UNIT-4

ELEMENTS OF INFORMATION THEORY

Summary of Concepts/theorems

"If the rate of **Information** is less than the **Channel capacity** then there exists a **coding technique** such that the information can be transmitted over it with very small probability of error despite the presence of noise."

What is Information?

For a layman, whatsoever may be the meaning of information but it should have following properties

- The amount of information (I_j) associated with any happening 'j' should be inversely proportional to its probability of occurrence.

- $I_{jk} = I_j + I_k$; if events i and j are independent.

Technical aspects of Information

- Shannon proved that the only mathematical function which can retain the previously stated properties of information for a symbol produced by a discrete source is

$$I_i = \log(1/P_i) \text{ bits}$$

The base of log (if 2) define the unit of information (then bits)

- A single binary digit (binit) may carry more/less than one bit (may not be integer) information depending upon its source probability.

Where is the difference?

- Human mind is more intelligent than any machine.
- Suppose a 8 month old child picks up the phone and pressed redial button if you are at the receiving end you will immediately realize that something like this has been happened and whatsoever he is saying it conveys no information to you.
- But for the system, which is less intelligent than us, it is a message with very small probability thus it is treated as most informative message.

Source Entropy

- Defined as average amount of information produced by the source, denoted by H(x).

- Find H(x) for a discrete source which can produce ‘n’ different symbols in a random fashion.

- There is a binary source with symbol probabilities ‘p’ and (1-p). Find the maximum and minimum value of H(x).

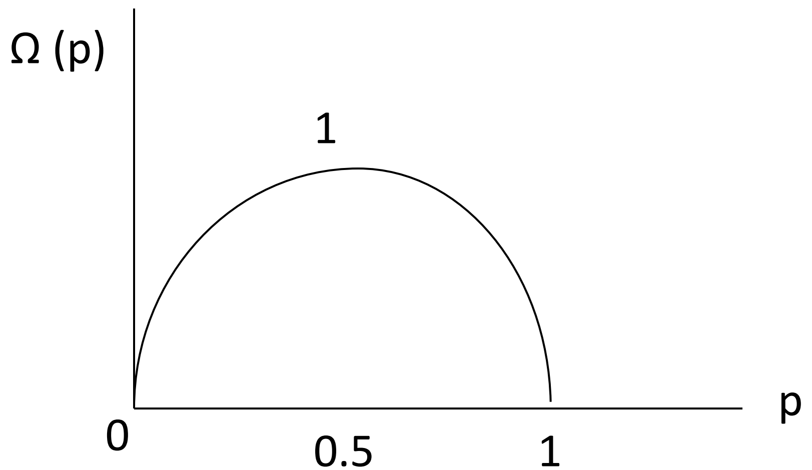
- $H(x) = \sum x_i * P(x_i)$; If X is discrete

  $\int x * p(x)\,dx$; If X is continuous.

  {H(x)= 1/N(N1*x1+N2*x2+ …….)}

- $H(x) = \Omega(p) = p*\log(1/p) + (1-p)*\log(1/(1-p))$

{can be solved as simple Maxima-Minima problem}



Entropy of a M-ary source

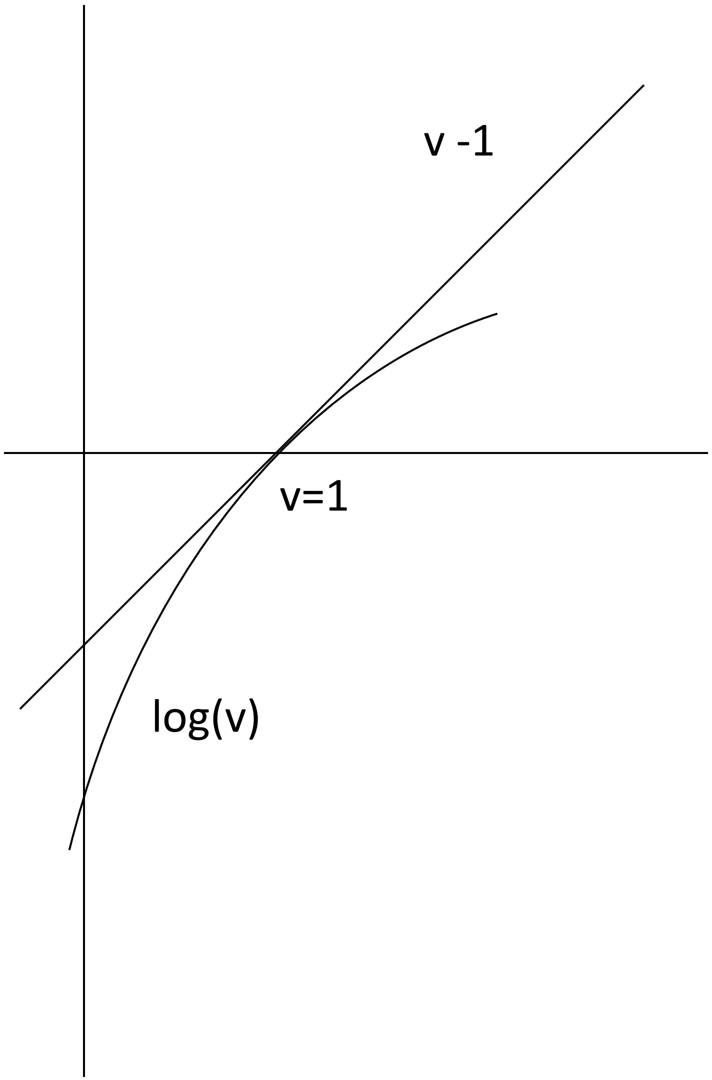
- There is a known mathematical inequality

  $(V-1) >= \log V$

  equality holds at V=1

- Let $V = (Q_i/P_i)$; such that $\sum Q_i = \sum P_i = 1$

  (P may be assumed as set of source symbol probabilities and Q is another independent set of probabilities having same number of elements)



- thus, $\{(Q_i/P_i) - 1\} \geq \log(Q_i/P_i)$
- $P_i * \{(Q_i/P_i) - 1\} \geq P_i * \log(Q_i/P_i)$
- $\sum P_i * \{(Q_i/P_i) - 1\} \geq \sum P_i * \log(Q_i/P_i)$
- $\{\sum Q_i - \sum P_i\} = 0 \geq \sum P_i * \log(Q_i/P_i)$
- $\sum P_i * \log(Q_i/P_i) \leq 0$
- Let $Q_i = 1/M$ (all events are equally likely)
- $\sum P_i * \log(1/M * P_i) \leq 0$

- $\sum Pi * \log(1/Pi) - \log(M) \sum Pi <= 0$

- $H(x) <= \log(M)$

- Equality holds when v=1 i.e. Pi=Qi i.e. P should also be a set of equally likely events.

- Conclusion-

  “A source which generates equally likely symbols will have maximum avg. information”

  “Source coding is done to achieve it”

Coding for Memoryless source

- Generally the information source is not of designers choice thus source coding is done such that it appears equally likely to the channel.

- Coding should neither generate nor destroy any information produced by the source i.e. the rate of information at I/P and O/P of a source coder should be same.

Rate of Information

- If the rate of symbol generation of a source, with entropy H(x), is r symbols/sec. then

$$R = r*H(x) \text{ and } R \le r*\log(M)$$

- If a binary encoder is used then

$$o/p \text{ rate} = rb*\ \Omega(p) \ \text{ and } \le rb$$

(if the 0’s and 1’s are equally likely in coded seq)

- Thus as per basic principle of coding theory

$$R \{= r*H(x)\} \le rb \ ;\ H(x) \le rb/r \ ;\ H(x) \le \bar{N}$$

- Code efficiency = $H(x)/\bar{N} \le 100\%$

Uniquely Decipherability (Kraft's inequality)

- A source can produce four symbols
  {A(1/2, 0); B(1/4, 1); C(1/8, 10); D(1/8, 11)}.

  [symbol (probability, code)]

Then H(x)= 1.75 and N = 1.25 so **efficiency > 1**

where is the problem?

- Kraft's inequality

$$K = \sum 2^{-N_i} <= 1$$

Source coding algorithms

- **Comma code**

  (each word will start with '0' and one extra '1' at the end. first code = 0)

- **Tree code**

  (no code word appears as prefix in another codeword, first code = 0)

- **Shannon – Fano**

  (Bi partitioning till last two elements. '0' in upper/lower part and '1' in lower/upper part)

- **Huffman**

  (adding two least symbol probabilities and rearrangement till two elements, back tracing for code.)

- **n^{th} extension**

  (form a group by combining 'n' consecutive symbols then code it.)

- **Lempel – Ziv**

  (Table formation for compressing binary data)

Source Coding Theorem

$H(x) <= \overline{N} < H(x) + \phi$; ϕ should be very small.

Proof:

- It is known that $\sum Pi * \log (Qi/Pi) <= 0$
- As per Kraft's inequality $1 = (1/K)\sum 2^{-Ni}$, thus it can be assumed that $Qi = 2^{-Ni}/K$ (so that addition of all Qi =1).
- Thus, $\sum Pi*\{\log(1/Pi) – Ni – \log (K)\} <=0$
- $H(x) – \overline{N} – \log (K)<=0$; $H(x)<= N + \overline{\log} (K)$
- since $\log (K)<=0$ (as $0<K<=1$) thus $H(x)<=\overline{N}$
- For optimum codes K=1 and Pi=Qi

Symbol Probability Vs code length

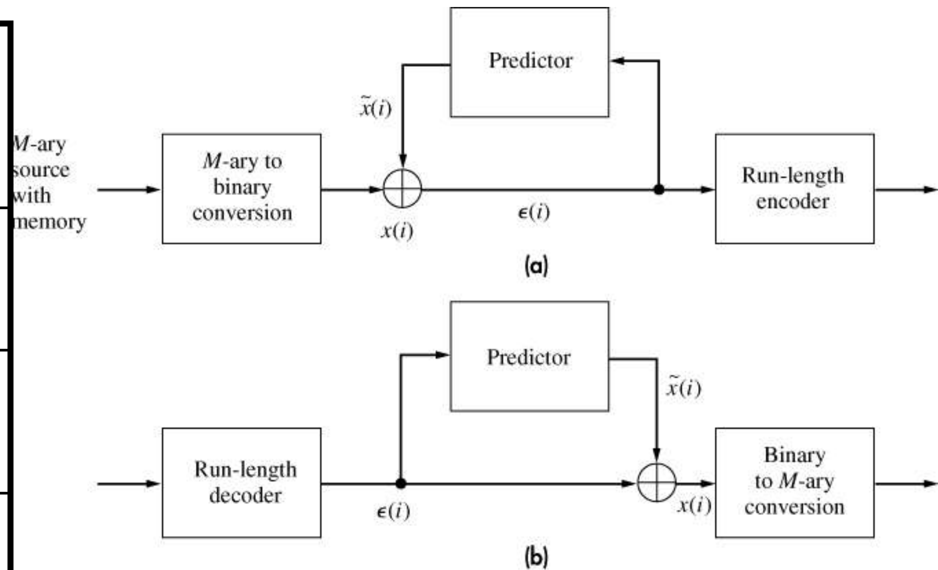
- We know that an optimum code requires K=1 and Pi=Qi
- Thus, Pi = Qi = 2^{-Ni}/K(=1) thus Ni = log(1/Pi)
- Ni = li

(the length of code should be (*<u>inversely</u>*) proportional to its information (*<u>probability</u>*))

Samuel Morse applied this principle long before Shannon has mathematically proved it

Predictive run encoding

n	Encoding	Decoding
0	00…00(k)	1
1	00…01	01
2	00...10	001
-	-	-
-	-	-
m-1	11…10	00..01
>=m	11…11	00..00 (m)



- ‘run of n’ means ‘n’ successive 0’s followed by a 1.
- m = 2^k-1
- k-digit binary codeword is sent in place of a ‘run of n’ such that 0<=n<=m-1

Designing parameters

- A *run of n* has total n+1 bits. If ‘p’ is the probability of correct prediction by the predictor then the probability of a *run of n* is $P(n)= p^n*(1-p)$.

- $E[n] = \overline{E} = \sum(n+1)*P(n)$; (for $0<=n<=infinity$) $= 1/(1-p)$

- The series $(1-v)^{-2}=1+2v+3v^2+$------; *for* $v^2<1$ *is used.*

- If n>m such that $(L-1)*m<=n<=L*m - 1$ then number of codeword bits required to represent it will be $N=L*k$

- Write an expression for avg. no. of code digits per run.

- $\overline{N} = k^* \sum P(n);_{0<=n<=(m-1)}$

  $+2k^*\sum P(n);_{(m-1)<=n<=(2m-1)}$

  $+3k^*\sum P(n);_{(m-1)<=n<=(2m-1)} + \ldots\ldots$

- It can be solved to **$\overline{N} = k/(1-p^m)^{-2}$**

- There is an optimal value of k which minimizes $\overline{N}$ for a given predictor.

- $\overline{N}/\overline{E} = r_b/r$; measures the compression ratio. It should be as low as possible.

Information Transmission

Channel Types:-

- Discrete Channel produces discrete symbols at the receiver. (source is implicitly assumed to be discrete)

- Definitely, the channel noise converts a discrete signal into continuous but it is assumed that the term ‘channel’ includes an pre processing section which will again convert it into discrete nature and it is supplied to the receiver.

- The continuous channel analysis does not involve above assumptions.

Discrete Channel Examples

- Binary Symmetric Channel (BSC)

  2 source and 2 receiver symbols.

  (single threshold detection)

- Binary Erasure Channel (BEC)

  2 source and 3 receiver symbols.

  (two threshold detection)

Discrete channel analysis

- $P(x_i)$; Probability that the source selects symbol x_i for Tx.
- $P(y_j)$; Probability that symbol y_j is received.
- $P(y_j|x_i)$ is called forward transition probability.
- Mutual information measures the amount of information transferred when x_i is transmitted and y_j is received.

Mutual Information (MI)

- If we happen to have an **ideal noiseless** channel then definitely each y_j uniquely identifies a particular x_i; then **$P(x_i|y_j)=1$** and MI is expected to be equal to self information of x_i.

- On the other hand if channel noise has such a large effect that y_j is totally unrelated to x_i then $P(x_i|y_j)=P(x_i)$ and MI is expected to be zero.

- All real channels falls between these two extremes.

- Shannon suggested following expression for MI which does satisfy both the above conditions

$$I(x_i;y_j) = \log \{P(x_i|y_j) / P(x_i)\} \text{ bits}$$

Discrete Channel Capacity

- Being a stochastic process, Instead of I(xi,yj) the quantity of interest is I(X;Y), the Avg MI, defined as the average amount of source information gained per received symbol.

- $I(X;Y) = \sum P(x_i,y_j) * I(x_i;y_j)$; (for all possible values of i and j)

- Discrete Channel Capacity (Cs) = max I(X;Y).

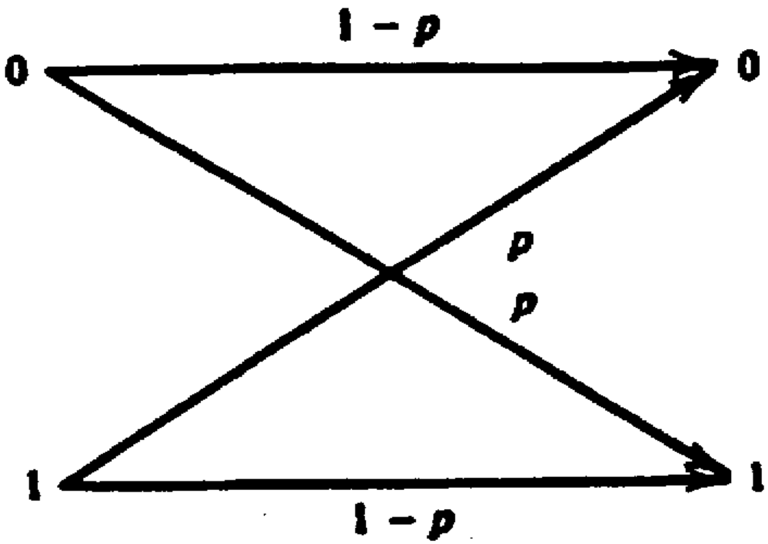
- If ‘s’ symbols/sec is the maximum symbol rate allowed by the channel then channel capacity (C) = s*Cs bits/sec i.e. maximum rate of information transfer.

Channel Capacity

- *Capacity* in the channel is defined as a intrinsic ability of a channel to convey information
- Using mutual information the channel capacity of a discrete memoryless channel is a maximum average mutual information in any single use of channel over all possible probability distributions

Discrete Memoryless Channels

- Example : Binary symmetric channel *revisited*
  - *capacity of a binary symmetric channel with given input probabilities*
  - *variability with the error probability*



Channel Coding Theorem

- Channel coding consists of mapping the incoming data sequence into a channel input sequence and vice versa via inverse mapping
  - *mapping operations performed by encoders*

Information Capacity Theorem

- A channel with noise and the signal are received is described as discrete time, memoryless Gaussian channel (with power-limitation)
  - *example : Sphere Packing*

Implications of the Information Capacity Theorem

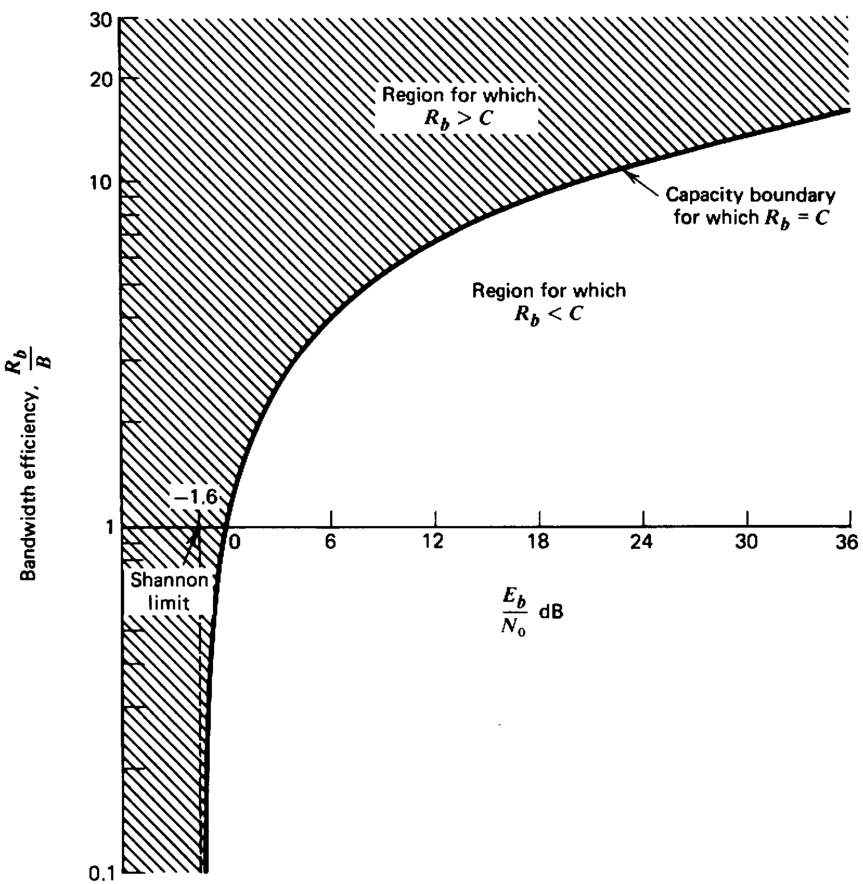


Figure 10.16 Bandwidth-efficiency diagram.

- Set of *M*-ary examples

Shannon's fundamental theorem

- It is intuitive that R <= C otherwise the channel will cause distortion which in turn will increase the error rate even if the channel is noiseless.

- Shannon combined the above result with source and channel coding theory and stated that

  "If R<C, then there exists a coding technique such that the O/P of a source can be transmitted over the channel with an arbitrarily small frequency of errors."

The general proof of theorem is well beyond the scope of this course but following cases may be considered to make it plausible –

(a) Ideal Noiseless Channel

- Let it has $m=2^k$ symbols then

$$C_s = \max I(X;Y) = \max H(x) = \log(m) = k \text{ and } \mathbf{C = s*k}.$$

- Errorless transmission rests on the fact that the channel itself is noiseless.

- If R is rate of information of source, r_b is rate of binary encoder then rate of symbol to the channel (o/p of binary to m-ary block) will be

$$s = r_b/\log(m) = r_b/k \text{ thus } \mathbf{r_b = s*k = C}$$

- We have already proved that $r_b >= R$ otherwise it will violate Kraft's inequality thus $C >= R$

(b) Binary Symmetric Channel

- I (X;Y) = Ω(α + p - 2*p*α) - Ω(α); Ω(α) being constant for a given α.
- Ω(α + p - 2*p*α) varies with source probability p and reaches a maximum value of unity at (α+p - 2*p*α)=1/2.
- Ω(α + p - 2*p*α) =1 if p=1/2; irrespective of α (it is already proved that Ω(1/2)=1).
- Using an optimum source coding technique p=1/2 can be achieved.
- Thus Cs = max I(X;Y) = 1- Ω(α) and C=s*{1- Ω(α)}.

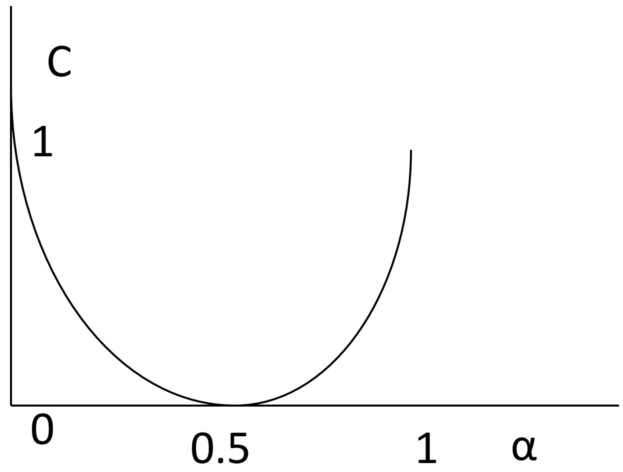


Fig.1

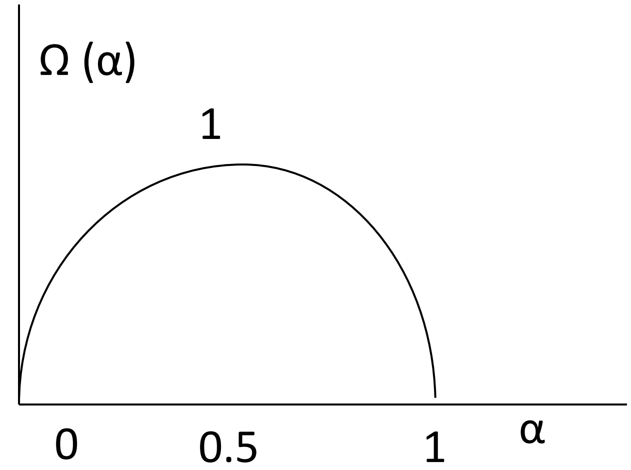


Fig.2

- **In figure 1, C decreases to zero and again it increases to one, as alpha varies from 0 to 1. Explain the reason.**

- **Please write it down in your notebook.**

- p=1/2 can be achieved by optimum source coding.
- Extra bits are required to be added for error control (concept of redundancy).
- If *q* redundant bits are added to a *k* bit message then code rate $R_c = k/(k+q) < 1$.

- Effect of decrease in Rc (by increasing ‘q’) -

(a) The value of α decreases thus the capacity will increase.

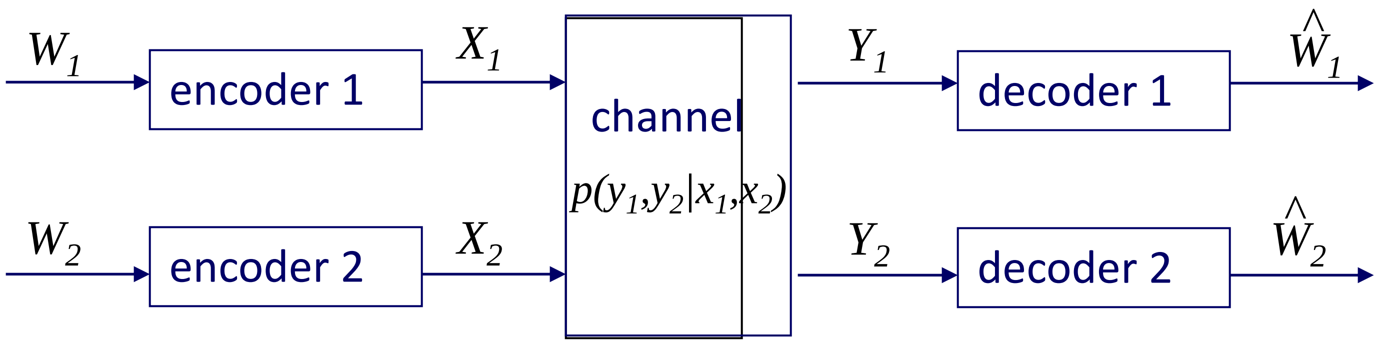
(b) Effective message digit rate rb = Rc*s and Information rate (R) <= rb thus the effective R will decreases.

GAUSSION CHANNLE

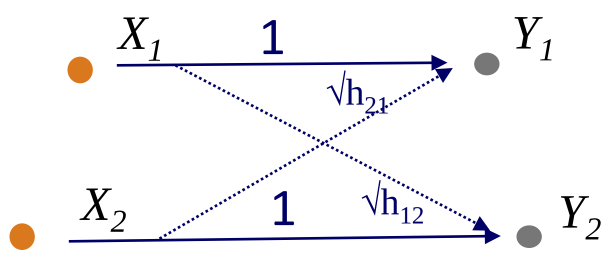
Interference Channel

How to cope with interference is not well understood

The interference channel capacity is a long standing open problem



Gaussian Channel in Standard Form



$$Y_1 = X_1 + \sqrt{h_{21}} X_2 + Z_1$$

$$Y_2 = \sqrt{h_{12}} X_1 + X_2 + Z_2$$

where $Z_1 \sim \mathcal{N}(0,1)$, $Z_2 \sim \mathcal{N}(0,1)$

Strong Interference

Focus of this work: *strong* interference

- Capacity region known if there is *strong* interference:

$$I(X_1;Y_1 \mid X_2) \le I(X_1;Y_2 \mid X_2)$$
$$I(X_2;Y_2 \mid X_1) \le I(X_2;Y_1 \mid X_1)$$

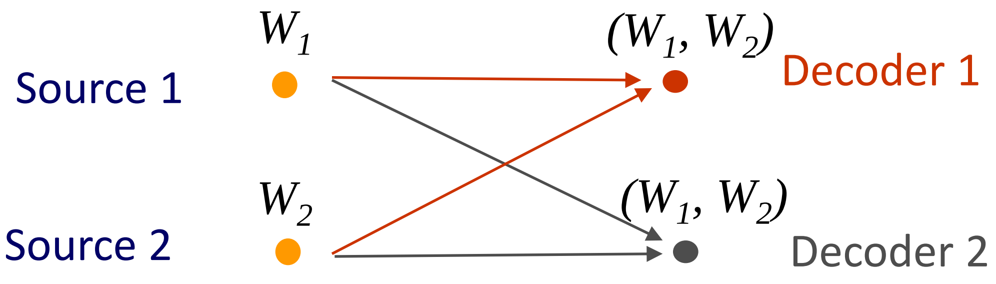
for all input product probability distributions *[Costa&El Gamal,1987]*

For the Gaussian channel in standard form, this means *[Sato, Han&Kobayashi, 1981]*

$$h_{12} \ge 1$$
$$h_{21} \ge 1$$

Capacity region = Capacity region of a compound MAC in which *both* messages are decoded at *both* receivers *[Ahlswede, 1974]*

Compound MAC



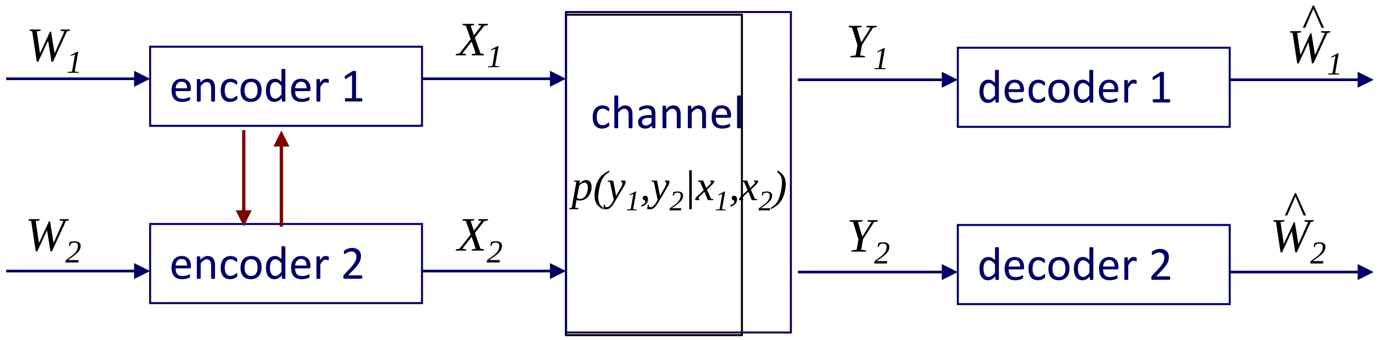
Cooperation in Interference Channel

Not captured in the interference channel model:

- Broadcasting: sources "overhear" each other's transmission
- Cooperation

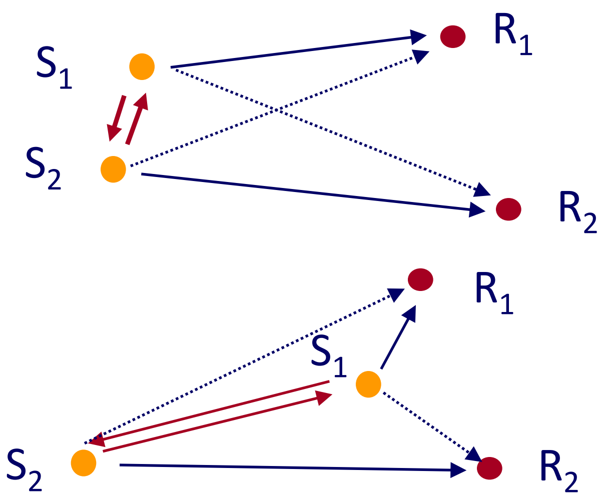
Our goal: consider cooperation and derive capacity results

- Work on more involved problems than already unsolved?
- What are insightful and tractable channel models?
- How do we model cooperation?



Transmitter Cooperation for Gaussian Channels

- Full cooperation: MIMO broadcast channel
  - DPC optimal *[Weingarten, Steinberg & Shamai, 04], [Caire & Shamai, 01], [Viswanath, Jindal & Goldsmith, 02]*
- Several cooperation strategies proposed *[Host-Madsen, Jindal, Mitra& Goldsmith, 03, Ng & Goldsmith, 04]*
- *[Jindal, Mitra& Goldsmith, Ng& Goldsmith]:*
  - Gain from DPC when sources close together
  - When apart, relaying outperforms DPC
  - **Assumptions**:
  - Dedicated orthogonal cooperation channels
  - Total power constraint



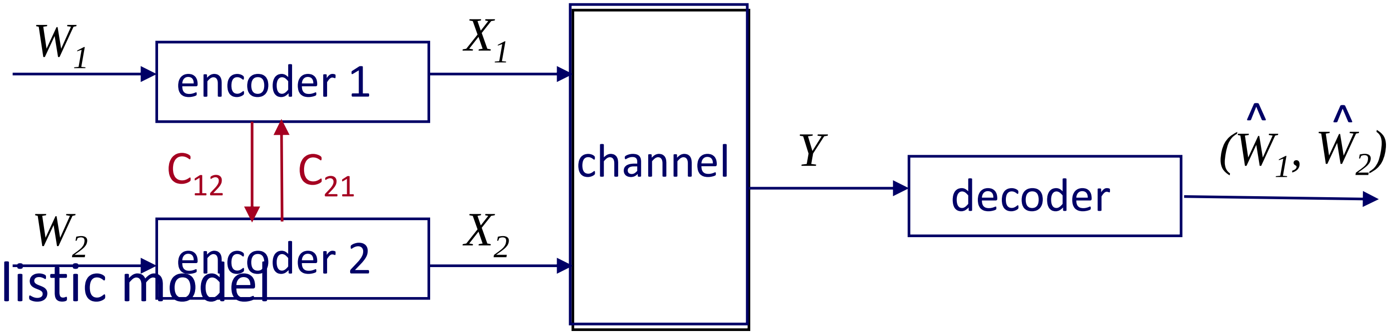
Transmitter Cooperation In Discrete Memoryless Channel

MAC with partially cooperating encoders *[Willems, 1983]*

Capacity region $C_{\text{MAC}}(C_{12}, C_{21})$ determined by Willems, 1983

Communication through *conference*

Realistic model



Cooperation through Conference

Each transmitter sends *K* symbols $(V_{t1},\ldots,V_{tK}),\quad t=1,2$

V_{1k} depends on previously received $V_2^{k-1}=(V_{21},\ldots,V_{2k-1})$

$$V_{1k}=h_{1k}(W_1,V_2^{k-1})$$
$$V_{2k}=h_{2k}(W_2,V_1^{k-1})\quad k=1,\ldots,K \qquad K\geq 1$$

Alphabet size of V_1^K is at most NC_{12}

$$\sum_{k=1}^{K}\log(\|\mathcal{V}_{1k}\|)\leq NC_{12}$$

Alphabet size of V_2^K is at most NC_{21}

$$\sum_{k=1}^{K}\log(\|\mathcal{V}_{2k}\|)\leq NC_{21}$$

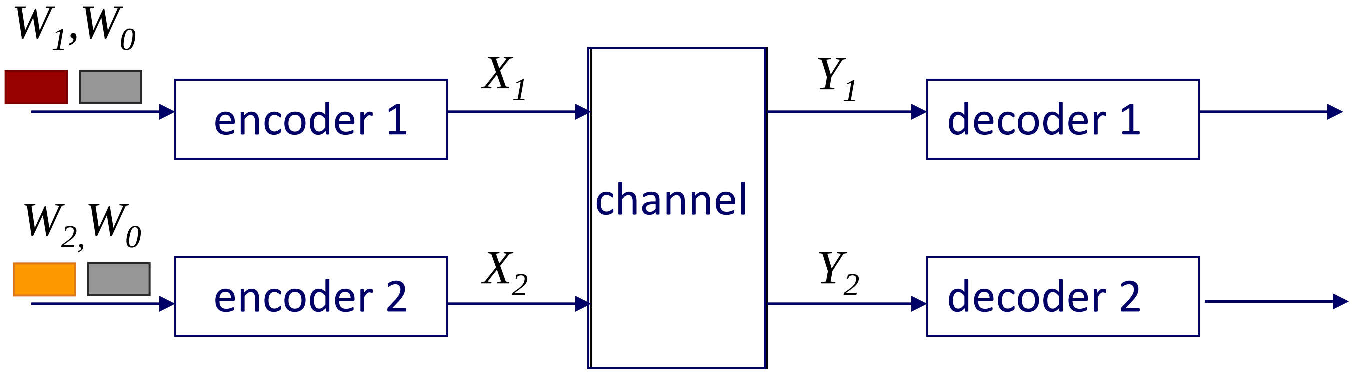
Partial Transmitter Cooperation

In the conference: partial information about messages W_1, W_2 exchanged between encoders

After the conference:

Common message at the encoders

Encompasses scenarios in which encoders have a partial knowledge about each other's messages



Transmitter Cooperation

Limited cooperation allows transmitters to exchange partial information about each other messages

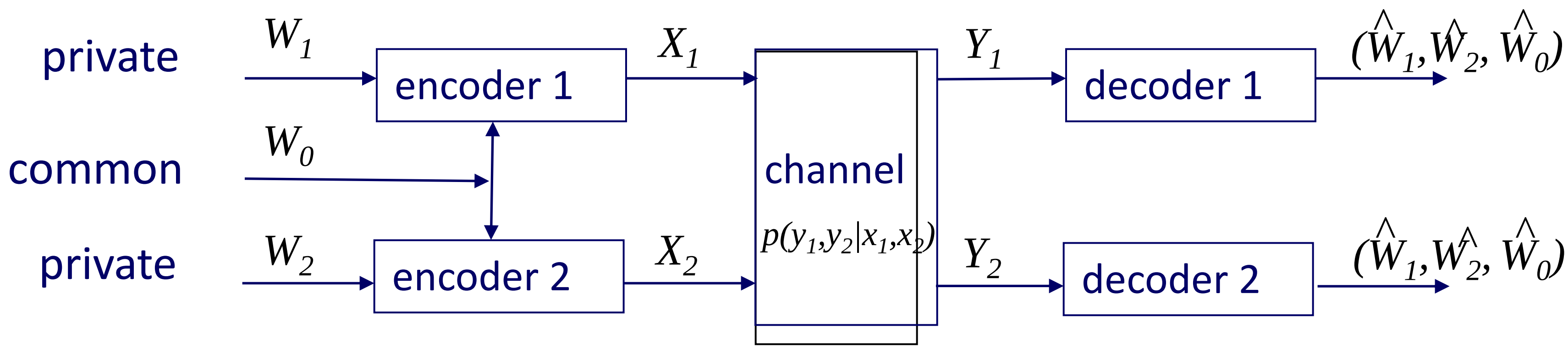
After cooperation

Common message known to both encoders

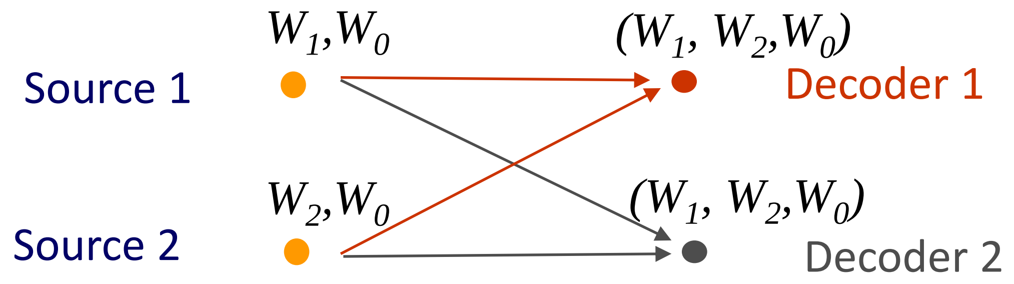
Private message at each encoder

In the strong interference, the capacity region tied to the capacity region of *The Compound MAC with Common Information*

Two MAC channels with a common and private messages



Compound MAC with Common Information



The Compound MAC with Common Information

- Encoding

$$\mathbf{x}_t = f_t(W_t, W_0)$$

- Decoding

$$\left(\hat{W}_1(t), \hat{W}_2(t), \hat{W}_0(t)\right) = g_t(\mathrm{Y}_t) \qquad t = 1,2$$

- The error probability

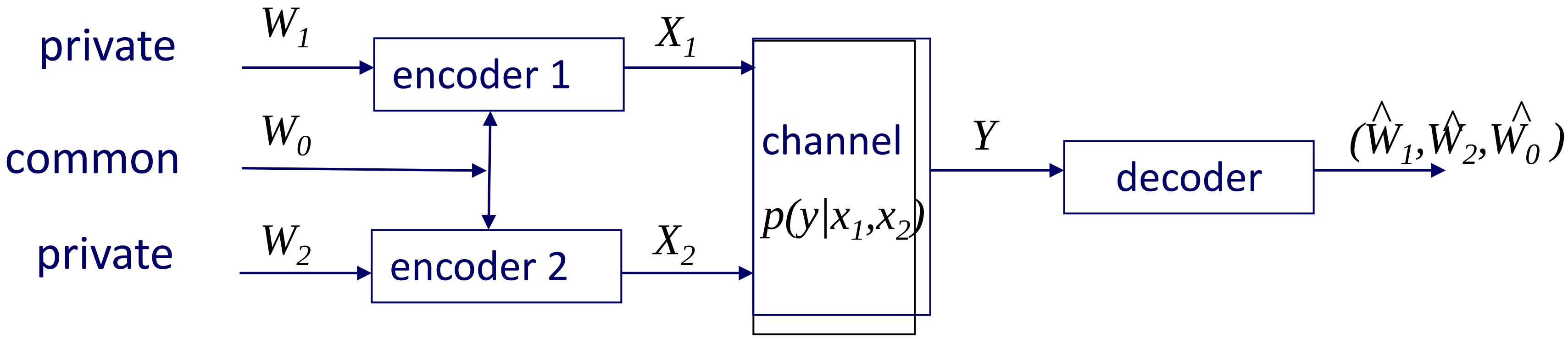
$$P_e = P\left[g_1(\mathrm{Y}_1) \neq (W_0, W_1, W_2) \cup g_2(\mathrm{Y}_2) \neq (W_0, W_1, W_2)\right]$$

- (R_0, R_1, R_2) achievable if, for any $\varepsilon > 0$, there is an (M_0, M_1, M_2, N, P_e) code such that

$$P_e \leq \varepsilon \qquad M_i \geq 2^{NR_i} \qquad i = 0,1,2$$

- The capacity region is the closure of the set of all achievable (R_0, R_1, R_2)
  - Easy to determine given the result by *[Slepian & Wolf, 1973]*

The MAC with Common Information

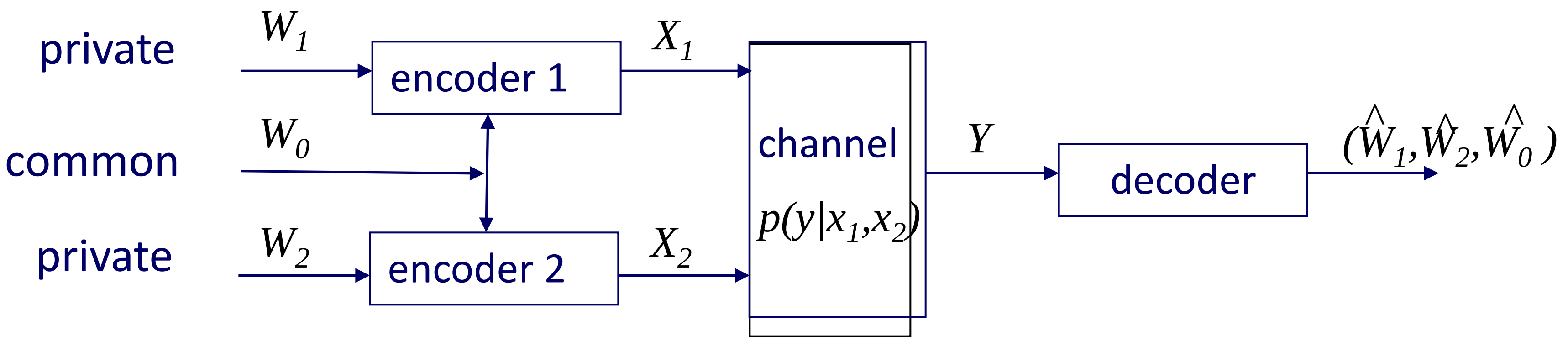


Capacity region *[Slepian & Wolf , 1973]*

$$
C_{MAC} = \bigcup \{(R_0, R_1, R_2): \quad
\begin{aligned}
&R_1 \le I(X_1;Y \mid X_2, U) \\
&R_2 \le I(X_2;Y \mid X_1, U) \\
&R_1 + R_2 \le I(X_1, X_2;Y \mid U) \\
&R_0 + R_1 + R_2 \le I(X_1, X_2;Y)\}
\end{aligned}
$$

union over $p(u)p(x_1|u)p(x_2|u)p(y|x_1,x_2)$

The MAC with Common Information



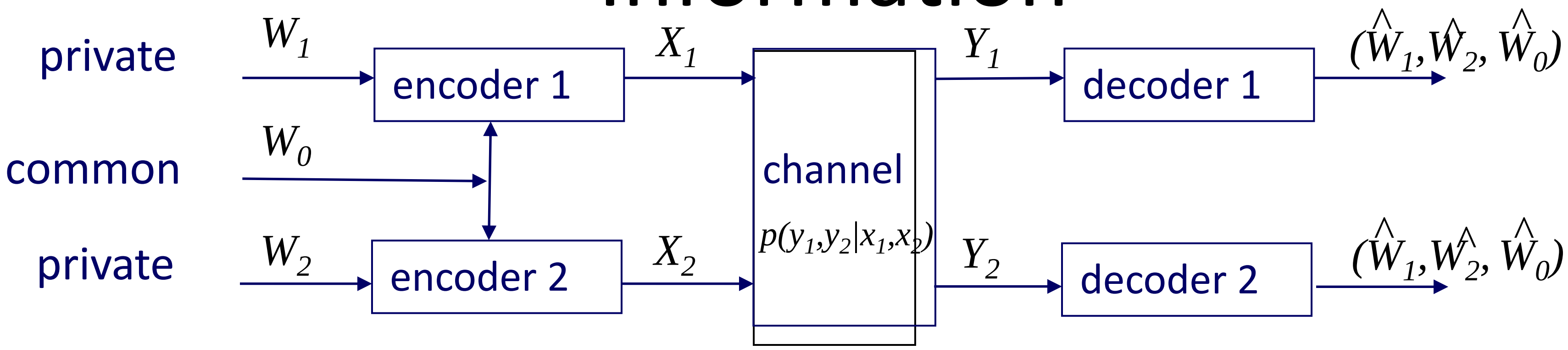
Capacity region *[Slepian & Wolf , 1973]*

$$C_{MAC} = \bigcup \{ (R_0, R_1, R_2) : \quad \begin{aligned} R_1 &\le I(X_1;Y \mid X_2, U) \\ R_2 &\le I(X_2;Y \mid X_1, U) \\ R_1 + R_2 &\le I(X_1, X_2;Y \mid U) \\ R_0 + R_1 + R_2 &\le I(X_1, X_2;Y) \} \end{aligned}$$

$\mathcal{R}_{MAC}(p')$

union over $p(u)p(x_1|u)p(x_2|u)p(y|x_1,x_2)$

The Compound MAC with Common Information



Capacity region *[MYK, ISIT 2005]*

$$C_{CMAC} = \bigcup \Big\{ (R_0, R_1, R_2): \quad \begin{aligned} R_1 &\le \min\{I(X_1;Y_1 \mid X_2,U), I(X_1;Y_2 \mid X_2,U)\} \\ R_2 &\le \min\{I(X_2;Y_1 \mid X_1,U), I(X_2;Y_2 \mid X_1,U)\} \\ R_1 + R_2 &\le \min\{I(X_1,X_2;Y_1 \mid U), I(X_1,X_2;Y_2 \mid U)\} \\ R_0 + R_1 + R_2 &\le \min\{I(X_1,X_2;Y_1), I(X_1,X_2;Y_2)\} \Big\} \end{aligned}$$

union over $p(u)p(x_1|u)p(x_2|u)p(y_1,y_2|x_1,x_2)$

The Compound MAC with Common Information

Capacity region

$$\boldsymbol{C}_{CMAC} = \bigcup_{\boldsymbol{p}} \left\{ \mathcal{R}_{MAC1}(\boldsymbol{p}) \cap \mathcal{R}_{MAC2}(\boldsymbol{p}) \right\}$$

union over $p(u)p(x_1|u)p(x_2|u)p(y_1,y_2|x_1,x_2)$

- For each $\boldsymbol{p}$: (R_0,R_1,R_2) is an intersection of rate regions $\mathcal{R}_{MACt}$ achieved in two MACs with common information:

$$p_1 = p(y_1 \mid x_1, x_2) = \sum_{y_2} p(y_1, y_2 \mid x_1, x_2) \qquad p_2 = p(y_2 \mid x_1, x_2) = \sum_{y_1} p(y_1, y_2 \mid x_1, x_2)$$

Converse

- Error probability in MAC$_t$ $$P_{et} = P\left[g_t(\mathbf{Y}_t) \neq \left(W_0, W_1, W_2\right)\right] \quad t = 1,2$$
- Error probability in CMAC $$P_e = P\left[g_1(\mathbf{Y}_1) \neq \left(W_0, W_1, W_2\right) \cup g_2(\mathbf{Y}_2) \neq \left(W_0, W_1, W_2\right)\right]$$

$$\Rightarrow \quad \boxed{\max\left\{P_{e1}, P_{e2}\right\} \leq P_e}$$

$$P_e \to 0 \qquad P_{e1} \to 0, \quad P_{e2} \to 0$$

→ Necessary condition for :

→ Rates confined to $\mathcal{R}_{MAC1}(\boldsymbol{p})$ and $\mathcal{R}_{MAC2}(\boldsymbol{p})$ for every $\boldsymbol{p}$

Achievability

The probability of error

$$
\begin{aligned}
P_e &= P\big[g_1(\mathbf{Y}_1) \neq (W_0, W_1, W_2) \cup g_2(\mathbf{Y}_2) \neq (W_0, W_1, W_2)\big] \\
&\leq P\big[g_1(\mathbf{Y}_1) \neq (W_0, W_1, W_2)\big] + P\big[g_2(\mathbf{Y}_2) \neq (W_0, W_1, W_2)\big] \\
&= P_{e1} + P_{e2}
\end{aligned}
$$

- From Slepian and Wolf result, choosing the rates (R_0, R_1, R_2) in

$$\left\{\mathcal{R}_{MAC\,1} \bigcap \mathcal{R}_{MAC\,2}\right\}$$

will guarantee that P_{e1} and P_{e2} can be made arbitrarily small

→ P_e will be arbitrarily small

Implications

- We can use this result to determine the capacity region of several channel with partial transmitter cooperation:

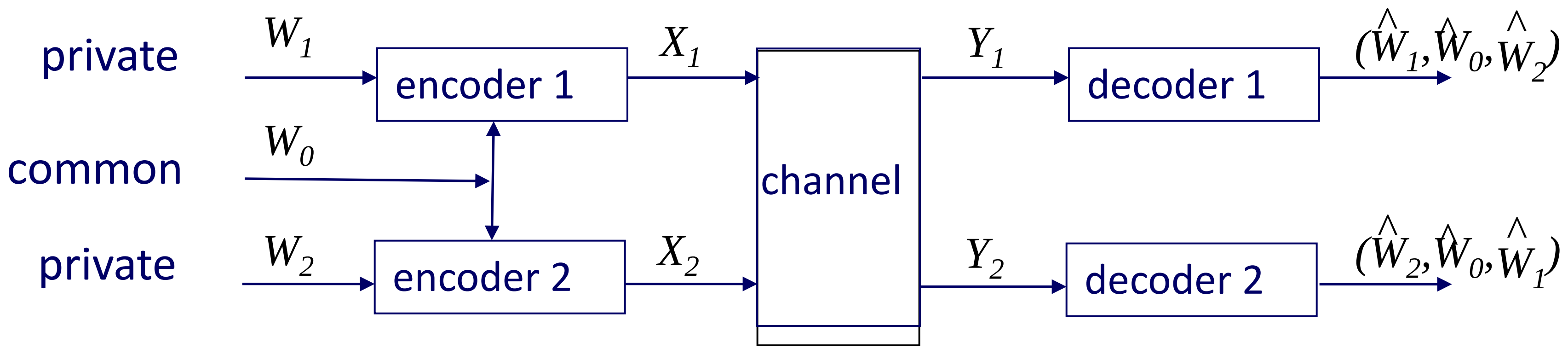
1. The Strong Interference Channel with Common Information

2. The Strong Interference Channel with Unidirectional Cooperation

3. The Compound MAC with Conferencing Encoders

After The Conference

Compound MAC with common information



Relax the decoding constraints:

Each receiver decodes only *one* private message

Theorem

- For the interference channel with common information satisfying

$$I(X_1;Y_1 \mid X_2) \le I(X_1;Y_2 \mid X_2)$$
$$I(X_2;Y_2 \mid X_1) \le I(X_2;Y_1 \mid X_1)$$

for all product input distributions the capacity region $\mathcal{C}$ is

$$\mathcal{C} = \bigcup \{ (R_0, R_1, R_2) : \begin{array}{l} R_1 \le I(X_1;Y_1 \mid X_2, U) \\ R_2 \le I(X_2;Y_2 \mid X_1, U) \\ R_1 + R_2 \le \min\{ I(X_1, X_2; Y_1 \mid U), I(X_1, X_2; Y_2 \mid U) \} \\ R_0 + R_1 + R_2 \le \min\{ I(X_1, X_2; Y_1), I(X_1, X_2; Y_1) \} \} \end{array}$$

union over $p(u)p(x_1|u)p(x_2|u)p(y_1,y_2|x_1,x_2)$

Capacity region = capacity region of the compound MAC with common information

Proof

- Achievability
  - Follows directly from the achievability in the Compound MAC with common information
    - Decoding constraint is relaxed
- Converse
  - Using Fano's inequality
  - In the interference channel with no cooperation: outer bounds rely on the independence of X_1 and X_2
  - Due to cooperation: Codewords not independent

- Theorem conditions obtained from the converse

Relationship to the Strong Interference Conditions

Strong interference channels conditions

$$I(X_1;Y_1 \mid X_2) \le I(X_1;Y_2 \mid X_2)$$
$$I(X_2;Y_2 \mid X_1) \le I(X_2;Y_1 \mid X_1)$$

for all product input distributions $\boldsymbol{p_{X_1}(x_1)p_{X_2}(x_2)}$

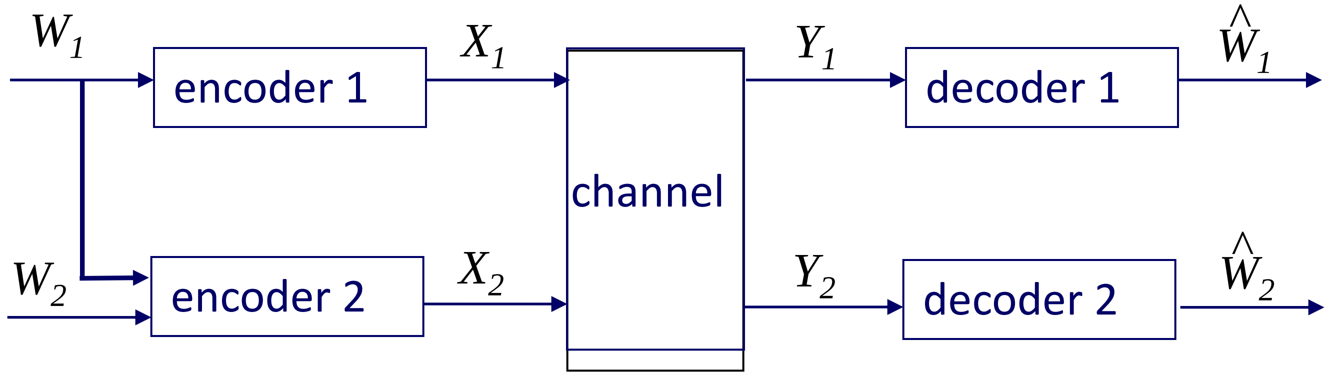
$\Leftrightarrow$

$$I(X_1;Y_1 \mid X_2,U) \le I(X_1;Y_2 \mid X_2,U)$$
$$I(X_2;Y_2 \mid X_1,U) \le I(X_2;Y_1 \mid X_1,U)$$

for all input distributions $\boldsymbol{p_U(u)p_{X_1|U}(x_1|u)p_{X_2|U}(x_2|u)}$

The same interference channel class satisfies two sets of conditions

Interference Channel With Unidirectional Cooperation



The difference from the interference channel: one encoder knows both messages

- Encoding functions

$$\mathbf{x}_1 = f_1(W_1)$$
$$\mathbf{x}_2 = f_2(W_1, W_2)$$

- Decoding functions

$$\hat{W}_1 = g_1(\mathbf{Y}_1)$$
$$\hat{W}_2 = g_2(\mathbf{Y}_2)$$

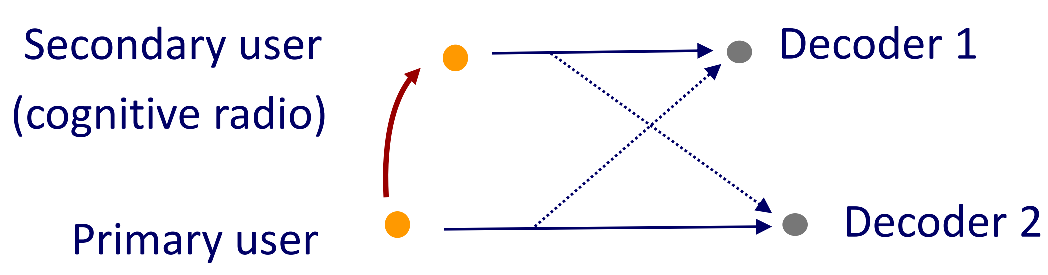
The error probability $P_e = \max\{P_{e,1}, P_{e,2}\}$ for $P_{e,t} = P[g_t(\mathbf{Y}_t) \neq W_t] \quad t = 1,2$

Cognitive Radio Settings

- Cognitive Radio Channel *[Devroye, Mitran, Tarokh, 2005]*
  - An achievable rate region
- Consider simple two-transmitter, two-receiver network:
  Assume that one transmitter is cognitive
  - It can "overhear" transmission of the primary user
  - It obtains *partially* the primary user's message ⇨ it can cooperate

⇨ Interference channel with unidirectional cooperation

The assumption that the full message W_1 is available at the cognitive user is an over-idealization



Interference Channel with Unidirectional Cooperation

- The Interference Channel with Unidirectional Cooperation *[Marić, Yates & Kramer, 2005]*
  - Capacity in very strong interference
- The Interference Channel with Degraded Message Set *[Wu, Vishwanath & Arapostathis, 2006]*
  - Capacity for weak interference and for Gaussian channel in weak interference
- Cognitive Radio Channel *[Jovičić& Viswanath, 2006]*
  - Capacity for Gaussian channel in weak interference
- The Interference Channel with a Cognitive Transmitter *[Marić, Goldsmith, Kramer& Shamai, 2006]*
  - New outer bounds and an achievable region

Theorem

- For the interference channel with unidirectional cooperation satisfying

$$I(X_2;Y_2 \mid X_1) \le I(X_2;Y_1 \mid X_1)$$
$$I(X_1,X_2;Y_1) \le I(X_1,X_2;Y_2)$$

for all joint input distributions $p(x_1,x_2)$, the capacity region $\mathcal{C}$ is

$$\mathcal{C} = \bigcup \{(R_1,R_2): \\ R_2 \le I(X_2;Y_2 \mid X_1) \\ R_1 + R_2 \le I(X_1,X_2;Y_1)\}$$

where the union is over $p(x_1,x_2)p(y_1,y_2|x_1,x_2)$.

- Capacity region = capacity region of the Compound MAC channel with Common Information

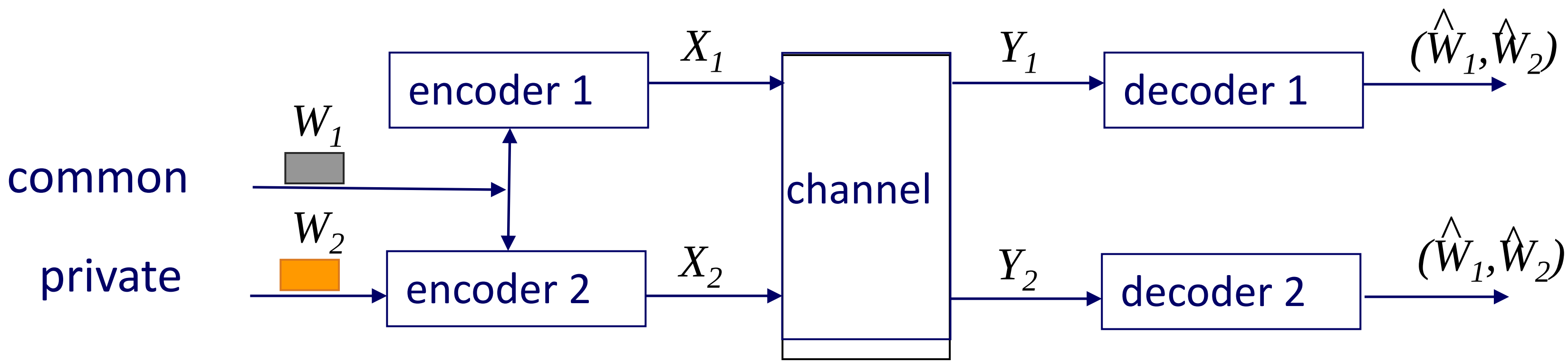
Achievability: Compound MAC with Common Information

Encode as for a Compound MAC with Common Information

Due to unidirectional cooperation:

- W_1 is a common message
- encoder 1 has no private message ⇨ $R'_0=R_1$, $R'_1=0$ choose: $U=X_1$

$$C_{CMAC}=\bigcup\{(R_1,R_2):\ R_2\le\min\{I(X_2;Y_1\mid X_1),I(X_2;Y_2\mid X_1)\}$$
$$R_1+R_2\le\min\{I(X_1,X_2;Y_1),I(X_1,X_2;Y_2)\}\}$$



Converse

- Using Fano's inequality
- Interference channel with no cooperation: outer bounds rely on the independence of X_1 and X_2
- Due to cooperation: Codewords not independent
- Theorem conditions obtained from the converse

Converse (*Continued*)

Bound $$N\left(R_1 + R_2\right) \le \sum_{n=1}^{N} I(X_{1n}, X_{2n}; Y_{1n})$$

Requires $$I(\mathbf{X}_2;\mathbf{Y}_2 \mid \mathbf{X}_1, W_1) \le I(\mathbf{X}_2;\mathbf{Y}_1 \mid \mathbf{X}_1, W_1)$$

- Lemma: If per-letter conditions $I(X_2;Y_2|X_1) \le I(X_2;Y_1|X_1)$ are satisfied for all $p(x_1,x_2)$, then

$$I(\mathbf{X}_2;\mathbf{Y}_2 \mid \mathbf{X}_1, W_1) \le I(\mathbf{X}_2;\mathbf{Y}_1 \mid \mathbf{X}_1, W_1)$$

Proof: Similar to proof by *[Costa& El Gamal,1987]*

with the changes

X_1, X_2 not independent

Conditioning on W_1

We would next like $$N\left(R_1 + R_2\right) \le \sum_{n=1}^{N} I(X_{1n}, X_{2n}; Y_{2n}) + N\delta_N$$

But because the situation is asymmetric, this seems difficult

Converse (*Continued*)

- Recall that the achievable rates are

$$\mathcal{C}_{CMAC} = \bigcup \Big\{ (R_1, R_2): \quad R_2 \le \min\{ I(X_2;Y_1 \mid X_1), I(X_2;Y_2 \mid X_1) \} \\ R_1 + R_2 \le \min\{ I(X_1,X_2;Y_1), I(X_1,X_2;Y_2) \} \Big\}$$

- By assumption, for all $p(x_1,x_2)$

$$I(X_2;Y_2 \mid X_1) \le I(X_2;Y_1 \mid X_1)$$
$$I(X_1,X_2;Y_1) \le I(X_1,X_2;Y_2)$$

- The rates

$$\bigcup \Big\{ (R_1, R_2): \; R_2 \le I(X_2;Y_2 \mid X_1) \\ R_1 + R_2 \le I(X_1,X_2;Y_1) \Big\}$$

union over all $p(x_1,x_2)$

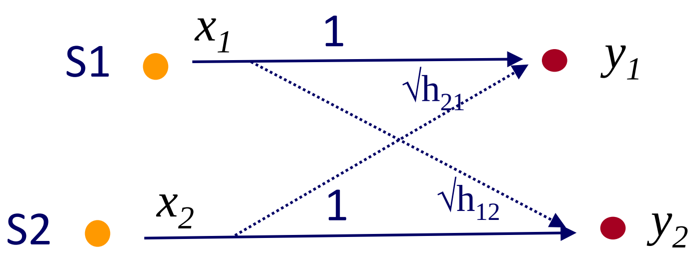
are thus achievable <u>and</u> are an outer bound

Gaussian Channel

- Channel outputs:

$$y_{1i} = x_{1i} + \sqrt{h_{21}}x_{2i} + z_{1i}$$

$$y_{2i} = \sqrt{h_{12}}x_{1i} + x_{2i} + z_{2i}$$



Power constraints:

$$\frac{1}{N}\sum_{i=1}^{N} x_{ti}^2 \le P_t \quad t = 1,2$$

Noise:

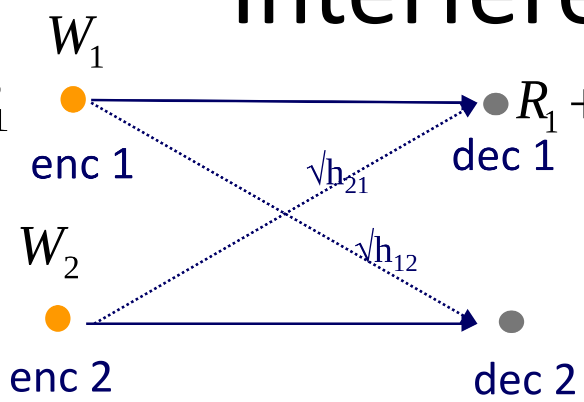
$$Z_1 \sim \mathcal{N}(0,1)$$

$$Z_2 \sim \mathcal{N}(0,1)$$

Capacity Achieving Scheme in Strong Interference

$X_1 \sim \mathrm{N}(0, P_1)$ $\quad P_1$

$X_2 = X_{20} + \sqrt{\frac{\alpha P_2}{P_1}} X_1$



$\alpha P_2 \quad (1-\alpha)P_2$

$$R_1 + R_2 = \frac{1}{2}\log\left(1 + h_{21}{}^2 P_2 + 2|h_{21}|\sqrt{\alpha P_1 P_2} + P_1\right)$$

$$R_2 = \frac{1}{2}\log\left(1 + (1-\rho^2)P_2\right)$$

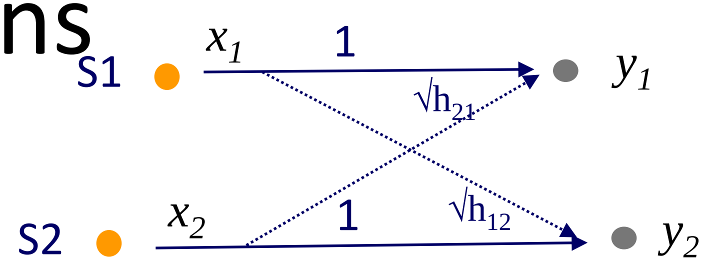
- Encoder 1: Codebook $\mathbf{x}_1(w_1)$ $\quad w_1 = 1, \ldots, 2^{NR_1}$ $\quad X_1 \sim \mathrm{N}(0, P_1)$

Encoder 2: - Dedicates a portion αP_2 to help

-Uses superposition coding $\quad X_{20} \sim \mathrm{N}(0, \alpha P_2)$

Decoders reduce interference as they can decode each other's messages

Gaussian Channel- Strong Interference Conditions



$$\alpha = \sqrt{P_1 / P_2}$$

Interference channel
– no cooperation:

$$h_{12} \geq 1$$
$$h_{21} \geq 1$$

Interference channel
with common information:

$$h_{12} \geq 1$$
$$h_{21} \geq 1$$

Unidirectional cooperation:
- more demanding conditions

$$|h_{21}| \geq 1$$
$$(1+\alpha h_{12})^2 \geq (\alpha + h_{21})^2$$
$$(1-\alpha h_{12})^2 \geq (\alpha - h_{21})^2$$

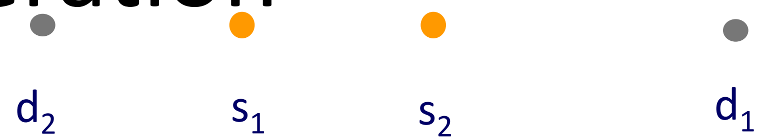
For $P_1 = P_2$ sufficient conditions:

$$h_{12} \geq h_{21}$$
$$h_{21} \geq 1$$

For $P_1 = 0$ conditions never satisfied
Channel reduces to a degraded broadcast channel from sender 2

Gaussian Channel with Unidirectional Cooperation

Strong interference scenario:

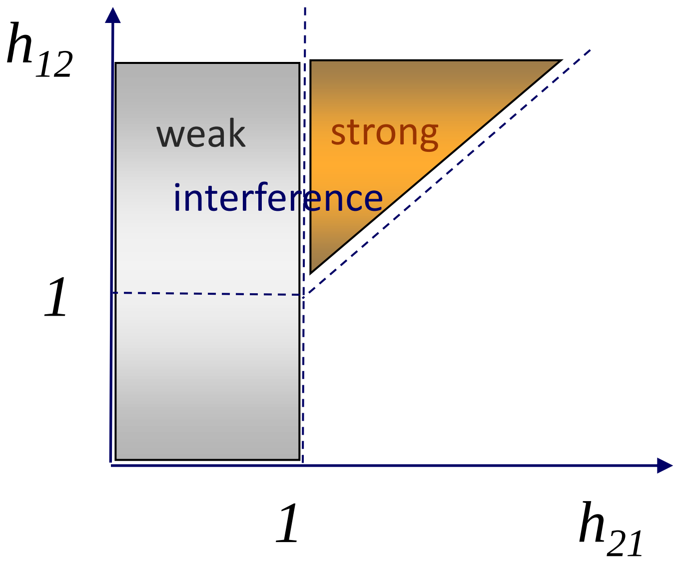


Let $\alpha = \sqrt{P_1 / P_2} = 1$

Sufficient conditions:

$$h_{12} \geq h_{21}$$

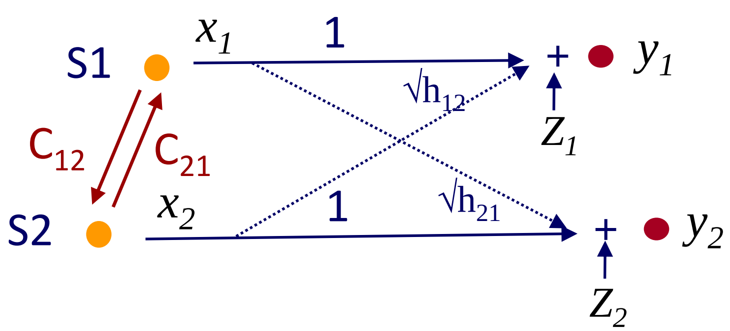
$$h_{21} \geq 1$$



Weak and strong interference regime solved

Our current work: in between regime

Gaussian Channel With Conferencing



Note: additional resources needed for the conference

- Channel outputs:

$$y_{1i} = x_{1i} + \sqrt{h_{12}}\, x_{2i} + z_{1i}$$

$$y_{2i} = \sqrt{h_{21}}\, x_{1i} + x_{2i} + z_{2i}$$

Power constraints:

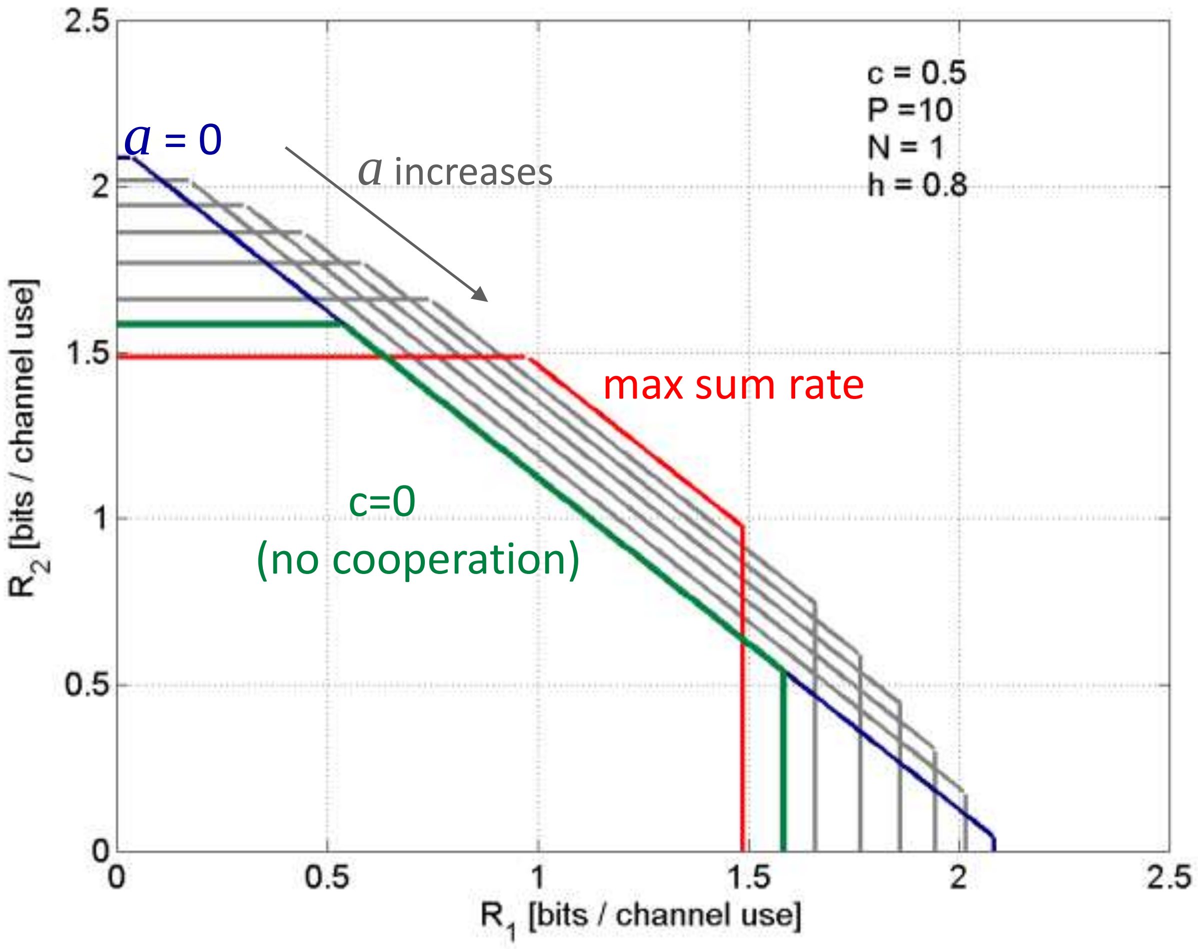
$$\frac{1}{N}\sum_{i=1}^{N} x_{ti}^2 \le P_t \quad t = 1,2$$

Noise:

$$Z_1 \sim \mathcal{N}(0,1)$$

$$Z_2 \sim \mathcal{N}(0,1)$$

Gaussian Channel With Conferencing



Symmetric channel

$$h_{12} = h_{21} = h$$
$$P_1 = P_2 = P$$
$$c = c_{12} = c_{21}$$
$$Z_t \sim \mathcal{N}(0,1) \quad t=1,2$$

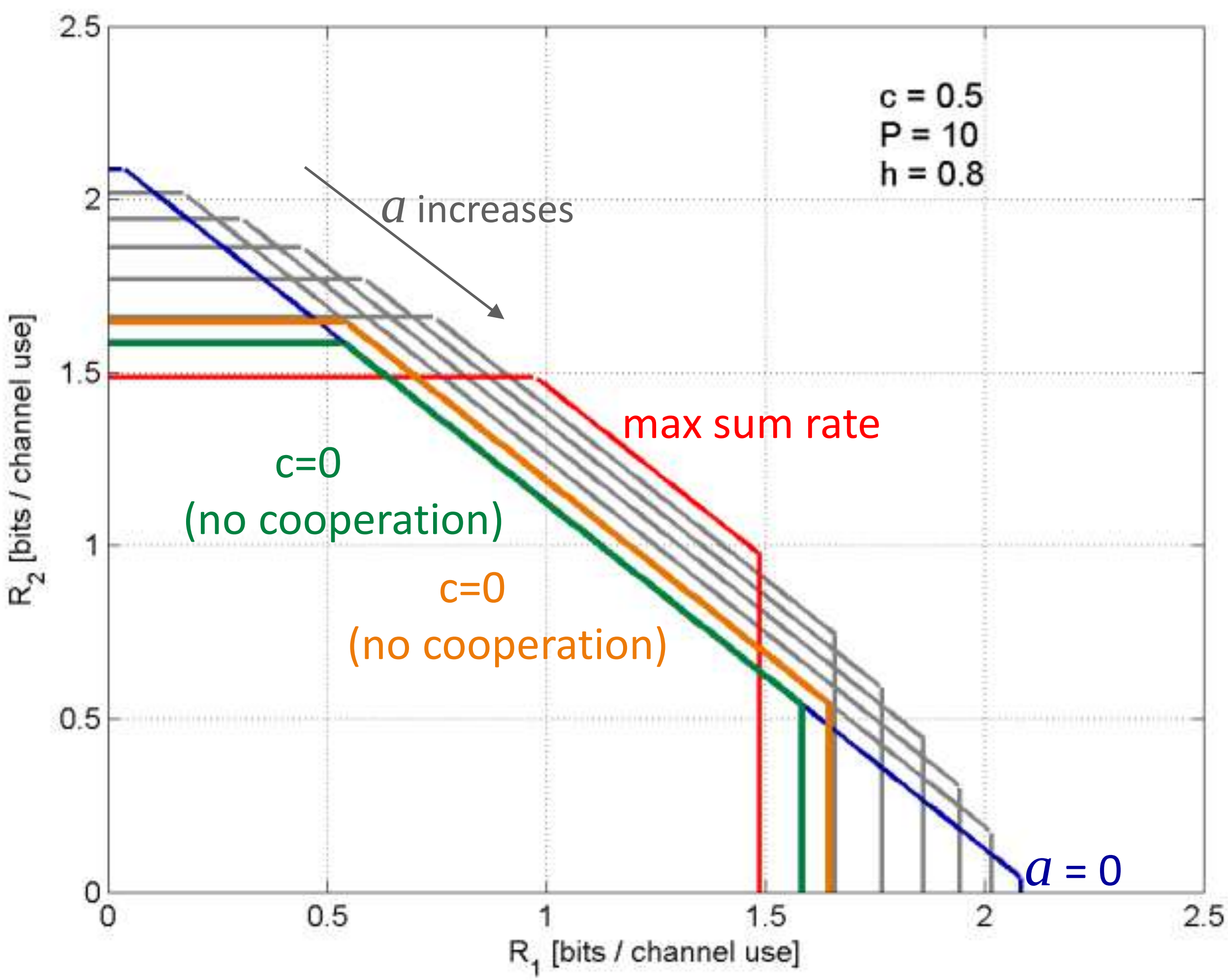
$$a = \frac{E[X_1 X_2]}{\sqrt{P_1 P_2}}$$

$a=0$: X_1, X_2 independent

Max sum-rate:

Two sum-rate bounds equal

Gaussian Channel With Conferencing



Symmetric channel

Sender t power, $t=1,2$

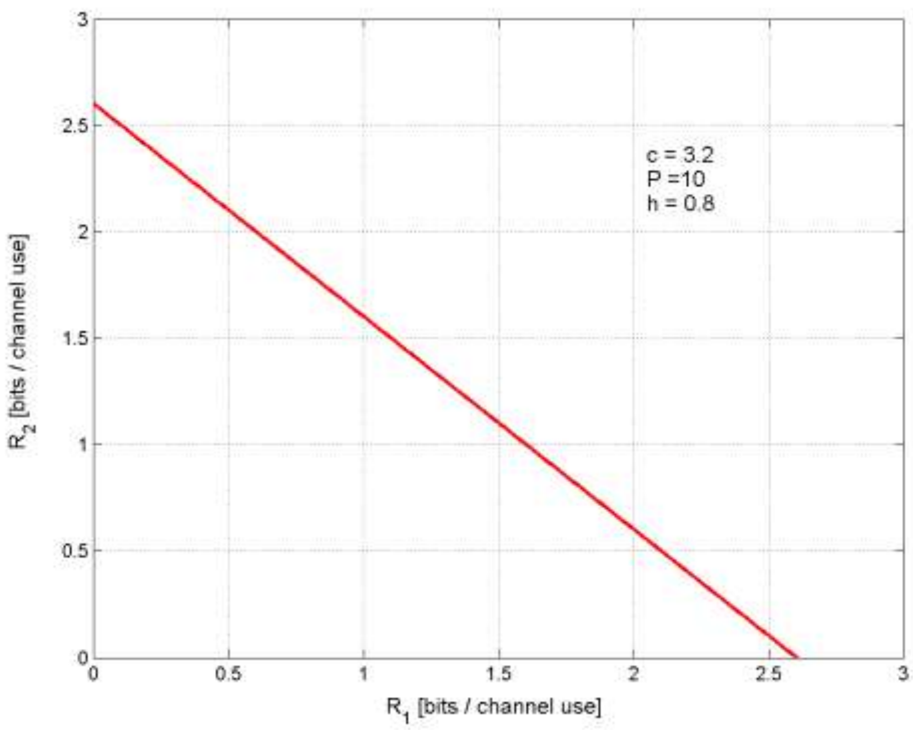
$$\frac{1}{N}\sum_{i=1}^{N} x_{ti}^2 \le P$$

$$\frac{1}{N}\sum_{i=1}^{N} x_{ti}^2 \le P + P_c$$

P_c-conference power

$$c = \frac{1}{2}\log(1+P_c)$$

Full Cooperation in Gaussian Channel



Large c:

Rate region

$$R_1 + R_2 \le C\left(\frac{P}{N}\left(h + 2a\sqrt{h}\right)\right)$$

→ Maximized for $a=1$

Senders able to exchange indexes W_1 and W_2

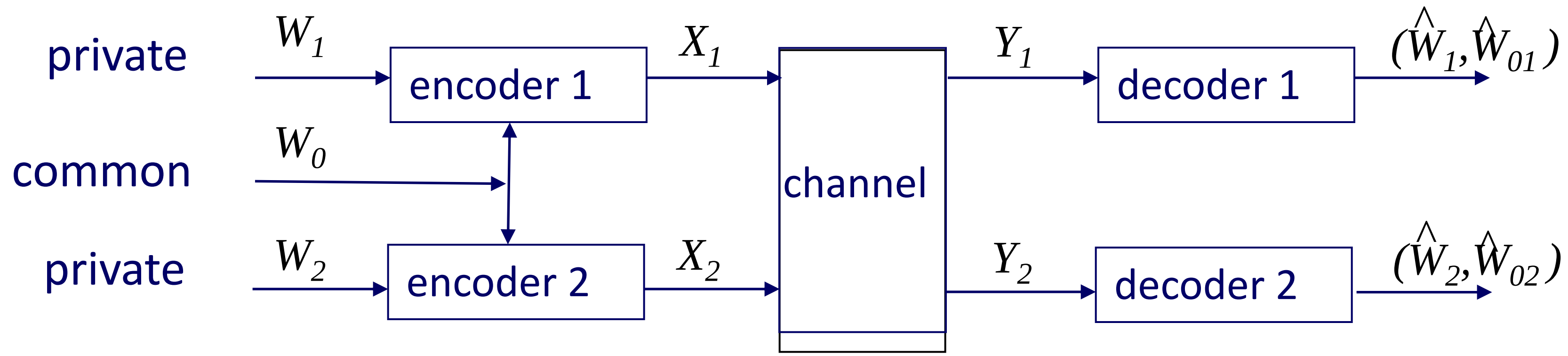
→ Full cooperation condition

$$c \ge C\left(\frac{P}{N}\left(1 + \sqrt{h}\right)^2\right)$$

→ Less cooperation needed as the receivers further away

Interference Channel With Conferencing

- Relax the constraint → Each user decodes only one message
  Strong interference conditions?
- After the conference



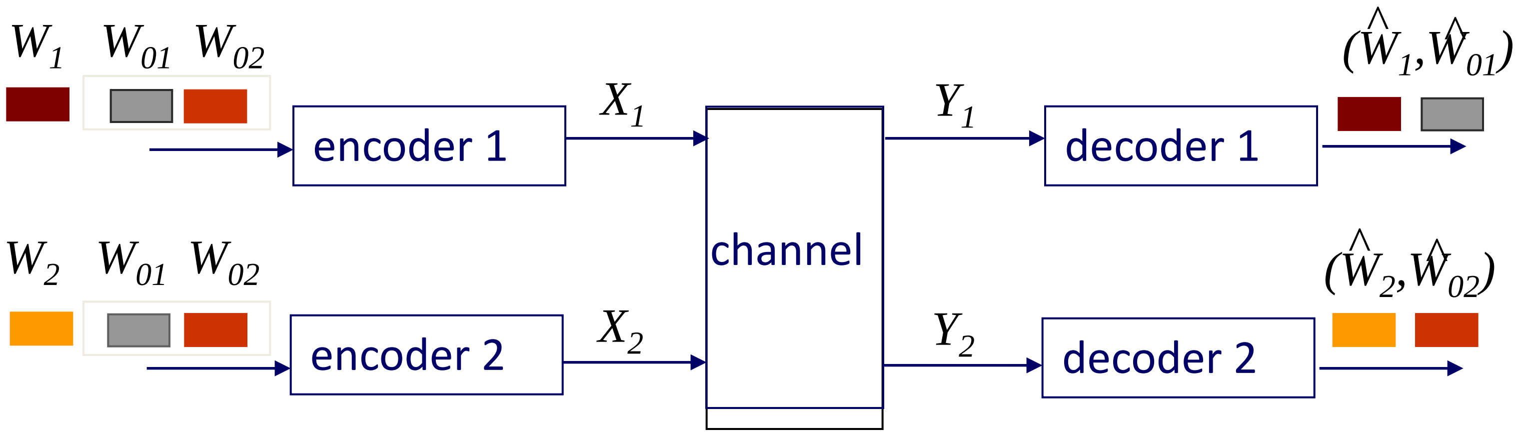
Common message contains information about *both* original messages

$$W_0=(W_{01},W_{02})$$

Decoders are interested in a *part* of the common message

Strong interference conditions?

Interference Channel With Conferencing



Encompasses various models:

- MIMO broadcast
- Unidirectional cooperation
- Partial cooperation

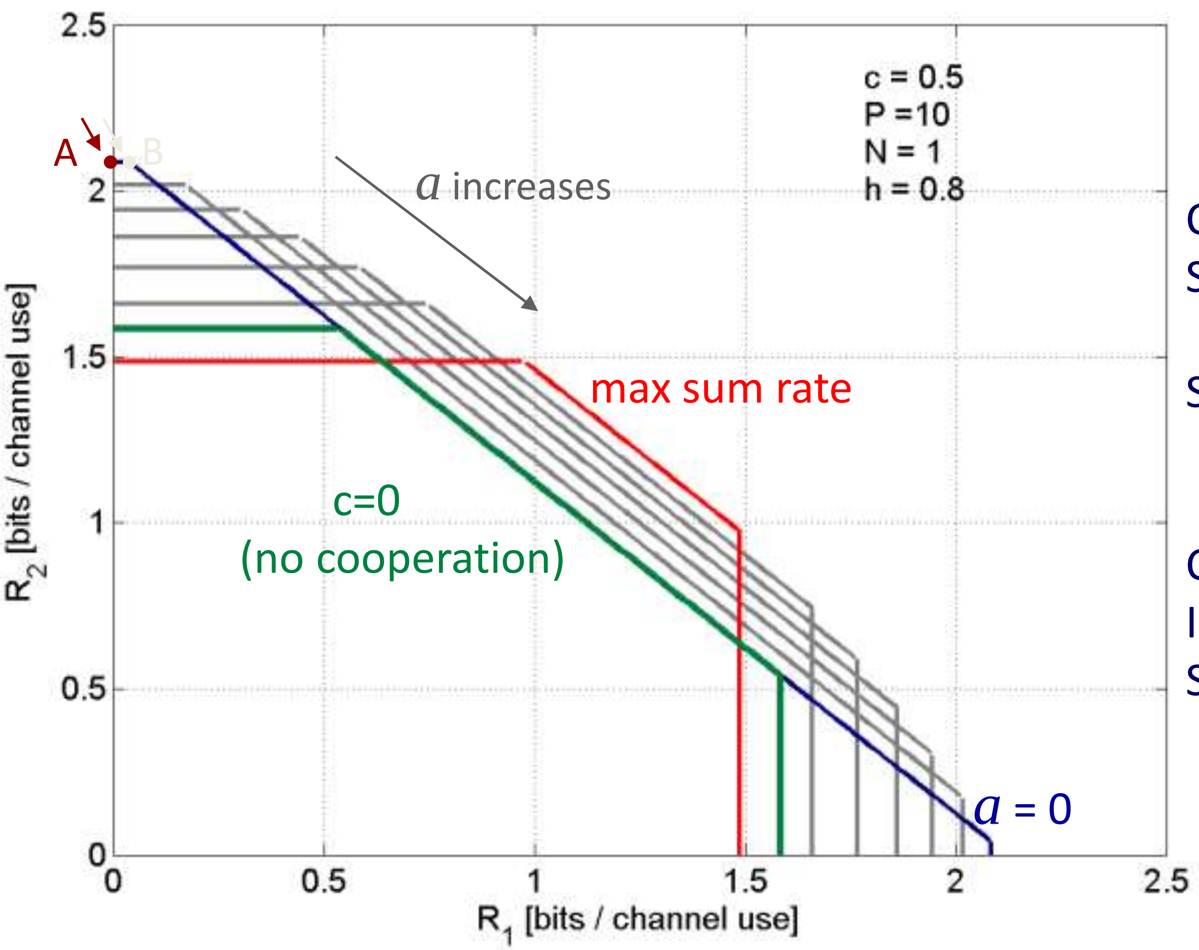
Strong Interference conditions?

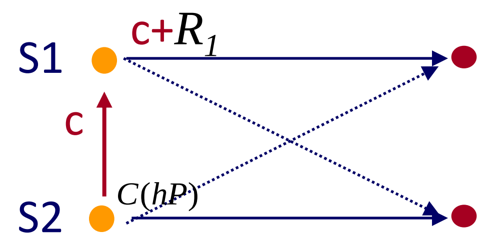
Discussion

- Introduced cooperation into the Interference Channel:
  - Capacity results for scenarios of *strong interference*
  - If the interference is strong : decoders can decode the interference
  - No partial decoding at the receivers → easier problem

  - Ongoing work:
  - Channel models that incorporate node cooperation
    - Capture characteristics of cognitive radio networks
  - Capacity results and cooperation gains for more general settings

The Gaussian Channel Capacity Region





Corner A:

S1 acts as a relay:

Sends at rate c

S2 sends independent information at rate $C(hP)$

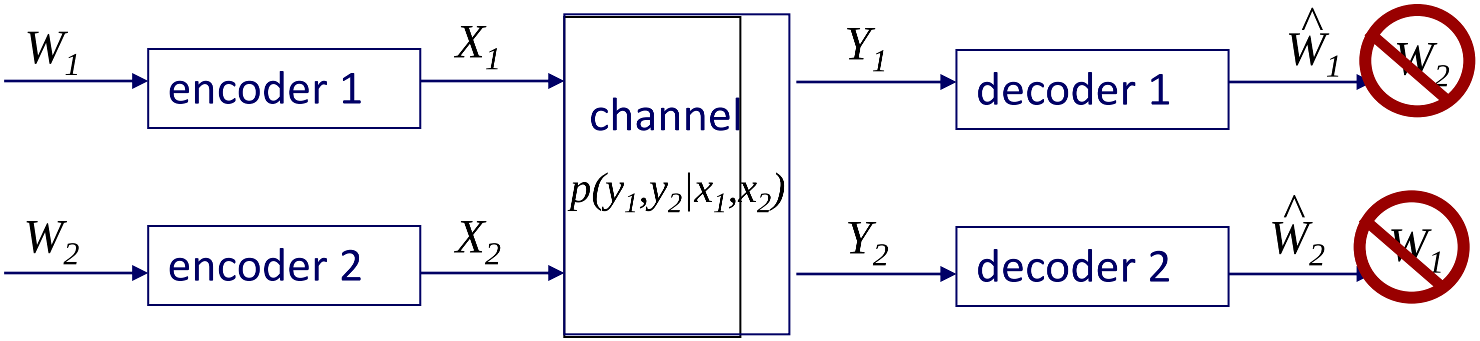
Corner B:

In addition, S1 sends own data

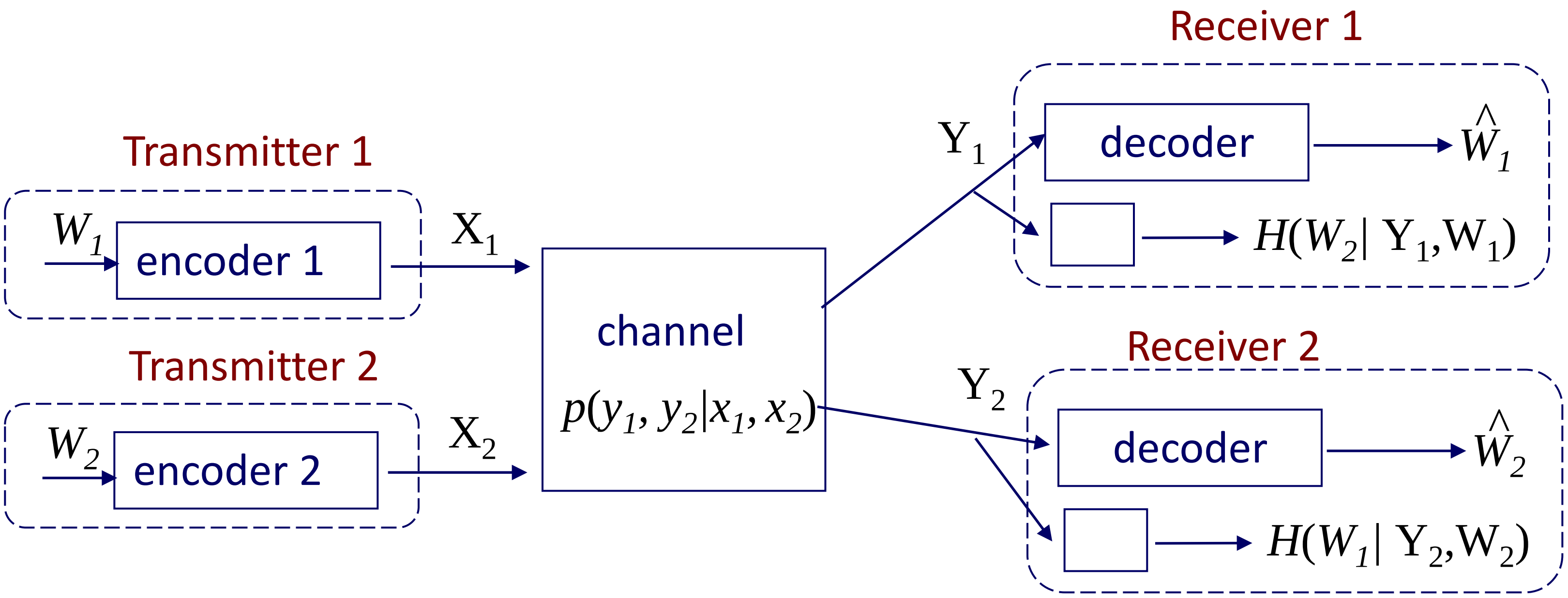
Sum-rate for $a=0$ achieved

$$R_1 + R_2 \le C(hP + P)$$

Interference Channel

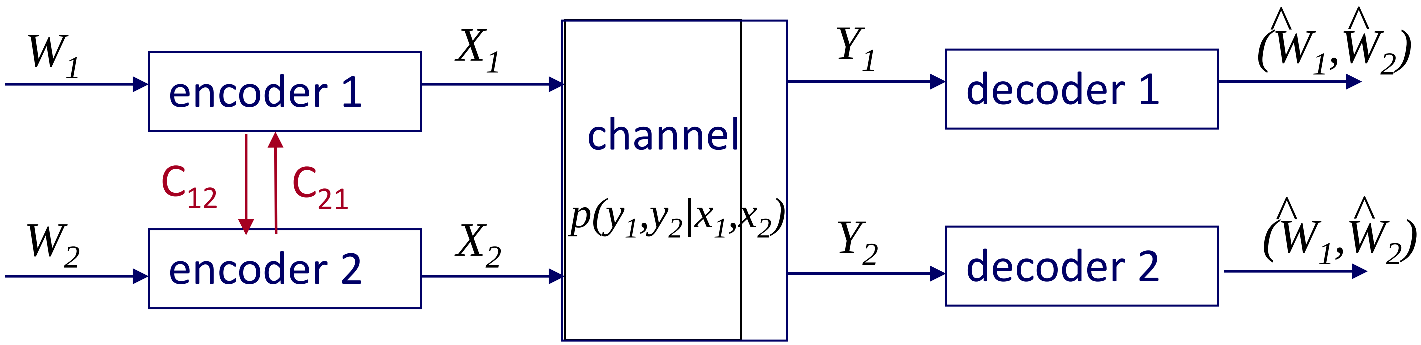


Interference Channel with Confidential Messages



- Joint work with: Ruoheng Liu, Predrag Spasojevic and Roy D. Yates
- Developed inner and outer bounds

The Compound MAC With Conferencing Encoders



Adopt Willems' cooperation model

Decoding constraint: Both messages decoded by both receivers

Encoding:

$$\mathbf{x}_1 = f_1\left(W_1, V_2^K\right)$$

$$\mathbf{x}_2 = f_2\left(W_2, V_1^K\right)$$

Decoding:

$$\left(\hat{W}_1, \hat{W}_2\right) = g_t(\mathbf{Y}_t)$$

The probability of error:

$$P_e = P\left[g_1(\mathbf{Y}_1) \neq (W_1, W_2) \cup g_2(\mathbf{Y}_2) \neq (W_1, W_2)\right]$$

Theorem

- The Compound MAC capacity region $\mathcal{C}(C_{12},C_{21})$ is

$$\mathcal{C}(C_{12},C_{21}) = \bigcup \{(R_1,R_2):$$

$$R_1 \le I(X_1;Y_1 \mid X_2,U) + C_{12} \qquad R_1 \le I(X_1;Y_2 \mid X_2,U) + C_{12}$$

union over $p(u)p(x_1|u)p(x_2|u)p(y_1,y_2|x_1,x_2)$ denoted $\boldsymbol{p}$

- For each $\boldsymbol{p}$: Intersection between rate regions of two MACs with partially cooperating encoders

$$R_2 \le I(X_2;Y_1 \mid X_1,U) + C_{21} \qquad R_2 \le I(X_2;Y_2 \mid X_1,U) + C_{21}$$

$$R_1 + R_2 \le I(X_1,X_2;Y_1 \mid U) + C_{12} + C_{21} \qquad R_1 + R_2 \le I(X_1,X_2;Y_2 \mid U) + C_{12} + C_{21}$$

$$\underbrace{R_1 + R_2 \le I(X_1,X_2;Y_1)}_{\mathcal{R}_{MAC1}(p,C_{12},C_{21})} \qquad \underbrace{R_1 + R_2 \le I(X_1,X_2;Y_2)\}}_{\mathcal{R}_{MAC2}(p,C_{12},C_{21})}$$

$$\mathcal{C}(C_{12},C_{21}) = \bigcup_{\boldsymbol{p}} \{\mathcal{R}_{MAC1}(\boldsymbol{p},C_{12},C_{21}) \cap \mathcal{R}_{MAC2}(\boldsymbol{p},C_{12},C_{21})\}$$

Gaussian Channel With Conferencing

Use the maximum entropy theorem

The Gaussian C-MAC capacity region

$$\mathcal{C}(C_{12}, C_{21}) = \bigcup \{(R_1, R_2):$$

$$0 \le R_1 \le \min_{j\in\{1,2\}} C(h_{j1}\bar{a}P_1) + C_{12}$$

$$0 \le R_2 \le \min_{j\in\{1,2\}} C(h_{j2}\bar{b}P_2) + C_{21}$$

$$0 \le R_1 + R_2 \le \min_{j\in\{1,2\}} C(h_{j1}\bar{a}P_1 + h_{j2}\bar{b}P_2) + C_{12} + C_{21}$$

$$0 \le R_1 + R_2 \le \min_{j\in\{1,2\}} C(h_{j1}P_1 + h_{j2}P_2 + 2\sqrt{h_{j1}aP_1 h_{j2}bP_2})\}$$

- Union over all
- $0 \le a \le 1, \quad 0 \le b \le 1$

$$h_{11} = h_{22} = 1, \quad \bar{a} = 1 - a, \quad \bar{b} = 1 - b$$

For $C_{12}=C_{21}=0$: optimum choice is $a=0$, $b=0$