

Lecture 11

Lecturer: Haim Permuter

Scribe: Elad Perez

I. FOR A GAUSSIAN CODE BOOK THE WORSE NOISE IS GAUSSIAN

In this section we will show that for a channel with a code book generated by a Gaussian distribution the worst noise is a Gaussian noise. For simplicity we will look at the channel with just one user and a power constraint $\frac{1}{n} \sum_{i=1}^n X_i \leq P$. At time i the output of the channel Y_i is given by $Y_i = X_i + Z_i$, where Z_i is the noise of the channel, with variance N , and X_i is the input of the channel.

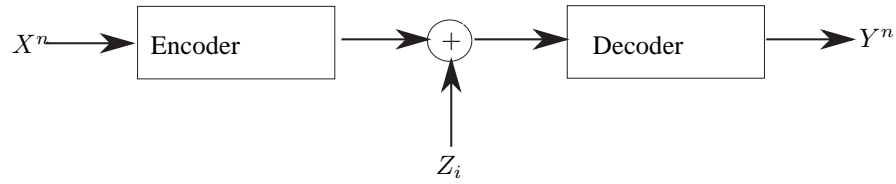


Fig. 1. Additive noise channel

Notation- when a random variable X has a Gaussian distribution we will denote it by X_G .

As we know for a channel $Y = X + Z_G$ where Z_G is a Gaussian noise with a variance N and the input X has a power constraint of P we have $I(X; Y) = I(X; X + Z_G) \leq I(X_G; X_G + Z_G)$, where $X_G \sim N(0, P)$. So if we want to maximize the mutual information between X and Y we will use a Gaussian encoder, we will now show that if X is Gaussian than the worst noise is a Gaussian noise.

Lemma 1 For a Gaussian input the worst noise is Gaussian For $Y = X_G + Z$ we have $I(X_G; Y) \geq \frac{1}{2} \log(1 + \frac{P}{N}) = I(X_G; X_G + Z_G)$ (where $Z_G \sim N(0, N)$).

Proof:

$$\begin{aligned}
 I(X_G; Y) &= h(X_G) - h(X_G|Y) \\
 &= \frac{1}{2} \log(2\pi e P) - h(X_G - \alpha Y|Y) \\
 &\stackrel{(a)}{\geq} \frac{1}{2} \log(2\pi e P) - h(X_G - \alpha Y)
 \end{aligned}$$

$$\stackrel{(b)}{\geq} \frac{1}{2} \log(2\pi e P) - \frac{1}{2} \log(2\pi e E((X_G - \alpha Y)^2)).$$

Where

(a) follows from the fact that conditioning reduces entropy.

(b) follows from the fact that entropy is the largest with a Gaussian distribution.

Note that the above is true for any $\alpha \in \mathbb{R}$ (any real α) we want to find α such that $E(X_G - \alpha Y)^2$ is minimal- we'll choose α based on the best linear approximation of X_G by Y so we want α such that:

$$\begin{aligned} E((X_G - \alpha Y)Y) &= 0 \\ E(X_G Y) &= \alpha E(Y^2) \\ E(X_G(X_G + Z)) &= \alpha E((X_G + Z)^2) \end{aligned} \tag{1}$$

Follows from (1) and the fact that X_G and Z have no correlation and $E(X_G^2) = P, E(X_G) = 0$ and $E(Z^2) = N$ we obtain:

$$P = \alpha(P + N) \tag{2}$$

and from (2) we obtain $\alpha = \frac{P}{P+N}$.

Now to place α in our earlier equation and we get:

$$I(X_G; Y) \geq \frac{1}{2} \log(2\pi e P) - \frac{1}{2} \log(2\pi e E \left[(X_G - \frac{P}{P+N} Y)^2 \right]) = \frac{1}{2} \log \left(\frac{P}{E \left[(X_G - \frac{P}{P+N} Y)^2 \right]} \right). \tag{3}$$

Now if we look at the denominator of the log we can see:

$$\begin{aligned} E \left[(X_G - \frac{P}{P+N} Y)^2 \right] &= E \left[(X_G(1 - \frac{P}{P+N}) - \frac{P}{P+N} Z)^2 \right] \\ &= E \left[(X_G(1 - \frac{P}{P+N}))^2 \right] + E \left[(\frac{P}{P+N} Z)^2 \right] \\ &= P(\frac{N}{P+N})^2 + N(\frac{P}{P+N})^2 \\ &= \frac{PN}{P+N}. \end{aligned} \tag{4}$$

Now placing (4) into (3) we obtain:

$$I(X_G; Y) \geq \frac{1}{2} \log \left(\frac{P(P+N)}{PN} \right) = \frac{1}{2} \log \left(1 + \frac{P}{N} \right).$$

■

II. NEAREST-NEIGHBOR DECODING FOR ADDITIVE NON-GAUSSIAN NOISE CHANNELS[1].

In this section we will consider a channel with noise that is not necessarily a Gaussian noise, for this channel we will use nearest neighbor decoding- the decoder receives Y^n and finds message m such that $\|Y^n - X^n(m)\| = \sqrt{\sum_{i=1}^n (Y_i - X_i(m))^2}$ is minimal. The importance of mutual information comes from the claim that the capacity of a channel is the maximum mutual information, but this is only true when we know the model of the noise. We will show that we can achieve the rate $\frac{1}{2} \log(1 + \frac{P}{N})$ even when the noise is unknown:

Theorem 1 [1, Thorem1] For a single user additive noise channel with nearest neighbor decoding we have that, irrespective of the noise distribution, the average probability of error, averaged over the ensemble of Gaussian codebooks of power P , approaches zero as the blocklength n tends to infinity when the rate is below $\frac{1}{2} \log(1 + \frac{P}{N})$.

Proof:

lets assume we are using Gaussian encoding meaning $X = X_G$, we will build the system assuming the noise is Gaussian and the decoder will find the message according to nearest neighbor. We will show that using this system we can achieve the rate $\frac{1}{2} \log(1 + \frac{P}{N})$.

Lets assume that the message is $m = 1$ so $Y^n = X_G^n(1) + Z^n$, note that with high probability $X_G^n(1)$ and Z^n will have no correlation so:

- $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_G^n(i) Z(i) = 0$
- $\frac{1}{n} \|Y^n\|^2 \approx P + N \Rightarrow \frac{1}{n} \|Y^n - X_G^n(1)\|^2 \approx N$

And the probability of an error will be:

$$P\{\exists m' \neq 1 : \frac{1}{n} \|Y^n - X_G^n(m')\|^2 < N\} = P\{\exists m' \neq 1 : \frac{1}{n} \|X_G^n(1) + Z^n - X_G^n(m')\|^2 < N\} \quad (5)$$

From [1, Thorem1,proof] we obtain $\lim_{n \rightarrow \infty} P\{\exists m' \neq 1 : \frac{1}{n} \|X_G^n(1) + Z^n - X_G^n(m')\|^2 < N\} = 0$ ■

We can also prove that for a channel with two users $Y = X_1 + X_2 + Z$ where Z is any noise with variance N and the power constraints on the inputs are P_1 and P_2 than we can achieve:

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

By using a Gaussian encoder and a nearest neighbor decoder similar to case of the single user we saw in this lecture, see [1, part v].

REFERENCES

- [1] A. Lapidoth Nearest-neighbor decoding for additive non-Gaussian noise channels IEEE Transactions on Information Theory, vol. IT-42, pp. 1520-1529, Sept. 1996.