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 \(C\): 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:
  \[ R_1 + R_2 \leq C_{12} = \max_{p(x)} I(X; Y_1, Y_2) \]
Examples

- **Symmetric DM-BC:**

  ![Diagram of Symmetric DM-BC]

  \[ p(y|x) \rightarrow Y_1 \]

  \[ p(y|x) \rightarrow Y_2 \]

- **DM-BC with orthogonal components:**

  ![Diagram of DM-BC with orthogonal components]

  \[ X_1 \rightarrow \begin{array}{c} \quad \downarrow \quad \end{array} \]

  \[ \begin{array}{c} p(y_1|x_1) \quad \rightarrow Y_1 \end{array} \]

  \[ X \]

  \[ X_2 \rightarrow \begin{array}{c} \quad \downarrow \quad \end{array} \]

  \[ \begin{array}{c} p(y_2|x_2) \quad \rightarrow Y_2 \end{array} \]

**Binary symmetric BC**

- **Binary symmetric BC:**

  ![Diagram of Binary symmetric BC]

  \[ Z_1 \sim \text{Bern}(p_1) \]

  \[ X \rightarrow \begin{array}{c} \quad \downarrow \quad \end{array} \]

  \[ \begin{array}{c} + \quad \rightarrow Y_1 \end{array} \]

  \[ Z_2 \sim \text{Bern}(p_2) \]

  \[ X \rightarrow \begin{array}{c} \quad \downarrow \quad \end{array} \]

  \[ \begin{array}{c} + \quad \rightarrow Y_2 \end{array} \]

  \[ 1 - H(p_2) \]

  \[ 1 - H(p_1) \]

  Assume \( p_1 < p_2 < 1/2 \)
• 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 \( y_2^n = u^n(m_2) \oplus (v^n(m_1) \oplus z_2^n) \):
  \[
  R_2 < I(U; Y_2) = 1 - H(\alpha \ast p_2)
  \]

- Decoder 1 uses successive cancellation decoding:
  - It recovers \( m_2 \) from \( y_1^n = u^n(m_2) \oplus (v^n(m_1) \oplus z_1^n) \):
    \[
    R_2 < I(U; Y_1) = 1 - H(\alpha \ast p_1) \ (> 1 - H(\alpha \ast p_2))
    \]
  - Then recovers \( m_1 \) from \( v^n(m_1) \oplus z_1^n \):
    \[
    R_1 < I(V; V \oplus Z_1) = H(\alpha \ast p_1) - H(p_1)
    \]

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

\[
R_1 < I(X; Y_1 | U),
R_2 < I(U; Y_2),
R_1 + R_2 < I(X; Y_1)
\]

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 $u^n(m_2) \sim \prod_{i=1}^{n} p_{U_i}(u_i), m_2 \in [1 : 2^{nR_2}]$
  - For each $m_2 \in [1 : 2^{nR_2}]$, conditionally independently generate $2^{nR_1}$ sequences $x^n(m_1, m_2) \sim \prod_{i=1}^{n} p_{X_j|U}(x_i|u_i(m_2)), m_1 \in [1 : 2^{nR_1}]$

- **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 $(u^n(\hat{m}_2), y^n_2) \in T^{(n)}_\varepsilon$ (by the packing lemma, $P(\mathcal{E}_2) \to 0$ as $n \to \infty$ if $R_2 < I(U; Y_2) - \delta(\varepsilon)$)
  - Decoder 1 finds the unique message $\hat{m}_1$ such that
    $$(u^n(m_2), x^n(\hat{m}_1, m_2), y^n_1) \in T^{(n)}_\varepsilon \quad \text{for some } m_2$$
Analysis of the probability of error for decoder 1

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

\[
\begin{array}{c|c|c}
   m_1 & m_2 & \text{Joint pmf} \\
   \hline
   1 & 1 & p(u^n, x^n)p(y_1^n|y^n) \\
   * & 1 & p(u^n, x^n)p(y_1^n|u^n) \\
   * & * & p(u^n, x^n)p(y_1^n) \\
   1 & * & p(u^n, x^n)p(y_1^n)
\end{array}
\]

- Error events:
  \[
  \mathcal{E}_{11} = \{(U^n(1), X^n(1, 1), Y_1^n) \notin T_{\epsilon}^{(n)}\}
  \]
  \[
  \mathcal{E}_{12} = \{(U^n(1), X^n(m_1, 1), Y_1^n) \in T_{\epsilon}^{(n)} \text{ for some } m_1 \neq 1\}
  \]
  \[
  \mathcal{E}_{13} = \{(U^n(m_2), X^n(m_1, m_2), Y_1^n) \in T_{\epsilon}^{(n)} \text{ for some } m_1 \neq 1, m_2 \neq 1\}
  \]

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 arbitrary distributed
- Let $X^n(m) \sim \prod_{i=1}^n p_{X|U}(x_i|\tilde{u}_i), m \in \mathcal{A}, |\mathcal{A}| \leq 2^{nR},$ be pairwise conditionally independent of $\tilde{Y}^n$ given $\tilde{U}^n$

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

\[
\lim_{n \to \infty} P\{(\tilde{U}^n, X^n(m), \tilde{Y}^n) \in T_{\epsilon}^{(n)} \text{ for some } m \in \mathcal{A}\} = 0,
\]

if $R < I(X; Y|U) - \delta(\epsilon)$
Analysis of the probability of error for decoder 1

- Error events:
  \[ \mathcal{E}_{11} = \{(U^n(1), X^n(1, 1), Y^n_1) \notin \mathcal{T}_e^{(n)} \} \]
  \[ \mathcal{E}_{12} = \{(U^n(1), X^n(m_1, 1), Y^n_1) \in \mathcal{T}_e^{(n)} \text{ for some } m_1 \neq 1 \} \]
  \[ \mathcal{E}_{13} = \{(U^n(m_2), X^n(m_1, m_2), Y^n_1) \in \mathcal{T}_e^{(n)} \text{ for some } m_1 \neq 1, m_2 \neq 1 \} \]

- By LLN, \( P(\mathcal{E}_{11}) \rightarrow 0 \)

- By the packing lemma,
  \[ P(\mathcal{E}_{12}) \rightarrow 0 \text{ if } R_1 < I(X; Y_1 | U) - \delta(\varepsilon), \]
  \[ P(\mathcal{E}_{13}) \rightarrow 0 \text{ if } R_1 + R_2 < I(U, X; Y_1) - \delta(\varepsilon) = I(X; Y_1) - \delta(\varepsilon) \]

- 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: \( X \rightarrow Y_1 \rightarrow Y_2 \) form a Markov chain

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

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

for some $p(u, x)$ with $|U| \leq \min\{|X|, |Y_1|, |Y_2|\} + 1$

Achievability: Superposition coding + degradedness ($I(U; Y_2) \leq I(U; Y_1)$)

For BS-BC, the capacity region simplifies to the set of $(R_1, R_2)$ such that

$$R_1 \leq H(\alpha \ast p_1) - H(p_1),$$
$$R_2 \leq 1 - H(\alpha \ast p_2)$$

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 $P_e^{(n)} \to 0,$

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

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

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

By Fano’s inequality

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

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

Hence

$$nR_j \leq I(M_j; Y_j^n) + n\epsilon_n, \quad j = 1, 2$$
Proof of the converse

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

\[
I(M_1; Y_1^n) \leq I(M_1; Y_1^n | M_2) = I(M_1; Y_1^n | U) \\
= \sum_{i=1}^{n} I(M_1; Y_1i | U, Y_1^{i-1}) \\
\leq \sum_{i=1}^{n} I(M_1, Y_1^{i-1}; Y_1i | U) \\
= \sum_{i=1}^{n} I(X_i, M_1, Y_1^{i-1}; Y_1i | U) \\
= \sum_{i=1}^{n} I(X_i; Y_1i | U)
\]

- 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 $\leq I(U; Y_{2i})$

Proof of the converse (Gallager 1974)

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

\[
I(M_2; Y_2^n) \leq \sum_{i=1}^{n} I(M_2, Y_2^{i-1}; Y_{2i}) \\
\leq \sum_{i=1}^{n} I(M_2, Y_2^{i-1}, Y_1^{i-1}; Y_{2i})
\]

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 $X \rightarrow Y_1 \rightarrow Y_2$, thus $Y_2^{i-1} \rightarrow (M_2, Y_1^{i-1}) \rightarrow Y_{2i}$ and

\[
I(M_2; Y_2^n) \leq \sum_{i=1}^{n} I(U_i; Y_{2i})
\]
Proof of the converse (Gallager 1974)

- Define time-sharing r.v. $Q \sim \text{Unif}[1 : n]$, independent of $(M_1, M_2, X^n, Y^n_1, Y^n_2)$
- Let $U = (Q, U_Q)$, $X = X_Q$, $Y_1 = Y_{1Q}$, $Y_2 = Y_{2Q}$
- Clearly, $U \rightarrow X \rightarrow (Y_1, Y_2)$; hence

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

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

- Bound on cardinality of $U$ (NIT Appendix C)

- Remark: Proof works also with $U_i = (M_2, Y_{i-1}^2)$ or $U_i = (M_2, Y_{i-1}^1, Y_{i-1}^2)$

Gaussian broadcast channel

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

$$Z_1 \sim N(0, N_1) \quad \tilde{Z}_2 \sim N(0, N_2 - N_1)$$

$$X \rightarrow + \rightarrow Y_1 \rightarrow + \rightarrow Y_2$$
The capacity of the Gaussian BC is the set of \((R_1, R_2)\) such that
\[
R_1 \leq C(\alpha S_1), \\
R_2 \leq C\left(\frac{\alpha S_2}{\alpha S_2 + 1}\right)
\]
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 \(U \sim N(0, \bar{\alpha}P), V \sim N(0, \alpha P)\) 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
  \(Y_1 = X + Z_1, \ Y_2 = X + Z_2 = Y_1 + \tilde{Z}_2\)
- 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,
  \[ nR_1 \leq I(M_1; Y_1^n | M_2) + ne_n, \]
  \[ nR_2 \leq I(M_2; Y_2^n) + ne_n. \]

- Need to show that there exists an \( \alpha \in [0, 1] \) such that
  \[ I(M_1; Y_1^n | M_2) \leq n \left( \frac{\alpha P}{N_1} \right), \]
  \[ I(M_2; Y_2^n) \leq n \left( \frac{\bar{\alpha} S_2}{\alpha S_2 + 1} \right) = n \left( \frac{\bar{\alpha} P}{\alpha P + N_2} \right). \]

- Consider
  \[ I(M_2; Y_2^n) = h(Y_2^n) - h(Y_2^n | M_2) \leq \frac{n}{2} \log (2\pi e (P + N_2)) - h(Y_2^n | M_2) \]

- Since
  \[ \frac{n}{2} \log (2\pi e N_2) = h(Z_2^n) = h(Y_2^n | M_2, X^n) \leq h(Y_2^n | M_2) \leq h(Y_2^n) \leq \frac{n}{2} \log (2\pi e (P + N_2)), \]
  there must exist an \( \alpha \in [0, 1] \) such that
  \[ h(Y_2^n | M_2) = \frac{n}{2} \log (2\pi e (\alpha P + N_2)) \]

Proof of the converse (Bergmans 1974)

- Next consider
  \[ I(M_1; Y_1^n | M_2) = h(Y_1^n | M_2) - h(Y_1^n | M_1, M_2) \]
  \[ = h(Y_1^n | M_2) - h(Y_1^n | M_1, M_2, X^n) \]
  \[ = h(Y_1^n | M_2) - \frac{n}{2} \log (2\pi e N_1) \]

- Using the conditional vector EPI,
  \[ h(Y_2^n | M_2) = h(Y_1^n + \bar{Z}_2^n | M_2) \]
  \[ \geq \frac{n}{2} \log \left( 2^{\frac{n}{2} h(Y_1^n | M_2)/n} + 2^{\frac{n}{2} h(\bar{Z}_2^n | M_2)/n} \right) = \frac{n}{2} \log \left( 2^{\frac{n}{2} h(Y_1^n | M_2)/n} + 2\pi e (N_2 - N_1) \right) \]

- But since \( h(Y_2^n | M_2) = \frac{n}{2} \log (2\pi e (\alpha P + N_2)), \)
  \[ 2\pi e (\alpha P + N_2) \geq 2^{\frac{n}{2} h(Y_1^n | M_2)/n} + 2\pi e (N_2 - N_1) \Rightarrow h(Y_1^n | M_2) \leq \frac{n}{2} \log (2\pi e (\alpha P + N_1)) \]

- Hence
  \[ I(M_1; Y_1^n | M_2) \leq \frac{n}{2} \log (2\pi e (\alpha P + N_1)) - \frac{n}{2} \log (2\pi e N_1) = n \left( \frac{\alpha P}{N_1} \right). \]
Less noisy and more capable broadcast channels

- **Less noisy** if \( I(U; Y_1) \geq I(U; Y_2) \) for all \( p(u, x) \)
- **More capable** if \( I(X; Y_1) \geq I(X; Y_2) \) 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

\[
\begin{align*}
R_1 &\leq I(X; Y_1 | U), \\
R_2 &\leq I(U; Y_2), \\
R_1 + R_2 &\leq I(X; Y_1)
\end{align*}
\]

for some \( p(u, x) \), where \(|U| \leq \min\{|X|, |Y_1| \cdot |Y_2|\} + 2\)

**Example: A BSC and a BEC**

- For \( 0 \leq \epsilon \leq 2p \): \( Y_1 \) is a **degraded** version of \( Y_2 \)
- For \( 2p < \epsilon \leq 4p(1 - p) \): \( Y_2 \) is **less noisy** than \( Y_1 \), but not degraded
- For \( 4p(1 - p) < \epsilon \leq H(p) \): \( Y_2 \) is **more capable** than \( Y_1 \), but not less noisy
- For \( H(p) < \epsilon \leq 1 \): 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

\[
\begin{align*}
R_1 &< I(X; Y_1|U), \\
R_0 + R_2 &< I(U; Y_2), \\
R_0 + R_1 + R_2 &< I(X; Y_1)
\end{align*}
\]

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

  \[
  \begin{align*}
  R_1 &< I(X; Y_1|U_2), \\
  R_2 &< I(U_2; Y_2|U_3), \\
  R_3 &< I(U_3; Y_3)
  \end{align*}
  \]

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

- Capacity region for less noisy is not known for \(k \geq 4\)

- Capacity region for more capable is not known for \(k \geq 3\)
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


