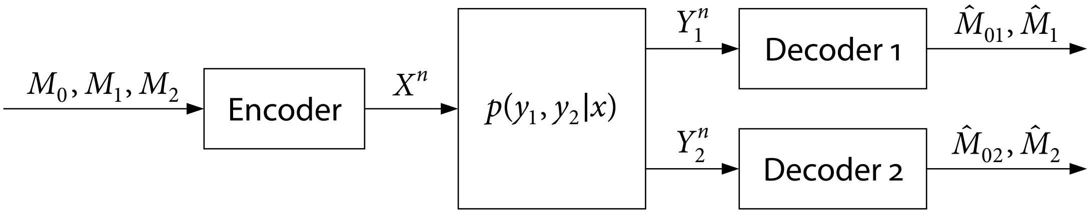
Lecture #5 Degraded Broadcast Channels

(Reading: NIT 5.1–5.7)

- Discrete memoryless broadcast channel
- Superposition coding inner bound
- Degraded broadcast channels
- Gaussian broadcast channel
- Less noisy and more capable broadcast channels
- Extensions

Broadcast communication system



- DM broadcast channel (BC) $(\mathcal{X}, p(y_1, y_2|x), \mathcal{Y}_1 \times \mathcal{Y}_2)$
- $(2^{nR_0}, 2^{nR_1}, 2^{nR_2}, n)$ code, $P_e^{(n)}$, achievability: Same as MAC
- Capacity region $\mathscr{C}$: Closure of the set of achievable (R_0, R_1, R_2)
- Useful fact:

Lemma 5.1

The capacity region depends on $p(y_1, y_2|x)$ only through $p(y_1|x)$ and $p(y_2|x)$

- The capacity region of the DM-BC is not known in general

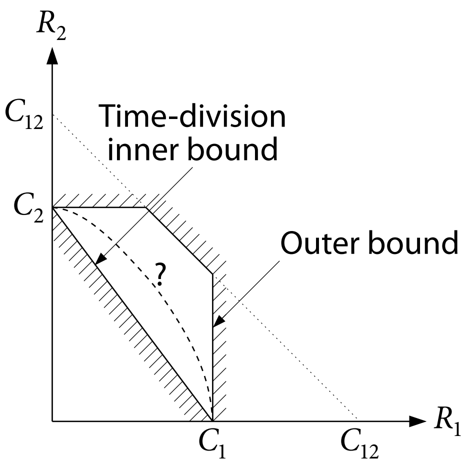
Overview

- There are inner and outer bounds that coincide in several cases
- Common message only ($R_1 = R_2 = 0$):

$$C_0 = \max_{p(x)} \min\{I(X; Y_1),\ I(X; Y_2)\}$$

- Degraded message sets ($R_1 = 0$ or $R_2 = 0$): see NIT 9.1
- Several classes of DM-BCs with restrictions on their channel structures, e.g.,
  - Degraded
  - Less noisy
  - More capable
  - Semideterministic
- Focus of this lecture:
  - Superposition coding inner bound
  - Special classes of BCs for which superposition coding is optimal
  - Private-message capacity region: $R_0 = 0$

Simple bounds on the capacity region



- Individual capacities:

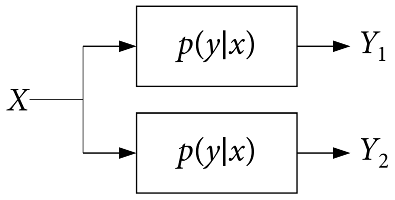
$$C_j = \max_{p(x)} I(X; Y_j), \quad j = 1, 2$$

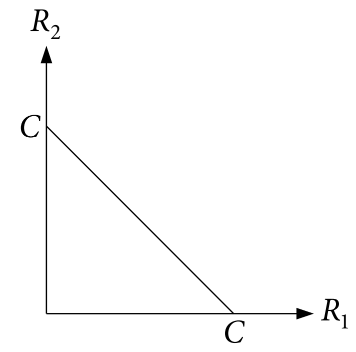
- Upper bound on the sum-rate:

$$R_1 + R_2 \le C_{12} = \max_{p(x)} I(X; Y_1, Y_2)$$

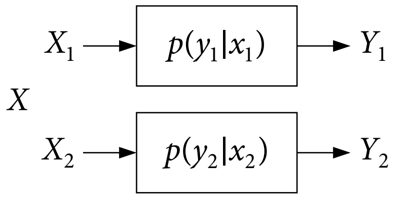
Examples

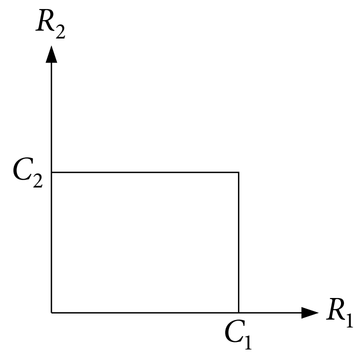
- Symmetric DM-BC:



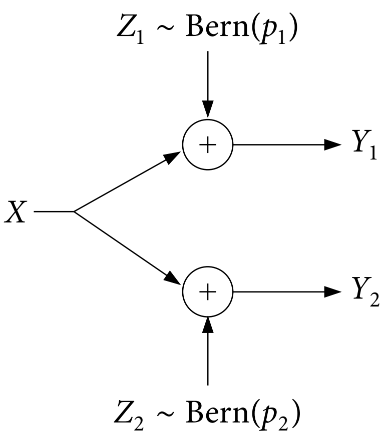


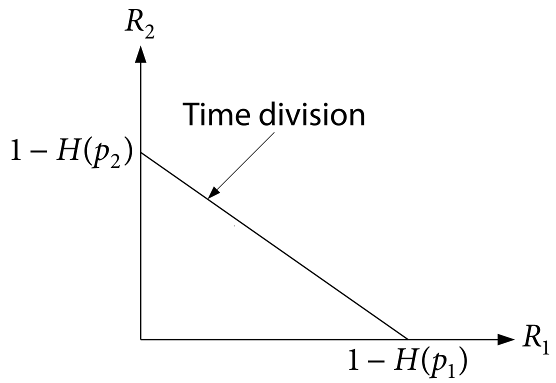
- DM-BC with orthogonal components:





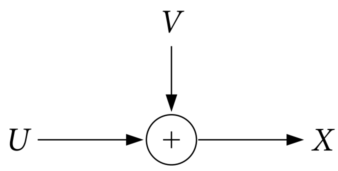
Binary symmetric BC





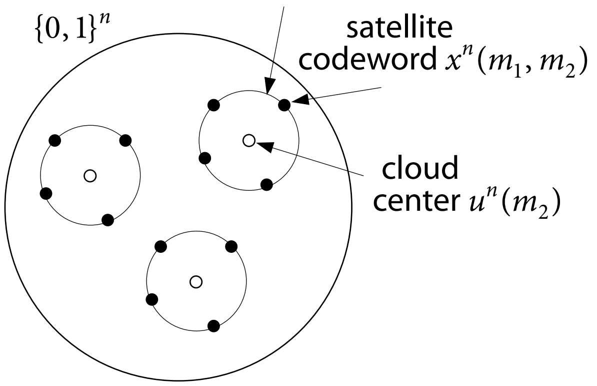
Assume $p_1 < p_2 < 1/2$

BS-BC: Superposition coding (Cover 1972)



- Codebook generation and encoding:
  - Let $U \sim \text{Bern}(1/2)$, $V \sim \text{Bern}(\alpha)$, $\alpha \in [0, 1/2]$, be independent, and $X = U \oplus V$
  - Independently generate 2^{nR_2} sequences $u^n(m_2) \sim \prod_{i=1}^n p_U(u_i)$, $m_2 \in [1:2^{nR_2}]$
  - Independently generate 2^{nR_1} sequences $v^n(m_1) \sim \prod_{i=1}^n p_V(v_i)$, $m_1 \in [1:2^{nR_1}]$
- To send (m_1, m_2), transmit $x^n(m_1, m_2) = u^n(m_2) \oplus v^n(m_1)$

BS-BC: Superposition coding (Cover 1972)



- Codebook generation and encoding:
  - Let $U \sim \text{Bern}(1/2)$, $V \sim \text{Bern}(\alpha)$, $\alpha \in [0, 1/2]$, be independent, and $X = U \oplus V$
  - Independently generate 2^{nR_2} sequences $u^n(m_2) \sim \prod_{i=1}^n p_U(u_i)$, $m_2 \in [1:2^{nR_2}]$
  - Independently generate 2^{nR_1} sequences $v^n(m_1) \sim \prod_{i=1}^n p_V(v_i)$, $m_1 \in [1:2^{nR_1}]$
- To send (m_1, m_2), transmit $x^n(m_1, m_2) = u^n(m_2) \oplus v^n(m_1)$

BS-BC: Superposition coding (Cover 1972)

- Decoder 2 recovers m_2 from $y_2^n = u^n(m_2) \oplus (v^n(m_1) \oplus z_2^n)$:

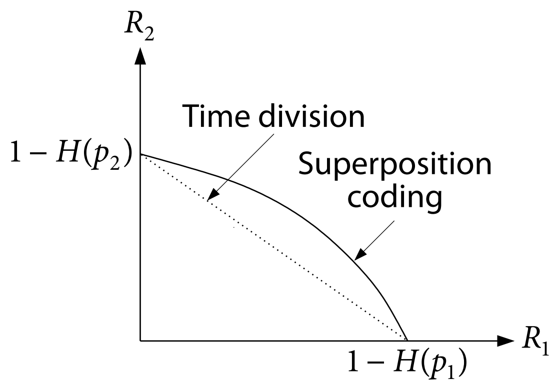
$$R_2 < I(U; Y_2) = 1 - H(\alpha * p_2)$$

- Decoder 1 uses successive cancellation decoding:
  - It recovers m_2 from $y_1^n = u^n(m_2) \oplus (v^n(m_1) \oplus z_1^n)$:

$$R_2 < I(U; Y_1) = 1 - H(\alpha * p_1) \quad (> 1 - H(\alpha * p_2))$$

  - Then recovers m_1 from $v^n(m_1) \oplus z_1^n$:

$$R_1 < I(V; V \oplus Z_1) = H(\alpha * p_1) - H(p_1)$$



Superposition coding bound (Cover 1972, Bergmans 1973)

Theorem 5.1

A rate pair (R_1, R_2) is achievable for the DM-BC $p(y_1, y_2|x)$ if

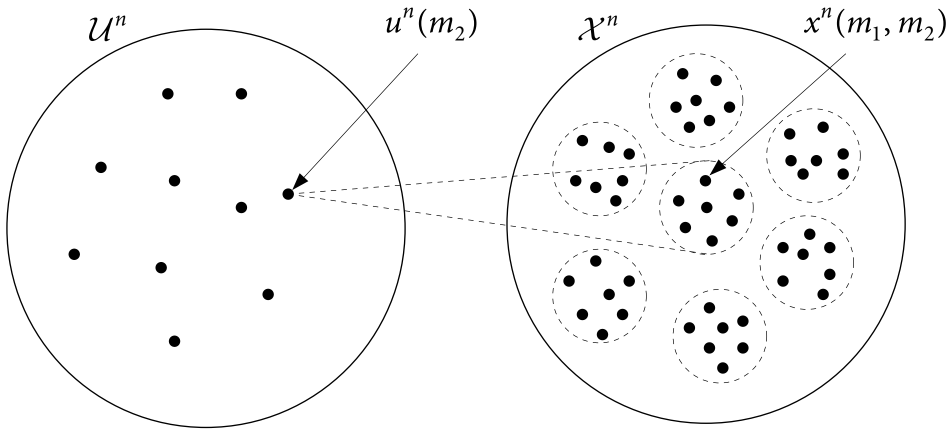
$$\begin{aligned} R_1 &< I(X; Y_1|U), \\ R_2 &< I(U; Y_2), \\ R_1 + R_2 &< I(X; Y_1) \end{aligned}$$

for some pmf $p(u, x)$

- U is an auxiliary random variable
- This inner bound is tight (capacity region) for several classes of BCs

Proof of achievability

- New ideas: Superposition coding and simultaneous nonunique decoding
- Codebook generation:
  - Independently generate 2^{nR_2} sequences $u^n(m_2) \sim \prod_{i=1}^n p_U(u_i)$, $m_2 \in [1:2^{nR_2}]$
  - For each $m_2 \in [1:2^{nR_2}]$, conditionally independently generate 2^{nR_1} sequences $x^n(m_1, m_2) \sim \prod_{i=1}^n p_{X|U}(x_i|u_i(m_2))$, $m_1 \in [1:2^{nR_1}]$



Proof of achievability

- New ideas: Superposition coding and simultaneous nonunique decoding
- Codebook generation:
  - Independently generate 2^{nR_2} sequences $u^n(m_2) \sim \prod_{i=1}^n p_U(u_i)$, $m_2 \in [1:2^{nR_2}]$
  - For each $m_2 \in [1:2^{nR_2}]$, conditionally independently generate 2^{nR_1} sequences $x^n(m_1, m_2) \sim \prod_{i=1}^n p_{X|U}(x_i|u_i(m_2))$, $m_1 \in [1:2^{nR_1}]$
- Encoding:
  - To send (m_1, m_2), transmit $x^n(m_1, m_2)$
- Decoding:
  - Decoder 2 finds the unique message $\hat{m}_2$ such that $(u^n(\hat{m}_2), y_2^n) \in \mathcal{T}_\epsilon^{(n)}$
    (by the packing lemma, $\mathrm{P}(\mathcal{E}_2) \to 0$ as $n \to \infty$ if $R_2 < I(U; Y_2) - \delta(\epsilon)$)
  - Decoder 1 finds the unique message $\hat{m}_1$ such that

$$(u^n(m_2), x^n(\hat{m}_1, m_2), y_1^n) \in \mathcal{T}_\epsilon^{(n)} \quad \text{for some } m_2$$

Analysis of the probability of error for decoder 1

- Consider $\mathsf{P}(\mathcal{E})$ conditioned on $(M_1, M_2) = (1, 1)$

m_1	m_2	Joint pmf
1	1	$p(u^n, x^n) p(y_1^n \mid x^n)$
*	1	$p(u^n, x^n) p(y_1^n \mid u^n)$
*	*	$p(u^n, x^n) p(y_1^n)$
1	*	$p(u^n, x^n) p(y_1^n)$

- Error events:

$$\mathcal{E}_{11} = \{(U^n(1), X^n(1,1), Y_1^n) \notin \mathcal{T}_\epsilon^{(n)}\}$$
$$\mathcal{E}_{12} = \{(U^n(1), X^n(m_1,1), Y_1^n) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } m_1 \neq 1\}$$
$$\mathcal{E}_{13} = \{(U^n(m_2), X^n(m_1,m_2), Y_1^n) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } m_1 \neq 1,\ m_2 \neq 1\}$$

Packing lemma

- Let $(U, X, Y) \sim p(u, x, y)$
- Let $(\tilde{U}^n, \tilde{Y}^n) \sim p(\tilde{u}^n, \tilde{y}^n)$ be arbitrarily distributed
- Let $X^n(m) \sim \prod_{i=1}^n p_{X|U}(x_i|\tilde{u}_i)$, $m \in \mathcal{A}$, $|\mathcal{A}| \le 2^{nR}$,

  be pairwise conditionally independent of $\tilde{Y}^n$ given $\tilde{U}^n$

Packing lemma

There exists $\delta(\epsilon) \to 0$ as $\epsilon \to 0$ such that

$$\lim_{n\to\infty} \mathsf{P}\{(\tilde{U}^n, X^n(m), \tilde{Y}^n) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } m \in \mathcal{A}\} = 0,$$

if $R < I(X; Y|U) - \delta(\epsilon)$

Analysis of the probability of error for decoder 1

- Error events:

$$\mathcal{E}_{11} = \{(U^n(1), X^n(1,1), Y_1^n) \notin \mathcal{T}_\epsilon^{(n)}\}$$
$$\mathcal{E}_{12} = \{(U^n(1), X^n(m_1,1), Y_1^n) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } m_1 \neq 1\}$$
$$\mathcal{E}_{13} = \{(U^n(m_2), X^n(m_1,m_2), Y_1^n) \in \mathcal{T}_\epsilon^{(n)} \text{ for some } m_1 \neq 1,\ m_2 \neq 1\}$$

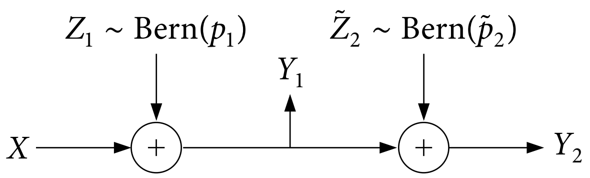
- By LLN, $\mathsf{P}(\mathcal{E}_{11}) \to 0$
- By the packing lemma,

$$\mathsf{P}(\mathcal{E}_{12}) \to 0 \text{ if } R_1 < I(X; Y_1 | U) - \delta(\epsilon),$$
$$\mathsf{P}(\mathcal{E}_{13}) \to 0 \text{ if } R_1 + R_2 < I(U, X; Y_1) - \delta(\epsilon) = I(X; Y_1) - \delta(\epsilon)$$

- Remark: The inner bound does not change if decoder 1 is required to recover M_2
- The superposition coding scheme is optimal for some classes of BCs

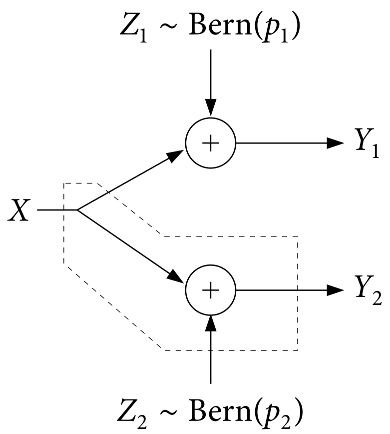
Degraded broadcast channels

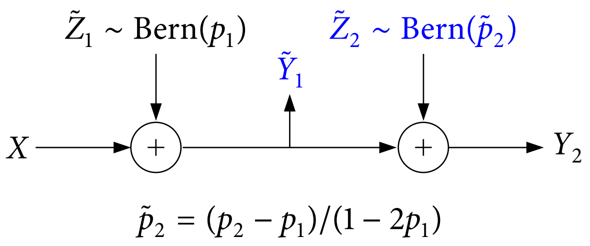
- Physically degraded: $X \to Y_1 \to Y_2$ form a Markov chain



$X \to Y_1 :$ BSC(p_1),
$X \to Y_2 :$ BSC$(p_1 \star \tilde{p}_2)$

- (Stochastically) degraded: $\exists\ \tilde{Y}_1 | \{X = x\} \sim p_{Y_1|X}(\tilde{y}_1 | x)$ such that $X \to \tilde{Y}_1 \to Y_2$





Capacity region of the degraded BC

Theorem 5.2 (Cover 1972, Bergmans 1973, Gallager 1974)

The capacity region of the degraded DM-BC $p(y_1, y_2|x)$ is the set of (R_1, R_2) such that

$$R_1 \le I(X; Y_1|U),$$
$$R_2 \le I(U; Y_2)$$

for some $p(u, x)$ with $|\mathcal{U}| \le \min\{|\mathcal{X}|, |\mathcal{Y}_1|, |\mathcal{Y}_2|\} + 1$

- Achievability: Superposition coding + degradedness ($I(U; Y_2) \le I(U; Y_1)$)
- For BS-BC, the capacity region simplifies to the set of (R_1, R_2) such that

$$R_1 \le H(\alpha * p_1) - H(p_1),$$
$$R_2 \le 1 - H(\alpha * p_2)$$

for some $\alpha \in [0, 1]$

Proof of the converse

- Need to show that for any sequence of $(2^{nR_1}, 2^{nR_2}, n)$ codes with $P_e^{(n)} \to 0$,

$$R_1 \le I(X; Y_1|U), \quad R_2 \le I(U; Y_2)$$

for some $p(u, x)$ such that $U \to X \to (Y_1, Y_2)$

- The key is to identify U
- Each code induces a joint pmf

$$(M_1, M_2, X^n, Y_1^n, Y_2^n) \sim 2^{-n(R_1+R_2)} p(x^n|m_1, m_2) \prod_{i=1}^{n} p_{Y_1,Y_2|X}(y_{1i}, y_{2i}|x_i)$$

- By Fano's inequality

$$H(M_j|Y_j^n) \le nR_j P_e^{(n)} + 1 \le n\epsilon_n, \quad j = 1, 2$$

for some $\epsilon_n \to 0$ as $n \to \infty$

- Hence

$$nR_j \le I(M_j; Y_j^n) + n\epsilon_n, \quad j = 1, 2$$

Proof of the converse

- Let's try $U = M_2$ (satisfies $U \to X_i \to (Y_{1i}, Y_{2i})$)

$$
\begin{aligned}
I(M_1; Y_1^n) \le I(M_1; Y_1^n | M_2) = I(M_1; Y_1^n | U) \\
&= \sum_{i=1}^n I(M_1; Y_{1i} | U, Y_1^{i-1}) \\
&\le \sum_{i=1}^n I(M_1, Y_1^{i-1}; Y_{1i} | U) \\
&= \sum_{i=1}^n I(X_i, M_1, Y_1^{i-1}; Y_{1i} | U) \\
&= \sum_{i=1}^n I(X_i; Y_{1i} | U)
\end{aligned}
$$

- Now consider the second inequality

$$
I(M_2; Y_2^n) = \sum_{i=1}^n I(M_2; Y_{2i} | Y_2^{i-1}) = \sum_{i=1}^n I(U; Y_{2i} | Y_2^{i-1})
$$

But $I(U; Y_{2i} | Y_2^{i-1})$ is not necessarily $\le I(U; Y_{2i})$

Proof of the converse (Gallager 1974)

- Let's try $U_i = (M_2, Y_1^{i-1})$ (satisfies $U_i \to X_i \to (Y_{1i}, Y_{2i})$), so

$$
I(M_1; Y_1^n | M_2) = \sum_{i=1}^n I(X_i; Y_{1i} | U_i)
$$

- Now consider the other term

$$
\begin{aligned}
I(M_2; Y_2^n) &\le \sum_{i=1}^n I(M_2, Y_2^{i-1}; Y_{2i}) \\
&\le \sum_{i=1}^n I(M_2, Y_2^{i-1}, Y_1^{i-1}; Y_{2i})
\end{aligned}
$$

- But $I(M_2, Y_2^{i-1}, Y_1^{i-1}; Y_{2i})$ is not necessarily equal to $I(M_2, Y_1^{i-1}; Y_{2i})$

- Key insight: Capacity region is the same as equivalent physically degraded BC

- Can assume that $X \to Y_1 \to Y_2$; thus $Y_2^{i-1} \to (M_2, Y_1^{i-1}) \to Y_{2i}$ and

$$
I(M_2; Y_2^n) \le \sum_{i=1}^n I(U_i; Y_{2i})
$$

Proof of the converse (Gallager 1974)

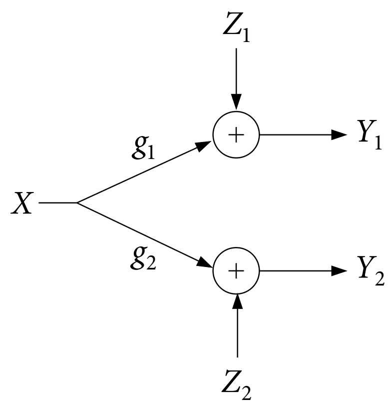
- Define time-sharing r.v. $Q \sim \mathrm{Unif}[1:n]$, independent of $(M_1, M_2, X^n, Y_1^n, Y_2^n)$
- Let $U = (Q, U_Q)$, $X = X_Q$, $Y_1 = Y_{1Q}$, $Y_2 = Y_{2Q}$
- Clearly, $U \to X \to (Y_1, Y_2)$; hence

$$nR_1 \le \sum_{i=1}^{n} I(X_i; Y_{1i} | U_i) + n\epsilon_n = nI(X; Y_1 | U) + n\epsilon_n,$$

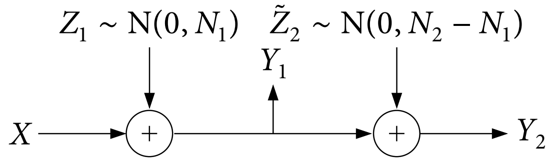
$$nR_2 \le \sum_{i=1}^{n} I(U_i; Y_{2i}) + n\epsilon_n = nI(U_Q; Y_2 | Q) + n\epsilon_n \le nI(U; Y_2) + n\epsilon_n$$

- Bound on cardinality of U (NIT Appendix C)
- Remark: Proof works also with $U_i = (M_2, Y_2^{i-1})$ or $U_i = (M_2, Y_1^{i-1}, Y_2^{i-1})$

Gaussian broadcast channel



- g_1, g_2: channel gains (wolog $|g_1| \ge |g_2|$); $Z_1, Z_2 \sim \mathrm{N}(0,1)$; $S_j = g_j^2 P$, $j = 1, 2$
- Average power constraint: $\sum_{i=1}^{n} x_i^2(m_1, m_2) \le nP$, $(m_1, m_2) \in [1:2^{nR_1}] \times [1:2^{nR_2}]$
- Channel is degraded
- Equivalent physically degraded model ($N_1 = 1/g_1^2$, $N_2 = 1/g_2^2 \ge N_1$):



Capacity region of the Gaussian BC

Theorem 5.3 (Cover 1972, Bergmans 1974)

The capacity of the Gaussian BC is the set of (R_1, R_2) such that

$$R_1 \leq \mathsf{C}(\alpha S_1),$$
$$R_2 \leq \mathsf{C}\left(\frac{\bar{\alpha} S_2}{\alpha S_2 + 1}\right)$$

for some $\alpha \in [0,1]$, where $S_j = g_j^2 P$, $j = 1, 2$

- Achievability: Consider DM-BC with cost and use discretization procedure
- More explicitly, let $U \sim \mathrm{N}(0, \bar{\alpha} P)$, $V \sim \mathrm{N}(0, \alpha P)$ are independent and $X = U + V$
- Follow similar steps to BS-BC scheme:
  - Send $x^n(m_1, m_2) = v^n(m_1) + u^n(m_2)$
  - Receiver Y_1 uses successive cancellation decoding

Proof of the converse (Bergmans 1974)

- Capacity region same as equivalent physically degraded Gaussian BC
- Hence, we assume the physically degraded Gaussian BC

$$Y_1 = X + Z_1, \quad Y_2 = X + Z_2 = Y_1 + \tilde{Z}_2$$

- We will need the following (Shannon 1948, Stam 1959, Blachman 1965)

Entropy power inequality (EPI)

- Vector EPI: Let $X^n \sim f(x^n)$ and $Z^n \sim f(z^n)$ be independent and $Y^n = X^n + Z^n$, then

$$2^{2h(Y^n)/n} \geq 2^{2h(X^n)/n} + 2^{2h(Z^n)/n}$$

  with equality if X^n and Z^n are Gaussian with $K_X = aK_Z$ for some $a > 0$

- Conditional EPI: Let X^n and Z^n be conditionally independent given an arbitrary U, with $f(x^n|u)$ and $f(z^n|u)$, and $Y^n = X^n + Z^n$, then

$$2^{2h(Y^n|U)/n} \geq 2^{2h(X^n|U)/n} + 2^{2h(Z^n|U)/n}$$

Proof of the converse (Bergmans 1974)

- By Fano's inequality,

$$nR_1 \le I(M_1; Y_1^n | M_2) + n\epsilon_n,$$
$$nR_2 \le I(M_2; Y_2^n) + n\epsilon_n$$

- Need to show that there exists an $\alpha \in [0,1]$ such that

$$I(M_1; Y_1^n | M_2) \le n\,\mathsf{C}(\alpha S_1) = n\,\mathsf{C}\left(\frac{\alpha P}{N_1}\right),$$
$$I(M_2; Y_2^n) \le n\,\mathsf{C}\left(\frac{\bar{\alpha} S_2}{\alpha S_2 + 1}\right) = n\,\mathsf{C}\left(\frac{\bar{\alpha} P}{\alpha P + N_2}\right)$$

- Consider

$$I(M_2; Y_2^n) = h(Y_2^n) - h(Y_2^n | M_2) \le \frac{n}{2}\log\left(2\pi e(P + N_2)\right) - h(Y_2^n | M_2)$$

- Since

$$\frac{n}{2}\log\left(2\pi e N_2\right) = h(Z_2^n) = h(Y_2^n | M_2, X^n) \le h(Y_2^n | M_2) \le h(Y_2^n) \le \frac{n}{2}\log(2\pi e(P + N_2)),$$

there must exist an $\alpha \in [0,1]$ such that

$$h(Y_2^n | M_2) = \frac{n}{2}\log\left(2\pi e(\alpha P + N_2)\right)$$

Proof of the converse (Bergmans 1974)

- Next consider

$$\begin{aligned} I(M_1; Y_1^n | M_2) &= h(Y_1^n | M_2) - h(Y_1^n | M_1, M_2) \\ &= h(Y_1^n | M_2) - h(Y_1^n | M_1, M_2, X^n) \\ &= h(Y_1^n | M_2) - \frac{n}{2}\log\left(2\pi e N_1\right) \end{aligned}$$

- Using the conditional vector EPI,

$$\begin{aligned} h(Y_2^n | M_2) &= h(Y_1^n + \tilde{Z}_2^n | M_2) \\ &\ge \frac{n}{2}\log\left(2^{2h(Y_1^n|M_2)/n} + 2^{2h(\tilde{Z}_2^n|M_2)/n}\right) = \frac{n}{2}\log\left(2^{2h(Y_1^n|M_2)/n} + 2\pi e(N_2 - N_1)\right) \end{aligned}$$

- But since $h(Y_2^n | M_2) = \frac{n}{2}\log\left(2\pi e(\alpha P + N_2)\right)$,

$$2\pi e(\alpha P + N_2) \ge 2^{2h(Y_1^n|M_2)/n} + 2\pi e(N_2 - N_1) \Rightarrow h(Y_1^n | M_2) \le \frac{n}{2}\log\left(2\pi e(\alpha P + N_1)\right)$$

- Hence

$$I(M_1; Y_1^n | M_2) \le \frac{n}{2}\log\left(2\pi e(\alpha P + N_1)\right) - \frac{n}{2}\log\left(2\pi e N_1\right) = n\,\mathsf{C}\left(\frac{\alpha P}{N_1}\right)$$

Less noisy and more capable broadcast channels

- Less noisy if $I(U;Y_1) \ge I(U;Y_2)$ for all $p(u,x)$
- More capable if $I(X;Y_1) \ge I(X;Y_2)$ for all $p(x)$
- Degraded $\Rightarrow$ less noisy $\Rightarrow$ more capable
- Superposition coding is optimal

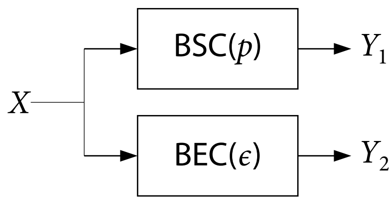
Capacity region of more capable BC (El Gamal 1979)

The capacity region of the more capable BC is the set of (R_1, R_2) such that

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

for some $p(u,x)$, where $|\mathcal{U}| \le \min\{|\mathcal{X}|, |\mathcal{Y}_1| \cdot |\mathcal{Y}_2|\} + 2$

Example: A BSC and a BEC



- For $0 \le \epsilon \le 2p$: Y_1 is a degraded version of Y_2
- For $2p < \epsilon \le 4p(1-p)$: Y_2 is less noisy than Y_1, but not degraded
- For $4p(1-p) < \epsilon \le H(p)$: Y_2 is more capable than Y_1, but not less noisy
- For $H(p) < \epsilon \le 1$: The channel does not belong to any of the three classes

Capacity region with common message

- If $(0, R_1, R_2)$ is achievable by superposition coding, so is $(R_0, R_1, R_2 - R_0)$

Superposition coding inner bound with common message

A rate triple (R_0, R_1, R_2) is achievable for the DM-BC $p(y_1, y_2|x)$ if

$$\begin{aligned} R_1 &< I(X; Y_1 | U), \\ R_0 + R_2 &< I(U; Y_2), \\ R_0 + R_1 + R_2 &< I(X; Y_1) \end{aligned}$$

for some pmf $p(u, x)$

- Tight for more capable BCs

Extensions to more than two receivers

- Capacity region for degraded can be easily extended

  For 3-receivers, the capacity region is the set of (R_1, R_2, R_3) such that

$$\begin{aligned} R_1 &\le I(X; Y_1 | U_2), \\ R_2 &\le I(U_2; Y_2 | U_3), \\ R_3 &\le I(U_3; Y_3) \end{aligned}$$

  for some $p(u_3)p(u_2|u_3)p(x|u_2)$

- Capacity region for less noisy is not known for $k \ge 4$

- Capacity region for more capable is not known for $k \ge 3$

Summary

- Discrete memoryless broadcast channel (DM-BC)
- Capacity region depends only on the channel marginal pmfs
- Superposition coding
- Simultaneous nonunique decoding
- Physically and stochastically degraded BCs
- Capacity region of degraded BCs is achieved by superposition coding
- Identification of the auxiliary random variable in the proof of the converse
- Gaussian BC is always degraded
- Use of EPI in converse for Gaussian BC
- Less noisy and more capable BCs:
  - Degraded $\Rightarrow$ less noisy $\Rightarrow$ more capable
  - Superposition coding is optimal

References

Bergmans, P. P. (1973). Random coding theorem for broadcast channels with degraded components. *IEEE Trans. Inf. Theory*, 19(2), 197–207.

Bergmans, P. P. (1974). A simple converse for broadcast channels with additive white Gaussian noise. *IEEE Trans. Inf. Theory*, 20(2), 279–280.

Blachman, N. M. (1965). The convolution inequality for entropy powers. *IEEE Trans. Inf. Theory*, 11(2), 267–271.

Cover, T. M. (1972). Broadcast channels. *IEEE Trans. Inf. Theory*, 18(1), 2–14.

El Gamal, A. (1979). The capacity of a class of broadcast channels. *IEEE Trans. Inf. Theory*, 25(2), 166–169.

Gallager, R. G. (1974). Capacity and coding for degraded broadcast channels. *Probl. Inf. Transm.*, 10(3), 3–14.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.*, 27(3), 379–423, 27(4), 623–656.

Stam, A. J. (1959). Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Inf. Control*, 2(2), 101–112.