LECTURE 19

Last time:

- Gaussian channels: parallel
- colored noise
- inter-symbol interference
- general case: multiple inputs and outputs

Lecture outline

- Gaussian channels with feedback
- Upper bound to benefit of capacity

Reading: Section 10.6.

Gaussian channels with feedback

In the case of a DMC that there is no benefit to feedback

The same arguments extend to the case where we have continuous inputs and outputs

What happens in the case when the noise is not white? We can garner information about future noise from past noise

$$Y_i = X_i + N_i$$

but now the X_i is also a function of the past Ys, within an energy per codeword constraint

Gaussian channels with feedback

A code is now a mapping $x_i(M, \underline{Y}^{i-1})$ from the messages in $\mathcal{M} = \{1, 2, \ldots, 2^{nR}\}$ and from $\underline{Y}^{i-1}$ onto reals under the constraint

$$E_{\underline{N}^n}\left[\frac{1}{n}\sum_{i=1}^{n} x_i(m, \underline{Y}^{i-1})\right] \leq \mathcal{P}$$

$$\forall m \in \{1, 2, \ldots, 2^{nR}\}$$

How do we define capacity? Let's try:

$$C_{n,FB} = \max_{\frac{1}{n}trace\left(\Lambda_{\underline{X}^n}\right) \leq \mathcal{P}} \left(\frac{1}{n} I\left(\underline{X}^n; \underline{Y}^n\right)\right)$$

moreover

$$\begin{aligned} & I\left(\underline{X}^n; \underline{Y}^n\right) \\ = \;& h(\underline{Y}^n) - h(\underline{Y}^n|\underline{X}^n) \\ = \;& h(\underline{X}^n) - h(\underline{X}^n|\underline{Y}^n) \end{aligned}$$

but then select $(X_1, X_2, \ldots, X_n) = (0, N_1, \ldots, N_{n-1})$

the mutual information blows up!

Gaussian channels with feedback

Let's try:

$$C_{n,FB} = \max_{\frac{1}{n} trace\left(\Lambda_{\underline{X}^n}\right) \leq \mathcal{P}} \left(\frac{1}{n} I\left(M; \underline{Y}^n\right) \right)$$

Note: in the case of no feedback, then M and $\underline{X}^n$ are equivalent

$$\begin{aligned}
& I\left(M; \underline{Y}^n\right) \\
= \; & h(\underline{Y}^n) - h(\underline{Y}^n | M) \\
= \; & h(\underline{Y}^n) - \sum_{i=1}^{n} h(Y_i | M, \underline{Y}^{i-1}) \\
= \; & h(\underline{Y}^n) - \sum_{i=1}^{n} h(Y_i | M, \underline{Y}^{i-1}, \underline{X}^i) \\
= \; & h(\underline{Y}^n) - \sum_{i=1}^{n} h(Y_i | M, \underline{Y}^{i-1}, \underline{X}^i, \underline{N}^{i-1}) \\
= \; & h(\underline{Y}^n) - \sum_{i=1}^{n} h(Y_i | X_i, \underline{N}^{i-1}) \\
= \; & h(\underline{Y}^n) - \sum_{i=1}^{n} h(N_i | \underline{N}^{i-1}) \\
= \; & h(\underline{Y}^n) - h(\underline{N}^n)
\end{aligned}$$

Gaussian channels with feedback

How do we maximize $I(M;\underline{Y}^n)$, or equivalently $h(\underline{Y}^n) - h(\underline{N}^n)$

Since a Gaussian distribution maximizes entropy,

$$h(\underline{Y}^n) \leq \frac{1}{2}\ln\left((2\pi e)^n|\Lambda_{\underline{X}^n+\underline{N}^n|}\right)$$

we can always achieve this by taking the Xs to be jointly Gaussian with the past Ys

$$X_i = \sum_{j=1}^{i-1} \alpha_{i,j} Y_j + V_i + c_i$$

where V_i is mutually independent from the Y_js, for $1 \leq j \leq i-1$ and any constant c_i will leave the autocorrelation matrix unchanged. Note that the past Xs are a constant, so in particular we can select $c_i = -\sum_{j=1}^{i-1} \alpha_{i,j} x_j$

so

$$X_i = \sum_{j=1}^{i-1} \alpha_{i,j} N_j + V_i$$

Gaussian channels with feedback

Do we have coding theorems?

Joint typicality between input and output hold as a means of decoding

WLLN of large numbers holds

Sparsity argument for having multiple identical mappings holds

Converse: Fano's lemma still holds, with M being directly involved in the bound

Question: how does this compare to the non-feedback capacity?

Gaussian channels with feedback

Non-feedback capacity is simply Gaussian colored noise channel:

$$C_n = \max_{\frac{1}{n} trace\left(\Lambda_{\underline{X}^n}\right) \leq \mathcal{P}} \left(\frac{1}{n} I\left(\underline{X}^n; \underline{Y}^n\right)\right)$$

In this case

$$\begin{aligned} & I\left(\underline{X}^n; \underline{Y}^n\right) \\ = \ & h(\underline{Y}^n) - h(\underline{Y}^n | \underline{X}^n) \\ = \ & h(\underline{X}^n + \underline{N}^n) - h(\underline{N}^n) \end{aligned}$$

which is maximized by taking $\underline{X}^n$ to be Gaussian colored noise determined using water-filling

so $C_n = \max_{\frac{1}{n} trace\left(\Lambda_{\underline{X}^n}\right) \leq \mathcal{P}} \left(\frac{1}{2n} \ln\left(\frac{|\Lambda_{\underline{X}^n} + \Lambda_{\underline{N}^n}|}{|\Lambda_{\underline{N}^n}|}\right)\right)$

From our previous discussion,

$$C_{n,FB} = \frac{1}{2n} \ln\left(\frac{|\Lambda_{\underline{X}^n + \underline{N}^n}|}{|\Lambda_{\underline{N}^n}|}\right)$$

we can find this if we determine the $\alpha_{i,j}$s, but this may not be easy

An upper bound

Fact 1:

$$\Lambda_{\underline{X}^n+\underline{N}^n} + \Lambda_{\underline{X}^n-\underline{N}^n} = 2\left(\Lambda_{\underline{X}^n} + \Lambda_{\underline{N}^n}\right)$$

Look at elements in the diagonal and the off-diagonals

Fact 2:

If $C = A - B$ is symmetric positive definite, when A and B are also symmetric positive definite, then $|A| \geq |B|$

Consider $V \sim \mathcal{N}(0, C), W \sim \mathcal{N}(0, B)$ independent random variables

Let $S = V + W$, then $S \sim \mathcal{N}(0, A)$

$h(S) \geq h(S|V) = h(W|V) = h(W)$ so $|A| \geq |B|$

An upper bound

From fact 1:

$$2(\Lambda_{\underline{X}^n} + \Lambda_{\underline{N}^n}) - \Lambda_{\underline{X}^n+\underline{N}^n} = \Lambda_{\underline{X}^n-\underline{N}^n}$$

hence $2(\Lambda_{\underline{X}^n} + \Lambda_{\underline{N}^n}) - \Lambda_{\underline{X}^n+\underline{N}^n}$ is positive definite

From fact 2:

$$|\Lambda_{\underline{X}^n+\underline{N}^n}| \leq |2(\Lambda_{\underline{X}^n} + \Lambda_{\underline{N}^n})| = 2^n|(\Lambda_{\underline{X}^n} + \Lambda_{\underline{N}^n})|$$

Hence

$$\begin{aligned} C_{n,FB} = & \max_{\frac{1}{n}trace\left(\Lambda_{\underline{X}^n}\right)\leq\mathcal{P}} \left(\frac{1}{2n}\ln\left(\frac{|\Lambda_{\underline{X}^n+\underline{N}^n}|}{|\Lambda_{\underline{N}^n}|}\right)\right) \\ \leq & \max_{\frac{1}{n}trace\left(\Lambda_{\underline{X}^n}\right)\leq\mathcal{P}} \left(\frac{1}{2n}\ln\left(2^n\frac{|\Lambda_{\underline{X}^n} + \Lambda_{\underline{N}^n}|}{|\Lambda_{\underline{N}^n}|}\right)\right) \\ = & C_n + \frac{ln(2)}{2} \end{aligned}$$

Writing on dirty paper

Suppose that the sender knows the degradation d exactly, what should he do? What should the receiver do?

May not always be able to subtract d at the sender.

Example: we try to send S uniformly distributed over $[-1, 1]$

select X such that $(X + d)$ mod 2 $= S$

$X = S - d$ mod 2 and the receiver takes mod 2