Cooperative Communication Fundamentals & Coding Techniques

2011.2.14

Electronics and Telecommunication Research Institute

Kiung Jung

Kiung Jung, kujung@etri.re.kr

Coverage

1. **Preparatory Overview of Information Theory Basics**
   - Information theory basics & channel capacity
   - Network information theory basics
   - Cooperative communication examples
     - Simulcast transmission
     - Relay transmission and iterative diversity

Break

2. **Coding Techniques for Cooperative Communication**
   - Principal design tools
   - Cooperative coding schemes in wireless cooperative channels

Kiung Jung, kujung@etri.re.kr

Section 1

Preparatory Overview of Information Theory Basics

Kiung Jung, kujung@etri.re.kr

What is cooperative communication?

- Classically, it refers to the relay channel
- We consider it in a broader aspect as **multi-terminal communications:**
  The terminals jointly use the medium as shared resource instead of one that is divided orthogonally among pairs of terminals.

Kiung Jung, kujung@etri.re.kr

Motivation

- Wireless networks increasingly take places important in modern communications.
- **Typical wireless channel** usage: Avoiding interference among users by using **orthogonally divided** channels.
- **Capacity approaching process** – Information theory based **resource sharing** among multiple terminals rather than using orthogonally divided channels.

Kiung Jung, kujung@etri.re.kr

Coverage of Cooperative Communication

- Classically, it refers to the relay channel
- We consider it in a broader aspect as **multi-terminal communications:**
  The terminals jointly use the medium as shared resource instead of one that is divided orthogonally among pairs of terminals.

Kiung Jung, kujung@etri.re.kr

An example: T.M. Cover's Broadcast Channel, 1972

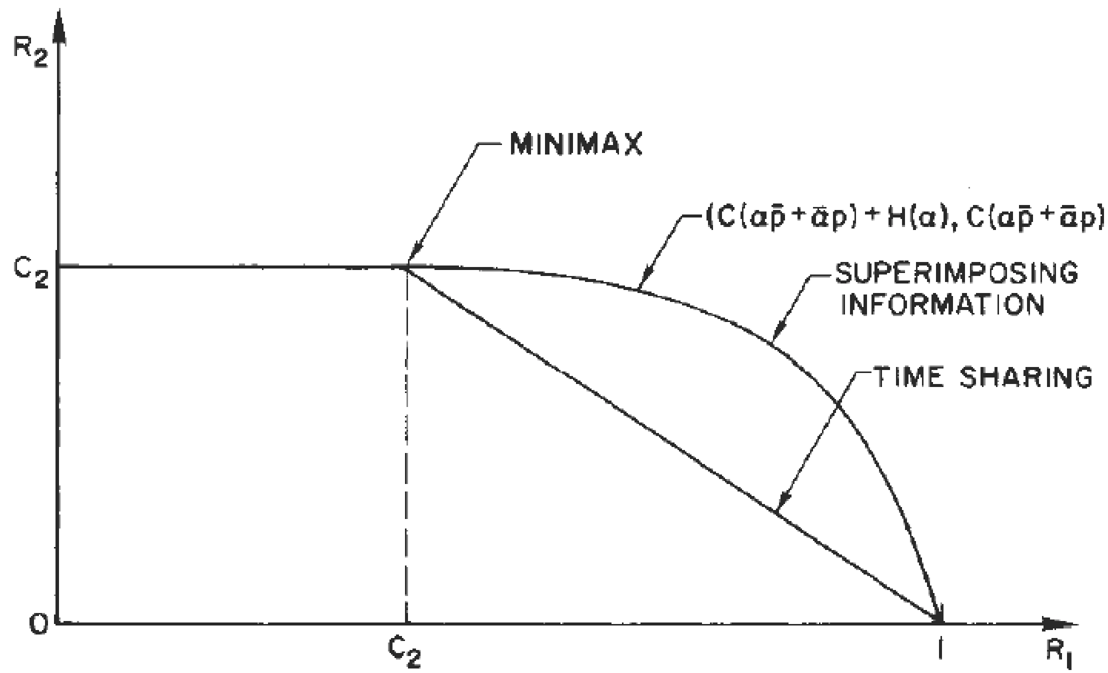


Fig. 5. Set of achievable rates for BSC.

From T. M. Cover, "Broadcast Channels," *IEEE Trans. Inform. Theory*, Jan. 1972

Why Now?

- Information theoretical consideration on cooperation has been started since the early 1970s.
- It did NOT attract significant interests until 2003 and beyond.

Kiung Jung, kujung@etri.re.kr

Why Now?

- Importance of mobile communication
  - Fading channel
- Importance of networked communication
  - Multiple communications
- Ability to implementation
  - Signal processing capability
  - Modern error-correction codes

Kiung Jung, kujung@etri.re.kr

Review: Information Theory Basics

- **Entropy**
  - Uncertainty
  - Average amount of information per source output

    (* *"You became a father!!"*)

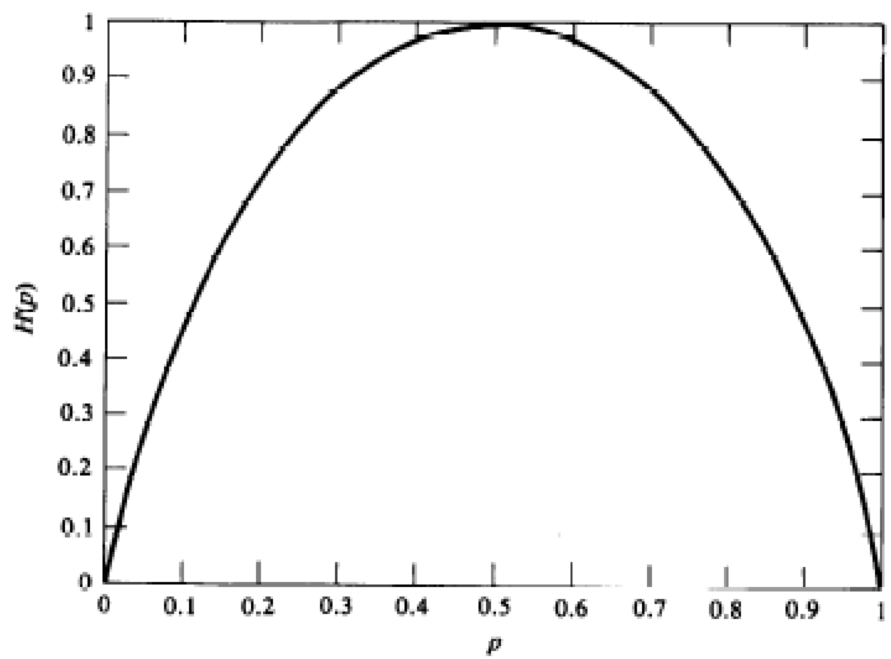


Figure 2.1. $H(p)$ versus p.

$$H(X) = -\sum_{x \in X} p(x) \log p(x)$$

$$= E_p \log \frac{1}{p(X)}$$

-> Function of the distribution X

** From Thomas Cover et al., Elements of Information Theory*

Review: Information Theory Basics

- **Entropy**
  - A measure of **uncertainty** of a random variable.
  - Function of distribution of a random variable, which depends only on the probabilities.
  - Average **amount of information** per source output.

    -> **Channel capacity**
  - Average length of the **shortest description** of the random variable

    -> Expected description length must be greater than or equal to the entropy.

    -> **Data compression**

Review: Information Theory Basics

- By the definition of the entropy
- Joint Entropy

$$H(X,Y) = -\sum_{x}\sum_{y} p(x,y)\log p(x,y)$$

$$= E_p \log \frac{1}{p(X,Y)}$$

- Conditional Entropy

$$H(X|Y) = -\sum_{x}\sum_{y} p(x,y)\log p(x|y)$$

$$= E_p \log \frac{1}{p(X|Y)}$$

Kiung Jung, kujung@etri.re.kr

Review: Information Theory Basics

- **Mutual Information**

A measure of the amount of information that one random variable contains about another random variable.

$$I(X;Y) = \sum_x \sum_y p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$$

$$= E_{p(x,y)} \log \frac{p(X,Y)}{p(X)p(Y)}$$

Then,

$$I(X;Y) = H(X) - H(X \mid Y)$$

-> **The reduction in the uncertainty of *X* due to the knowledge of *Y*.**

- **Channel Capacity**

$$C = \max_{p(x)} I(X;Y)$$

Review: Information Theory Basics

- **The Asymptotic Equipartition Property (AEP)**
  - The analog of the law of large numbers.
  - The law of large numbers states that for i.i.d. random variables,

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to EX, \quad \text{for large } n$$

  - The AEP states that;

$$\frac{1}{n}\log\frac{1}{p(X_1, X_2, \ldots, X_n)} \to H, \text{ where } X_1, X_2, \ldots, X_n \text{ are } i.i.d.$$

$$p(X_1, X_2, \ldots, X_n): \text{ Probability of observing the sequence } X_1, X_2, \ldots, X_n$$

$$\text{So, } p(X_1, X_2, \ldots, X_n) = 2^{-nH}$$

  - This arises the **typical set** in all sequences **where the sample entropy is close to true entropy.**

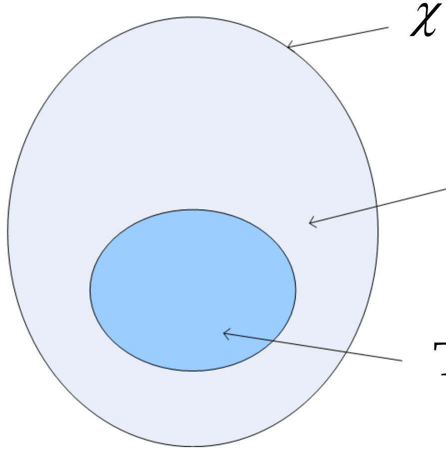
Kiung Jung, kujung@etri.re.kr

Review: Information Theory Basics

- **Typical Set,** $A_{\varepsilon}^{(n)}$
  - Typical sequences determine the average behavior of a large sample with high probability.
  - The number of elements in the typical set is nearly 2^{nH}

$$\left|A_{\varepsilon}^{(n)}\right| \le 2^{n(H(X)+\varepsilon)}$$

where ε is a arbitrary small number according to an appropriate choice of *n*.



Review: Information Theory Basics

- **Theorem 3.2.1 (T. M. Cover, p54)**
  - Let X^n be $i.i.d. \sim p(x)$. Let $\varepsilon > 0$. Then there exists a code which maps sequences x^n of length n into binary strings such that the mapping is one-to-one (and therefore invertible) and

$$E\left[\frac{1}{n}l\left(X^n\right)\right] \le H\left(X\right) + \varepsilon$$

For n sufficiently large,

x^n : denote a sequence $x_1, x_2, \ldots\ldots, x_n$

$l\left(x^n\right)$: length of th codeword corresponding to x^n

- Thus we **can represent sequences** X^n using $nH\left(X\right)$ bits on the average.

Review: Channel Capacity

- **Information Channel Capacity**

$$C = \max_{p(x)} I(X;Y)$$

- **Operational Channel Capacity**

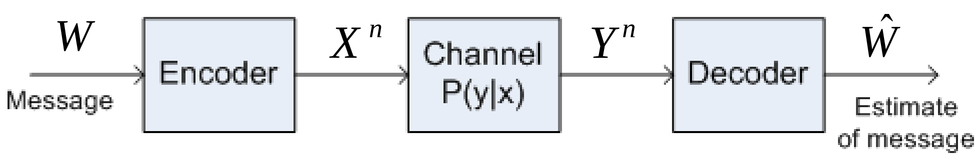
  - **Highest rate in bits per channel use** at which information can be sent with arbitrarily low probability of error.

- **Shannon's second theorem** establishes

  Inform. Channel Capacity = Oper. Channel Capacity

Review: Channel Capacity

- **Shannon's second theorem**
  - Operational meaning to the definition of capacity as the number of bits we can transmit reliably over the channel.
- **Basic Idea**
  - For each (typical) input n-sequences, there are approximately $2^{nH(Y|X)}$ possible Y sequences, all of them equally likely.



  - The total number of possible (typical) Y sequences is $\approx 2^{nH(Y)}$.
  - This set has to be divided into sets of size $2^{nH(Y|X)}$ corresponding to the different input X sequences.

Review: Channel Capacity

- **Basic Idea (continued)**

  - The total number of disjoint sets is $\leq 2^{n(H(Y)-H(Y|X))} = 2^{nI(X;Y)}$.

  - Hence we can send at most $\approx 2^{nI(X;Y)}$ distinguishable sequences of length n.

- **Theorem 8.7.1 (T. M. Cover, p198)**

  The channel coding theorem: All rates below capacity C are achievable. Specifically, for every rate $R < C$, there exists a sequence of $(2^{nR}, n)$ codes with maximum probability of error $\lambda^n \to 0$

  Conversely, any sequence of $(2^{nR}, n)$ codes with $\lambda^n \to 0$ must have

$$R \leq C = \max_{p(x)} I(X;Y)$$

Review: Channel Capacity

- **Gaussian Channel**
  - The most common limitation on the input is **power constraint**.
  - Input X and output Y random variables are **continuous**.

$$Y_i = X_i + Z_i, \quad Z_i \sim N\left(0, \sigma^2\right)$$

$$\frac{1}{n}\sum_{i=1}^{n} x_i^2 \le P$$

- **Differential entropy** $h(X)$ of a continuous random variable;

$$h(X) = -\int_S f(x) \log f(x) dx, \quad f(x) \sim \text{pdf for } X$$

- For a **normal distribution**, (T. M. Cover, p225)

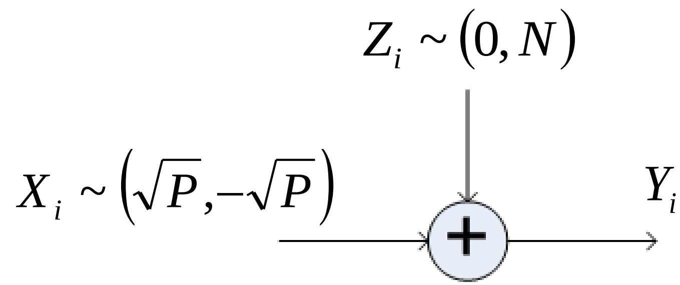
$$X \sim \phi(X) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2} \rightarrow h(\phi) = \frac{1}{2}\log 2\pi e\sigma^2 \text{ bits}$$

Kiung Jung, kujung@etri.re.kr

Review: Channel Capacity

- **Information Capacity in Gaussian Channel**

$$C = \max_{p(x):EX^2 \le P} I(X;Y)$$



$$\begin{aligned} I(X;Y) &= h(Y) - h(Y \mid X) \\ &= h(Y) - h(X + Z \mid X) \\ &= h(Y) - h(Z \mid X) \\ &= h(Y) - h(Z), \quad X, Z \sim \text{independent} \end{aligned}$$

where, $h(Z) = \frac{1}{2} \log 2\pi e N$, $EY^2 = E(X+Z)^2 = EX^2 + EZ^2 = P + N$

With power constraint $EX^2 \le P$,

$$\begin{aligned} I(X;Y) &= h(Y) - h(Z) \\ &\le \frac{1}{2} \log 2\pi e (P + N) - \frac{1}{2} \log 2\pi e N \\ &= \frac{1}{2} \log \left(1 + \frac{P}{N} \right). \end{aligned}$$

Kiung Jung, kujung@etri.re.kr

Review: Channel Capacity

- **Information Capacity in Gaussian Channel**

$$C = \max_{p(x):EX^2 \le P} I(X;Y) = \frac{1}{2}\log\left(1+\frac{P}{N}\right) \text{ bits per transmission}$$

- **A rate *R* is said to be achievable for a Gaussian channel with a power constraint *P* if there exists a sequence of $(2^{nR}, n)$ codes with code words satisfying the power constraint such that the maximal probability of error tends to zero.**

$$R \le C$$

Kiung Jung, kujung@etri.re.kr

Review: Channel Capacity

- **Band-Limited Gaussian Channel**
  - A common model for communication over a radio network is a band-limited with white noise.
- **Nyquist-Shannon sampling theorem**
  - Sampling a band-limited signal at a sampling rate $\frac{1}{2W}$ is sufficient to reconstruct the signal from the samples.

Kiung Jung, kujung@etri.re.kr

Review: Channel Capacity

- **So, for a Band-Limited Gaussian Channel,**

if Nose PSD ~ $\frac{N_0}{2}$ **, Bandwidth** W **, Time interval** $(0,T)$

$$C = \frac{1}{2}\log\left(1+\frac{P}{N}\right) \text{bits per transmission}$$

$$= \frac{1}{2}\log\left(1+\frac{P}{2W}\Big/\frac{N_0}{2}\right) = \frac{1}{2}\log\left(1+\frac{P}{N_0 W}\right) \text{ bits per samples.}$$

Since there are $2W$ **samples each second,**

$$C = W\log\left(1+\frac{P}{N_0 W}\right) \text{bits per second.}$$

Kiung Jung, kujung@etri.re.kr

Review: Channel Capacity

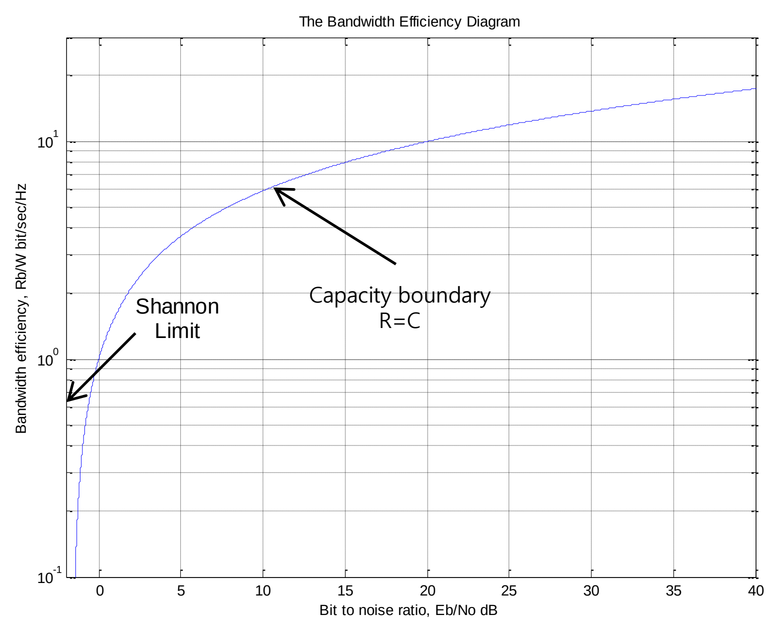
- **Shannon limit**

  The limiting value of Eb/No below which there can be no error-free communication at any information rate

- **Bandwidth Efficiency**

  A measure of how much data can be communicated in a specified bandwidth within a given time.

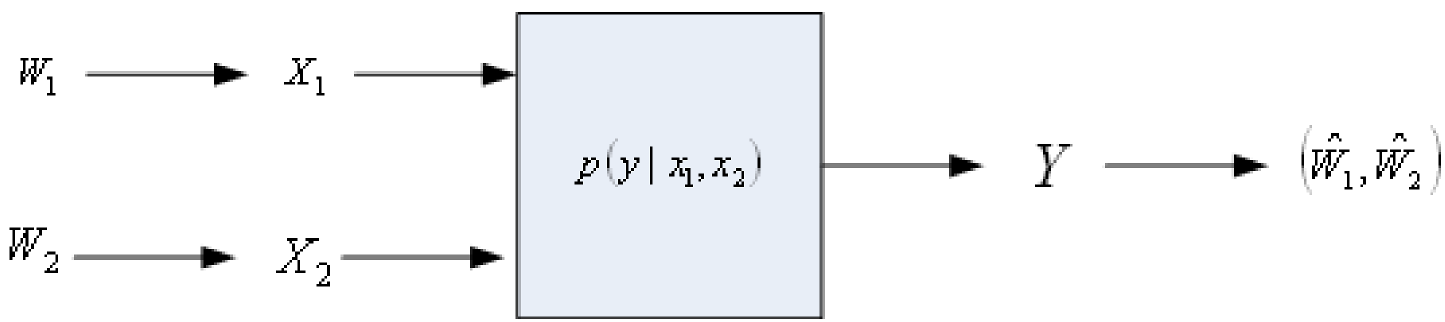
$$\eta = \frac{R}{W}$$



Kiung Jung, kujung@etri.re.kr

Multiple Access Channel (MAC)

- Two or more senders and a common receiver.
- **Definition**: A rate pair (R_1, R_2) is said to be **achievable** for the multiple access channel if there exists a sequence of $((2^{nR_1}, 2^{nR_2}), n)$ codes with $P_e^{(n)} \to 0$.

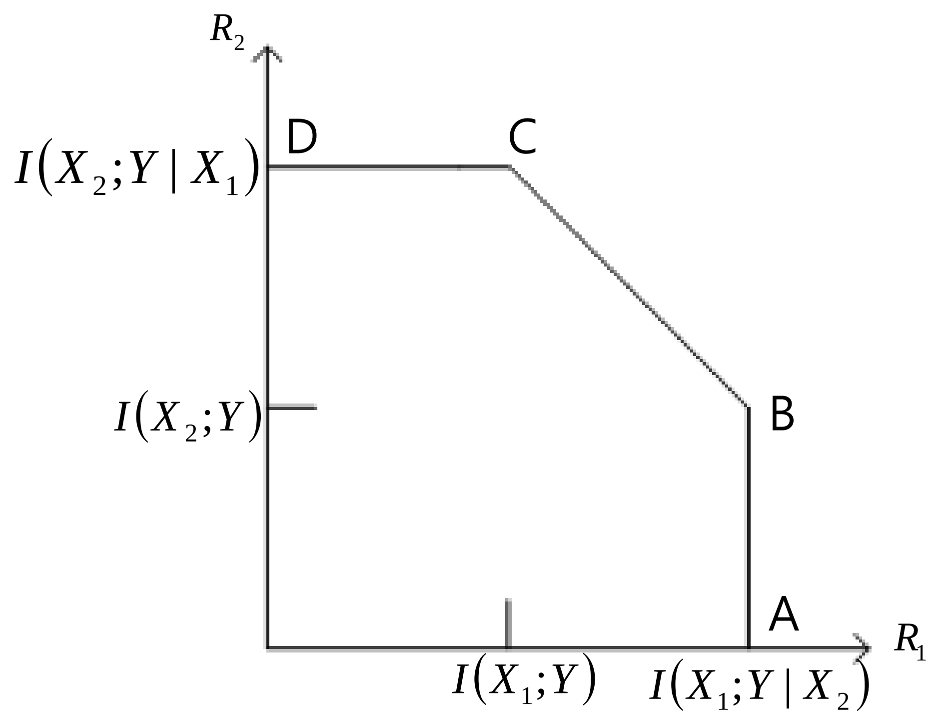


Multiple Access Channel (MAC)

- **Capacity region**

  The closure of the convex hull of the set of points (R_1, R_2) satisfying;

$$R_1 < I(X_1;Y \mid X_2)$$
$$R_2 < I(X_2;Y \mid X_1)$$
$$R_1 + R_2 < I(X_1, X_2;Y)$$



Kiung Jung, kujung@etri.re.kr

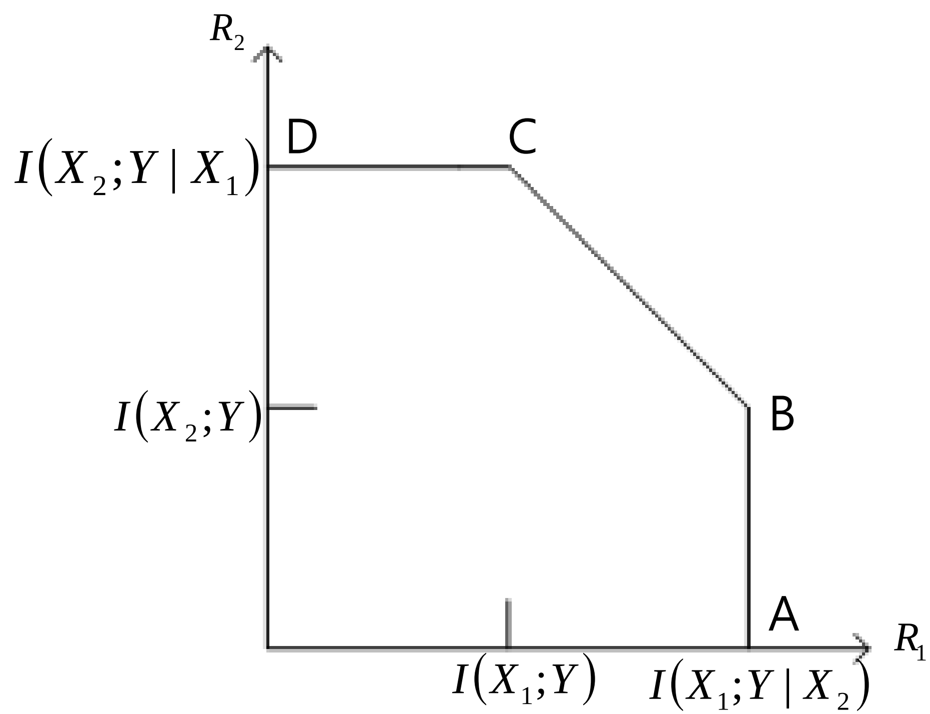
Multiple Access Channel (MAC)

- **Point A**

  Maximum rate achievable from sender 1 when sender 2 is not sending any information.

$$\max R_1 = \max_{p_1(x_1)p_2(x_2)} I(X_1;Y \mid X_2)$$

- **Point B**
  - Maximum rate sender 2 can send as long as sender 1 sends at this maximum rate.
  - X_1 is considered as noise for the channel $X_2 \rightarrow Y$



Multiple Access Channel (MAC)

- **Gaussian Multiple Access channel**
  - For some input distribution $f_1(x_1)f_2(x_2)$, $EX_1^2 \le P_1^2$ and $EX_2^2 \le P_2^2$
  - Define channel capacity function with signal to noise ratio x

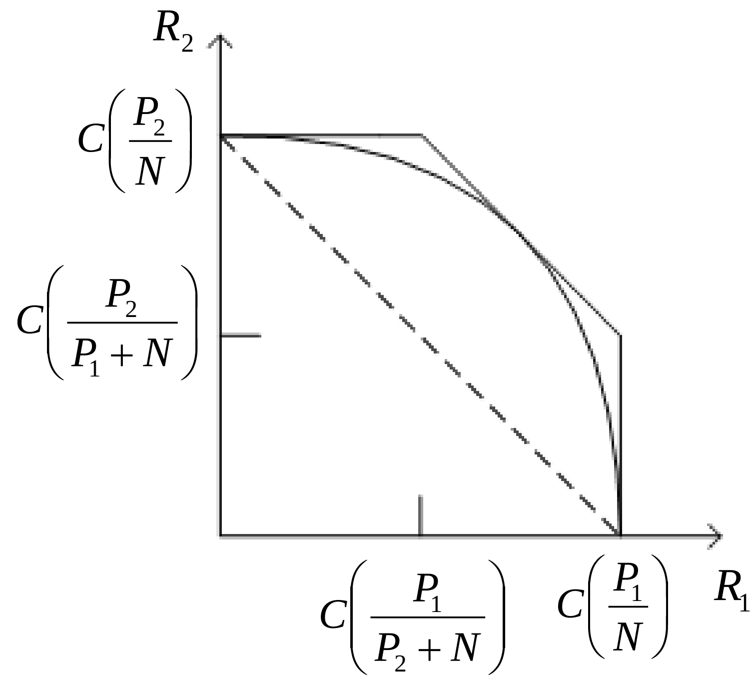
$$C(x) \equiv \frac{1}{2}\log(1+x),$$

Then,

$$R_1 \le C\left(\frac{P_1}{N}\right),$$

$$R_2 \le C\left(\frac{P_2}{N}\right),$$

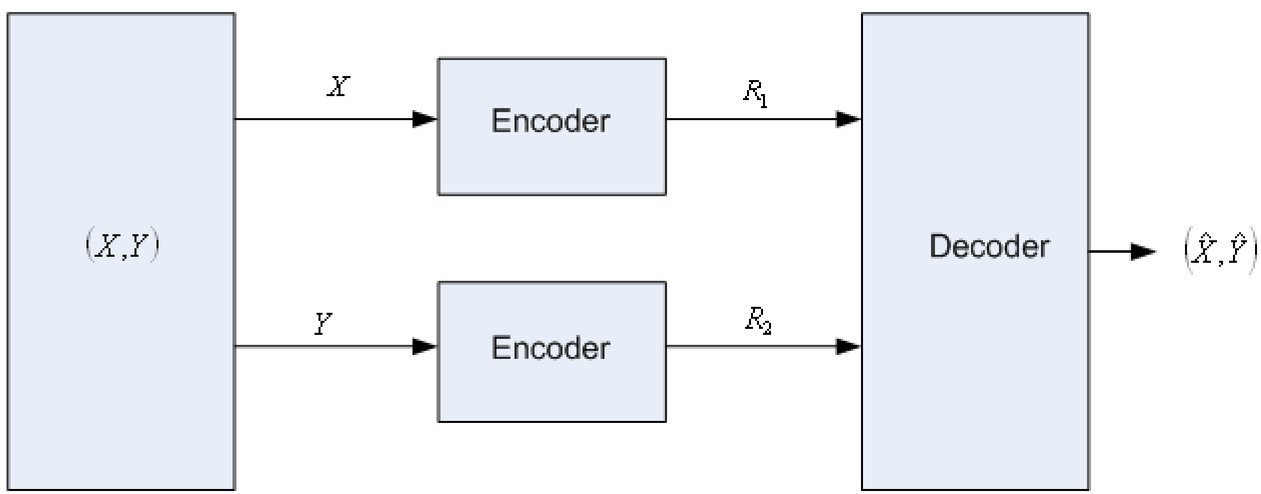
$$R_1 + R_2 \le C\left(\frac{P_1+P_2}{N}\right).$$



Encoding of Correlated Sources

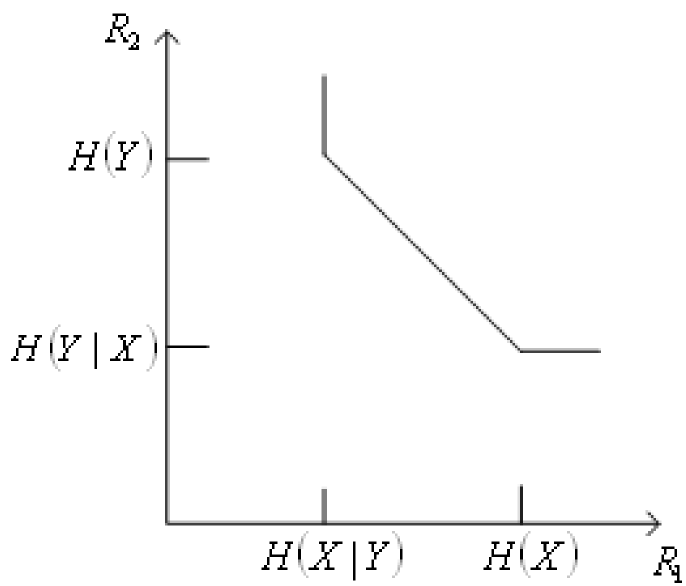
- Suppose there are two sources
- What if the X-source and the Y-source must be separately?

$$R_1 \geq H(X \mid Y),$$
$$R_2 \geq H(Y \mid X),$$
$$R_1 + R_2 \geq H(X, Y)$$



Slepian-Wolf Encoding

- **Random binning procedure**.
- Very similar to hash functions.
- With high probability, different source sequences have different indices, and we can recover the source sequence from the index.

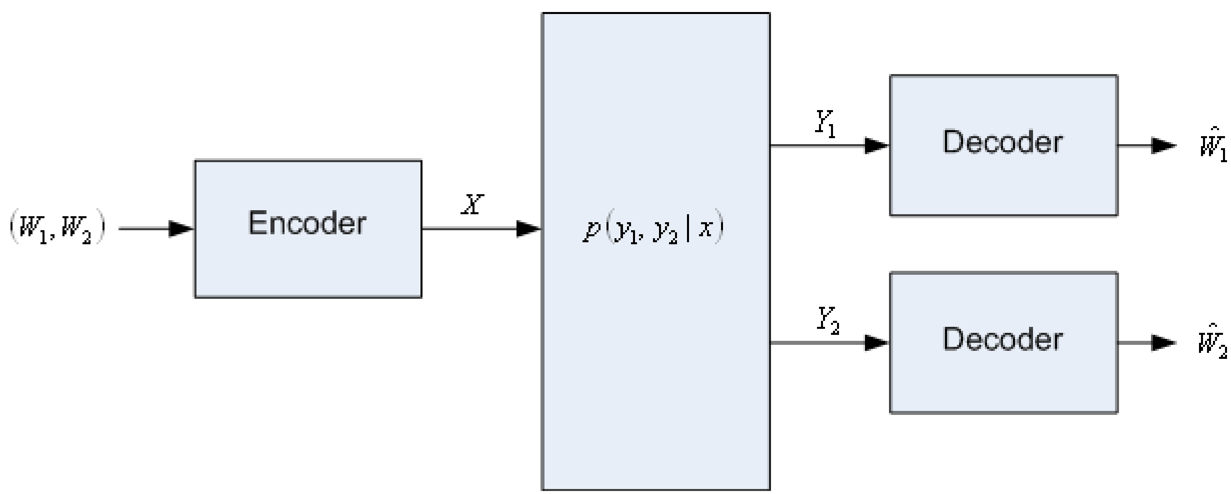


Rate region for Slepian-Wolf encoding.

Broadcast Channel

- **One sender** and **two or more receivers**.
- TV station – **Superposition** of information

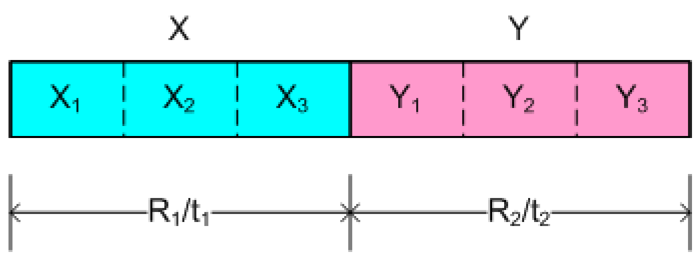
  We may wish to arrange the information in such a way that the **better receivers receive extra information**, which produces a better picture or sound.



Broadcast Diversity

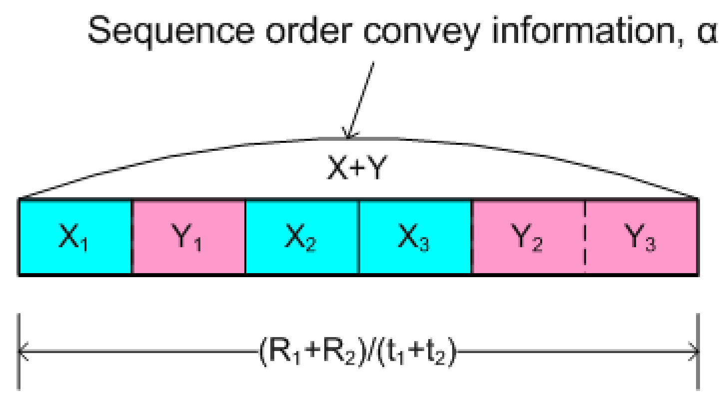
- Conceptual example

- Time sharing -



R=R1+R2

- Broadcast channel code -



R=R1+R2+α

Kiung Jung, kujung@etri.re.kr

Broadcast Diversity

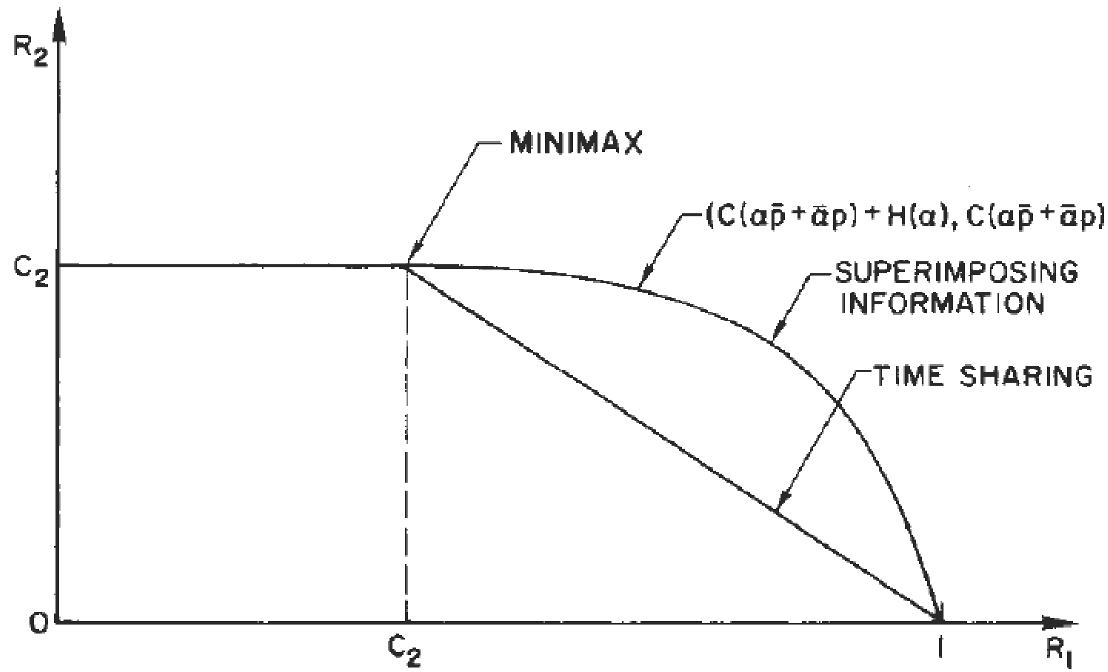
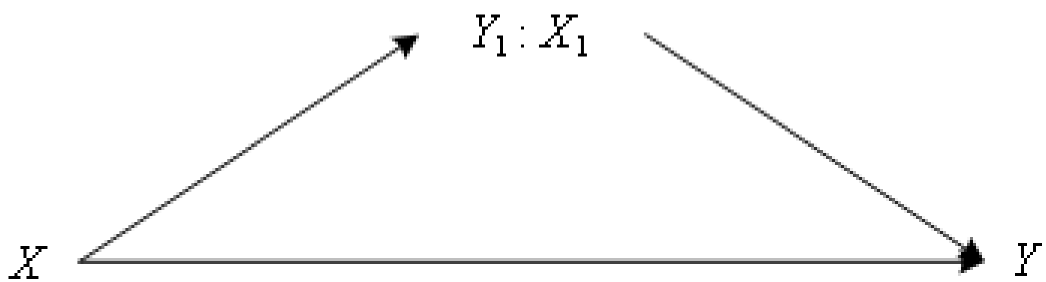


Fig. 5. Set of achievable rates for BSC.

From T. M. Cover, "Broadcast Channels," *IEEE Trans. Inform. Theory*, Jan. 1972

Relay Channel

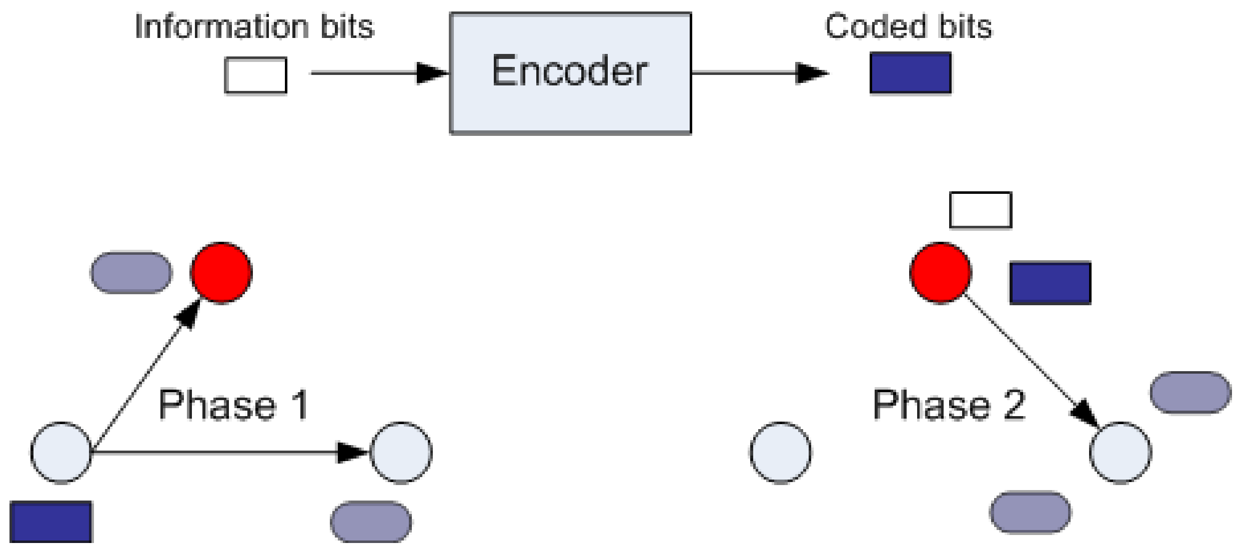
- One sender and one receiver with a number of **intermediate nodes**.
- The relay channel combines a broadcast channel and a multiple access channel.



$$C \leq \sup_{p(x,x_1)} \min \left\{ I\left(X, X_1; Y\right), I\left(X; Y, Y_1 \mid X_1\right) \right\}$$

Example I: Relay Transmission

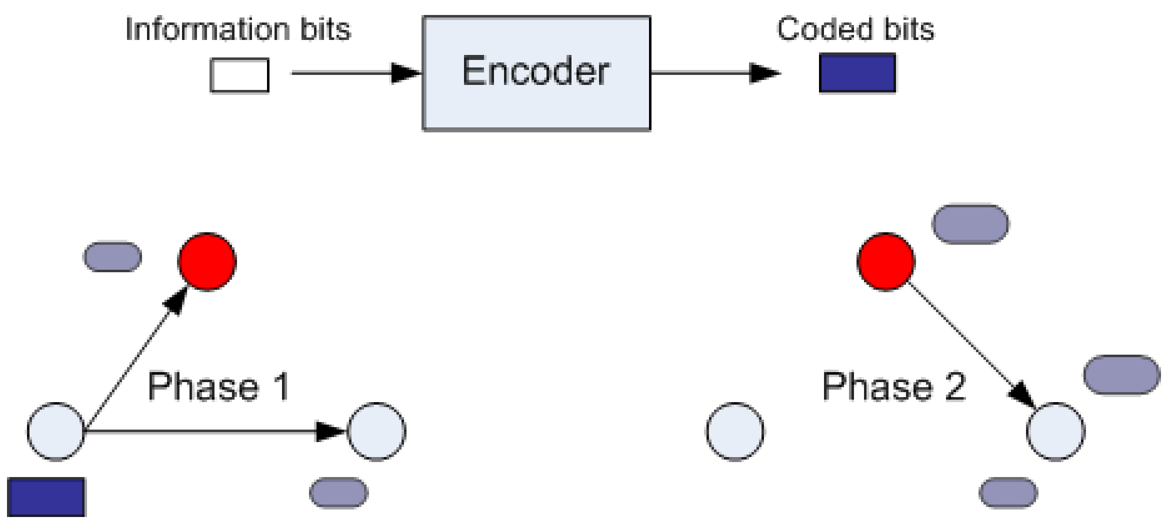
- **Decode-and-forward**, Laneman, 2002



* **Diversity channel is created only when the relay decodes successfully.**

Example I: Relay Transmission

- **Amplify-and-forward,** Laneman, 2002

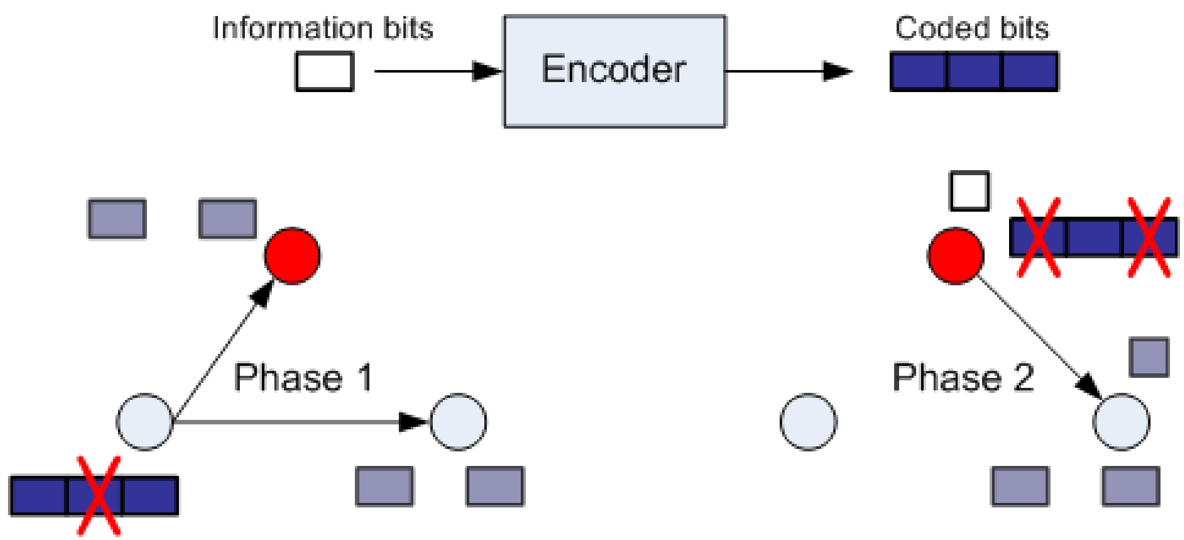


*** The relay also amplifies its own receiver noise.**

Kiung Jung, kujung@etri.re.kr

Example I: Relay Transmission

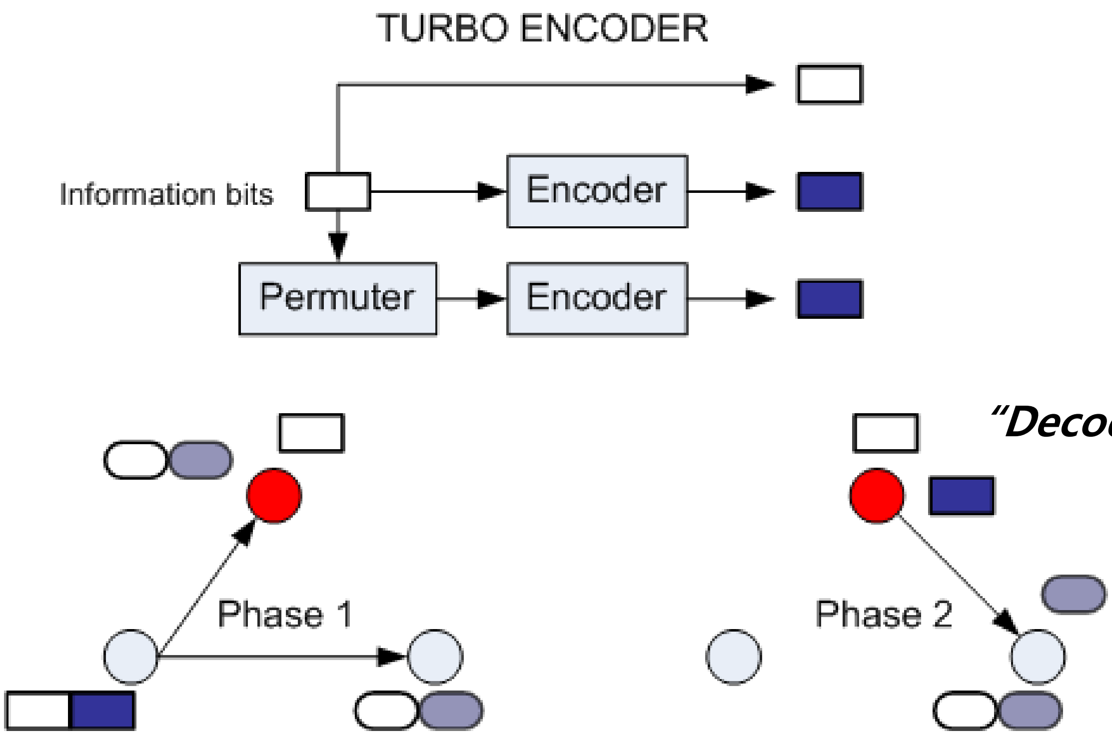
- **Coded cooperation,** T. E. Hunter, 2002



*** Rate-Compatible Punctured Convolutional code (RCPC) based**

Example I: Relay Transmission

- **Cooperation using turbo codes,** Zhao and Valenti, 2003



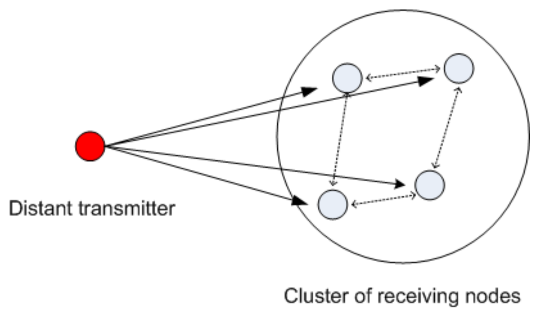
"Decode and permute"

* **Two separate Recursive Systematic Convolutional (RSC) encoders constitutes turbo encoder followed by iterative decoding between the relay and the destination**

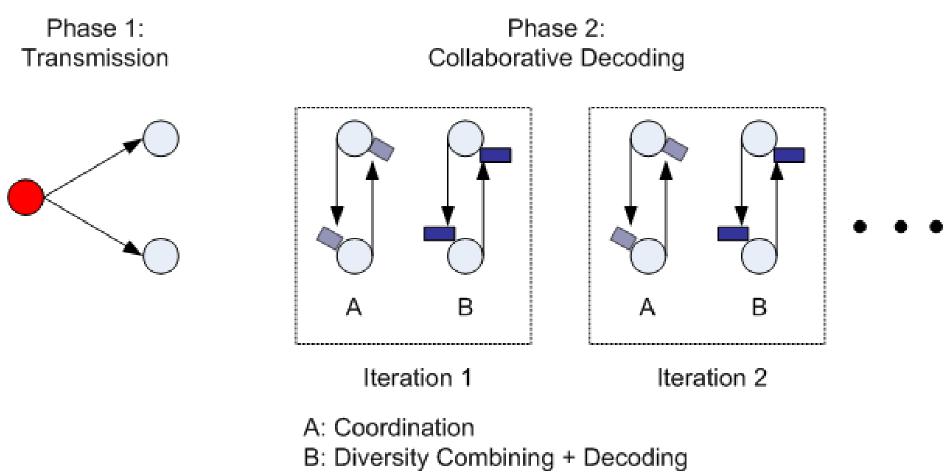
Example I: Iterative Diversity

- **Collaborative decoding,** Arun and J. M. Shea, 2006

  - System model for collaborative decoding

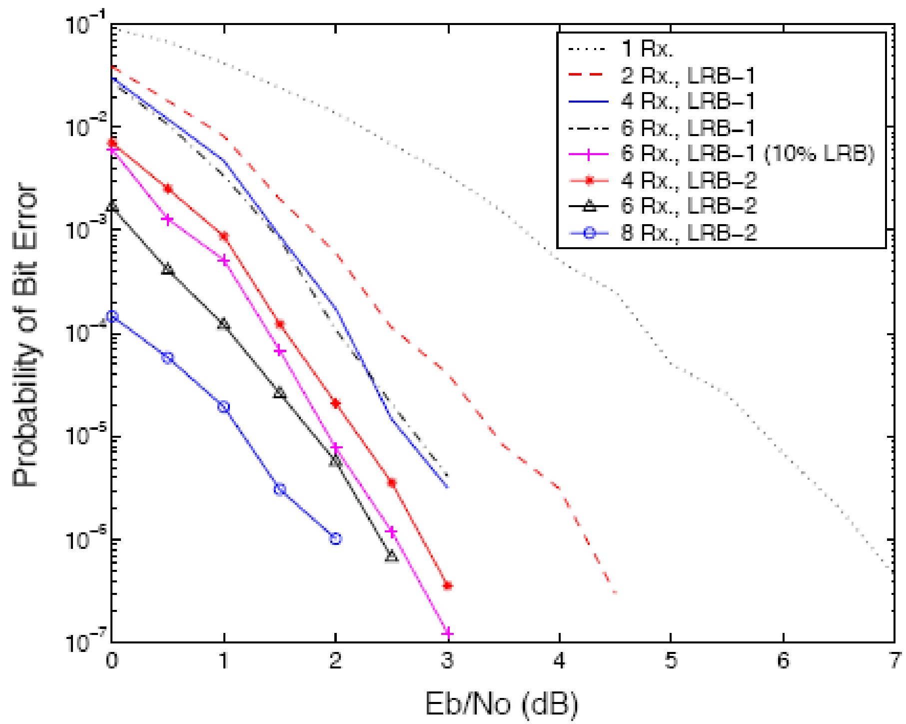


  - Principle of collaborative decoding with two nodes



Example I: Iterative Diversity

- 3 iterations, 5% LRB exchange, 900 packet size

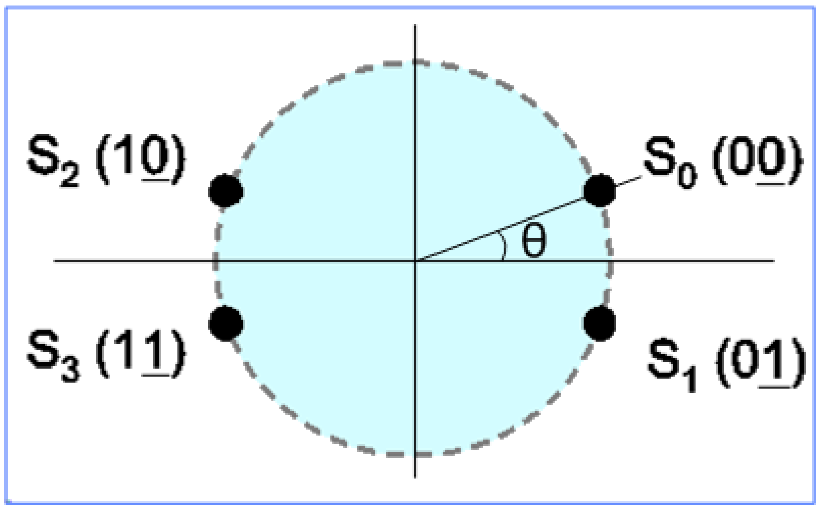


Performance of two collaborative decoding schemes in which receivers request information for a set of least-reliable bits.

Example II: Simulcast Transmission

- **Simulcast by using Non-Uniform QPSK UEP signaling,**

  K. Jung & J. Shea, 2005



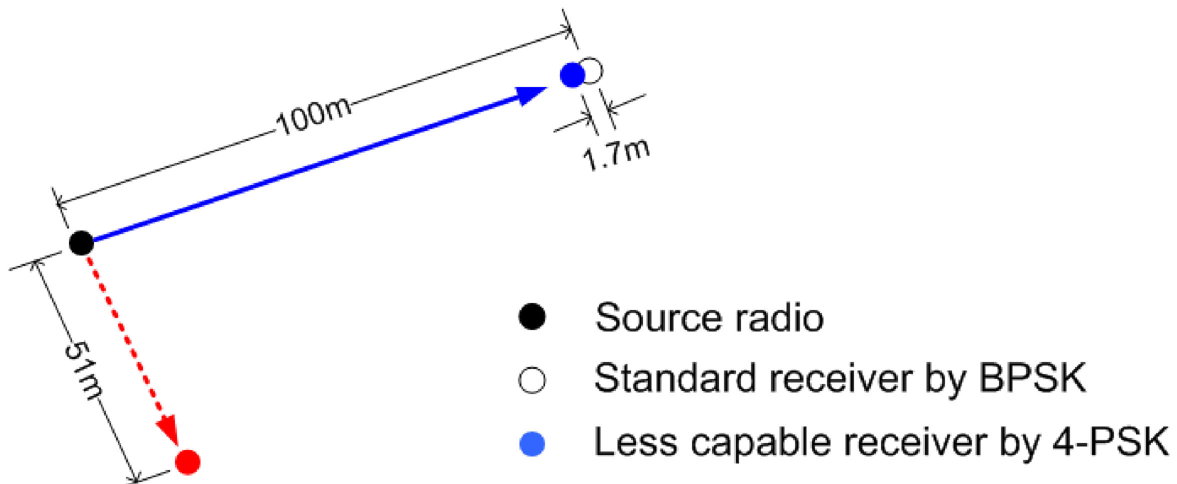
b_1: Basic message
b_2: Additional message

- Typical required bit error probability
  - for voice: 10^{-2}
  - for data: 10^{-4}

Kiung Jung, kujung@etri.re.kr

Example II: Simulcast Transmission

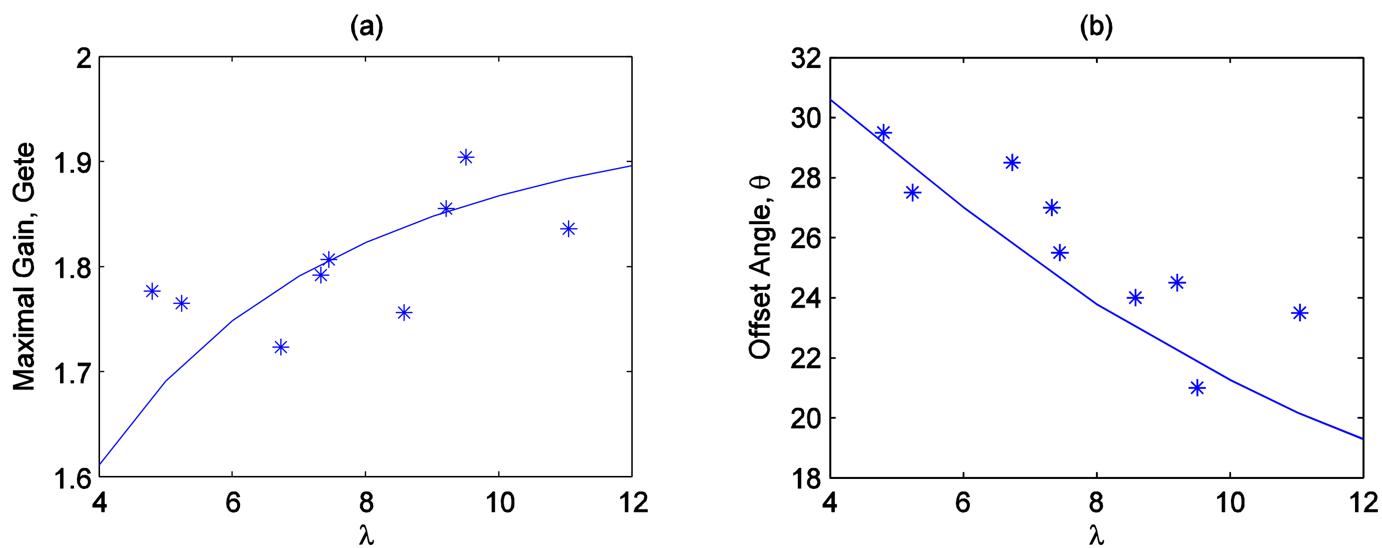
- **Example of transmission range calculation**
  - Signaling: 4-PSK
  - Degradation: 0.3dB (θ=15°)
  - Disparity: 11.4dB (P_B=P_A)
  - Original transmission range by BPSK: 100m
  - Propagation constant, n: 4



Example II: Simulcast Transmission

- **Simulcast by using Non-Uniform QPSK UEP signaling**

$$G(\theta) = \frac{S_{ete}(\theta)}{S_{ete}(0)}$$



Simulation results of ETE throughput at n=4, Poisson distribution of H(θ)

Kiung Jung, kujung@etri.re.kr