Mathematical methods in communication 27.05.2011

lecture 9

Lecturer: Haim Permuter *Scribe: Racheli Hayun*

In this lecture we will present:

1) Intuition for Gaussian channel capacity $C = \frac{1}{2}\log(\frac{P+N}{N})$.
2) Channel with state information.

I. Intuition for Gaussian channel capacity

The volume of sphere in R^n with radius r is proportional to r^n.

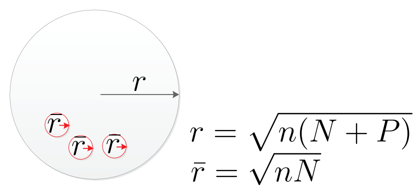


Fig. 1. Sphere with radius r that contains small spheres with radius $\bar{r}$, each represent a codeword X with noise

Reminder:

- Gaussian channel with input X has a power constraint $\frac{1}{n}\sum_i X_i^2 \leq P$, therefore the radius of the sphere is $r = \sqrt{\sum_i X_i^2} \leq \sqrt{nP}$.
- Z is a Gaussian noise and is independent of the input signal X.

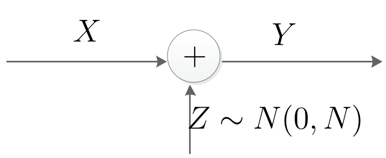


Fig. 2. Gaussian channel with noise

If we send a signal X^n the output will be in a small sphere with radius $\sqrt{nN}$, the variance of Y^n:

$$var(Y^n) = \sqrt{n(N+P)} \tag{1}$$

The sphere where the output take place has a radius $r = \sqrt{n(N+P)}$.
How many codeword we can broadcast without errors?

$$\#codewords = \frac{vol(output)}{vol(noise)} = \frac{\alpha(\sqrt{n(N+P)})^n}{\alpha(\sqrt{nN})^n} = \left(\sqrt{\frac{N+P}{N}}\right)^n \tag{2}$$

R is defined:

$$R = \frac{1}{n}\log(\#codewords) = \frac{1}{n}\log\left(\frac{(P+N)^{\frac{n}{2}}}{N^{\frac{n}{2}}}\right) = \frac{1}{2}\log\left(\frac{P+N}{N}\right) \tag{3}$$

In achievability proof we saw that we generate the codewords according to $N(0,P)$, therefore all codewords are expected to be at the edge of the sphere.

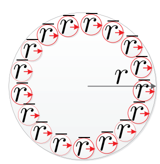


Fig. 3. Sphere with codewords on the edge. For large dimension it turns out that we do not loos anything by placing the codewords on the edge.

Example 1 : (Ratio between volumes - R^2)
Consist two balls - one with radius R and one with radius $r-\epsilon$. Lets calculate the ratio between the volumes for $n=2$:

$$\left(\frac{r-\epsilon}{r}\right)^2 \to 1 \tag{4}$$

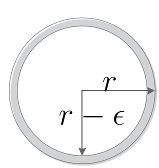


Fig. 4. Two balls with about the same volume for $n = 2$

The internal ball has the same volume as the external ring therefore it can't include more codewords.

Lets calculate the ratio between the volumes for a sphere of dimension n, where n is very large:

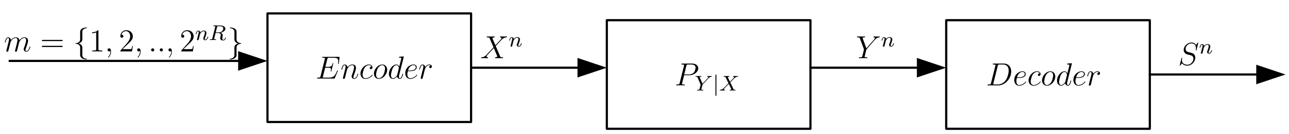
$$\left(\frac{r-\epsilon}{r}\right)^n \rightarrow 0 \quad \forall\epsilon > 0 \tag{5}$$

In this case all the volume is close to the edge, therefore all the codewords are close to the edge and if we will put some inside it will not improve the rate.

II. Channel With State Information

In most cases, particularly wireless communication, the channels we use is varying with time. Denote the channel state $S \sim i.i.d$ with $P(s)$ and independent of the messages.

Channel coding:



$$C = \max_{p(x)} I(X;Y) \tag{6}$$

Fig. 5. Communication system without state information

Channel coding with Side Information:

Assuming:

$$P(s_{i+1} \mid x^i, y^{i-1}, s^i) = P(s_{i+1}) \tag{7}$$

$$P(y_i \mid x^i, s^i, y^{i-1}) = P(y_i \mid x_i, s_i) \tag{8}$$

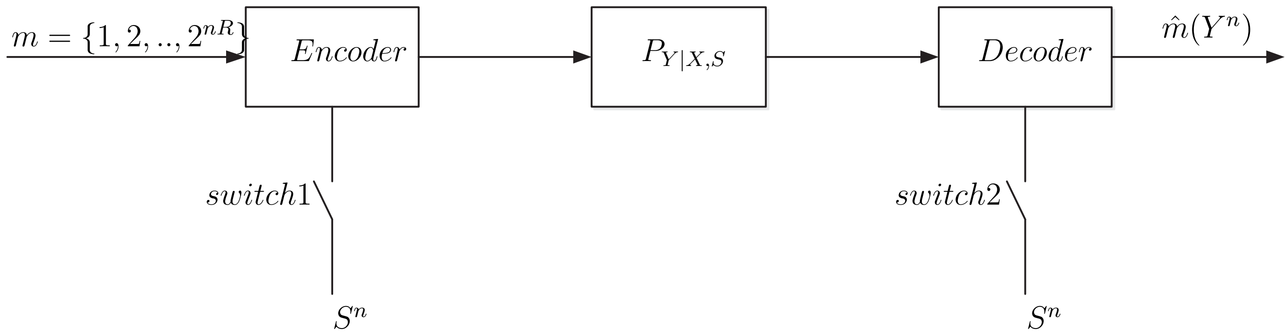


Fig. 6. Communication system with side information available either at the encoder, or the decoder or both

which implies:

$$\begin{aligned} P(y_i, s_{i+1} \mid x^i, y^{i-1}, s^i) &\overset{(a)}{=} P(s_{i+1} \mid x^i, y^{i-1}, s^i) \cdot P(y_i \mid x^i, s_i, s_{i+1}) \\ &= P(s_{i+1}) \cdot P(y_i \mid x_i, s_i) \end{aligned} \tag{9}$$

where

(a) follows from chain rule.

discuss the following cases:

1) No state information is available - CSI (both switches in Fig. 6 are open).
2) CSI is available to the decoder (Only to the decoder, switch (1) in Fig. 6 is open).
3) CSI is available to the decoder and encoder (Both switches in Fig. 6 are closed).

Case I : No State Information Is Available - CSI

Theorem 1: The capacity of a channel with no state information is

$$C = \max_{p(x)} I(X;Y) \tag{10}$$

Proof: This is the regular channel we discussed so far. A channel is memoryless if $P(y_i|x^i, y^{i-1}) = P(y_i|x_i)$. Now let us show that a channel with $i.i.d$ state is memoryless:

$$P(y_i \mid x^i, y^{i-1}) = \sum_{s_i} P(y_i, s_i \mid x^i, y^{i-1})$$

$$\begin{aligned}
&= \sum_{s_i} P(s_i \mid x^i, y^{i-1}) \cdot P(y_i \mid s_i, x^i, y^{i-1}) \\
&\stackrel{(a)}{=} \sum P(s_i) \cdot P(y_i \mid s_i, x_i) \\
&\stackrel{(b)}{=} \sum P(s_i \mid x_i) \cdot P(y_i \mid s_i, x_i) \\
&= P(y_i \mid x_i) \qquad (11)
\end{aligned}$$

where

(a) follows from $S \sim i.i.d$ and the second assumption.

(b) follows from S_i, X_i are independent.

■

Case II : CSI Available For The Decoder We define the problem:

$$\begin{aligned}
Encoder &: \quad f : \{1, 2, ..., 2^{nR}\} \ \rightarrow \ X^n \\
Decoder &: \quad g : Y^n, S^n \ \rightarrow \ \{1, 2, ..., 2^{nR}\} \\
& S^n \sim p(s) \quad i.i.d
\end{aligned}$$

Theorem 2: The capacity of a CSI available for the decoder channel is:

$$C = \max_{p(x)} I(X; Y, S) = \max_{p(x)} I(X; Y|S) \qquad (12)$$

Proof: The channel output is the vector $(Y, S)^n$, denote a new output as $\tilde{Y} = (Y, S)^n$. The channel capacity is $C = \max_{p(x)} I(X; \tilde{Y}) = \max_{p(x)} I(X; Y, S)$. The joint distribution is $P(x)P(s)P(y|x, s)$ and therefore X is independent at S:

$$\begin{aligned}
C &= \max_{p(x)} I(X; Y, S) \\
&\stackrel{(a)}{=} \max_{p(x)} [I(X; S) + I(X; Y \mid S)] \\
&\stackrel{(b)}{=} \max_{p(x)} I(X; Y|S) \qquad (13)
\end{aligned}$$

where

(a) follows from chain rule.

(b) follows from X, S independent.

■

Case III : CSI available for the Decoder and Encoder We define the problem:

$$\begin{aligned} Encoder &: \; f : \{1, 2, ..., 2^{nR}\}, S^n \;\rightarrow\; X^n \\ Decoder &: \; g : Y^n, S^n \;\rightarrow\; \{1, 2, ..., 2^{nR}\} \end{aligned} \tag{14}$$

Theorem 3: The capacity where CSI is available at the decoder and encoder channel is:

$$\begin{aligned} C &= max_{p(x|s)} I(X; Y \mid S) \\ &= max_{p(x|s)} \sum_{s_i} P(S = s) \cdot I(X; Y \mid S = s) \qquad (15) \end{aligned} \tag{16}$$

Proof: We will split the message $\{1, 2, ..., 2^{nR}\} \;:\; \{1, 2, ..., 2^{nR_0}\}$ x $\{1, 2, ..., 2^{nR_1}\}$

$$R_0 = I(X; Y|S = 0) \tag{17}$$

$$R_1 = I(X; Y|S = 1) \tag{18}$$

$$\tag{19}$$

Doing that we split the channel into two separate channels:

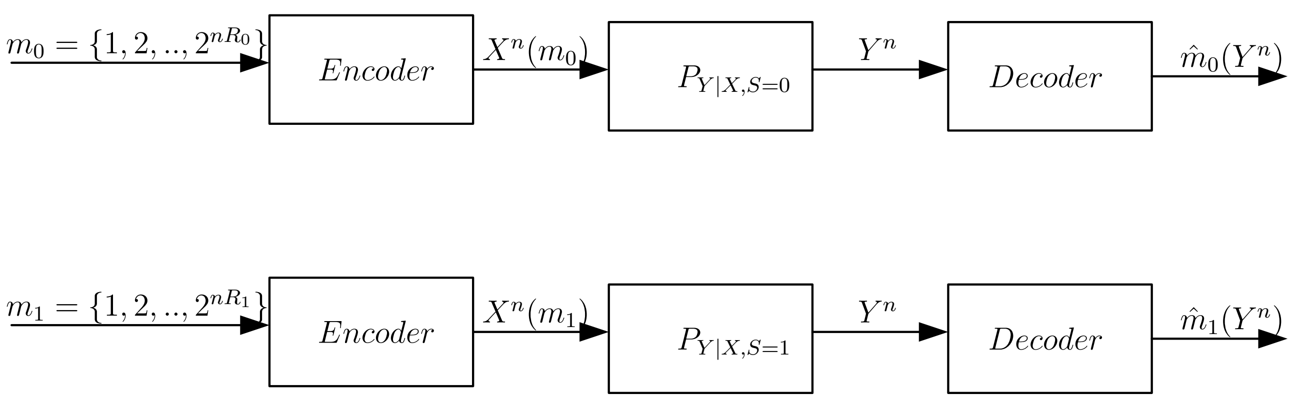


Fig. 7. Splitted Channel

such that:

$$\begin{aligned}
R_0 &= n \cdot p(S=0) I(X;Y|S=0) \\
BlockSize(S=0) &= n \cdot p(S=0) \\
R_1 &= n \cdot (S=1) I(X;Y|S=1) \\
BlockSize(S=1) &= n \cdot p(S=1)
\end{aligned}$$

Calculating the total rate of the channel:

$$R = \frac{1}{n}[np(S=0)I(X;Y|S=0) + np(S=1)I(X;Y|S=1)] = I(X;Y|S) \tag{20}$$

Converse :

$$\begin{aligned}
nR &= H(M) \qquad (21)\\
&\overset{(a)}{=} H(M|S^n) \\
&= H(M|S^n) - H(M|S^n, Y^n) + H(M|S^n, Y^n) \\
&\overset{(b)}{\leq} I(M;Y^n|S^n) + n\epsilon_n \\
&= I(M, X^n(M, S^n); Y^n|S^n) + n\epsilon_n \\
&= H(Y^n|S^n) - H(Y^n|S^n, X^n, M) + n\epsilon_n \\
&\overset{(c)}{=} \sum_{i=1}^{n} H(Y_i|Y^{i-1}, S^n) - H(Y_i|Y^{i-1}, S^n, X^n, M) + n\epsilon_n \\
&\leq \sum_{i=1}^{n} H(Y_i|S_i) - H(Y_i|S_i, X_i) + n\epsilon_n \\
&= \sum_{i=1}^{n} I(Y_i; X_i|S_i) + n\epsilon_n \\
&\leq [\max_{p(x|s)} I(Y;X|S) + \epsilon_n]n
\end{aligned}$$

(a) $S^n \perp M$

(b) Fano inequality

(c) Chain rule

■

Example 2 : Channel with side information

$S \sim B(\frac{1}{2})$

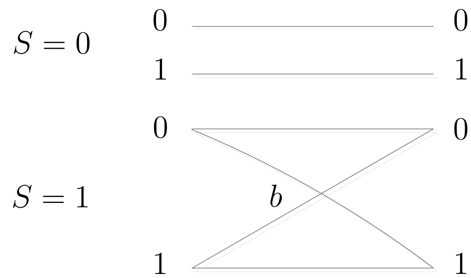


Fig. 8. Example - channel with CSI. For $S = 0$ the channel is perfect and for $S = 1$ we get a BSC channel with probability b

- If the decoder and encoder don't know the state, the equivalent channel is BSC:

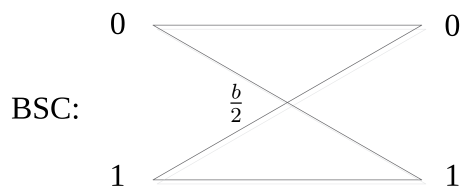


Fig. 9. Example - equivalent channel where CSI is not available for the decoder and encoder

Therefore, $C = \max_{p(x)} I(X;Y) = 1 - H(\frac{b}{2})$

- If the state is available only for the decoder and X, S independent

$$\begin{aligned} C_d &= \max_{p(x)} I(X;Y \mid S) \\ &= \max_{p(x)} [P(S=0) \cdot I(X;Y \mid S=0) + P(S=1) \cdot I(X;Y \mid S=1)] \\ &= \frac{1}{2} \cdot 1 + \frac{1}{2}(1 - H(b)) \\ &= 1 - \frac{1}{2} H(b) \end{aligned} \tag{22}$$

Who is bigger $H(\frac{b}{2})$ or $\frac{1}{2}H(b)$?

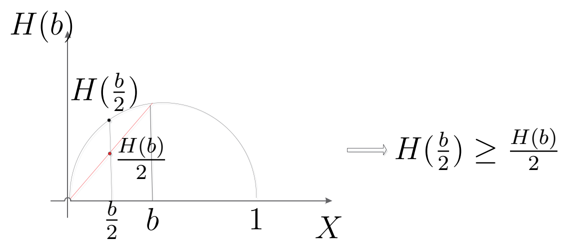


Fig. 10. Calculating the entropy at $X = b$, $X = \frac{b}{2}$

We know$H(\cdot)$ is convex with $\bar{\lambda} = \frac{1}{2}, \lambda = \frac{1}{2}$

Therefore, $H(\frac{1}{2}b + \frac{1}{2}0) \geq \frac{1}{2}H(b) + \frac{1}{2}H(0) = \frac{1}{2}H(b)$