Electrical Engineering 229A Lecture 23 Notes

Daniel Raban

November 16, 2021

1 Two Receiver Broadcast Channels

1.1 Degraded two receiver broadcast channels

The two receiver broadcast channel (for a discrete memoryless channel) is defined via

- $p(y_1, y_2 \mid x)$, which is nonnegative with $\sum_{y_1,y_2} p(y_1, y_2 \mid x) = 1$ for all x,
- Input alphabet $x \in \mathscr{X}$,
- Output alphabet $\mathscr{Y}_1$ of receiver 1,
- Output alphabet $\mathscr{Y}_2$ of receiver 2,
- Memorylessness of the channel, given by

$$p(y_{1,[1:n]}, y_{2,[1:n]} \mid x_{[1:n]}) = \prod_{i=1}^{n} p(y_{1,i}, y_{2,i} \mid x_i),$$

  where $y_{1,[1:n]}$ is new notation for $(y_1)_1^n$,
- Encoding map $e_n : [M_n^{(1)}] \times [M_n^{(2)}] \to \mathscr{X}^n$ of block length n,
- Decoding map $d_n : \mathscr{Y}^n \to [M_n^{(1)}] \times [M_n^{(2)}]$ of block length n,
- Rate region given by the closure of the set

$$\begin{aligned}\{(R_1, R_2) : \exists((e_n, d_n), n \geq 1) \text{ s.t. } & \liminf_n \tfrac{1}{n} \log M_n^{(1)} \geq R_1, \\ & \liminf_n \tfrac{1}{n} \log M_n^{(2)} \geq R_2, \\ & \lim_{n\to\infty} P(d_n(e_n(W_{1,n}, W_{2,n})) \neq (W_{1,n}, W_{2,n})) = 0, \}\end{aligned}$$

  where $W_{1,n} \sim \mathrm{Unif}([M_n^{(1)}]), W_{2,n} \sim \mathrm{Unif}([M_n^{(2)}])$.

The bad news is that finding the rate region has been an open problem for about 50 years. A special case where the rate region is known is called the *stochastically degraded* case.

Definition 1.1. $p(y_1, y_2 \mid x)$ is called **physically degraded** if

$$p(y_1, y_2 \mid x) = p(y_1 \mid x)p(y_2 \mid y_1).$$

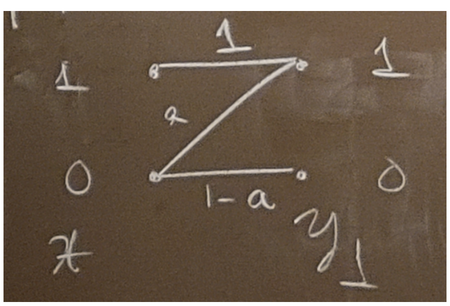
It is called **stochastically degraded** if there exists some distribution $p'(y_2 \mid y_1)$ such that

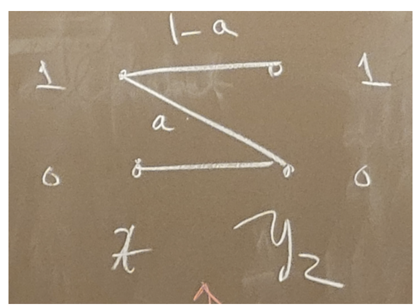
$$p(y_2 \mid x) = \sum_{y_1} p(y_1 \mid x)p'(y_2 \mid y_1).$$

The physical degradation condition means that we have the Markov chain $X - Y_1 - Y_2$. The stochastic degradation condition does not require $X - Y_1 - Y_2$ but is "equivalent" since the rate region only depends on $p(y_1 \mid x)$ and $p(y_2 \mid x)$.

Example 1.1 (Stochastically but not physically degraded channel)**.** Let $\mathscr{X} = \mathscr{Y}_1 = \mathscr{Y}_2 = \{0, 1\}$, and suppose that $Y_1 = X \oplus Z$, where $Z \in \{0, 1\}$, $\mathbb{P}(Z = 1) = a = 1 - \mathbb{P}(Z = 0)$. Here, $0 < a < 1$. Also, let $Y_2 = Z$, where $Z \perp\!\!\!\perp X$. This is not a physically degraded channel, since $X - Y_1 - Y_2$ is false (e.g. knowing both X and Y_1 determines Y_2). But it is stochastically degraded because we can replace Y_2 by Z', where $Z' \overset{d}{=} Z$, $Z' \perp\!\!\!\perp (X, Z)$.

Example 1.2 (Broadcast channel that is not stochastically degraded)**.** Let $\mathscr{X} = \mathscr{Y}_1 = \mathscr{Y}_2 = \{0, 1\}$ with $p(y_1 \mid x)$ given by a Z-channel and $p(y_2 \mid x)$ given by a different Z-channel.





We claim that there cannot be any $p'(y_2 \mid y_1)$ such that the stochastic degradation condition holds, i.e.

$$p(y_2 \mid x) = sum_{y_1} p(y_1 \mid x)p'(y_2 \mid y_1).$$

If such a p' existed, then

$$\begin{aligned} 0 &= p_{Y_2|X}(1 \mid 0) \\ &= p_{Y_1|X}(1 \mid 0)p'(1 \mid 1) + p_{Y_1|X}(0 \mid 0)p'(1 \mid 0). \end{aligned}$$

That is,

$$0 = p'(1 \mid 1) + (1-a)p'(1 \mid 0),$$

so

$$p'(1 \mid 1) = p'(1 \mid 0) = 0,$$

which makes

$$p'(0 \mid 1) = p(0 \mid 0) = 1.$$

Then $p(y_2 \mid x) = \sum_{y_1} p(y_1 \mid x)p'(y_2 \mid y_1)$ gives the the wrong channel.

1.2 Capacity region for a stochastically degraded broadcast channel

Theorem 1.1. *The capacity region for independent private messages over a stochastically degraded broadcast channel is the closure of the convex hull of*

$$\{(R_1, R_2) : R_2 \leq I(U; Y_2), R_1 \leq I(X; Y_1 \mid U)\}$$

for some $p(x)p(x \mid u)p(y_1, y_2 \mid x)$*, where* $U \in \mathscr{U}$ *and* $|\mathscr{U}| \leq \max\{|\mathscr{X}|, |\mathscr{Y}_1|, |\mathscr{Y}_2|\}$.

Think of these U variables as information that receiver 1, the stronger receiver, can use to get a better signal.

Proof. We will use a random coding achievability argument. The codebook is going to be comprised of $2^{n(R_1-\delta)}2^{n(R_2-\delta)}$ codewords in $\mathscr{X}^n$, organized as $2^{n(R_2-\delta)}$ clusters, each with $2^{n(R_1-\delta)}$ codewords.

Generate $2^{n(R_2-\delta)}$ independent sequences $(U_1(m_2), \ldots, W_m(m_2))$ with $1 \leq m_2 2^{n(R_2-\delta)}$, and entries $\overset{\text{iid}}{\sim} p(u)$. For each m_2, generate $2^{n(R_1-\delta)}$ sequences $(X_1(m_1, m_2), \ldots, X_n(m_1, m_2))$ with $1 \leq m_1 \leq 2^{n(R_1-\delta)}$ and, for each m_1, joint law $\prod_{i=1}^n p(x_i \mid U_i(m_2))$ (independently over m_1).

To send (m_1, m_2) the transmitter sends $(X_1(m_1, m_2), \ldots, X_n(m_1, m_2))$. Receiver 2, receiving $(Y_{2,1}, \ldots, Y_{2,n})$, determines all m_2 such that $(U_{[1:n]}(m_2), Y_{2,[1:m]})$ is ε-jointly weakly typical. If there is only one such message, it decodes as m_2. If there are none or more than one such message, it decodes arbitrarily. Receiver 1, receiving $(Y_{1,1}, \ldots, Y_{1,n})$, finds all (m_1, m_2) such that $(U_{[1:n]}(m_2), X_{[1:n]}(m_1, m_2), Y_{1,[1:n]})$ is ε-jointly weakly typical. If there is only one such message, it decodes as m_1. If there are none or more than one such message, it decodes arbitrarily.

If we take the probability over the random codebook, W_1, and W_2, symmetry gives us

$$\mathbb{P}(d_n(e_n(W_{1,n}, W_{2,n})) \neq (W_{1,n}, W_{2,n}))] = \mathbb{P}(d_n(e_n(1,1)) \neq (1,1)),$$

so we can condition on the message pair $(m_1, m_2) = (1, 1)$ being sent.

The error events for receiver 2 are

$$E_n^{(2)} = \{(U_{[1:n]}(1), Y_{2,[1:n]}) \notin A^{(n)}_{\varepsilon,(U,Y_2)}\}, \qquad E_{n,i}^{(2)} = \{(U_{[1:n]}(i), Y_{2,[1:n]}) \in A^{(n)}_{\varepsilon,(U,Y_2)}\}$$

for $i \neq 1$. By the weak law of large numbers,

$$\mathbb{P}(E_n^{(2)}) \xrightarrow{n\to\infty} 0$$

On the other hand,

$$\mathbb{P}(E_{n,i}^{(2)}) \leq 2^{-nI(U;Y_2)} 2^{3n\varepsilon},$$

so we want $2^{n(R_2-\delta)} 2^{-nI(U;Y_2)} 2^{n3\varepsilon} \to 0$, i.e. $R_2 < U(U;Y_2) - 3\varepsilon + \delta$.

The error events for receiver 1 are

$$E_n^{(1)} = \{(U_{[1:n]}(1), X_{[1:n]}(1,1), Y_{1,[1:n]}) \notin A^{(n)}_{\varepsilon,(U,X,Y_1)}\}, \qquad E_{n,i}^{(1)} = \{(U_{[1:n]}(i) Y_{1,[1:n]}) \in A^{(n)}_{\varepsilon,(U,Y_2)}\}$$

for $i \neq 1$. By the weak law of large numbers,

$$\mathbb{P}(E_n^{(1)}) \xrightarrow{n\to\infty} 0.$$

On the other hand,

$$\mathbb{P}(E_{n,i}^{(1)}) \leq 2^{-nI(U;Y_1)} 2^{3n\varepsilon}.$$

There are $2^{n(R_2-\delta)}$, and $I(U;Y_1) \geq I(U;Y_2)$, so the earlier condition on R_2 ensures $\sum_{i\neq 1} \mathbb{P}(E_{n,i}^{(1)}) \to 0$.

For $j \neq 1$, we also have the error event

$$E_{n,1,j}^{(1)} = \{(U_{[1:n]}(1), X_{[1:n]}(j,1), Y_{1,[1:n]}) \in A^{(n)}_{\varepsilon,(U,X,Y_1)}\}.$$

Then

$$\mathbb{P}(E_{n,1,j}^{(1)}) = \sum_{u_{[1:n]}, x_{[1:n]}, y_{1,[1:n]} \in A_\varepsilon^{(n)}} 2^{-nH(U,Y_1)} 2^{n\varepsilon} 2^{-nH(X|U)} 2^{n\varepsilon}$$

The size of $A_\varepsilon^{(n)}$ is $\leq 2^{nH(U,X,Y_1)} 2^{n\varepsilon}$.

$$\leq 2^{-nI(X;Y_1|U)} 2^{n4\varepsilon}.$$

The converse part of the proof is homework. $\square$

1.3 Capacity region for a stochastically degraded Gaussian broadcast channel

The Gaussian case (with power constrained to P, receiver 1 noise $\mathcal{N}(0, \sigma_1^2)$, and receiver noise $\mathcal{N}(0, \sigma_2^2)$ with $\sigma_2^2 > \sigma_1^2$) is automatically stochastically degraded.

Theorem 1.2. *The rate region is the union of the sets of the form*

$$\{(R_1, R_2) : R_2 \le C((1-\alpha)P, \alpha P + \sigma_2^2), R_1 \le C(\alpha P, \sigma_1^2)\}$$

over $0 < \alpha < 1$, *where*

$$C(P, \sigma^2) = \frac{1}{2} \log \left(1 + \frac{P}{\sigma^2}\right).$$