LECTURE 22

Last time:

- Broadcast channel
- Gaussian degraded broadcast channel

Lecture outline

- Finite-state channels
- Lower capacity and upper capacities
- Indecomposable channels
- Markov channels

Finite-state channels

The state of the channel at time i is S_i

The state determines the transition probability in the following way:

$$P_{\underline{Y}^n, S_n | \underline{X}^n, S_0}\left(\underline{y}^n, s_n | \underline{x}^n, s_0\right)$$

we assume that the channel is stationary

There is memory that is different than in the case of ISI channel - memory in state rather than in input

The transition probabilities are determined recursively from the one-step transition probabilities:

$$P_{\underline{Y}^n, S_n | \underline{X}^n, S_0}\left(\underline{y}^n, s_n | \underline{x}^n, s_0\right)$$

$$= \sum_{s_{n-1}} P_{Y_n, S_n | X_n, S_{n-1}}\left(y_n, s_n | x_n, s_{n-1}\right)$$

$$P_{\underline{Y}^{n-1}, S_{n-1} | \underline{X}^{n-1}, S_0}\left(\underline{y}^{n-1}, s_{n-1} | \underline{x}^{n-1}, s_0\right)$$

Upper and lower capacities

$P_{\underline{Y}^n|\underline{X}^n,S_0}\left(\underline{y}^n, s_n|\underline{x}^n, s_0\right)$

$= \sum_{s_n} P_{\underline{Y}^n,S_n|\underline{X}^n,S_0}\left(\underline{y}^n, s_n|\underline{x}^n, s_0\right)$

Upper capacity:

$\overline{C} = \lim_{n\to\infty} \overline{C}^n$

where

$\overline{C}^n = \frac{1}{n} \max_{P_{\underline{X}^n}(\underline{x}^n)} \max_{s_0} I\left(\underline{X}^n; \underline{Y}^n|s_0\right)$

Lower capacity:

$\underline{C} = \lim_{n\to\infty} \underline{C}^n$

where

$\underline{C}^n = \frac{1}{n} \max_{P_{\underline{X}^n}(\underline{x}^n)} \min_{s_0} I\left(\underline{X}^n; \underline{Y}^n|s_0\right)$

Neither is the "capacity"

Indecomposable channels

Suppose that we have finite input alphabet $\mathcal{X}$ and a finite set of states $\mathcal{S}$

Indecomposable means that the initial state does not matter $\forall \epsilon > 0, \exists n_0 s.t. \forall n > n_0$

$|P_{S_n|\underline{X}^n, S_0}(s_n|\underline{x}^n, s_0) - P_{S_n|\underline{X}^n, S_0}(s_n|\underline{x}^n, s_0')| \leq \epsilon$

For such channels: $\underline{C} = \overline{C}$

Proof: pick some integer number μ

let s_0' be an initial state such that the maximum average mutual information over n samples starting from s_0' yields $\overline{C}^n$

let s_0'' be an initial state such that the maximum average mutual information over n samples starting from s_0'' yields $\underline{C}^n$

Indecomposable channels

Applying twice the chain rule on mutual information, we obtain (with the input distribution that yields $\overline{C}^n$)

$$
\begin{aligned}
& \overline{C}^n \\
= \ & \frac{1}{n}\Big[I\left(\underline{X}^{\mu};\underline{Y}^{n}|S_0=s_0'\right) \\
& +I\left(\underline{X}_{\mu+1}^{n};\underline{Y}^{\mu}|S_0=s_0',\underline{X}^{\mu}\right) \\
& +I\left(\underline{X}_{\mu+1}^{n};\underline{Y}_{\mu+1}^{n}|S_0=s_0',\underline{X}^{\mu},\underline{Y}^{\mu}\right)\Big]
\end{aligned}
$$

let us bound each element in the sum:

the first term is upper bounded by $\mu \log(|\mathcal{X}|)$

the second term is 0

the third term, using the chain rule on mutual information and the cardinality bound on the information about S_n, can be upper bounded by:

$$\log(|\mathcal{S}|)+I\left(\underline{X}_{\mu+1}^{n};\underline{Y}_{\mu+1}^{n}|S_0=s_0',\underline{X}^{\mu},\underline{Y}^{\mu},S_n\right)$$

Indecomposable channels

Similarly, we obtain the bound:

$$\begin{aligned}&\underline{C}^n \geq \frac{1}{n}\left[-\log(|\mathcal{S}|)\right.\\&\left.+I\left(\underline{X}_{\mu+1}^n;\underline{Y}_{\mu+1}^n|S_0=s_0'',\underline{X}^\mu,\underline{Y}^\mu,S_n\right)\right]\end{aligned}$$

then

$$\begin{aligned}&\overline{C}^n-\underline{C}^n\\=&\ \frac{1}{n}\left[\mu\log(|\mathcal{X}|)+2\log(|\mathcal{S}|)\right.\\&+I\left(\underline{X}_{\mu+1}^n;\underline{Y}_{\mu+1}^n|S_0=s_0',\underline{X}^\mu,\underline{Y}^\mu,S_n\right)\\&\left.-I\left(\underline{X}_{\mu+1}^n;\underline{Y}_{\mu+1}^n|S_0=s_0'',\underline{X}^\mu,\underline{Y}^\mu,S_n\right)\right]\\=&\ \frac{1}{n}\left[\mu\log(|\mathcal{X}|)+2\log(|\mathcal{S}|)\right.\\&+I\left(\underline{X}_{\mu+1}^n;\underline{Y}_{\mu+1}^n|S_0=s_0',\underline{X}^\mu,S_n\right)\\&\left.-I\left(\underline{X}_{\mu+1}^n;\underline{Y}_{\mu+1}^n|S_0=s_0'',\underline{X}^\mu,S_n\right)\right]\\\leq&\ \frac{1}{n}\left[\mu\log(|\mathcal{X}|)+2\log(|\mathcal{S}|)\right.\\&\left.+\epsilon(n-\mu)\log(|\mathcal{X}|)\right]\end{aligned}$$

where ϵ can be made arbitrarily small by appropriately large choice of μ (which entails sufficiently large n)

Markov channels

A special case of indecomposable channels arise when we have a Markov description for the channel states, the states are independent of the inputs conditioned on other states

Assume the states form a single class of recurrent states, aperiodic

Then we have indecomposable channel (see lecture 4)

How might we establish a coding theorem for such channels with finite input alphabet and no other constraint on input?

Markov channels

Suppose the current state is known at the sender and the receiver

For each state, establish capacity as though that were the only state, let C_s be the capacity if we were to remain in state s always

Average over all the states, the capacity is:

$\sum_{s \in \mathcal{S}} \pi_s C_s$

The coding theorem relies on interleaving codewords over the different states

Pick one codebook per state, send portions of the codeword in a codebook only when that state occurs

Because of recurrence, each state is always guaranteed to reappear eventually

The average transmission rate is the average of the transmissions over all the codebooks

Markov channels

If receiver and sender both have Markovian channel side information (CSI) and sender is a function of receiver CSI, then this concept extends

Example: delayed knowledge at the sender of state of channel (because of delays in feedback, for example)

Let U_i be the sender CSI (SCSI), which is a deterministic function g of the perfect receiver CSI (RCSI) V_i (thus, the receiver knows the channel exactly $V_i = S_i$) such that U_i remains Markovian

Then the capacity is given by (using stationarity to eliminate time subscripts):

$C = \sum_{u \in \mathcal{U}} P_U(u) \max_{P_{X|U=u}} I(X;Y|S,U = u)$

Markov channels

Other cases:

- Known at the receiver but not the sender: then the computation is more involved, but the theorem for indecomposable channels still holds. Sender does not change strategy.

- Known at the sender but not the receiver: strategies for the receiver to estimate the channel are important.

- Channel statistics known but channel not known: sender does not change strategy. Hypothesis testing at receiver. Recall Gilbert-Eliot channel. More states: different approach is required.

Markov channels

What happens when we have for instance different types of AWGN channels?

The theorems need to be revisited. The finiteness of states is still an important aspect.

Suppose we have two channels, one (channel 1) with noise psd N_0 and the other(channel 2) with noise psd $N_0' > N_0$

Perfect CSI at the sender and the receiver

Does water-filling work?

With a modification: make several channels with noise energy WN_0 and several channels with noise energy WN_0' in proportion $\pi_1 : \pi_2$

Then do water-filling on these channels

This is *power control*

Markov channels

What happens when we do not have perfect CSI at sender and receiver?

Consider the case of AWGN channels with different states where the channel state represents an amplification factor (this view is clearly equivalent to changing the psd of the AWGN)

We have that:

$$Y = \sqrt{S}X + N$$

when channel S is in force

The receiver has perfect CSI

The sender has some function of S

We then have that:

$$C = \max_{\gamma} E\left[\frac{1}{2}\ln\left(1 + \frac{S\gamma(U)}{\sigma_N^2}\right)\right]$$

γ is a mapping from U to $\mathbf{R}^+$ such that $E\left[\gamma(U)\right] \leq \mathcal{P}$

Markov channels

How can we take ISI into account?

Block fading model: within each block we have a particular ISI profile

Transmit optimal scheme for that ISI profile

Problem: what happens at junction of channels?

For underspread channels, coherence time is much longer than delay spread

Time in a channel state is of the order of a coherence time

Time during which "echo" from previous state persists is small with respect to coherence time

Markov channels

How can we take ISI into account?

Perform waterfilling in the following way:

- decompose along a basis for all channels in the set of possible channels
- replicate channels to have channels repeated in proportion to their frequency of occurrence
- perform water-filling

What if the block approach does not work? Use partial codebook approach.

Markov channels

How about multiple access?

When the CSI is perfect at the senders and the receiver, simply consider that the channel state is the cartesian product of the channel states for all the users (note: the users must know each other's channel)

The capacity region for multiple-access channels and the coding theorem arguments show that the arguments for the single user case extend well

Take expectation over all states of all users

$R_1 + R_2 \leq$

$E\left[\max_{P_{X^1|S^1,S^2}, P_{X^2|S^1,S^2}}((X^1, X^2); Y|S^1, S^2)\right]$

$R_2 \leq \left[\max_{P_{X^2|S^2}}(X^2; Y|X^1, S^1, S^2)\right]$

$R_1 \leq \left[\max_{P_{X^1|S^1}}(X^1; Y|X^2, S^1, S^2)\right]$