1 Broadcast Channels:

![Broadcast Channel Diagram](image)

**Figure 1:** Broadcast Channel.

**Example 1** [Television broadcast.] MAC($J, 1, L$). Every one needs to get everything.

**Example 2** [Spanish and Dutch simultaneous newscast.]
- Transmitter can emit 1 word/second in either language.
- $2^{20}$ words $\Rightarrow$ 20 bits/second to an user and 0 to the other.
- Timesharing gets all rate pairs in $R_1 + R_2 \leq 20$ bits/second.
- We can get more. Suppose $R_1 = R_2 = 10$ bits/second. By ordering Spanish and Dutch: \( n/2 \) $\Rightarrow$ 1 bit/second extra (common, or can go to either user).

**Example 3** [Blackwell Channel.] We anticipate $C_{BC} \subseteq \mathcal{S}$, where

\[
\mathcal{S} = \text{clo} \{ (R_1, R_2) \in \mathbb{R}_+^2 : R_1 \leq H(Y_1|Q), R_2 \leq H(Y_2|Q), R_1 + R_2 \leq H(Y_1Y_2|Q), \text{ for some } Z = QXY_1Y_2 \in \mathcal{S} \}
\]

\[
= \text{clo} \text{ conv} \{ (R_1, R_2) \in \mathbb{R}_+^2 : R_1 \leq H(Y_1), R_2 \leq H(Y_2), R_1 + R_2 \leq H(Y_1Y_2), \text{ for some } Z = XY_1Y_2 \in \mathcal{S} \}
\]

**Example 4** [Scalar Gaussian Channel.]

\[
\begin{align*}
Y_1 &= X + \xi_1 \\
Y_2 &= X + \xi_2
\end{align*}
\]

- $X, Y_k, \xi_k \in \mathbb{R}, k = 1, 2$
- $\mathbb{E}X^2 \leq P$
- $\xi_k \sim N(0, \sigma^2), k = 1, 2$

**Example 5** [Vector Gaussian Channel.]

\[
\begin{align*}
Y_1 &= H_1X + \xi_1 \\
Y_2 &= H_2X + \xi_2
\end{align*}
\]

- $X, \xi_k \in \mathbb{R}^n, Y_k \in \mathbb{R}^{n_{rc}}, k = 1, 2$
\begin{itemize}
  \item $\text{E}\|X\|^2 \leq P$
  \item $\xi_k \sim \mathcal{N}(0, C_k), k = 1, 2$
\end{itemize}

**Definition 1 (DM–BC).** A (two user) discrete memoryless broadcast channel (DM–BC) denoted by $(\mathcal{X}, \mathcal{Y}_1, \mathcal{Y}_2, p_{Y_1Y_2|X}(y_1y_2|x))$, consists of three finite sets $\mathcal{X}, \mathcal{Y}_1$, and $\mathcal{Y}_2$, and a collection of probability mass functions $p_{Y_1Y_2|X}(\cdot|\cdot)$ on $\mathcal{Y}_1 \times \mathcal{Y}_2$, one for each $x \in \mathcal{X}$, with the interpretation that $X$ is the input and $Y_k$ is the output for user $k$. For $n \in \mathbb{N}$, with $X^n = (X_1, X_2, \cdots, X_n)$ as inputs, the output sequence $Y^n_k$ has pmf

\begin{equation}
  p_{Y^n_k|X^n}(y^n_k|x^n) = \prod_{i=1}^{n} p_{Y_k|x}(y_k|x_i)
\end{equation}

**Definition 2 (Code).** An $(n, M_1, M_2)$ code for the channel $(\mathcal{X}, \mathcal{Y}_1, \mathcal{Y}_2, p_{Y_1Y_2|X}(y_1y_2|x))$ consists of the following:

1. An index set of messages for each user $k$, $\mathcal{W}_k = [M_k]$.
2. An encoder $f$, $f : [M_1] \times [M_2] \rightarrow \mathcal{X}$. The codebook can be represented by an ordered set $c = \{f(\{1, 1\}), f(\{1, 2\}), \cdots, f(\{1, M_2\}); f(\{2, 1\}), f(\{2, 2\}), \cdots, f(\{M_1, M_2\})\}$.
3. A decoding rule, $g_k : \mathcal{Y}_k^n \rightarrow \phi \cup \mathcal{W}_k, k = 1, 2$, i.e., $y^n_k \mapsto g_k(y^n_k) = \tilde{w}_k \in \phi \cup \mathcal{W}_k$.

**Definition 3 (Probability of error).** Let $W_1W_2$ be the message transmitted and let $Y^n_k$ be the signal received. The conditional probability of error of user $k$ when $(W_1W_2) = (w_1w_2)$ was transmitted is given by

\[ P_{e,w_1,w_2}^{(n)}(c; k) = \text{Pr}\left\{ g_k(Y^n_k) \neq W_k | W_1W_2 = w_1w_2 \right\}, k = 1, 2. \]

The average probability of error of user $k$ for the code $c$ is given by

\[ P_{e}^{(n)}(c; k) = \frac{1}{M_1M_2} \sum_{w_1w_2} P_{e,w_1,w_2}^{(n)}(c; k), k = 1, 2 \]

Note that the above equation assumes that all messages are equally likely and the users choose their messages independently.

**Definition 4 (Achievability).** The rate pair $(R_1, R_2)$ is achievable, if for every $\eta > 0, \lambda \in (0, 1)$, there exists a sequence of $(n, M_1, M_2)$ codes that satisfy

1. $P_{e}^{(n)}(c; k) \leq \lambda, k = 1, 2, \text{ and}$
2. $\frac{\log_2 M_k}{n} \geq R_k - \eta, k = 1, 2.$

**Definition 5 (Capacity region).** The capacity region is the set of all achievable rate pairs, denoted by $\mathcal{C}_{BC}$.

**Remark 1.**
\begin{itemize}
  \item $\mathcal{C}_{BC}$ is closed and convex. Moreover, $R \in \mathcal{C}_{BC}$, $R$ dominates $r \implies r \in \mathcal{C}_{BC}$.
  \item $\mathcal{C}_{BC}$ depends on $p_{Y|X}$ only through $p_{Y_1|X}$ and $p_{Y_2|X}$.
  \item Yet again, we are looking for a single letter characterisation (open).
  \item Solved:
    \begin{enumerate}
      \item Degraded BC, component-wise degradation (Bergmans, Gallager).
    \end{enumerate}
\end{itemize}

Lecture 12 : Broadcast channels-2
2. Vector MIMO GBC (degraded or not) (Weingarten, Steinberg, Shamai).
4. One user’s channel is deterministic (Gelfand–Pinsker–Marton).
5. Less noisy, more capable (Körner–Marton, ElGamal).

Definition 6 (Degraded). \( p_{Y_2|X} \) is a degraded form of \( p_{Y_1|X} \) if there exists \( Y_1 \) such that
\[
p_{Y_2|X}(y_2|x) = \sum_{y_1} p_{Y_1|X}(y_1|x)p_{Y_1|Y_2}(y_1|y_2)
\]

Definition 7 (Less noisy). \( p_{Y_2|X} \) is less noisy than \( p_{Y_1|X} \) if \( U \rightarrow X \rightarrow Y_2 \) implies \( I(U;Y_1) \geq I(U;Y_2) \).

Definition 8 (More capable). \( p_{Y_1|X} \) is more capable than \( p_{Y_2|X} \) if \( I(U;Y_1) > I(U;Y_2) \) for some \( p_X \).

Remark 2.
- Degraded \( \implies \) less noisy \( \implies \) more capable.

Exercise: 1) Identify a less noisy channel that is not degraded. 2) Also, identify a more capable channel that is not less noisy.

Example 6 [Degraded and therefore less noisy and more capable channel]
\[
\begin{align*}
Y_1 &= X + \xi_1, \quad \xi_1 \sim N(0, \sigma_1^2) \\
Y_2 &= X + \xi_2, \quad \xi_2 \sim N(0, \sigma_2^2), \sigma_2 \geq \sigma_1 \\
&= X + \tilde{\xi}_1 + \tilde{\xi}_2
\end{align*}
\]

Example 7 [BSC(\( p_1 \)), BSC(\( p_2 \)), \( p_1 \leq p_2 < 1/2 \)]

Suppose we write \( p_2 = (1-p_1)\alpha + (1-\alpha)p_1 = p_1 + \alpha(1-2p_1) > p_1 \)

2 Coding for the degraded BC

We expect user 1 should decode user 2’s information. Let \( U \) be an auxiliary random variable representing user 2’s message.

Theorem 1. If \( X \rightarrow Y_1 \rightarrow Y_2 \) (degraded BC) then
\[
\mathcal{C} = \text{clo. conv. } \left\{ (R_1, R_2) \in \mathbb{R}_+^2 : \begin{array}{c}
R_1 \leq I(X;Y_1|U) \\
R_2 \leq I(U;Y_2)
\end{array} \text{ for some } U \rightarrow X \rightarrow Y_1 \rightarrow Y_2 \text{ with } |U| < \infty. \right\}
\]

- Achievability can be proved by random-coding argument.
- Converse is an exercise.
- We will prove both as specialisations of a more general result.
- Converse is unnecessary. The set with union over all \( U \rightarrow X \rightarrow Y_1 \rightarrow Y_2 \) is already converse.

Example 8 [D–GBC]

Achievability:
\[
\begin{align*}
U &\sim N(0, (1-t)P) \\
X &= U + V, \ V \sim N(0, tP), V \perp U \\
I(U;Y_2) &= \frac{1}{2} \log \left( 1 + \frac{(1-t)P}{tP + \sigma_2^2} \right) \\
I(X;Y_1|U) &= \frac{1}{2} \log \left( 1 + \frac{tP}{\sigma_1^2} \right)
\end{align*}
\]

Lecture 12 : Broadcast channels-3
\[ C' = \text{clo. conv.} \left\{ (R_1, R_2) \in \mathbb{R}_+^2 : \begin{array}{c}
R_1 \leq \frac{1}{2} \log \left( 1 + \frac{tP}{\sigma^2} \right), \\
R_2 \leq \frac{1}{2} \log \left( 1 + \frac{(1-t)P}{\sigma^2} \right), \end{array} \right\} \]

Direct part is clear. Converse requires entropy power inequality (along with the more general GBC).

**Example 9 [BSC(p_1), BSC(p_2), p_1 \leq p_2 < 1/2]**

Let \( U \in \{0,1\} \), equiprobable; \( U \xrightarrow{\text{BSC(\beta)}} X \xrightarrow{\text{BSC(p_1)}} Y_1 \xrightarrow{\text{BSC(\alpha)}} Y_2 \)

\[
I(U; Y_2) = \begin{cases} 1 - H(\beta \ast p_2), & \text{where } \beta \ast p_2 = \beta(1 - p_2) + (1 - \beta)p_2 \\ H(Y_1 | U) - H(Y_1 | X), & \text{where } H(Y_1 | X) \\ H(\beta \ast p_1) - H(p_1). & \end{cases}
\]

\[ C' = \text{clo. conv.} \left\{ (R_1, R_2) \in \mathbb{R}_+^2 : \begin{array}{c}
R_1 \leq H(\beta \ast p_1) - H(p_1) \\
R_2 \leq 1 - H(\beta \ast p_2), \\
\beta \in [0,1/2] \end{array} \right\} \]

Direct part is clear. Converse requires showing that uniform \( U \), symmetric \( \text{BSC(\beta)} \) are sufficient to get the largest region.

## 3 DBC

- \( \mathcal{P} := \left\{ Z = U_{[3]} XY_{[2]} \text{ such that satisfy (0)–(3) hold} \right\} : \\
  (0) U_{[3]} \in U_{[3]} \text{ an arbitrary finite set;}
  (1) X \in \mathcal{X}, Y_k \in \mathcal{Y}_{k}, k = 1, 2 \text{ where } \mathcal{X}, \mathcal{Y}_k \text{ are arbitrary finite sets;}
  (2) p_Z = pu_{[3]}p_{X|U_{[3]}}p_{Y_{[2]}|X}, \text{i.e., } U_{[3]} \rightarrow X \rightarrow Y_{[2]};
  (3) p_{Y_{[2]}|X} \text{ is the given channel.} \]

- \( \mathcal{R}(Z) := \left\{ (R_1, R_2) \in \mathbb{R}_+^2 : \begin{array}{c}
R_1 \leq I(U_0U_1; Y_1) \\
R_2 \leq I(U_0U_2; Y_2) \\
R_1 + R_2 \leq \min \{ I(U_0; Y_1), I(U_0; Y_2) \} + I(U_1; Y_1U_0) + I(U_2; Y_2U_0) - I(U_1; U_2|U_0) \end{array} \right\} \]

- \( \mathcal{R} := \text{clo.} \bigcup_{Z \in \mathcal{P}} \mathcal{R}(Z) \)

**Remark 3.**

- \( U_0U_1U_2 \) independent (appear to decrease the capacity region, but other components may be large to compensate for \( I(U_1; U_2|U_0) \)).
- \( U_0 \) is decoded by both, \( U_k \) by user \( k \).

**Lemma 2.** \( \mathcal{R} \) is convex.

**Proof.** We will prove \( \bigcup_{Z \in \mathcal{P}} \mathcal{R}(Z) \) is convex. Let \( R^{(\ell)} \in \bigcup_{Z \in \mathcal{P}} \mathcal{R}(Z), \ell = 1, 2. \) Then \( \exists Z^{(\ell)}, \ell = 1, 2 \) such that \( R^{(\ell)} \in \mathcal{R}(Z^{(\ell)}). \) Fix \( \lambda \in (0,1) \), arbitrary.
• Define \( Z = \mathcal{U}_0 U_1 U_2 X Y_1 Y_2 \) as follows:
\[
\mathcal{U}_0 U_1 U_2 X Y_1 Y_2 = \left( (Q U_0^{(Q)}), U_1^{(Q)}, U_2^{(Q)} \right), \quad Q \in \{1, 2\}, \quad p_Q(1) = \lambda.
\]
Therefore, \( p_{\mathcal{U}_0 U_1 U_2} = p_Q(\ell)p_{U_0^{(Q)}}p_{U_1^{(Q)}}p_{U_2^{(Q)}} \).

Since \( Z^{(\ell)} \in \mathcal{P} \), so does \( Z \) since 1) given \( U_0^{(\ell)} U_1^{(\ell)} U_2^{(\ell)} \), \( X \) does not depend on \( \ell \), and 2) given \( X \), \( Y_{[2]} \) does not depend on \( \mathcal{U}_0 U_1 U_2 \).

• Consider terms with \( I(\cdot;U_0^{(\ell)}) \). Averaging, we get \( I(\cdot;\mathcal{U}_0) \).

• \( I(U_0^{(\ell)}; Y_k^{(\ell)}) \) : Averaging, we get \( I(U_0; Y_k | Q) \leq I(Q U_0; Y_k) = I(\mathcal{U}_0; Y_k) \).

\[
\lambda \min \left\{ I(U_0^{(1)}; Y_1^{(1)}), I(U_0^{(1)}; Y_2^{(1)}) \right\} + (1 - \lambda) \min \left\{ I(U_0^{(2)}; Y_1^{(2)}), I(U_0^{(2)}; Y_2^{(2)}) \right\}
\leq \min \left\{ I(U_0; Y_1 | Q), I(U_0; Y_2 | Q) \right\}
\leq \min \left\{ I(U_0 Q; Y_1), I(U_0 Q; Y_2) \right\}
= \min \left\{ I(\mathcal{U}_0; Y_1), I(\mathcal{U}_0; Y_2) \right\}
\]

• So \( \lambda R^{(1)} + (1 - \lambda) R^{(2)} \in \mathcal{R}(Z) \).

• Closure of a convex set is convex.

\[\square\]

Remark 4. Convexification is not necessary.