LECTURE 14

Last time:

- Sphere packing and Sphere covering
- Joint Source-Channel Coding
- Achievable performance

Lecture outline

- Converse of Source-Channel coding theorem.
- Channel with Feedback.

Review

- Source coding: reduce the redundancy in the source, sphere covering.
- Channel coding: add redundancy to ensure reliable communication, sphere packing.
- Joint Source Channel Coding

$$\underline{V}^n \to \underline{X}^n \to \underline{Y}^n \to \widehat{\underline{V}}^n$$

Goal: find a map $\underline{V}^n \to \underline{X}^n$, such that

$$P(\underline{V}^n \neq \widehat{\underline{V}}^n) \to 0$$

- As long as $H(V) < C$, find an R such that

$$H(V) < R < C$$

and maps

$$\underline{V}^n \to \underline{U}^* \to \underline{X}^n$$

Source-Channel coding Theorem

Theorem (simple version)

For an i.i.d. Source $\underline{V}^n$ and the DMC $P_{Y|X}$, as long as $H(V) < C$, there exists a souce channel code with $P_e^{(n)} \to 0$.

Conversely, if $H(V) > C$, the probability of error is bounded away from 0. It is not possible to send the source over the channel with arbitrarily low probability of error.

Proof of the converse

Fano's inequality:

$$H(\underline{V}^n|\underline{\widehat{V}}^n) \leq 1 + P_e^{(n)} n \log |\mathcal{V}|$$

Now

$$\begin{aligned}
H(V) &= \frac{1}{n} H(\underline{V}^n) \\
&= \frac{1}{n} H(\underline{V}^n|\underline{\widehat{V}}^n) + \frac{1}{n} I(\underline{V}^n; \underline{\widehat{V}}^n) \\
\text{Fano's} \quad &\leq \frac{1}{n}(1 + P_e^{(n)} n \log |\mathcal{V}|) + \frac{1}{n} I(\underline{V}^n; \underline{\widehat{V}}^n) \\
&\leq \frac{1}{n}(1 + P_e^{(n)} n \log |\mathcal{V}|) + \frac{1}{n} I(\underline{X}^n; \underline{Y}^n) \\
&\leq \frac{1}{n} + P_e^{(n)} \log |\mathcal{V}| + C
\end{aligned}$$

$$P_e^{(n)} \to 0 \Rightarrow H(V) \leq C$$

Generalize the Theorem

- what if the source has memory (random process, not i.i.d.)

Replace $H(V)$ by the entropy rate $H(\mathcal{V})$

Recall entropy rate

$$H(\mathcal{V}) = \lim_{n \to \infty} \frac{1}{n} H(\underline{V}^n)$$

if the limit exists.

- what if the channel is also varying over the time? Replace the capacity by

$$C = \lim_{n \to \infty} \frac{1}{n} I(\underline{X}^n; \underline{Y}^n)$$

When Separation Theorem does NOT hold

- Channel is correlated with the source
- multiple access with correlated sources
- non-degraded broadcasting channel
- The limit in the entropy rate and capacity of time-varying channel does not exist.
- ...

Example

Question Even if the separation theorem holds, does this mean we would always use separation-based source-channel codes?

Channel with Feedback

- Channel with Feedback: does feedback help us to transmit higher data rate?
- Perfect feedback: assume all previously received signals $\underline{Y}_1^{i-1}$ is available at the transmitter at time i.
  - why is causality in feedback important?
  - why do we ignore the possible errors in the feedback link?

Feedback Capacity

- Feedback code: a $(2^{nR}, n)$ feedback code is a sequence of mapping:

$$X_i = X_i(W, \underline{Y}^{i-1}), i = 1, \ldots n$$

a decoder recover W as $\hat{W} = g(\underline{Y}^n)$.

- A rate is achievable if there exists a sequence of $(2^{nR}, n)$ codes such that the probability of error

$$P_e^{(n)} = P(\hat{W} \neq W) \to 0$$

- The capacity with feedback C_{FB} is the supremum of the rates achievable with feedback codes.

Obviously:

$$C_{FB} \geq C$$

Does feedback increase the capacity?

Theorem For a DMC channel

$$C_{FB} = C = \max_{P_X} I(X;Y)$$

Feedback does NOT improve the capacity at all.

Repeat the converse of the coding theorem without feedback:

$$\begin{aligned} nR &= H(W) = H(W|\underline{Y}^n) + I(W;\underline{Y}^n) \\ &\leq H(W|\underline{Y}^n) + I(\underline{X}^n;\underline{Y}^n) \\ &\leq 1 + P_e^{(n)} nR + I(\underline{X}^n;\underline{Y}^n) \\ &\leq 1 + P_e^{(n)} nR + nC \end{aligned}$$

What is wrong?

$$\begin{aligned} I(\underline{X}^n;\underline{Y}^n) &= H(\underline{Y}^n) - H(\underline{Y}^n|\underline{X}^n) \\ &= H(\underline{Y}^n) - \sum_{i=1}^{n} H(Y_i|\underline{Y}_1^{i-1}, \underline{X}^n) \\ &= H(\underline{Y}^n) - \sum_{i=1}^{n} H(Y_i|X_i) \end{aligned}$$

with feedback, the last step does not hold since Y_i depends on the future X's.

Proof of the converse

$$
\begin{aligned}
I(W;\underline{Y}^n) &= H(\underline{Y}^n) - H(\underline{Y}^n|W) \\
&= H(\underline{Y}^n) - \sum_{i=1}^{n} H(Y_i|\underline{Y}_1^{i-1}, W) \\
(*) \quad &= H(\underline{Y}^n) - \sum_{i=1}^{n} H(Y_i|\underline{Y}_1^{i-1}, W, X_i) \\
&= H(\underline{Y}^n) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&\leq \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i|X_i) \\
&\leq nC
\end{aligned}
$$

The trick is in $(*)$, I only further condition on X_i, instead of $\underline{X}^n$.

Example: BEC with Feedback

- Recall the capacity of BEC is $C = 1 - \epsilon$.
- Consider a binary erasure channel with perfect feedback. Channel feedback reveals if a transmission is successful. Design a feedback code as:
  - transmit a new bit if the last transmission is successful.
  - retransmit if it failed.
- The average number of channel uses is

$$1 + \epsilon + \epsilon^2 + \ldots = \frac{1}{1 - \epsilon}$$

- Average number of bits transmitted per channel use is $1 - \epsilon$.
- The capacity is not improved at all, but the code is greatly simplified.

Discussions

Question Is this the only reason that feedback is used?

- DMC does not benefit from feedback.
- General channels with memory
  - wireless channels.
  - network protocols.