Lecture #6  Broadcast Channels  
(Reading: NIT 5.1–5.7, 8.3, 9.6)

- Discrete memoryless broadcast channel
- Superposition coding inner bound
- Degraded broadcast channels
- Gaussian broadcast channel
- Less noisy and more capable broadcast channels
- Extensions
- Gaussian vector broadcast channel
- Marton’s inner bound with common message
- Nair–El Gamal outer bound

© Copyright 2002–2019 Abbas El Gamal and Young-Han Kim

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
  - Marton inner bound
  - Special classes of BCs for which the capacity region is known
  - 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:

- DM-BC with orthogonal components:

\[
p(y_i|x_i) \quad \text{for } i = 1, 2
\]

Binary symmetric BC

\[
X \sim \text{Bern}(p_1), \quad Z \sim \text{Bern}(p_2)
\]

Assume \( p_1 < p_2 < \frac{1}{2} \)

BS-BC: Superposition coding (Cover 1972)

\[
\{0, 1\}^n \quad \text{cloud (radius } \approx n\alpha)
\]

\[
\text{satellite codeword } x^n(m_1, m_2)
\]

\[
\text{cloud center } u^n(m_2)
\]

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 \( 2^n \) sequences \( u^n(m_1), m_1 \in [1 : 2^n] \)
- Independently generate \( 2^n \) sequences \( v^n(m_2), m_2 \in [1 : 2^n] \)
- To send \((m_1, m_2)\), transmit \( x^n(m_1, m_2) = u^n(m_2) \oplus v^n(m_1) \)

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 \( 2^n \) sequences \( u^n(m_1), m_1 \in [1 : 2^n] \)
- Independently generate \( 2^n \) sequences \( v^n(m_2), m_2 \in [1 : 2^n] \)
- 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 $y_2^n = u^n(m_2) \oplus (v^n(m_2) \oplus z^n_2)$:
  \[
  R_2 < I(U; Y_2) = 1 - H(\alpha * p_2)
  \]
- Decoder 1 uses successive cancellation decoding:
  - It recovers $m_2$ from $y^n_1 = u^n(m_2) \oplus (v^n(m_1) \oplus z^n_1)$:
    \[
    R_2 < I(U; Y_1) = 1 - H(\alpha * p_1) \quad (> 1 - H(\alpha * p_2))
    \]
  - Then recovers $m_1$ from $v^n(m_1) \oplus z^n_1$:
    \[
    R_1 < I(V; V \oplus Z_1) = H(\alpha * p_1) - H(p_1)
    \]

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
  \[
  R_1 < I(X; Y_1 | U),
  R_2 < I(U; Y_2),
  R_1 + R_2 < I(X; Y_1)
  \]
  for some pmf $p(u, x)$

- **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 $x^n(m_1, m_2) \sim \prod_{i=1}^{n} p_{X|U}(x_i|u(m_2)), m_1 \in [1: 2^{nR_1}]$
  - **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 $(u^n(\hat{m}_2), y^n_2) \in T^{(u)}_{\epsilon}$
      (by the packing lemma, $P(\mathcal{E}_2) \to 0$ as $n \to \infty$ if $R_2 < I(U; Y_2) - \delta(\epsilon)$)
    - Decoder 1 finds the unique message $\hat{m}_1$ such that
      \[
      (u^n(m_1), x^n(\hat{m}_1, m_2), y^n_1) \in T^{(u)}_{\epsilon} \quad \text{for some } m_2
      \]

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 $x^n(m_1, m_2) \sim \prod_{i=1}^{n} p_{X|U}(x_i|u(m_2)), m_1 \in [1: 2^{nR_1}]$
- **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 $(u^n(\hat{m}_2), y^n_2) \in T^{(u)}_{\epsilon}$
    (by the packing lemma, $P(\mathcal{E}_2) \to 0$ as $n \to \infty$ if $R_2 < I(U; Y_2) - \delta(\epsilon)$)
  - Decoder 1 finds the unique message $\hat{m}_1$ such that
    \[
    (u^n(m_1), x^n(\hat{m}_1, m_2), y^n_1) \in T^{(u)}_{\epsilon} \quad \text{for some } m_2
    \]
Analysis of the probability of error for decoder 1

- Consider $P(E)$ conditioned on $(M_1, M_2) = (1, 1)$

<table>
<thead>
<tr>
<th>$m_1$</th>
<th>$m_2$</th>
<th>Joint pmf</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>$p(u^n, x^n)p(y^n_1</td>
</tr>
<tr>
<td>*</td>
<td>1</td>
<td>$p(u^n, x^n)p(y^n_2</td>
</tr>
<tr>
<td>*</td>
<td>*</td>
<td>$p(u^n, x^n)p(y^n_2)$</td>
</tr>
<tr>
<td>1</td>
<td>*</td>
<td>$p(u^n, x^n)p(y^n_2)$</td>
</tr>
</tbody>
</table>

- Error events:
  \[ E_{11} = \{(u^n(1), x^n(1), y^n_1) \notin \mathcal{T}_\varepsilon^{(n)}\} \]
  \[ E_{12} = \{(u^n(1), x^n(m_1, 1), y^n_1) \in \mathcal{T}_\varepsilon^{(n)} \text{ for some } m_1 \neq 1\} \]
  \[ E_{13} = \{(u^n(m_2), x^n(m_1, m_2), y^n_1) \in \mathcal{T}_\varepsilon^{(n)} \text{ for some } m_1 \neq 1, m_2 \neq 1\} \]

- Analysis of the probability of error for decoder 1
  - Error events:
    \[ E_{11} = \{(u^n(1), x^n(1), y^n_1) \notin \mathcal{T}_\varepsilon^{(n)}\} \]
    \[ E_{12} = \{(u^n(1), x^n(m_1, 1), y^n_1) \in \mathcal{T}_\varepsilon^{(n)} \text{ for some } m_1 \neq 1\} \]
    \[ E_{13} = \{(u^n(m_2), x^n(m_1, m_2), y^n_1) \in \mathcal{T}_\varepsilon^{(n)} \text{ for some } m_1 \neq 1, m_2 \neq 1\} \]

- By LLN, $P(E_{11}) \to 0$

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

- 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

Packing lemma

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

Packing lemma

- There exists $\delta(\varepsilon) \to 0$ as $\varepsilon \to 0$ such that
  \[ \lim_{n \to \infty} P((\hat{U}^n, X^n(m), \hat{Y}^n) \notin \mathcal{T}_\varepsilon^{(n)} \text{ for some } m \in \mathcal{A}) = 0, \]

Degraded broadcast channels

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

  \[ Z_1 \sim \text{Bern}(p_1), \quad Z_2 \sim \text{Bern}(\tilde{p}_2), \quad X \to Y_1: \text{BSC}(p_1), \quad X \to Y_2: \text{BSC}(p_1 \ast \tilde{p}_2) \]

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

  \[ Z_1 \sim \text{Bern}(p_1), \quad Z_2 \sim \text{Bern}(p_2), \quad \tilde{Z}_1 \sim \text{Bern}(\tilde{p}_1), \quad \tilde{Z}_2 \sim \text{Bern}(\tilde{p}_2), \quad \tilde{p}_2 = (p_2 - p_1)/(1 - 2p_1) \]

- By LLN, $P(E_{11}) \to 0$

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

- 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
Achievability: Superposition coding + degradedness \( I(U; Y_2) \leq I(U; Y_1) \) for some \( p(u, x) \) with \( |U| \leq \min\{|X_1|, |Y_1|, |Y_2|\} + 1 \)

- **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 \( |U| \leq \min\{|X_1|, |Y_1|, |Y_2|\} + 1 \)

- For BS-BC, the capacity region simplifies to the set of \( (R_1, R_2) \) such that
  \[
  R_1 \leq H(\alpha * p_1) - H(p_1), \\
  R_2 \leq 1 - H(\alpha * p_2)
  \]
  for some \( \alpha \in [0, 1] \)

### Proof of the converse

- Let's try \( U = M_2 \) (satisfies \( U \rightarrow X_i \rightarrow (Y_{1i}, Y_{2i}) \))
  \[
  I(M_1; Y_1^n) \leq I(M_1; Y_1^n | M_2) = I(M_1; Y_1^n | U) = \sum_{i=1}^{n} I(M_1; Y_{1i}|U, Y_{1i}^{i-1}) \]
  \[
  \leq \sum_{i=1}^{n} I(M_1, Y_{1i}^{i-1}; Y_{1i}|U) = \sum_{i=1}^{n} I(X_i; M_1, Y_{1i}^{i-1}; Y_{1i}|U) = \sum_{i=1}^{n} I(X_i; Y_{1i}|U)
  \]

- Now consider the second inequality
  \[
  I(M_2; Y_2^n) = \sum_{i=1}^{n} I(M_2; Y_{2i}|Y_{2i}^{i-1}) = \sum_{i=1}^{n} I(U; Y_{2i}|Y_{2i}^{i-1})
  \]
  But \( I(U; Y_{2i}|Y_{2i}^{i-1}) \) 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
  \[
  I(M_2; Y_2^n) \leq \sum_{i=1}^{n} I(M_2, Y_1^{i-1}; Y_{2i}) \leq \sum_{i=1}^{n} I(M_2, Y_1^{i-1}, Y_{2i}^{i-1}; Y_{2i})
  \]
  But \( I(M_2, Y_1^{i-1}, Y_{2i}^{i-1}; Y_{2i}) \) 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 $(M_1, M_2, X^n, Y_1^n, Y_2^n)$
- 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
  \[ nR_1 \leq \sum_{i=1}^{n} I(X_i; Y_{1i}|U_i) + ne_n = nI(X; Y_1|U) + ne_n, \]
  \[ nR_2 \leq \sum_{i=1}^{n} I(U_i; Y_{2i}) + ne_n = nI(U_Q; Y_2|Q) + ne_n \leq nI(U; Y_2) + ne_n \]
- Bound on cardinality of $U$ (NIT Appendix C)
- Remark: Proof works also with $U_i = (M_2, Y_i^{-1}), i = 1, 2$

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

\[ R_1 \leq C(\alpha S_1), \]
\[ R_2 \leq C\left(\frac{\tilde{\alpha}S_2}{\alpha S_2 + 1}\right) \]

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 $U \sim N(0, \tilde{\alpha}P), V \sim N(0, aP)$ 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
  \[ Y_1 = X + Z_1, \quad Y_2 = X + Z_2 = Y_1 + \hat{Z}_2 \]
- 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
  \[ 2^{h(Y^n)/n} \geq 2^{h(X^n)/n} + 2^{h(Z^n)/n} \]
  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
  \[ 2^{h(Y^n|U)/n} \geq 2^{h(X^n|U)/n} + 2^{h(Z^n|U)/n} \]
By Fano's inequality,
\[ nR_1 \leq I(M_1; Y_1^n | M_2) + n\epsilon_n, \]
\[ nR_2 \leq I(M_2; Y_2^n) + n\epsilon_n. \]

Need to show that there exists an \( \alpha \in [0, 1] \) such that
\[ I(M_1; Y_1^n | M_2) \leq nC(\alpha S_1) = nC\left(\frac{\alpha P}{N_1}\right), \]
\[ I(M_2; Y_2^n) \leq nC\left(\frac{\alpha P}{\alpha S_2 + 1}\right) = nC\left(\frac{\alpha P}{\alpha P + N_2}\right). \]

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

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

Proof of the converse (Bergmans 1974)

Next consider
\[ I(M_1; Y_1^n | M_2) = h(Y_1^n | M_2) - h(Y_1^n | M_1, M_2) \]
\[ = h(Y_1^n | M_2) - h(Y_1^n | M_1, M_2, X^n) \]
\[ = h(Y_1^n | M_2) - \frac{n}{2} \log (2\pi e N_2) \]

Using the conditional vector \( \mathbf{E}_x \),
\[ h(Y_2^n | M_2) = h(Y_2^n + Z_2^n | M_2) \]
\[ \geq \frac{n}{2} \log \left( 2^{h(Y_2^n | M_2)} + 2^{h(Z_2^n | M_2) / n} \right) = \frac{n}{2} \log \left( 2^{h(Y_2^n | M_2) / n} + \pi e^2 N_2 - N_1 \right) \]

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

Hence
\[ I(M_1; Y_1^n | M_2) \leq \frac{n}{2} \log (2\pi e(aP + N_1)) - \frac{n}{2} \log (2\pi e N_1) = nC\left(\frac{\alpha P}{N_1}\right) \]

Example: A BSC and a BEC

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

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
\[ R_1 \leq I(X; Y_1 | U), \]
\[ R_2 \leq I(U; Y_2), \]
\[ R_1 + R_2 \leq I(X; Y_1) \]
for some \( p(u, x) \), where \(|U| \leq \min\{|X|, |Y_1|, |Y_2|\} + 2 \)
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

\[
R_1 < I(X; Y_1), \quad R_0 < I(U; Y_2), \quad R_0 + R_1 < I(X; Y_1)
\]

for some pmf \(p(u, x)\)

**Marton’s inner bound**

A simple inner bound: \((R_1, R_2)\) is achievable for the DM-BC \(p(y_1, y_2|x)\) if

\[
R_1 < I(U_1; Y_1), \quad R_2 < I(U_2; Y_2)
\]

for some pmf \(p(u_1)p(u_2)\) and function \(x(u_1, u_2)\)

**Theorem 8.3 (Marton 1979)**

\((R_1, R_2)\) is achievable if

\[
R_1 < I(U_1; Y_1), \quad R_2 < I(U_2; Y_2), \quad R_1 + R_2 < I(U_1; Y_1) + I(U_2; Y_2) - I(U_1; U_2)
\]

for some pmf \(p(u_1, u_2)\) and function \(x(u_1, u_2)\)

Region is not convex in general; can be convexified via \(Q\)

**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
  
  \[
  R_1 \leq I(X; Y_1|U_2), \quad R_2 \leq I(U_2; Y_2|U_3), \quad R_3 \leq I(U_3; Y_3)
  \]
  
  for some \(p(u_3)p(u_2|x|u_2)\)

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

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

**Semideterministic BC**

- Marton inner bound tight for semideterministic BC \((Y_1 = y_1(X))\): Set \(U_1 = Y_1\)

  The capacity region is the set of \((R_1, R_2)\) such that
  
  \[
  R_1 \leq H(Y_1), \quad R_2 \leq I(U; Y_2), \quad R_1 + R_2 \leq H(Y_1|U) + I(U; Y_2)
  \]
  
  for some \(p(u, x)\)

- Deterministic BC \((Y_1 = y_1(X), Y_2 = y_2(X))\): Further set \(U_2 = Y_2\)

  The capacity region is the set of \((R_1, R_2)\) such that
  
  \[
  R_1 \leq H(Y_1), \quad R_2 \leq H(Y_2), \quad R_1 + R_2 \leq H(Y_1, Y_2)
  \]
  
  for some \(p(x)\)
Example: Blackwell channel

Proof of achievability

- Use multicoding and the mutual covering lemma

Proof of achievability

- Codebook generation: Fix \( p(u_1, u_2) \) and \( x(u_1, u_2) \)
  - For each \( m_1 \in [1 : 2^{nR_1}] \) generate a subcodebook \( C_1(m_1) \) consisting of \( 2^{n(R_1-R_2)} \) sequences
    \[ u_1^T(l_1) = \prod_{i=1}^n p_{C_1}(u_{i1}), l_1 \in \{m_1 - 1 : 2^{nR_1-R_2} + 1 : m_1 2^{nR_1-R_2}\} \]
  - Similarly, generate \( C_2(m_2), m_2 \in [1 : 2^{nR_2}] \)

- Codebook generation: Fix \( p(u_1, u_2) \) and \( x(u_1, u_2) \)
  - For each \( (m_1, m_2) \), find \( (l_1, l_2) \) such that \( u_1^T(l_1) \in C_1(m_1), u_2^T(l_2) \in C_2(m_2) \)
  - If no such pair exists, choose \( (l_1, l_2) = (1, 1) \)
  - Generate \( x^T(m_1, m_2) \) as \( x^T(m_1, m_2) = x(u_{11}(l_1), u_{21}(l_2)), i \in [1 : n] \)
Proof of achievability

- **Encoding:**
  - To send a message pair \((m_1, m_2)\), transmit \(x^n(m_1, m_2)\).

- **Decoding:**
  - Decoder \(j = 1, 2\) finds unique \(\hat{m}_j\) such that \((u^n_j(l_j), y^n_j) \in T^{(n)}_{c_j}\) for some \(u^n_j(l_j) \in C_j(\hat{m}_j)\).

Analysis of the probability of error

- Consider \(P(\mathcal{E})\) conditioned on \((M_1, M_2) = (1, 1)\).
- Let \((L_1, L_2)\) denote the pair of chosen indices.
- Error events for decoder 1:

\[
\begin{align*}
\mathcal{E}_0 &= \{(U^n_1(l_1), U^n_2(l_2)) \notin T^{(n)}_{c_1} \text{ for all } (U^n_1(l_1), U^n_2(l_2)) \in C_1(1) \times C_2(1)\}, \\
\mathcal{E}_{11} &= \{(U^n_1(l_1), Y^n_1) \notin T^{(n)}_{c_1}\}, \\
\mathcal{E}_{12} &= \{(U^n_1(l_1), Y^n_1) \in T^{(n)}_{c_1}(U_1, Y_1) \text{ for some } l_1 \notin [1: 2^{n(R_1 - R_1)}]\}.
\end{align*}
\]

Thus, by the union of events bound,

\[
P(\mathcal{E}) \leq P(\mathcal{E}_0) + P(\mathcal{E}_0 \cap \mathcal{E}_{11}) + P(\mathcal{E}_{12})
\]
Mutual covering lemma ($U_0 = 0$)

- Let $(U_1, U_2) \sim p(u_1, u_2)$ and $\epsilon' < \epsilon$
- For $j = 1, 2$, let $U_j^n(m_j) \sim \prod_{i=1}^{j} p(u_i(u_i))$, $m_j \in [1:2^{m_j}]$, be pairwise independent
- Assume that $\{U_1^n(m_1)\}$ and $\{U_2^n(m_2)\}$ are independent

Lemma 8.1 (Mutual covering lemma)

There exists $\delta(\epsilon) \to 0$ as $\epsilon \to 0$ such that
\[
\lim_{n \to \infty} P[U_1^n(m_1), U_2^n(m_2) \notin \mathcal{T}_\epsilon^{(n)} \text{ for all } m_1 \in [1:2^{n_1}], m_2 \in [1:2^{n_2}]] = 0
\]
if $r_1 + r_2 > I(U_1; U_2) + \delta(\epsilon)$

- Proof: See NIT Appendix 8A

This lemma extends the covering lemma:

- For a single $U_1^n$ sequence ($r_1 = 0$), $r_2 > I(U_1; U_2) + \delta(\epsilon)$ as in the covering lemma
- Pairwise independence: linear codes for finite field models

Analysis of the probability of error

- Error events for decoder 1:
  \[
  \mathcal{E}_0 = \{(U_1^n(l_1), U_2^n(l_2)) \notin \mathcal{T}_\epsilon^{(n)} \text{ for all } (U_1^n(l_1), U_2^n(l_2)) \in C_1(1) \times C_2(1)\},
  \mathcal{E}_{11} = \{(U_1^n(l_1), Y_1^n) \notin \mathcal{T}_\epsilon^{(n)}\},
  \mathcal{E}_{12} = \{(U_1^n(l_1), Y_1^n) \in \mathcal{T}_\epsilon^{(n)}(U_1, Y_1) \text{ for some } l_1 \notin [1:2^{(R_1 - R_1^*)}]\}
  \]
  - By the mutual covering lemma (with $r_1 = R_1 - R_1$ and $r_2 = R_2 - R_2$), $P(\mathcal{E}_0) \to 0$ if $(R_1 - R_1) + (R_2 - R_2) > I(U_1; U_2) + \delta(\epsilon')$
  - By the conditional typicality lemma, $P(\mathcal{E}_{11} \cap \mathcal{E}_{11}) \to 0$
  - By the packing lemma, $P(\mathcal{E}_{12}) \to 0$ if $\hat{R}_1 < I(U_1; Y_1) - \delta(\epsilon)$
  - Similarly, $P(\mathcal{E}_2) \to 0$ if $\hat{R}_2 < I(U_2; Y_2) + \delta(\epsilon)$
  - Using Fourier–Motzkin to eliminate $\hat{R}_1$ and $\hat{R}_2$ completes the proof

Application: Gaussian BC

- Consider the Marton coding scheme
- Fix $p(u_1, u_2)$ and $x(u_1, u_2)$. This defines a pentagon region
- Consider corner point ($R_1 = I(U_1; Y_1) - I(U_1; U_2), R_2 = I(U_2; Y_2)$)
- Marton scheme for communicating $M_1$ is equivalent to G–P for $p(y_1|u_1, u_2)p(u_2)$

Decompose $X$ into the sum of independent $X_1 \sim N(0, aP)$ and $X_2 \sim N(0, \hat{a}P)$

- Send $M_2$ to $Y_2 = X_2 + X_1 + Z_2; R_2 < C(\hat{a}P/(aP + N_2))$ (treat $X_1$ as noise)
- Send $M_1$ to $Y_1 = X_1 + X_2 + Z_1; R_1 < C(aP/N_1)$ (writing on dirty paper)
  - Substitute $U_1 = X_2$ and $U_1 = \beta U_2 + X_1, \beta = aP/(aP + N_1)$ in Marton
- This coding scheme works even when $N_1 > N_2$ (unlike superposition coding)
Theorem 9.4 (Weingarten–Steinberg–Shamai 2006)
\( C \) is the convex closure of \( R_1 \cup R_2 \)

Marton’s inner bound with common message

\( (R_0, R_1, R_2) \) is achievable if\( R_0 + R_1 < I(U_0; U_1; Y_1), \)
\( R_0 + R_2 < I(U_0; U_2; Y_2), \)
\( R_0 + R_1 + R_2 < I(U_0; U_1; Y_1) + I(U_2; Y_2; U_0) < I(U_1; U_2; U_0), \)
\( 2R_0 + R_1 + R_2 < I(U_0; U_1; Y_1) + I(U_0; U_2; Y_2) < I(U_1; U_2; U_0) \)

for some \( p(u_0, u_1, u_2) \) and function \( x(u_0, u_1, u_2) \)

Proof of achievability for \( R_2 \)

• Decompose \( X \) into the sum of independent \( X_1 \sim N(0, K_1) \) and \( X_2 \sim N(0, K_2) \)

• Send \( M_2 \) to \( Y_2 = G_2 X_2 + G_2 X_1 + Z_2; R_2 < \frac{1}{2} \log \frac{|G_2 K_2 G_2^T + G_1 K_1 G_2^T + I_1|}{|G_2 K_2 G_2^T + I_1|} \)

• Send \( M_1 \) to \( Y_1 = G_1 X_1 + G_1 X_2 + Z_1; R_1 < \frac{1}{2} \log |G_1 K_1 G_1^T + I_1| \) (vector WDP)

\( R_1 \) is achieved similarly

Capacity region

\( R_1: (R_1, R_2) \) such that
\( R_1 < \frac{1}{2} \log \frac{|G_1 K_1 G_1^T + G_1 K_2 G_1^T + I_1|}{|G_1 K_1 G_1^T + I_1|} \),
\( R_2 < \frac{1}{2} \log \frac{|G_2 K_2 G_2^T + G_1 K_1 G_2^T + I_1|}{|G_2 K_2 G_2^T + I_1|} \)

for some \( K_1, K_2 \geq 0 \) with \( \text{tr}(K_1 + K_2) \leq P \)

\( R_2 \): \( (R_1, R_2) \) such that
\( R_1 < \frac{1}{2} \log |G_1 K_1 G_1^T + I_1| \),
\( R_2 < \frac{1}{2} \log \frac{|G_2 K_2 G_2^T + G_1 K_1 G_2^T + I_1|}{|G_2 K_2 G_2^T + I_1|} \)

for some \( K_1, K_2 \geq 0 \) with \( \text{tr}(K_1 + K_2) \leq P \)

Gaussian vector broadcast channel

\( Z_1, Z_2 \sim N(0, I_2) \)

Average power constraint: \( \sum_{i=1}^{n} x_i(m_1, m_2, i) x(m_1, m_2, i) \leq nP \)

Channel is not degraded in general (superposition coding not optimal)

Marton coding (vector writing on dirty paper) is optimal, however

Proof of achievability for \( R_2 \)

(R_0, R_1, R_2) is achievable if
\( R_0 + R_1 < I(U_0; U_1; Y_1), \)
\( R_0 + R_2 < I(U_0; U_2; Y_2), \)
\( R_0 + R_1 + R_2 < I(U_0; U_1; Y_1) + I(U_2; Y_2; U_0) < I(U_1; U_2; U_0), \)
\( 2R_0 + R_1 + R_2 < I(U_0; U_1; Y_1) + I(U_0; U_2; Y_2) < I(U_1; U_2; U_0) \)

for some \( p(u_0, u_1, u_2) \) and function \( x(u_0, u_1, u_2) \)

Proof of achievability: Superposition coding + Marton coding

Tight for all classes of BCs with known capacity regions

Even for \( R_0 = 0 \), larger than Marton’s inner bound with \( U_0 = 0 \) (Theorem 8.3)
## Theorem 8.6 (Nair–El Gamal 2007)

If $(R_0, R_1, R_2)$ is achievable, then

\[
R_0 \leq \min\{I(U_0; Y_1), I(U_0; Y_2)\},
R_0 + R_1 \leq I(U_0, U_1; Y_1),
R_0 + R_2 \leq I(U_0, U_2; Y_2),
R_0 + R_1 + R_2 \leq I(U_0, U_1; Y_1) + I(U_2; Y_2 | U_0, U_1),
R_0 + R_1 + R_2 \leq I(U_1; Y_1 | U_0, U_2) + I(U_0, U_2; Y_2)
\]

for some $p(u_1, u_2)p(u_1, u_1, u_2)$ and function $x(u_0, u_1, u_2)$

- Tight for all BCs with known capacity regions that we discussed so far
- Does not coincide with Marton's inner bound (Jog–Nair 2010)
- Not tight in general (Geng–Gohari–Nair–Yu 2011)

## 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


