LECTURE 22

Last time:

- Gaussian multiple access channel
- Discussions on X-DMA

Lecture outline

- Broadcast channel
- Capacity region for degraded broadcast channel
- Distributed Source Coding,
- Slepian-Wolf Theorem and Random binning

Review

- Multiple Access channels
  - Time Sharing
  - Successive Cancellation
- Capacity Region

$$\begin{aligned} R_1 &\leq I(X_1;Y|X_2) \\ R_2 &\leq I(X_2;Y|X_1) \\ R_1 + R_2 &\leq I(X_1,X_2;Y) \end{aligned}$$

- Dividing users into subspaces
  - In general suffers from a capacity loss
  - Optimal bandwidth allocation in Gaussian channel $W_i \propto P_i$
  - non-orthogonal signatures: MMSE and sufficient statistics

Broadcast Channel

One sender, many receivers.

Examples TV, radio, lecture, downlink from the base station, etc.

- The information to different receivers can be divided into separate sub-channels, depends on the protocol.
- Common information to different receivers: less than the worst channel.
- Independent information streams sharing the same medium

Definition A broadcast channel $(\mathcal{X}, P_{Y_1,Y_2|X}, \mathcal{Y}_1, \mathcal{Y}_2)$.

Memoryless:

$$p(\underline{y}_1, \underline{y}_2|\underline{x}) = \prod_i p(y_{1,i}, y_{2,i}|x_i)$$

A $(2^{nR_1}, 2^{nR_2}, n)$ code is

$$2^{nR_1} \times 2^{nR_2} \to \mathcal{X}^n$$

embed two messages in one codeword.

Separation Theorem

- Common information (R_0, R_1, R_2).
- Independently compress the incoming data by each user may not be optimal.
- Compare with multiple access channel with partial information of the interferer.
- In general, the capacity region is not known

Definition Degraded Broadcast channel

$$p(y_1, y_2|x) = p(y_1|x)p(y_2|y_1)$$

Example Gaussian broadcast channel

$$\begin{aligned} Y_1 &= X + W_1 \\ Y_2 &= X + W_2 \end{aligned}$$

assume $\sigma_1^2 < \sigma_2^2$.

Observation The capacity region only depends on the conditional marginal distributions $p(y_1|x), p(y_2|x)$.

Rewrite

$$\begin{aligned} Y_1 &= X + W_1 \\ Y_2 &= Y_1 + W_2' \end{aligned}$$

Degraded Broadcast Channel

Theorem The capacity region of sending independent information over the degraded BC $X \to Y_1 \to Y_2$ is the convex hull of the closure of all (R_1, R_2) with

$$\begin{aligned} R_2 &\leq I(U; Y_2) \\ R_1 &\leq I(X; Y_1 | U) \end{aligned}$$

for some auxiliary random variable U jointly distributed with X.

- Auxiliary random variable U: the message sent to user 2.
- Transmit 2^{nR_2} clouds, each containing 2^{nR_1} codewords.
- U defines the cloud centers that can be observed by both user 1 and 2; while user 1 as the better receiver can distinguish individual codewords within the clouds.

Outline of the Coding Theorem

- Encoding:
  - generate 2^{nR_2} codewords $\underline{U}(m_2)$ according to $p(u)$.
  - For each codeword $\underline{U}(m_2)$, generate 2^{nR_1} codewords $\underline{X}(m_1, m_2)$ according to $p(x|u)$ (conditioning on the value $u_i(m_2)$).
- Transmit $\underline{X}(m_1, m_2)$.
- Decoding:
  - Receiver 2 find $\underline{U}$ that is jointly typical with $\underline{Y}_2$.
  - Receiver 1 does the same, then find the $\underline{X}$ that is jointly typical with $\underline{U}$ and $\underline{Y}_1$.

Proof by AEP read the book

key difference from before: user 1 can have two different types of error, decoding the cloud center wrong and decoding the codeword wrong given the correct cloud center.

Example: Gaussian Channels

$$\begin{aligned} Y_1 &= X + W_1 \\ Y_2 &= X + W_2 \overset{d}{=} X + W_1 + W_2' \end{aligned}$$

Let the noise power be N_1, N_2 and power constraint on X be P, W_2' has power $N_2 - N_1$.

Capacity region

$$\begin{aligned} R_1 &\leq C\left(\frac{\alpha P}{N_1}\right) \\ R_2 &\leq C\left(\frac{(1-\alpha)P}{N_2 + \alpha P_1}\right) \end{aligned}$$

- Generate 2^{nR_2} codewords i.i.d. from $N(0, (1-\alpha)P)$, denote as $\underline{u}_j$,
- For each of these codewords, generate 2^{nR_1} satellite codewords i.i.d. from $N(0, \alpha P)$, denote as $\underline{v}_i$,
- Transmit $\underline{v}_i + \underline{u}_j$

Example: BSC

X passes through two binary symmetric channels with crossover probability $p_1 < p_2 < 1/2$.

- Rewrite in the form of a degraded channel: Y_1 pass through another BSC to get Y_2, the crossover probability is α with
$$p_1(1-\alpha)+(1-p_1)\alpha = p_2$$
- The input U (to be transmitted to user 2) must be binary, equiprobable, but jointly distributed with X. Can be thought as U passing through a BSC to get X.

$$\begin{aligned} R_2 &\leq I(U;Y_2) = 1 - H(\beta * p_2) \\ R_1 &\leq I(X;Y_1|U) = H(Y_1|U) - H(Y_1|X,U) \\ &= H(\beta * p_1) - H(p_1) \end{aligned}$$

- $\beta \to 0$, only transmitting information for Y_2, $R_1 = 0, R_2 = 1 - H(p_2)$.
- $\beta \to 1/2$, only transmitting information to Y_1, $R_2 = 0, R_1 = 1 - H(p_1)$.

Distributed Source Coding

Motivation Want to represent two correlated sources efficiently. Dual problem to the general broadcast channel

- Recall for coding for single source. X can be represented by $H(X)$ bits per symbol.
- Alternative proof: random binning.
  - lay down 2^{nR} bins
  - Randomly throw each possible sequence of $\underline{X}$ into one of these pins.
  - Use the bin number to represent the source, i.e., for a typical sequence $\underline{x} \in A_\epsilon$, use the bin number as the codeword
  - Observing the bin number, decode the typical sequence within that bin. Error occurs only if more than 1 typical sequence is thrown into the same bin.
  - If the number of bins $\ll$ the number of typical sequences, error probability $\rightarrow 0$.

Slepian-Wolf

Consider U, V as two correlated random variables. The source generates a sequence of (u_i, v_i) i.i.d. according the the joint distribution $P_{U,V}$.

How many bits about U and V separately is required so that the combined description will recover $\underline{U}, \underline{V}$ with small probability of error.

- If we can observe both U and V, then we need in total $nH(U, V)$ bits.
- If we code for U and V separately, we need $nH(U) + nH(V)$ bits in total.

Theorem There exists coding map $\mathcal{U}^n \rightarrow 2^{nR_1}, \mathcal{V}^n \rightarrow 2^{nR_2}$, with vanishing error probability, as long as

$$\begin{aligned} R_1 &> H(U|V) \\ R_2 &> H(V|U) \\ R_1 + R_2 &> H(U, V) \end{aligned}$$

We don't lose any rate by coding U and V separately.

Intuition

Can we achieve the rate pair

$$R_1 = H(U), R_2 = H(V|U)$$

- with $R_1 = H(U)$, $\underline{u}$ can be fully described.
- if encoder 2 observes $\underline{u}$, she can have a code of rate R_2 to describe all the sequences $\underline{v}$ that are jointly typical with this $\underline{u}$.
- Now she doesn't observe $\underline{u}$, so she assigns all possible $\underline{v}$'s into 2^{nR_2} bins, and inform this assignment to the decoder.
- The decoder decode $\underline{u}$ first, and find out the jointly typical set $A_\epsilon(\underline{V}|\underline{u})$.
- These jointly typical sequences $\underline{v} \in A_\epsilon(\underline{V}|\underline{u})$ are assigned into 2^{nR_2} bins, so can be distinguished.

Conclusion Encoder 2 does not need to observe $\underline{u}$.

Outline of the Proof

- Random code:
  - randomly assign each possible $\underline{u}$ into one of 2^{nR_1} bins, and each $\underline{v}$ into one of 2^{nR_2} bins.
  - Upon observing a typical sequence $\underline{U}$, encoder 1 send the corresponding bin number with nR_1 bits; encoder 2 does the same for $\underline{V}$,
- Decoding, look into the product of the two specified bins,decode if there exists unique jointly typical $(\underline{U}, \underline{V})$

Idea form $2^{nR_1} \times 2^{nR_2}$ product partition in the set $\mathcal{U}^n \times \mathcal{V}^n$, if the partition is fine enough, the chance that two typical sequences lie in the same bin is small.

Achievable rate region

Back to Broadcast Channel

Consider a deterministic broadcast channel with $y_1 = f(x)$, $y_2 = g(x)$.

Fix an input distribution p_X,

- the distribution p_{Y_1,Y_2} is then fixed.
- any typical sequence $\underline{x}$ corresponds to a jointly typical sequence $(\underline{y}_1, \underline{y}_2)$, and vice versa.

Forget about Y_2 for the moment, consider a point-to-point channel from $X \to Y_1$. Can communicate with any rate $R_1 < H(Y_1)$.

- divide $\mathcal{Y}_1^n$ into 2^{nR_1} bins
- To transmit a message $M_1 \in \{1, \ldots, 2^{nR_1}\}$, find a typical sequence $\underline{y}_1$ in the corresponding bin, and hope $\underline{y}_1$ be received.
- This can be done if the number of bins $2^{nR_1} < |A_\epsilon^{(n)}(Y_1)|$.
- transmit the corresponding $\underline{x}$ that produces $\underline{y}_1$.

Capacity Region for Deterministic Broadcast Channel

Theorem Fix an input distribution, any rate pair (R_1, R_2) that satisfies

$$\begin{aligned} R_1 &< H(Y_1) \\ R_2 &< H(Y_2) \\ R_1 + R_2 &< H(Y_1, Y_2) \end{aligned}$$

can be achieved.

- Divide the output space $\mathcal{Y}_1^n \times \mathcal{Y}_2^n$ into $2^{nR_1} \times 2^{nR_2}$ bins,
- For an input message $M_1 \times M_2 \in \{1, \ldots, 2^{nR_1}\} \times \{1, \ldots, 2^{nR_2}\}$, find a typical sequence $\underline{y}_1 \times \underline{y}_2$ in the corresponding bin. This is the desired output
- Transmit $\underline{x}$ to produce $\underline{y}_1, \underline{y}_2$.
- $R_1 < H(Y_1)$ ensures that exists typical $\underline{y}_1$ per bin
- $R_1 + R_2 < H(Y_1, Y_2)$ ensures that exists a typical $\underline{y}_1, \underline{y}_2$ per product bin

General Broadcast Channel

Now if channel has randomness $P_{Y_1,Y_2|X}$.

- Introduce auxiliary random variables U, V.
- For fixed distribution of U, V, from the channel output, the number of distinguishable sequences $\underline{u}$ and $\underline{v}$ are $2^{nI(U;Y_1)}$ and $2^{nI(V;Y_2)}$.

Theorem(Marton 75') The broadcast channel capacity region is given by

$$\begin{aligned} R_1 &\leq I(U;Y_1) \\ R_2 &\leq I(V;Y_2) \\ R_1 + R_2 &\leq I(U;Y_1) + I(V;Y_2) - I(U;V) \end{aligned}$$

for a fixed distribution $P_{U,V,X}$.

Outline of Proof

- generate $2^{nI(U;Y_1)}$ typical sequence $\underline{u}$'s and throw into 2^{nR_1} bins, generate $2^{n(I(V;Y_2)}$ typical $\underline{v}$'s throw into 2^{nR_2} bins.
- upon receiving $\underline{Y}_1$, $\underline{u}$ can be uniquely determined. similar to $\underline{v}$.
- There are $2^{n(I(U;Y_1)+I(V;Y_2))}$ possible $(\underline{u},\underline{v})$ pairs. Each pair being jointly typical with probability $2^{-nI(U;V)}$.
- If $R_1+R_2 \leq I(U;Y_1)+I(V;Y_2)-I(U;V)$, then there exists a jointly typical pair $(\underline{u},\underline{v})$ in each product bin. This is the desired received sequences.
- To transmit that bin number, simply transmit $\underline{x}$ that is jointly typical with $(\underline{u},\underline{v})$.