};\hat{M}^{K_n}\right) \tag{11.23}$$

$$\leqslant H_2(P_e)+P_e\log_2(2^{M^{K_n}})+I(M^{K_n}:M'^{K_n}) \tag{11.24}$$

(using Fano's inequality)

$$\leqslant H_2\left(P_e\right)+P_eK_n+I\left(M^{K_n};\hat{M}^{K_n}\right) \tag{11.25}$$

$$\leqslant H_2\left(P_e\right)+P_eK_n+I\left(X^n;Y^n\right) \tag{11.26}$$

(using data processing inequality)

$$\leqslant H_2\left(P_e\right)+P_eK_n+nC\quad\text{(From eq 11.21)} \tag{11.27}$$

$$K_n\leqslant H_2(P_e)+P_eK_n+nC\Rightarrow\frac{H_2(P_e)}{n}\geqslant\frac{K_n(1-P_e)}{n}-C$$

$$\lim_{n\to\infty}\left(\frac{K_n(1-P_e)}{n}-C\right)\leqslant\lim_{n\to\infty}\frac{H_2(P_e)}{n}$$

Where, $\mathrm{P}_e=\limsup_{n\to\infty}Pr\left[\hat{M}_i^{K_n}\neq M_i^{K_n}\right]$

Let $\frac{K_n}{n} = R$,

$$\lim_{n \to \infty} P_e \geqslant \frac{R - C}{R} \tag{11.28}$$

$$\limsup_{n \to \infty} P_e \geqslant \frac{R - C}{R} \tag{11.29}$$

If we operate at the rate more than capacity of the channel say R = C + ϵ

$$\limsup_{n \to \infty} \; P_e \geqslant \frac{\epsilon}{R} \tag{11.30}$$

That is probability or error is always non-zero.But if we operate the channel below the capacity and for

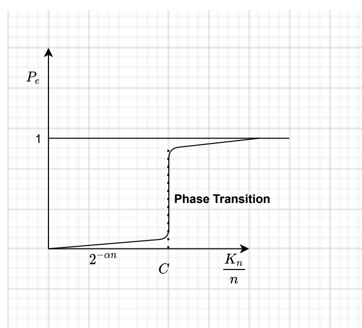

Figure 11.2: Probability of error with capacity constraint

large n, in (11.29) we can observe P$_e$ approaches to 0.

11.3 Mrs Gerber's Lemma:

Suppose we have X = $(X_i, \ldots X_n)$ where X$_i$ $\in \mathbb{M}^n \in \{0, 1\}$ be a binary random n-vector. Let P$_X(x)$ = $P\,(X = x)$, where $x \in \mathbb{M}^n$ define its probability distribution Let say X $\sim$ Ber(q). Let us consider that the random vector X is the input to a binary symmetric channel with crossover probability p, where $0 < p < \frac{1}{2}$. Let be Y = $(Y_i, \ldots Y_n)$ where Y$_i$ $\in \mathbb{M}^n \in \{0, 1\}$ the corresponding channel output n-vector. The probability

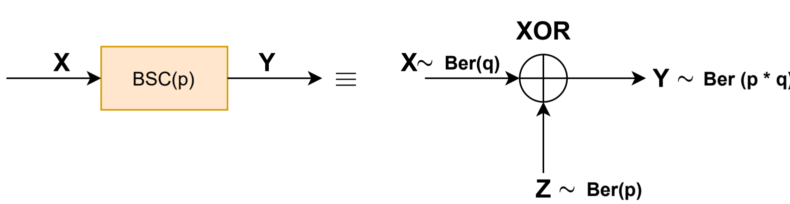

Figure 11.3: BSC channel

distribution of Y is defined using transition probability of channel and X (described in figure 11.3). We will use notation **p⋆q = p(1-q) + q(1-p)**. Then ,

$$H(Y) \geqslant H_2\left(p \star H_2^{-1}\left(H(X)\right)\right) \tag{11.31}$$

with equality if and only if the ${\{X_i\}_1}^n$ are independent. and H$\{X^k\} = kp$

Lemma : Suppose we are generating X which depends on some distribution U, and it passes through BSC(p) (with transition probability p) with output distribution Y. Then,

$$H\left(\frac{Y}{U}\right) \geqslant H_2\left(H_2^{-1}\left(H\left(\frac{X}{U}\right)\right) \star p\right) \tag{11.32}$$

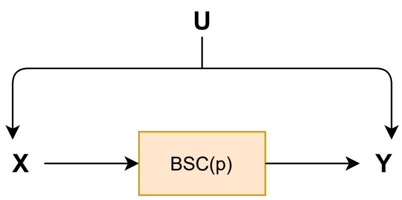

Figure 11.4: Binary symmetric channel

In vector form,

$$\frac{H\left(Y^n|U\right)}{n} \geqslant H_2\left(H_2^{-1}\left(H\left(\frac{X^n|U}{n}\right)\right) \star p\right) \tag{11.33}$$

Proof. We have claim ,

$$H\left(\frac{Y}{U}\right) \geqslant H_2\left(H_2^{-1}\left(H\left(\frac{X}{U}\right)\right) \star p\right) \tag{11.34}$$

Considering R.H.S. of the inequality we have,

$$H_2\left(H_2^{-1}\left(H\left(\frac{X}{U}\right)\right) \star p\right) \tag{11.35}$$

Suppose we have $f(u) = H_2\left(H_2^{-1}(v) \star p\right)$.
Now, eq (11.35) can be written as,

$$f\left(\sum_u P_U(u).H\left(X|U=u\right)\right) \tag{11.36}$$

Claim : f is convex function.

Proof. We have $f(u) = H_2\left(H_2^{-1}(v) \star p\right)$. Let consider, $g(u) = H_2^{-1}(v)$,
Then, we can write,

$$\begin{aligned} f(u) &= H_2\left(g(u) \star p\right), \\ f(u) &= H_2\left(g(u)(1-p) + (1-g(u))\,p\right) \end{aligned}$$

put $\alpha = g(u)(1-p) + (1-g(u))\,p$, so we have,

$$f(u) = H_2(\alpha) = -\left[\alpha \log_2(\alpha) + (1-\alpha)\log_2(1-\alpha)\right]$$

$$f'(u) = -\left[\frac{\alpha}{\alpha} + \log_2 \alpha \frac{1-\alpha}{1-\alpha}(-1) + (-1)\log_2(1-\alpha)\right]\frac{1}{\ln(2)}.\frac{\partial \alpha}{\partial u}$$

$$f'(u) = -\frac{1}{\ln(2)}.g'(u)(1-2p)\left[\log(\frac{\alpha}{1-\alpha})\right]$$

$$f''(u) = \frac{1}{\ln(2)}.g''(u)(1-2p)^2\left[\frac{2-\alpha}{\alpha(1-\alpha)}\right]$$

Here $f''(u)$ is always positive for $0 < \alpha < 1$, hence f is convex function. □

Now coming to main proof of lemma. Since f is convex function, we can write,

$$f\left(\sum_u P_U(u).H\left(X|U=u\right)\right) \leqslant \sum_u P_U(u).f\left(H\left(X|U=u\right)\right) \tag{11.37}$$

Again using the expression of $f(u)$ we can have,

$$f\left(H\left(X|U=u\right)\right) = H_2\left(H_2^{-1}\left(H\left(X|U=u\right)\right) \star p\right) \tag{11.38}$$

For $U = u, X \sim \text{Ber}(q_u)$ and Y $\sim (q_u \star p)$.

$$H\left(X|U=u\right) = H_2\left(q-u \star p\right) = H_2\left(H_2^{-1}\left(H\left(X|U=u\right)\right) \star p\right) = f\left(H\left(X|U=u\right)\right) \tag{11.39}$$

Combining all above expressions, R.H.S. becomes,

$$\text{R.H.S} \leqslant \sum_u P_U(u).H\left(X|U=u\right) \tag{11.40}$$

Hence,

$$H\left(\frac{Y}{U}\right) \geqslant H_2\left(H_2^{-1}\left(H\left(\frac{X}{U}\right)\right) \star p\right) \tag{11.41}$$

and this concludes the proof. □