EECS 290S: Network Information Flow **Fall 2008**

Lecture 19 — November 4

Lecturer: Anant Sahai and David Tse *Scribe: Sudeep Kamath*

19.1 Overview of the lecture

This lecture shall cover

- The Converse for the Gaussian Broadcast Channel
- Cutset bound for the Relay Channel

19.2 Recap of previous lecture

We had established that the capacity region for the degraded broadcast channel $X \to Y_1 \to Y_2$ is given by the convex hull of the closure of all rate pairs (R_1, R_2) satisfying

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

for some joint probability distribution $p(u)p(x|u)p(y_1, y_2|x)$ and where the auxiliary random variable U has cardinality bounded by $|\mathcal{U}| \leq \min\{|\mathcal{X}|, |\mathcal{Y}_1|, |\mathcal{Y}_2|\}$

19.3 The Gaussian Broadcast Channel

The Gaussian Broadcast channel is always degraded (stochastically if not physically). We can always represent it as:

$$\begin{aligned} Y_1 &= X + Z_1 \\ Y_2 &= Y_1 + Z_2 \end{aligned}$$

where X is the input, $Z_1 \sim \mathcal{N}(0, N_1)$ and $Z_2 \sim \mathcal{N}(0, N_2 - N_1)$ are the noises with X, Z_1 and Z_2 being mutually independent.

The cardinality bound over $|\mathcal{U}|$ does not hold when the input and output have continuous distributions. Recall that U was an auxiliary variable that represented the cloud center.

19.3.1 Achievability for the Gaussian Broadcast Channel

Consider the following encoding scheme. The encoder's codebook consists of two codebooks, one for the message of each user. The encoder uses Gaussian codebooks with powers αP and $(1-\alpha)P$ for users 1 and 2 respectively. Then, it adds up the individual codewords corresponding to the messages that users 1 and 2 need to communicate and transmits this sum over the channel. Let $X = U + V$ where $U \sim \mathcal{N}(0, (1-\alpha)P)$ and $V \sim \mathcal{N}(0, \alpha P)$ with U, V being independent.

The following rate pair can be achieved by this scheme:

$$\begin{aligned} R_1 &= I(X; Y_1|U) \\ &= \frac{1}{2}\log\left(1 + \frac{\alpha P}{N_1}\right) \\ R_2 &= I(U; Y_2) \\ &= \frac{1}{2}\log\left(1 + \frac{(1-\alpha)P}{\alpha P + N_2}\right) \end{aligned}$$

This achievability was established by Tom Cover and this scheme is called superposition coding. We keep this achievable scheme at the back of our minds while attempting to prove the converse.

19.3.2 Converse for the Gaussian Broadcast Channel

We want to show that if a rate pair (R_1, R_2) is achievable, then $\exists \alpha \in [0, 1]$ such that

$$\begin{aligned} R_1 &\leq \frac{1}{2}\log(1 + \frac{\alpha P}{N_1}) \\ R_2 &\leq \frac{1}{2}\log(1 + \frac{(1-\alpha)P}{\alpha P + N_2}) \end{aligned}$$

The following is the proof by Bergman.

Recall that if $W \sim \mathcal{N}(\mu, \sigma^2)$, then $h(W) = \frac{1}{2}\log(2\pi e \sigma^2)$.

$$\begin{aligned} I(X; Y_1|U) &= h(Y_1|U) - h(Y_1|X, U) \\ &\stackrel{a}{=} h(Y_1|U) - \frac{1}{2}\log(2\pi e N_1) \end{aligned}$$

where (a) follows because conditioned on X_1, U, the distribution of Y_1 is a Gaussian with variance N_1.

We are stuck at this point. So, we move to the other mutual information expression.

$$\begin{aligned} I(U; Y_2) &= h(Y_2) - h(Y_2|U) \\ &\stackrel{b}{\leq} \frac{1}{2}\log(2\pi e(P + N_2)) - h(Y_2|U) \\ &\stackrel{c}{=} \frac{1}{2}\log(2\pi e(P + N_2)) - h(Y_1 + Z_2|U) \end{aligned}$$

where (b) is true because $\text{Var}(Y_2) = P + N_2$ and the Gaussian distribution maximizes the differential entropy for a given variance. We do step (c) because we need a relation between $h(Y_1|U)$ and $h(Y_2|U)$ namely that for a given value of $h(Y_1|U)$, the value of $h(Y_2|U)$ is not too small.

The Entropy Power Inequality (EPI)

If A and B are independent continuous r.v.'s with probability densities, then

$$2^{2h(A+B)} \geq 2^{2h(A)} + 2^{2h(B)}$$

(Proof in Thomas and Cover's text on Information Theory)

Eg. If $A \sim \mathcal{N}(0, \sigma_A^2), B \sim \mathcal{N}(0, \sigma_B^2)$ and A, B are independent, then $A + B \sim \mathcal{N}(0, \sigma_A^2 + \sigma_B^2)$. We then, have $2^{2h(A)} = 2\pi e \sigma_A^2$, $2^{2h(B)} = 2\pi e \sigma_B^2$ and $2^{2h(A+B)} = 2\pi e(\sigma_A^2 + \sigma_B^2)$ Here, the EPI is met with equality.

$2^{2h(X)}$ is called the entropy power of X. If X is Gaussian, then the entropy power of X is proportional to its variance.

The Gaussian distribution turns out to be the solution to several interesting extremal optimization problems:

- For a given variance, the Gaussian distribution maximizes the differential entropy.
- Equality holds in EPI iff both A, B are Gaussian.

Now, let A, B be given r.v.'s. Let A', B' be such that $A' \sim \mathcal{N}(0, \sigma_1^2), B' \sim \mathcal{N}(0, \sigma_2^2)$ be such that $h(A') = h(A)$ and $h(B') = h(B)$. EPI says that

$$\begin{aligned} 2^{2h(A+B)} &\geq 2^{2h(A)} + 2^{2h(B)} \\ &= 2^{2h(A')} + 2^{2h(B')} \\ &= 2^{2h(A'+B')} \end{aligned}$$

Hence, EPI says that $h(A + B) \geq h(A' + B')$.

Thus, $\min_{h(A)=\alpha, h(B)=\beta} h(A+B) = h(A'+B')$ where A' and B' are Gaussian distributions such that $h(A') = \alpha, h(B') = \beta$.

Brunn-Minkowski Theorem

Let S_1, S_2 be sets in $\mathcal{R}^n$. The sum set of S_1, S_2 is defined as:

$$S_1 + S_2 = \{x + y : x \in S_1, y \in S_2\}$$

Then, $\min_{\text{Vol}(S_1)=V_1, \text{Vol}(S_2)=V_2} \text{Vol}(S_1+S_2) = \text{Vol}(S_1'+S_2')$ where $\text{Vol}(S_1') = V_1, \text{Vol}(S_2') = V_2$ and S_1' and S_2' are spheres.

The analogy here is that *Entropy : Gaussian distribution :: Volume : Sphere*

Now, we are ready to lower bound $h(Y_2|U)$.

$$\begin{aligned} h(Y_2|U) &= h(Y_1 + Z_2|U) \\ &\stackrel{d}{=} \sum_{u \in \mathcal{U}} p_U(u) h(Y_1 + Z_2|U = u) \\ &\stackrel{e}{\geq} \sum_{u \in \mathcal{U}} p_U(u) [\frac{1}{2} \log(2^{2h(Y_1|U=u)} + 2^{2h(Z_2)})] \\ &\stackrel{f}{\geq} \frac{1}{2} \log(2^{2h(Y_1|U)} + 2^{2h(Z_2)}) \end{aligned}$$

where in (d), Y_1 has a conditional distribution given $U = u$ while Z_2 is conditionally independent of $U = u$. (e) follows from EPI and (f) is true because the function given by $f(x) = \frac{1}{2} \log(2^x + c)$ for a positive constant c is a convex function in x. (Exercise: verify this by differentiating the function twice).

So, we have

$$\begin{aligned} I(X; Y_1|U) &= h(Y_1|U) - \frac{1}{2} \log(2\pi e N_1) \\ I(U; Y_2) &\leq \frac{1}{2} \log(2\pi e (P + N_2)) - \frac{1}{2} \log(2^{2h(Y_1|U)} + 2^{2h(Z_2)}) \end{aligned}$$

Equality holds everywhere iff U, X are Gaussian and the proof is complete.

19.4 Summary of results discussed and some new results

- We have obtained the capacity region of the MAC exactly.
- We have obtained the capacity region of the degraded broadcast channel.
- Other cases of broadcast channels whose capacity regions have been obtained exactly:
  - Parallel Broadcast Channel with Degraded Components

    This channel has input (X_A, X_B) and the outputs being (Y_{1A}, Y_{2A}) and (Y_{1B}, Y_{2B}) which satisfy $X_A \to Y_{1A} \to Y_{1B}$ and $X_B \to Y_{2A} \to Y_{2B}$ although $(X_A, X_B) \to (Y_{1A}, Y_{2A}) \to (Y_{1B}, Y_{2B})$ may not be true.

    This is an important model for subcarrier frequencies.
  - MIMO broadcast channel

    $\overline{Y_1} = H_1 \overline{X}, \overline{Y_2} = H_2 \overline{X}$ where H_1, H_2 are arbitrary matrices. The Parallel Broadcast Channel with Degraded Components is a special case of this channel.

19.5 The Relay Channel

The relay channel has an input X and an output Y. Further, it has a helper node called a relay node which observes Y_r and assists the sender to communicate its message by transmitting X_r. The memoryless relay channel is specified by specifying the probability distribution $p(y, y_r|x, x_r)$

- The source S encodes the message $W = w$ into $X^n(w)$.
- The relay R has a set of encoding functions $\{f_i\}$ which it uses to decide the symbol to transmit according to $X_{ri} = f_i(Y_{r1}, Y_{r2}, \ldots, Y_{r(i-1)})$.
- The destination T decodes Y^n to $\hat{W}$.

The maximum rate of reliable communication is called the capacity of the channel.

19.5.1 The Gaussian Relay Channel

The Gaussian Relay Channel is given by

$$Y_r = h_{sr}X + Z_1$$

$$Y = h_{st}X + h_{rt}X_r + Z_2$$

where there is a unit power constraint on X, X_r and Z_1 and Z_2 are independent Gaussians with zero mean and unit variance.

19.5.2 The Cutset Bound for the Relay Channel

The cutset bound for the relay channel is given by

$$C \leq \sup_{p(x_1,x_r)} \min[I(X; Y_1, Y_r|X_r), I(X, X_r; Y)]$$

This will be discussed in the next class.