**E2–301 Topics in Multiuser Communication** October 03, 2007

# Lecture 12 : Broadcast channels

*Instructor: Rajesh Sundaresan* *Scribe: Premkumar K.*

## 1 Broadcast Channels:

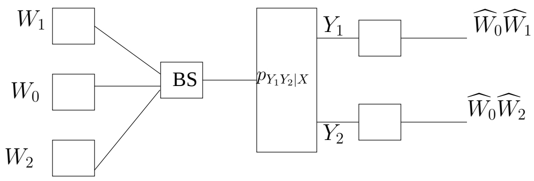


**Figure 1**: Broadcast Channel.

**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 $\implies$ 20 bits/second to an user and 0 to the other.
- Timesharing gets all rate pairs in $R_1 + R_2 \leqslant 20$ bits/second.
- We can get more. Suppose $R_1 = R_2 = 10$ bits/second. By ordering Spanish and Dutch : $\binom{n}{n/2} \doteq 2^{nh(1/2)} \implies$ 1 bit/second extra (common, or can go to either user).

**Example 3** [Blackwell Channel.] We anticipate $\mathscr{C}_{BC} \subseteq \mathscr{R}$, where

$$
\begin{aligned}
\mathscr{R} &= \text{clos.}\Big\{(R_1, R_2) \in \mathbb{R}_+^2 : R_1 \leqslant H(Y_1|Q), R_2 \leqslant H(Y_2|Q), R_1 + R_2 \leqslant H(Y_1Y_2|Q), \quad \text{for some } Z = QXY_1Y_2 \in \mathscr{P}^*\Big\} \\
&= \text{clos. conv.}\Big\{(R_1, R_2) \in \mathbb{R}_+^2 : R_1 \leqslant H(Y_1), R_2 \leqslant H(Y_2), R_1 + R_2 \leqslant H(Y_1Y_2), \quad \text{for some } Z = XY_1Y_2 \in \mathscr{P}\Big\}
\end{aligned}
$$

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

$$
\begin{aligned}
Y_1 &= X + \xi_1 \\
Y_2 &= X + \xi_2
\end{aligned}
$$

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

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

$$
\begin{aligned}
Y_1 &= H_1X + \xi_1 \\
Y_2 &= H_2X + \xi_2
\end{aligned}
$$

- $X, \xi_k \in \mathbb{R}^{n_t}, Y_k \in \mathbb{R}^{n_r}, k = 1, 2$

- $\mathbb{E}\|X\|^2 \leqslant P$
- $\xi_k \sim \mathcal{N}(\mathbf{0}, C_k), k = 1, 2$

**Definition 1 (DM–BC).** *A (two user) discrete memoryless broadcast channel (DM–BC) denoted by* $\left(\mathbb{X}, \mathbb{Y}_1, \mathbb{Y}_2, p_{Y_1Y_2|X}(y_1y_2|x)\right)$*, consists of three finite sets* $\mathbb{X}, \mathbb{Y}_1$ *and* $\mathbb{Y}_2$*, and a collection of probability mass functions* $p_{Y_1Y_2|X}(\cdot|x)$ *on* $\mathbb{Y}_1 \times \mathbb{Y}_2$*, one for each* $x \in \mathbb{X}$*, 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* $Y_{[2]}^n$ *has pmf*

$$p_{Y_{[2]}^n|X^n}(y_{[2]}^n|x^n) \quad = \quad \prod_{i=1}^{n} p_{Y_{[2]}|X}(y_{[2]i}|x_i) \tag{1}$$

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

1. *An index set of messages for each user* $k$*,* $\mathbb{W}_k = [M_k]$*.*
2. *An encoder* $f$*,* $f : [M_1] \times [M_2] \to \mathbb{X}^n$*. The codebook can be represented by an ordered set*

$$c \quad = \quad \left\{f(\{1,1\}), f(\{1,2\}), \cdots, f(\{1,M_2\}); f(\{2,1\}), f(\{2,2\}), \cdots, f(\{M_1,M_2\})\right\}.$$

3. *A decoding rule,* $g_k : \mathbb{Y}_k^n \to \phi \cup \mathbb{W}_k, k = 1, 2$*, i.e.,* $y_k^n \mapsto g_k(y_k^n) = \widehat{w}_k \in \phi \cup \mathbb{W}_k$*.*

**Definition 3 (Probability of error).** *Let* $W_1W_2$ *be the message transmitted and let* $Y_{[2]}^n$ *be the signal received. The conditional probability of error of user* $k$ *when* $(W_1W_2) = (w_1w_2)$ *was transmitted is given by*

$$P_{e,w_1w_2}^{(n)}(c; k) \quad = \quad Pr\left\{g_k(Y_k^n) \neq W_k \middle| W_1W_2 = w_1w_2\right\}, k = 1, 2.$$

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

$$P_e^{(n)}(c; k) \quad = \quad \frac{1}{M_1M_2} \sum_{w_1w_2} P_{e,w_1w_2}^{(n)}(c; k), k = 1, 2$$

*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. $P_e^{(n)}(k) \underset{n}{\leqslant} \lambda, k = 1, 2$*, and*
2. $\frac{\log_2 M_k}{n} \underset{n}{\geqslant} R_k - \eta, k = 1, 2.$

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

**Remark 1.**

- $\mathscr{C}_{BC}$ *is closed and convex. Moreover,* $R \in \mathscr{C}_{BC}$*,* $R$ *dominates* $r \implies r \in \mathscr{C}_{BC}$*.*
- $\mathscr{C}_{BC}$ *depends on* $p_{Y_{[2]}|X}$ *only through* $p_{Y_1|X}$ *and* $p_{Y_2|X}$*.*
- *Yet again, we are looking for a single letter characterisation (open).*
- *Solved:*
  1. *Degraded BC, component-wise degradation (Bergmans, Gallager).*

2. *Vector MIMO GBC (degraded or not) (Weingarten, Steinberg, Shamai).*
3. *Deterministic BC, (Pinsker–Marton).*
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 $\exists p_{Y_2|Y_1}$ such that*

$$p_{Y_2|X}(y_2|x) = \sum_{y_1} p_{Y_2|Y_1}(y_2|y_1)p_{Y_1|X}(y_1|x)$$

**Definition 7 (Less noisy).** *$p_{Y_2|X}$ is less noisy than $p_{Y_1|X}$ if $U \to X \to Y_{[2]}$ implies $I(U;Y_1) \geqslant I(U;Y_2)$.*

**Definition 8 (More capable).** *$p_{Y_1|X}$ is more capable than $p_{Y_2|X}$ if $I(U;Y_1) \geqslant I(U;Y_2), \forall p_X$.*

**Remark 2.**
- *Degraded $\Longrightarrow$ less noisy $\Longrightarrow$ 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]

$$\begin{aligned} Y_1 &= X + \xi_1, \quad \xi_1 \sim \mathcal{N}(0, \sigma_1^2) \\ Y_2 &= X + \xi_2, \quad \xi_2 \sim \mathcal{N}(0, \sigma_2^2), \sigma_2 \geqslant \sigma_1 \\ &= X + \widetilde{\xi}_1 + \widetilde{\xi}_2 \end{aligned}$$

**Example 7** [BSC($p_1$), BSC($p_2$), $p_1 \leqslant p_2 < 1/2$]
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 $X \to Y_1 \to Y_2$ (degraded BC) then*

$$\mathscr{C} = \text{clo. conv.} \left\{ (R_1, R_2) \in \mathbb{R}_+^2 : \begin{array}{lll} R_1 & \leqslant & I(X;Y_1|U) \\ R_2 & \leqslant & I(U;Y_2) \\ & & \text{for some } U \to X \to Y_1 \to Y_2 \text{ with } |\mathbb{U}| < \infty. \end{array} \right\}$$

- 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 $U \to X \to Y_1 \to Y_2$ is already converse.

**Example 8** [D–GBC]
Achievability:

$$\begin{aligned} U &\sim \mathcal{N}(0, (1-t)P) \\ X &= U + V, \ V \sim \mathcal{N}(0, tP), V \amalg U. \\ I(U;Y_2) &= \frac{1}{2}\log\left(1 + \frac{(1-t)P}{tP + \sigma_2^2}\right) \\ I(X;Y_1|U) &= \frac{1}{2}\log\left(1 + \frac{tP}{\sigma_1^2}\right) \end{aligned}$$

$$\mathscr{C} = \text{clo. conv.}\left\{ (R_1, R_2) \in \mathbb{R}_+^2 : \begin{array}{rcl} R_1 & \leqslant & \frac{1}{2}\log\left(1+\frac{tP}{\sigma_1^2}\right) \\ R_2 & \leqslant & \frac{1}{2}\log\left(1+\frac{(1-t)P}{tP+\sigma_2^2}\right), \\ & & t \in [0,1] \end{array} \right\}$$

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

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

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

$$\begin{aligned} I(U;Y_2) &= 1 - H(\beta \star p_2), \quad \text{where } \beta \star p_2 = \beta(1-p_2) + (1-\beta)p_2 \\ I(X;Y_1|U) &= H(Y_1|U) - H(Y_1|X,U) \\ &= H(\beta \star p_1) - H(Y_1|X) \\ &= H(\beta \star p_1) - H(p_1). \end{aligned}$$

$$\mathscr{C} = \text{clo. conv.}\left\{ (R_1, R_2) \in \mathbb{R}_+^2 : \begin{array}{rcl} R_1 & \leqslant & H(\beta \star p_1) - H(p_1) \\ R_2 & \leqslant & 1 - H(\beta \star p_2), \\ & & \beta \in [0, 1/2] \end{array} \right\}$$

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

# 3 DBC

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

$$\mathscr{R}(Z) := \left\{ (R_1, R_2) \in \mathbb{R}_+^2 : \begin{array}{rcl} R_1 & \leqslant & I(U_0U_1;Y_1) \\ R_2 & \leqslant & I(U_0U_2;Y_2) \\ R_1 + R_2 & \leqslant & \min\{I(U_0;Y_1), I(U_0;Y_2)\} \\ & & +I(U_1;Y_1|U_0) + I(U_2;Y_2|U_0) - I(U_1;U_2|U_0) \end{array} \right\}$$

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

**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.** *$\mathscr{R}$ is convex.*

*Proof.* We will prove $\bigcup_{Z \in \mathscr{P}} \mathscr{R}(Z)$ is convex. Let $R^{(\ell)} \in \bigcup_{Z \in \mathscr{P}} \mathscr{R}(Z), \ell = 1,2$. Then $\exists Z^{(\ell)}, \ell = 1,2$ such that $R^{(\ell)} \in \mathscr{R}(Z^\ell)$. Fix $\lambda \in (0,1)$, arbitrary.

- Define $Z = \overline{U}_0 U_1 U_2 X Y_1 Y_2$ as follows:
$\overline{U}_0 U_1 U_2 X Y_1 Y_2 = \Big((Q U_0^{(Q)}), U_1^{(Q)}, U_2^{(Q)}\Big)$, $Q \in \{1,2\}$, $p_Q(1) = \lambda$. Therefore, $p_{\overline{U}_0 U_1 U_2} = p_Q(\ell) p_{U_0^{(\ell)}} p_{U_1^{(\ell)}} p_{U_2^{(\ell)}}$.
Since $Z^{(\ell)} \in \mathscr{P}$, 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 $\overline{U}_0 U_1 U_2$.
- Consider terms with $I(\cdot^{(\ell)}; \cdot^{(\ell)} | U_0^{(\ell)})$. Averaging, we get $I(\cdot, \cdot; \overline{U}_0)$.
- $I(U_0^{(\ell)}; Y_k^{(\ell)})$ : Averaging, we get $I(U_0; Y_k | Q) \leqslant I(Q U_0; Y_k) = I(\overline{U}_0; Y_k)$.
- $$\begin{aligned} & \lambda \min\left\{I(U_0^{(1)}; Y_1^{(1)}), I(U_0^{(1)}; Y_2^{(1)})\right\} + (1-\lambda) \min\left\{I(U_0^{(2)}; Y_1^{(2)}), I(U_0^{(2)}; Y_2^{(2)})\right\} \\ \leqslant\ & \min\left\{I(U_0; Y_1|Q), I(U_0; Y_2|Q)\right\} \\ \leqslant\ & \min\left\{I(U_0 Q; Y_1), I(U_0 Q; Y_2)\right\} \\ =\ & \min\left\{I(\overline{U}_0; Y_1), I(\overline{U}_0; Y_2)\right\} \end{aligned}$$
- So $\lambda R^{(1)} + (1-\lambda) R^{(2)} \in \mathscr{R}(Z)$.
- Closure of a convex set is convex.

□

**Remark 4.** *Convexification is not necessary.*