LECTURE 21

Last time:

- Multiple access channels
- Coding theorem
- Capacity region for Gaussian channels

Lecture outline

- Broadcast channel
- Gaussian degraded broadcast channel

Reading: Section 14.6.

Broadcast channel

Single source, several receivers

The information to the receivers may be the same, or user 2 may have a subset of user 1, or the users have different information

Example where the receivers have the same information: radio

The rate is then no better than the worst-case channel

Broadcast channel

Example where user 2 may have a subset of user 1: video

user with better SNR obtains better resolution than user with worse SNR

different rates over the Internet

Does the separation theorem hold?

Broadcast channel

Users have different information

Example of person speaking two languages to two receivers who speak different languages (orthogonal signals)

Can do better than speaking to one half the time and to the other the rest of the time

Communicate in Esperanto using binary code (of course!)

By choosing arrangement of Language 1 and Language 2 words, communicate an extra bit (1 is Language 1 and 0 is language 2) to both users

Esperanto is shared by both users (common information)

Can we find a general framework to encompass all of these different cases?

Broadcast channel

General model of a stationary memoryless broadcast channel for one input and two outputs:

$$f_{\underline{Y_1}^n,\underline{Y_2}^n|\underline{Y}^n}\left(\underline{y_1}^n,\underline{y_2}^n|\underline{x}^n\right) = \prod_{i=1}^n p_{Y_1,Y_2|X}(y_{1i},y_{2i}|x_i)$$

Receiver 1 has rate R_1 and receiver 2 has rate R_2

The encoding maps the information to the two receivers to a single codeword:

$$\{1,\ldots,2^{nR_1}\}\times\{1,\ldots,2^{nR_2}\}\mapsto\underline{\mathcal{X}}^n$$

$$(m_1,m_2)\rightarrow\underline{x}^n$$

Broadcast channel

The decoding is done independently at each receiver i

$\underline{\mathcal{Y}}^n \mapsto \{1, \ldots, 2^{nR_i}\}$

$\underline{y}^n \to \widehat{m}_i$

An error occurs whenever $\widehat{m}_i \neq m_i$ for $i = 1$ or $i = 2$

What is the drawback here?

If we assume that for each user the messages are uniformly distributed and that the two users have independent transmissions, then the above works well

There is no requirement that the information for the different users be uncorrelated, so we are not considering IID over $\{1, \ldots, 2^{nR_1}\}$ for user 1 and $\{1, \ldots, 2^{nR_1}\}$ for user 2

Broadcast channel

How do we take the common information into account? Look at common information with rate R_0 (remember the Esperanto)

The encoding maps the information to the two receivers to a single codeword:

$\{1, \ldots, 2^{nR_0}\} \times \{1, \ldots, 2^{nR_1}\} \times \{1, \ldots, 2^{nR_2}\} \mapsto \underline{\mathcal{X}}^n$

$(m_1, m_2, m_0) \rightarrow \underline{x}^n$

The decoding is done independently at each receiver i

$\underline{\mathcal{Y}}^n \mapsto \{1, \ldots, 2^{nR_i}\} \times \{1, \ldots, 2^{nR_0}\}$

$\underline{y}^n \rightarrow (\widehat{m}_i, \widehat{m}_0^i)$

An error occurs whenever $\widehat{m}_i \neq m_i$ or $\widehat{m}_0^i$ for $i = 1$ or $i = 2$

Broadcast channel

What happens to the separation theorem?

By compressing for each user independently, then the messages are uniformly distributed over all possible messages for each user, but possibly this does not hold for the two messages jointly

If we compress independently for each user, AEP no longer holds, so in general may not be optimum in terms of minimizing description of what we want to transmit

Equivalently: if there is common information, because for instance the users are correlated in the information they need to receive, then we have been wasteful

If we do the compression of all the data jointly, then not clear how the decoders work, because they are independent.

Broadcast channel

A rate triplet (R_0, R_1, R_2) is achievable iff there exists a sequence of $((2^{nR_0}, 2^{nR_1}, 2^{nR_2}), n)$ codes that have probability of error vanish as $n \to \infty$

The capacity region is in general not known

Not surprising when we consider the language example

Exception: degraded broadcast channel, to be seen later

For $R_0 \leq \min\{R_1, R_2\}$ for a broadcast channel, $(R_0, R_1 - R_0, R_2 - R_0)$ is an achievable triplet of rates with common information

Broadcast channel

Although the broadcast channel capacity region is not known in general, we do know that the region depends only $f_{Y_1|X}(y_1|x)$ and on $f_{Y_2|X}(y_2|x)$

Type 1 error: user 1 has an error, type 2 error: user 2 has an error, type 3 error: both have errors

Over the channel for each user, the probability of errors of type 1 and 2 depends on the marginal pdf only

The probability of error of type 3 is lower bounded by the probability that at least one user has an error and upper bounded by the sum of the probabilities that both have an error

Degraded broadcast channel

$$f_{Y_1,Y_2|X}(y_1, y_2|x) = f_{Y_1|X}(y_1|x) f_{Y_2|Y_1}(y_2|y_1)$$

$$X \to Y_1 \to Y_2$$

Consider at first independent transmissions to the two receivers

The capacity region is the closure of the (R_1, R_2) such that

$$R_2 \leq I(U; Y_2)$$

$$R_1 \leq I(X; Y_1|U)$$

for some auxiliary U whose cardinality is less than that of any of the input and the outputs

Degraded broadcast channel

Method:

we generate codewords for user 2 by selecting IID sequences $\underline{U}^n$ using $\prod f_U(u_i)$ and mapping each of the messages m_2 in $\{1, \ldots, 2^{nR_2}\}$ onto some $\underline{u}^n(m_2)$

for each possible $\underline{u}^n(m_2)$, we generate a codeword for user 1 mapping (m_1, m_2) onto $\underline{x}^n(m_1, m_2)$ using $\prod f_{X|U}(x_i|u_i(m_2))$

Note: X is transmitted, U is not

Decoding:

user 1 decodes by looking at jointly typical $(\underline{U}^n(m_2), \underline{Y_2}^n)$ pairs

user 2 can first look at typical $(\underline{U}^n(m_2), \underline{Y_1}^n)$ pairs

having thus decoded m_2, it can reconstitute $\underline{U}^n(m_2)$ and then look at jointly typical $(\underline{U}^n(m_2), \underline{X}^n(m_1, m_2), \underline{Y_1}^n)$ triplet

Degraded broadcast channel

The results for receivers with dependent information can be obtained by using the fact that if (R_1, R_2) is an achievable rate pair when we have independent information, then:

For $R_0 \leq R_2$ for a *degraded* broadcast channel, $(R_0, R_1, R_2 - R_0)$ is an achievable triplet of rates with common information

U is decoded by both and carries the information of user 1, that is also received by user 2

Important example: degraded Gaussian broadcast channel

$Y_1 = X + Z_1$

$Y_2 = X + Z_1 + Z_2$

Consider R_0 is R_2

Degraded broadcast channel

Energy constraint $\mathcal{E}$ on input

The second user decodes m_2 from U

Rate R_0 is clearly upper bounded by the case where we communicate with only user 2 in mind: $\frac{1}{2}\ln\left(1+\frac{\mathcal{E}}{\sigma_{N_2}^2+\sigma_{N_1}^2}\right)$

In particular, we can always express R_0 as

$$\frac{1}{2}\ln\left(1+\frac{(1-\alpha)\mathcal{E}}{\alpha\mathcal{E}+\sigma_{N_2}^2+\sigma_{N_1}^2}\right)$$

with the upper bound being achieved for $\alpha = 0$

Degraded broadcast channel

For user 1: consider that we are trying to communicate the independent X, U over a channel with noise N_1

From our coding theorem, user 1 first decodes m_2 and is thus able to reconstitute U

Knowing U, user 1 now tries to recover $X|U$

Subtracting U, there is $\alpha\mathcal{E}$ left for $X|U$, which is $R_1 = \frac{1}{2}\ln\left(1 + \frac{\mathcal{E}}{\sigma_{N_1}^2}\right)$

Degraded broadcast channel

Note: the total rate at user 1 is less than if we had no broadcast

Hark back to our multiple-access channel:

the sum of the rates is upper bounded by $\frac{1}{2}\ln\left(1+\frac{\mathcal{E}}{\sigma^2_{N_1}}\right)$

$$
\begin{aligned}
& \frac{1}{2}\ln\left(1+\frac{\mathcal{E}}{\sigma^2_{N_1}}\right)-\frac{1}{2}\ln\left(1+\frac{(1-\alpha)\mathcal{E}}{\alpha\mathcal{E}+\sigma^2_{N_2}+\sigma^2_{N_1}}\right) \\
= \; & \frac{1}{2}\ln\left(\frac{\mathcal{E}+\sigma^2_{N_1}}{\sigma^2_{N_1}}\frac{\alpha\mathcal{E}+\sigma^2_{N_2}+\sigma^2_{N_1}}{\alpha\mathcal{E}+\sigma^2_{N_2}+\sigma^2_{N_1}+(1-\alpha)\mathcal{E}}\right) \\
= \; & \frac{1}{2}\ln\left(\frac{\mathcal{E}+\sigma^2_{N_1}}{\sigma^2_{N_1}}\frac{\alpha\mathcal{E}+\sigma^2_{N_2}+\sigma^2_{N_1}}{\sigma^2_{N_2}+\sigma^2_{N_1}+\mathcal{E}}\right) \\
= \; & \frac{1}{2}\ln\left(1+\frac{\alpha\mathcal{E}}{\sigma^2_{N_1}}+(1-\alpha)\frac{\mathcal{E}\sigma^2_{N_2}}{\sigma^2_{N_2}+\sigma^2_{N_1}+\mathcal{E}}\right)
\end{aligned}
$$

6.441 Information Theory
Spring 2010