COMPSCI 650 Applied Information Theory	Mar 24, 2016

Lecture 15

Instructor: Arya Mazumdar | *Scribe: Names Redacted*

1 Review

1.1 Channel Capacity

In the communication system (Figure 1),we define the channel capacity:

$$C = \max_{p(x)} I(X;Y) \tag{1}$$

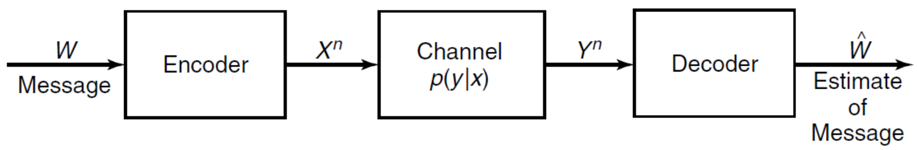

Figure 1: Communication System

1.2 Binary Symmetric Channel

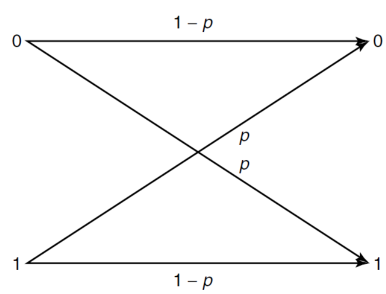

Figure 2: Binary Symmetric Channel

The information capacity of a binary symmetric channel with parameter p is:

$$C = 1 - h(p)\, bits \tag{2}$$

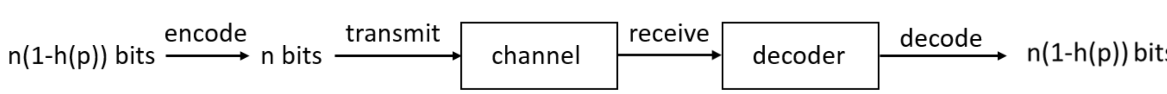

Figure 3: communication system for Binary Symmetric Channel

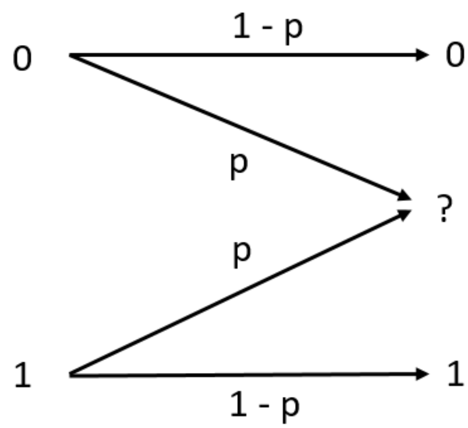

Figure 4: Binary Erasure Channel

2 Binary Erasure Channel (BEC)

In this channel, a fraction p of bits are erased (Figure 4).

2.1 Capacity of Binary Erasure Channel

$$\begin{aligned} C &= \max_{p(x)} I(X;Y) && (3) \\ &= \max_{p(x)} H(Y) - H(Y|X) && (4) \\ &= \max_{p(x)} H(Y) - \sum_x p(x)H(Y|X=x) && (5) \\ &= \max_{p(x)} H(Y) - h(p) && (6) \end{aligned}$$

First guess for $\max_{p(x)} H(Y)$ is $\log 3$, but we cannot achieve this by any choice of input distribution p. Assume, $p(x=1)=\pi$, $p(x=1)=1-\pi$.
let E be the event {Y= ?}, so $H(E)=h(p)$, $H(Y|E=1)=0$, $H(Y|E=0)\leq 1$. Thus

$$\begin{aligned} H(Y) = H(Y|E) &= H(E) + H(Y|E) && (7) \\ &= h(p) + p(E=0)H(Y|E=0) + p(E=1)H(Y|E=1) && (8) \\ &= h(p) + (1-p)H(Y|E=0) && (9) \\ &\leq h(p) + (1-p) && (10) \end{aligned}$$

Hence

$$\begin{aligned} C &= \max_{p(x)} H(Y) - h(p) && (11) \\ &= h(p) + 1 - p - h(p) && (12) \\ &= 1-p && (13) \end{aligned}$$

2.2 example

If $p(x=1)=p(x=0)=\frac{1}{2}$, then H(Y)= ?

$p(Y=0)=\frac{1}{2}(1-p)$

$p(Y=1)=\frac{1}{2}(1-p)$

$p(Y=?)=\frac{1}{2}p+\frac{1}{2}p=p$

$$\begin{aligned}
H(Y) &= \frac{1}{2}(1-p)\log\frac{1}{2}(1-p)-p\log p-\frac{1}{2}(1-p)\log\frac{1}{2}(1-p) \\
&= -p\log p-(1-p)\log(1-p)-(1-p)\log\frac{1}{2} \\
&= h(p)+(1-p) \text{ (this is } \max_{p(x)} H(Y) \text{ in BEC channel)} \\
I(X;Y) &= H(Y)-h(p)=1-p
\end{aligned}$$

Actually, the expression (13) for the capacity of BEC channel has some intuitive meaning:Since a proportion p of the bits are lost in the channel, we can recover (at most) a proportion (1-p) of the bits. Hence the capacity is at most 1-p.

3 Binary Deletion Channel

In this channel, each bit is deleted with probability p. However, the capacity of deletion channel is still an open problem.

4 Gaussian Channel

4.1 Definition

Gaussian channel is the most important continuous alphabet channel. It has the output Y, input X and noise Z. The noise Z is drawn i.i.d from a Gaussian distribution with variance σ^2.

$$Y=X+Z \qquad Z\sim\mathcal{N}(0,\sigma^2) \tag{14}$$

The limitation on input $(x_1, x_2, \cdots, x_n)$ is that:

$$\frac{1}{n}\sum_{i=1}^{n}x_i^2\leq p \rightarrow E[x^2]\leq p \tag{15}$$

4.2 Capacity of Gaussian Channel

define:

$$C=\sum_{f(x):E[x^2]\leq p} I(X;Y) \tag{16}$$

Here

$$\begin{aligned} I(X;Y) &= h(Y) - h(Y|X) && (17)\\ &= h(Y) - h(X+Z|X) && (18)\\ &= h(Y) - h(Z|X) && (19)\\ &= h(Y) - h(Z) && (20)\\ &= h(Y) - \frac{1}{2}\log 2\pi e\sigma^2 && (21) \end{aligned}$$

To find the maximum value of I(X;Y), we need to find the maximum value of h(Y). In lecture 12, we have proved the Theorem: A Gaussian random variable has the highest entropy of all random variables x with fixed variance σ^2. That is:

$$\sigma^2 \geq \frac{1}{2\pi e} e^{2h(x)} \tag{22}$$

And

$$\begin{aligned} E[Y^2] = E[(X = Z)^2] &= E[X^2] + E[Z^2] + 2E[X]\cdot E[Z] \quad (E[Z] = 0) && (23)\\ &= E[X^2] + E[Z^2] && (24) \end{aligned}$$

Since $E[X^2] \leq p$ and $E[Z^2] = \sigma^2$

$$E[Y^2] \leq p + \sigma^2 \tag{25}$$

from inequality(22), we can get:

$$h[Y] \leq \frac{1}{2}\log 2\pi e(p + \sigma^2) \tag{26}$$

Hence

$$I(X;Y) \leq \frac{1}{2}\log 2\pi e(p+\sigma^2) - \frac{1}{2}\log 2\pi e\sigma^2 = \frac{1}{2}\log(1 + \frac{p}{\sigma^2}) \tag{27}$$

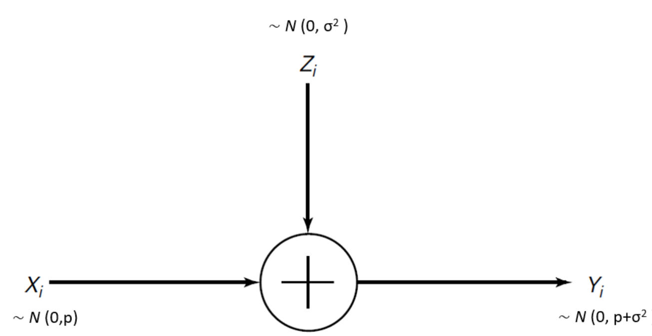

Figure 5: Gaussian Channel

Thus, in the Gaussian Channel (Figure 5), the capacity is:

$$C = \sum_{f(x):E[x^2]\leq p} I(X;Y) = \frac{1}{2}\log(1 + \frac{p}{\sigma^2}) \tag{28}$$

define

$\frac{p}{\sigma^2}$: signal to noise ratio(SNR)

Hence

$$C = \frac{1}{2}\log(1 + SNR) \tag{29}$$

In general, capacity can be computed for arbitrary channels using numerical algorithms such as Arimoto-Blahut algorithm.

5 Random codes achieve capacity

5.1 recall

index set $W = \{1, 2, \cdots, M\}$ messages:

$$\begin{aligned} 1 &\to X_1(1)\cdots X_n(1) \\ 2 &\to X_1(2)\cdots X_n(2) \\ &\cdots\cdots \\ M &\to X_1(M)\cdots X_n(M) \end{aligned}$$

$R = \frac{\log M}{n} = 1 - h(p)$ and $p_e^{(n)} \to 0$

Find all codewords that are within $n(p+\epsilon)$ bits flip away from y. If there is only one codeword found, then output that; otherwise, declare failure. In this algorithm, let A: X(i) is more than $n(p+\epsilon)$ bits flip away from y and B: $\exists j \neq 1$: X(j) is within $n(P+\epsilon)$ bits flip away from y. So the probability of error is:

$$p_e^{(n)} = Pr(A + B) \leq P(A) + P(B)$$

From chernoff bond, we know that:

$$P\left(\sum_{i=1}^{n} X(i) \geq n(p+\epsilon)\right) = 2^{-nD(p+\epsilon||p)} \to 0 \tag{30}$$

Thus $P(A) \to 0$ and $p_e^{(n)} \approx P(B)$
And $P(B) \leq (M-1)\cdot Pr(X(i)$ is within $n(p+\epsilon)$ bits flip away from y)
So

$$p_e^{(n)} \approx P(B) \leq (M-1)2^{-nD(p+\epsilon||p)} < M \cdot 2^{-nD(p+\epsilon||\frac{1}{2})} \tag{31}$$

say

$$M = 2^{n(1-h(p+2\epsilon))}$$

Hence

$$\begin{aligned} p_e^{(n)} &< 2^{n(1-h(p+2\epsilon))}\cdot 2^{-nD(p+\epsilon||\frac{1}{2})} && (32) \\ &= 2^{-n[h(p+2\epsilon)-h(p+\epsilon)]} && (33) \\ &= 2^{-nh\epsilon} \to 0 && (34) \end{aligned}$$

And

$$R = \frac{n(1-h(p+2\epsilon))}{n} = 1 - h(p+2\epsilon) \tag{35}$$

If $\epsilon \to 0$, then $R \to 1 - h(p)$.