Interactive Communication

Jie Ren

2012/8/14

ASPITRG, Drexel University

Outline

- Problem Setup (Two-Way Source Coding)
- Interaction in Lossless Expected Length Case
- Interaction in Zero-error Worst Length Case
- Interaction in Function Computation

Outline

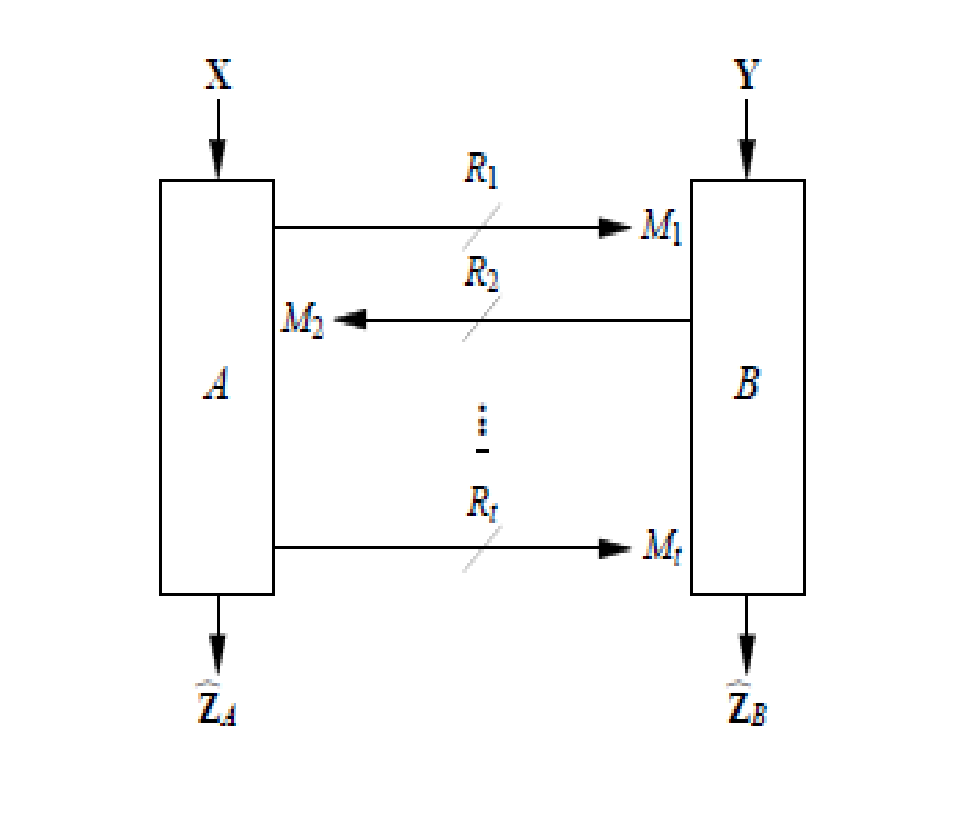
- Problem Setup
  - Two-way Source Coding model
  - Why Interested in It
  - Mathematical Description
  - Sum-rate-distortion Function
  - Proof(Achievability and Converse)
  - Problems Remain Open
- Interaction in Lossless Expected Length Case
- Interaction in Zero-error Worst Length Case
- Interaction in Function Computation

Outline

- Problem Setup
  - Two-way Source Coding model
  - Why Interested in It
  - Mathematical Description
  - Sum-rate-distortion Function
  - Proof(Achievability and Converse)
  - Problems Remain Open
- Interaction in Lossless Expected Length Case
- Interaction in Zero-error Worst Length Case
- Interaction in Function Computation

Two-way Source Coding model

- Two-terminal distributed source coding problem
- Reconstruct X/Y on both sides
- Alternating messages scheme (concurrent scheme)

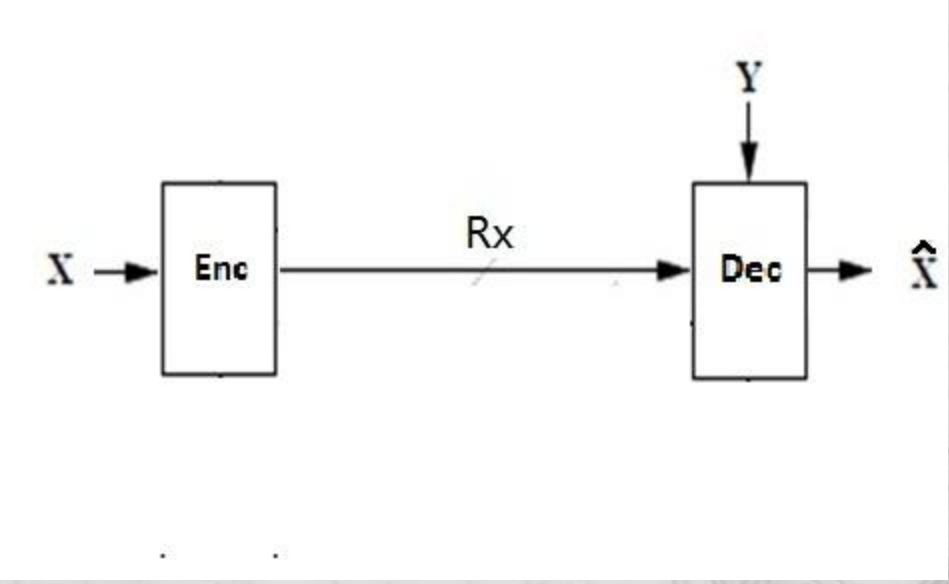


Outline

- Problem Setup
  - Two-way Source Coding model
  - Why Interested in It
  - Mathematical Description
  - Sum-rate-distortion Function
  - Proof(Achievability and Converse)
  - Problems Remain Open
- Interaction in Lossless Expected Length Case
- Interaction in Zero-error Worst Length Case
- Interaction in Function Computation

Why interested in it?

Recall Wyner-Ziv problem



$$R_X = \min_{P(U|X)|} [I(X;U|Y)]$$

U forms $U \leftrightarrow X \leftrightarrow Y$

$E[d(X,g(U,Y))] \leq D$

Question is, Can we save more rate?

Outline

- Problem Setup
  - Two-way Source Coding model
  - Why Interested in It
  - Mathematical Description
  - Sum-rate-distortion Function
  - Proof(Achievability and Converse)
  - Problems Remain Open
- Interaction in Lossless Expected Length Case
- Interaction in Zero-error Worst Length Case
- Interaction in Function Computation

Mathematical Description

A K-round scheme of two-way source coding

1-round: The X codec starts by sending R_X^1 bits

then the Y codec replies R_y^1 bits

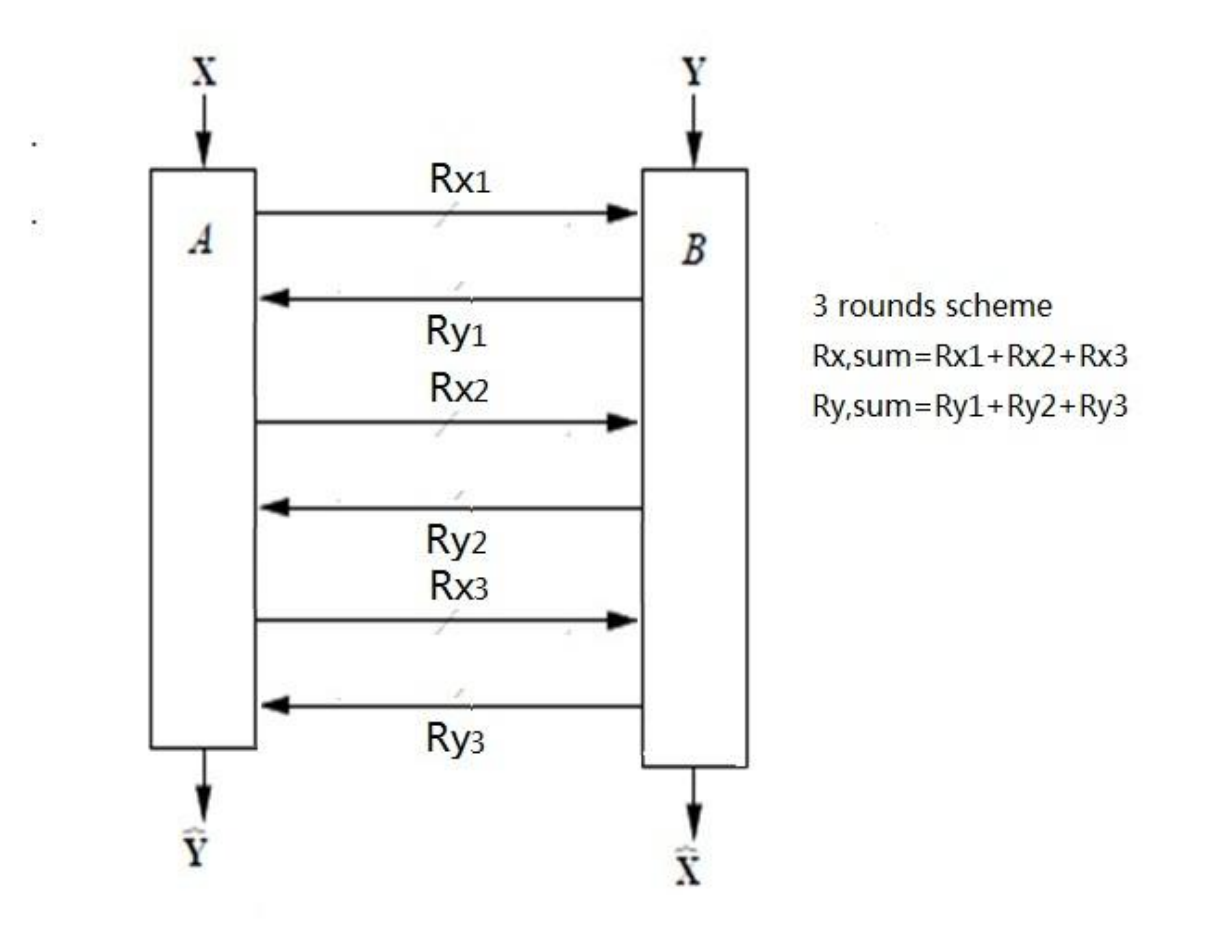
The process repeats K times

$$R_x = \sum_{k=1}^{K} R_x^k$$

$$R_y = \sum_{k=1}^{K} R_y^k$$

Mathematical Description

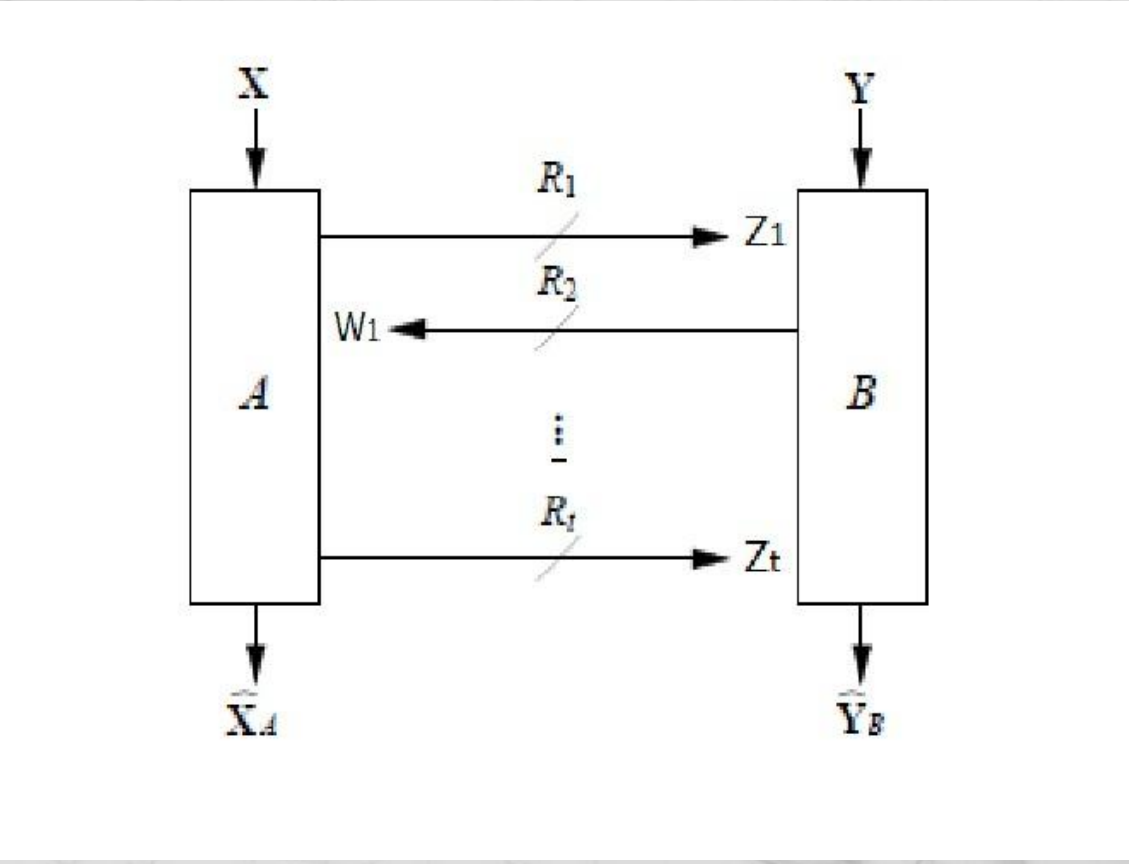
An example of 3-round scheme



Mathematical Description

Denote Z^k as the Kth-round forward message (X to Y)

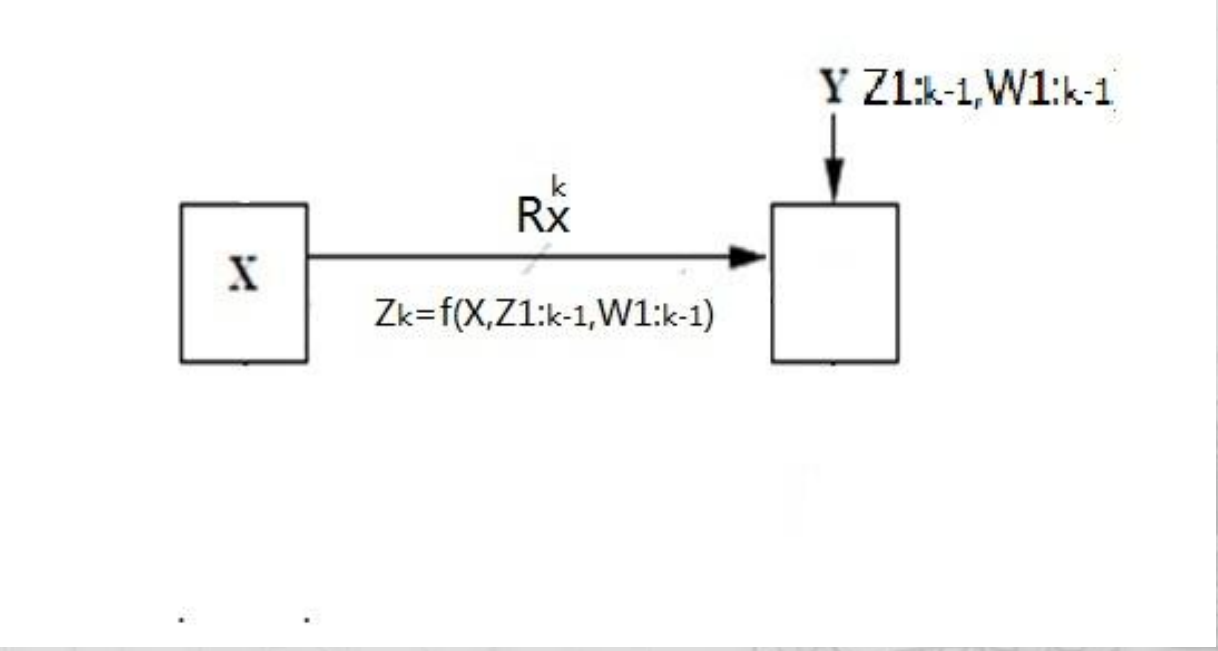
Denote W^k as the Kth-round backward message



Mathematical Description

The kth step of forward/backward passing

Both the Enc and Dec will consider all the former messages and the source

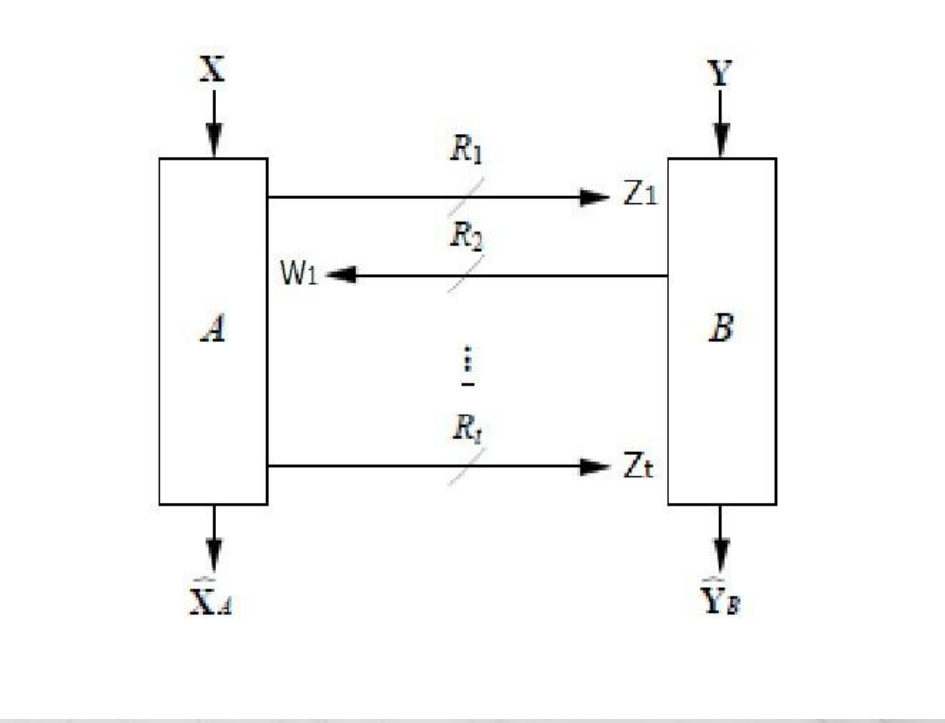


Mathematical Description

X,Y, $Z^{1:K}$ $W^{1:K}$ forms markov chains shown as follows:

$$Z^k \leftrightarrow X, Z^{1:k-1}, W^{1:k-1} \leftrightarrow Y$$

$$W^k \leftrightarrow Y, Z^{1:k}, W^{1:k-1} \leftrightarrow X$$



Outline

- Problem Setup
  - Two-way Source Coding model
  - Why Interested in It
  - Mathematical Description
  - Sum-rate-distortion Function
  - Proof(Achievability and Converse)
  - Problems Remain Open
- Interaction in Lossless Expected Length Case
- Interaction in Zero-error Worst Length Case
- Interaction in Function Computation

Sum-rate-distortion Function

$$R^{K}_{sum,x} = \min_{\substack{X,Y,Z^{1:k},W^{1:k} \text{ satisfy the} \\ \text{markov chain property} \\ E[n^{-1}\sum_{i=1}^{n} d_x(X_i,\hat{X}_i) \le D_X \\ E[n^{-1}\sum_{i=1}^{n} d_y(Y_i,\hat{Y}_i)] \le D_y}} I(X; Z^{1:K}, W^{1:K}|Y)$$

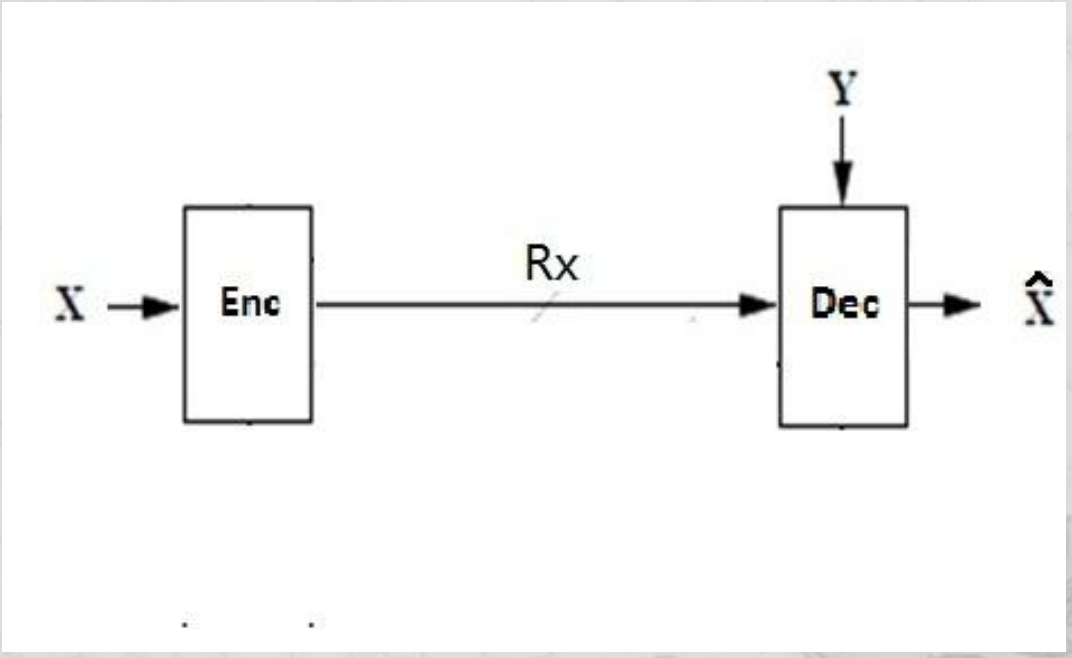
$$R^{K}_{sum,y} = \min_{\substack{X,Y,Z^{1:k},W^{1:k} \text{ satisfy the} \\ \text{markov chain property} \\ E[n^{-1}\sum_{i=1}^{n} d_x(X_i,\hat{X}_i) \le D_X \\ E[n^{-1}\sum_{i=1}^{n} d_y(Y_i,\hat{Y}_i)] \le D_y}} I(Y; Z^{1:K}, W^{1:K}|X)$$

Outline

- Problem Setup
  - Two-way Source Coding model
  - Why Interested in It
  - Mathematical Description
  - Sum-rate-distortion Function
  - Proof(Achievability and Converse)
  - Problems Remain Open
- Interaction in Lossless Expected Length Case
- Interaction in Zero-error Worst Length Case
- Interaction in Function Computation

Proof(Achievability)

Recall Achievability proof of Wyner-Ziv problem



$$R_X = \min_{\substack{P(U|X) \\ U \text{ forms } U \leftrightarrow X \leftrightarrow Y \\ E[d(X,g(U,Y))] \le D}} [I(X;U|Y)]$$

Proof(Achievability)

Recall Achievability proof of Wyner-Ziv problem

Jointly Strong Typicality

"Bin" method

Encoder:

find in $i \in \{1 \dots 2^{N\hat{R}}\}$ s. t. $(\vec{x}, \overrightarrow{u(i)}) \in T_\epsilon^N(X, U)$

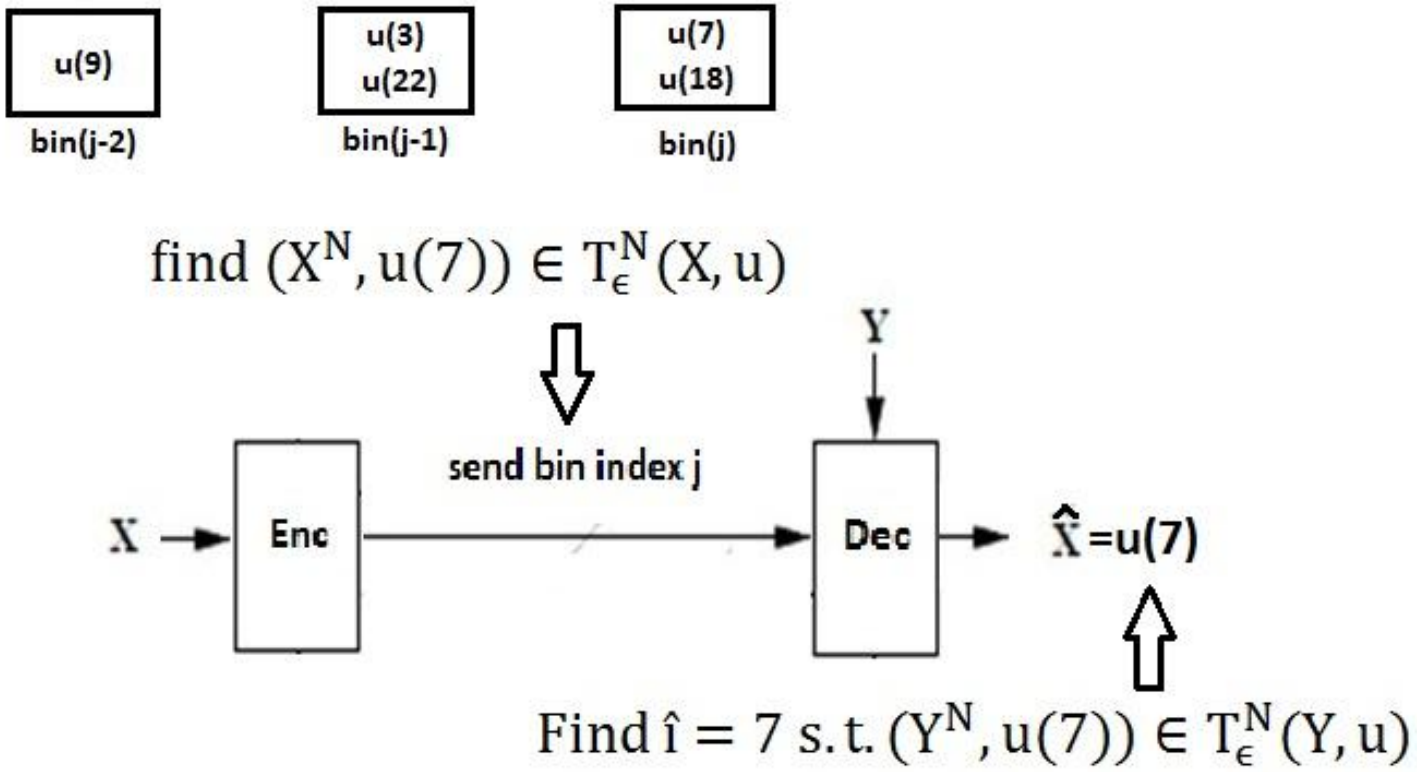
send the index j of the bin that i is in

Decoder:

find an $\hat{i} \in B(j)$ s. t. $(\vec{y}, \overrightarrow{u(i)}) \in T_\epsilon^N(Y, U)$

See figure in the next slide

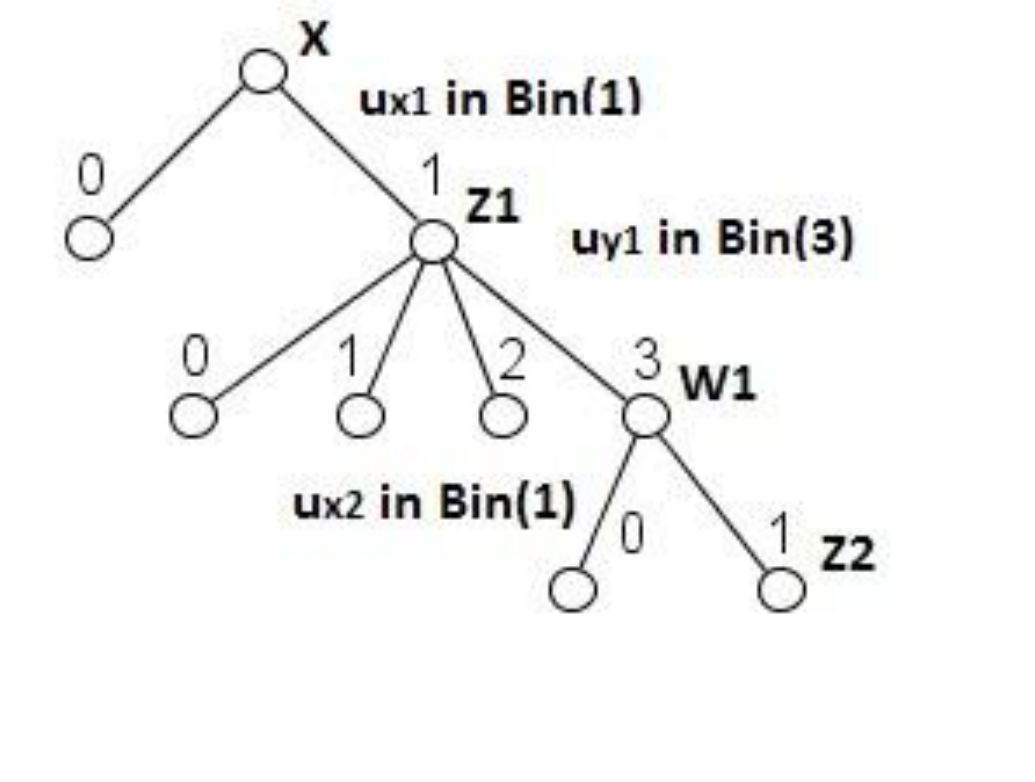
Proof(Achievability)



Proof(Achievability)

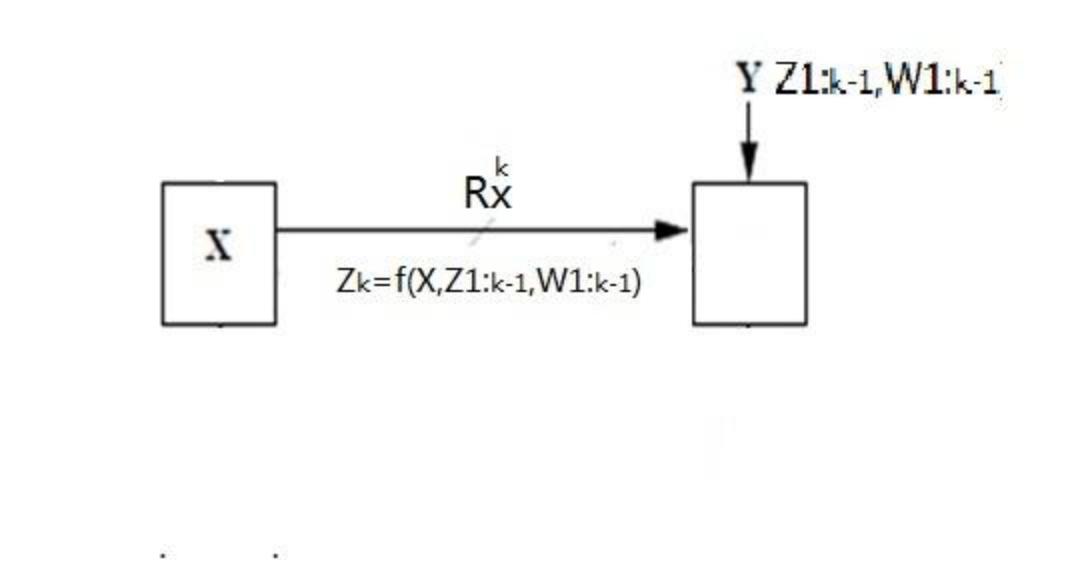
Similar to Wyner-Ziv's proof

A codebook tree instead of codebook



Proof(Achievability)

Consider one single step of message passing



$$R_x^k = I(X; Z^k | Y, Z^{1:k-1}, W^{1:k-1})$$

Proof(Achievability)

The random variables X Y Z W will satisfy the markov property

$$Z^k \leftrightarrow X, Z^{1:k-1}, W^{1:k-1} \leftrightarrow Y$$

$$W^k \leftrightarrow Y, Z^{1:k}, W^{1:k-1} \leftrightarrow X$$

Can show

$$\sum_{k=1}^{K} R_x^k = I(X; Z^{1:K}, W^{1:K} | Y)$$

Proof (Converse)

Recall converse proof of Wyner-Ziv problem

$$NR \geq \sum_{n=1}^{N} H(X_n|Y_n) - H(X_n|X^{n-1}, \underbrace{J, \bar{Y}_{N\setminus n}}_{U_n}, Y_n) \qquad J = f_N(X^N)$$

Group R.V.s
Define auxilary R.V.

$$= \sum_{n=1}^{N} H(X_n|Y_n) - H(X_n|X^{n-1}, U_n, Y_n)$$

$$\geq \sum_{n=1}^{N} H(X_n|Y_n) - H(X_n|U_n, Y_n)$$

Change to per-symbol rate

Proof (Converse)

Recall converse proof of Wyner-Ziv problem

One can prove

$$NR \geq \sum_{n=1}^{N} R_D(D_n)$$

By the convexity of mutual information

$$R \geq \frac{1}{N}\sum_{n=1}^{N} R_D(D_n) \geq R_D\left(\frac{1}{N}\sum_{n=1}^{N} D_n\right) \geq R_D(D)$$

Proof (Converse)

Given an achievable point $s=(r_x,r_y,d_x,d_y)$, prove that

$$r_x + \varepsilon \geq I(X; Z^{1:K}, W^{1:K}|Y)$$

$$r_{y+\varepsilon} \geq I(Y; Z^{1:K}, W^{1:K}|X)$$

$$d_x + \epsilon \geq D_x$$

$$d_y + \epsilon \geq D_y$$

Proof (Converse)

There exists a system

$$(n, K, \{f^k\}_{k=1}^K, \{g^k\}_{k=1}^K, F, G, D_x, D_y)$$

Specified by the encoding functions

$$f^1(X), g^1(Y, f^1), f^2(X, g^1), \ldots g^k(Y, f^1, f^2, \ldots f^k)$$

And decoding functions F,G satisfy

$$r_x + \varepsilon \geq R_x = n^{-1} \sum_{k=1}^{K} \log I_x^K \geq n^{-1} \sum_{k=1}^{K} H(f^k)$$

Proof (Converse)

Can show:

$$r_x + \varepsilon \geq n^{-1} H(X|Y) - n^{-1} H(X|Y, f^{1:K})$$

Denote X(-) as $X_{1:i-1}$ Y(+) as $Y_{i+1:n}$

Then can show:

$$r_x + \varepsilon \geq n^{-1} \sum_{i=1}^{n} I(X_i; f^{1:K}, g^{1:K}, Y(+), X(-)|y_i)$$

Proof (Converse)

Define auxiliary random variables

$$Z_i^1 = \left(f^1, X(-), Y(+)\right)$$
$$Z_i^k = f^k, k = 2{:}K$$
$$W_i^k = g^k, k = 1{:}K$$

We have

$$r_x + \varepsilon \geq n^{-1} \sum_{i=1}^{n} I(X_i; Z_i^{1:K}, W_i^{1:K} | y_i)$$

Proof (Converse)

Prove

(1) $$n^{-1}\sum_{i=1}^{n} I(X_i; Z_i^{1:K}, W_i^{1:K}|y_i) = I(X; Z_i^{1:K}, W_i^{1:K}|Y)$$

(2) $$f^k \leftrightarrow (X_i, f_i^{1:k-1}, g_i^{1:k-1}, X(-), Y(+)) \leftrightarrow Y_i$$

Outline

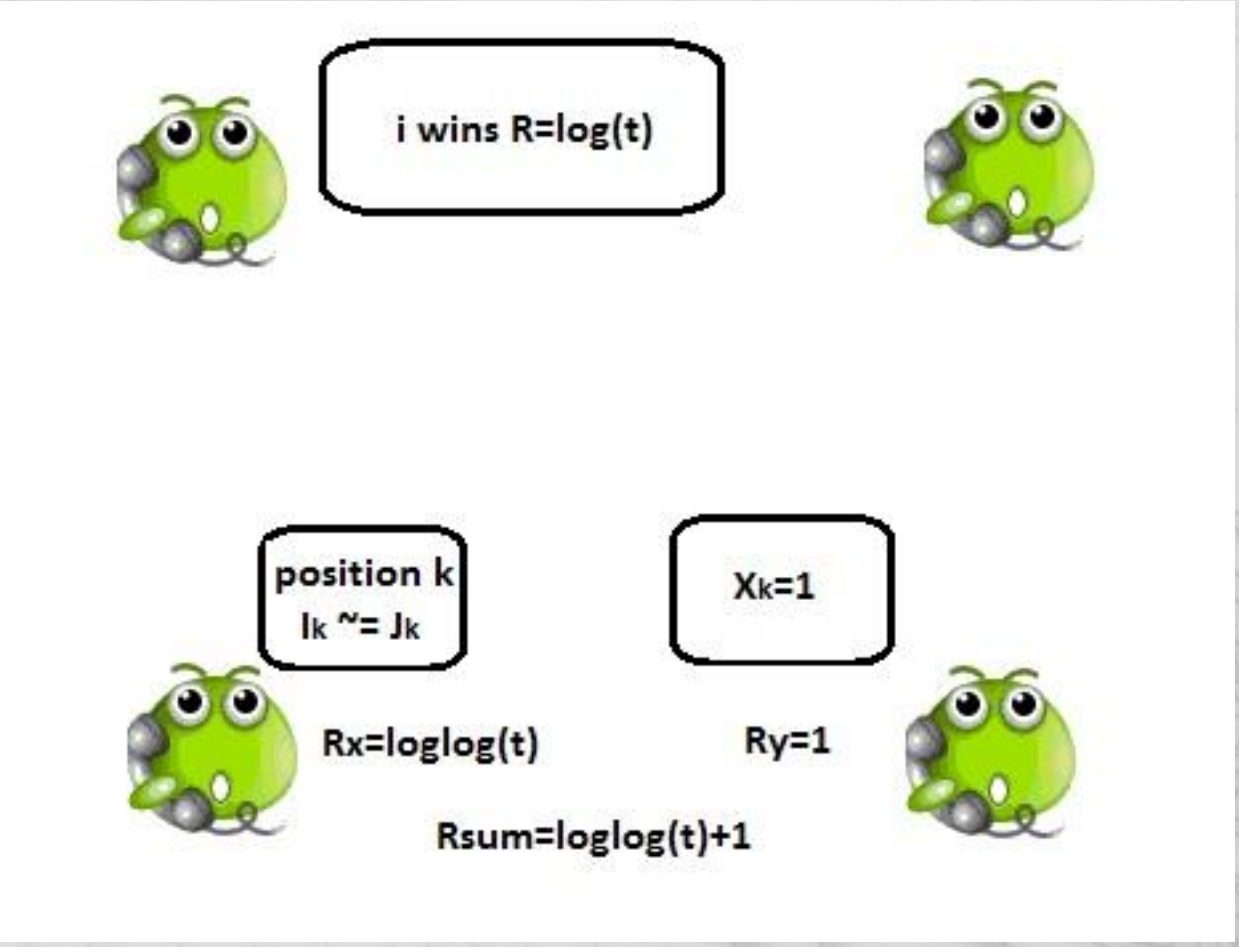
- Problem Setup
  - Two-way Source Coding model
  - Why Interested in It
  - Mathematical Description
  - Sum-rate-distortion Function
  - Proof(Achievability and Converse)
  - Problems Remain Open
- Interaction in Lossless Expected Length Case
- Interaction in Zero-error Worst Length Case
- Interaction in Function Computation

Problems Remain Open

1. Does interaction strictly improves rate-distortion function?
2. Does an unbounded K helps?
3. The existence of an optimal K* s.t. $K^* < \infty$
4. Zero-error worst-length case
5. Probability of block error for lossless reproduction
6. How many bits we can save?
7. Interaction in function computation case

Problems Remain Open

An example of interaction in zero-error case



Problems Remain Open

An example of interaction in function computation

X~Uniform{1…L} Y~Ber(p)

$f_a(x,y):=0$ $f_b(x,y):=xy$

The benefit of interaction can be arbitrarily large

$$R_{sum,1} = H(X|Y) = H(X)$$

$$R_1 = H(Y) = h(p)$$

$$R_2 = H(f_b(X,Y)|Y) = p\log_2 L$$

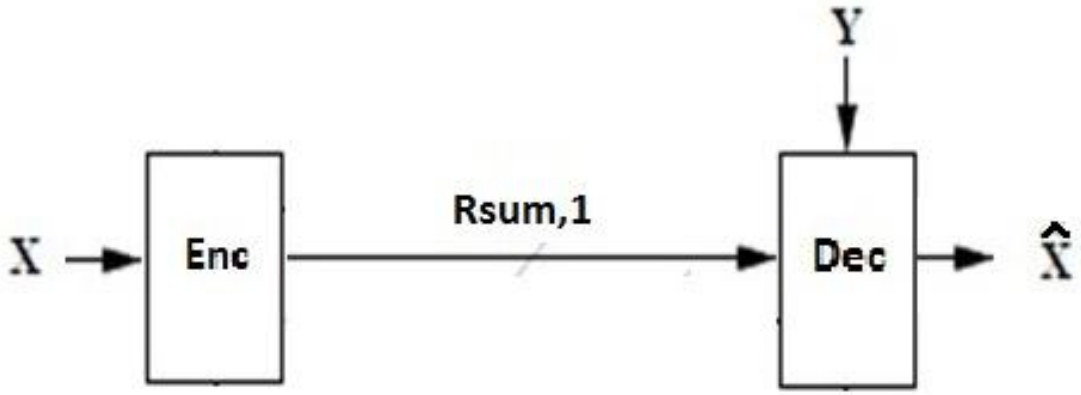
Conclusion

- Idea of Interaction
- K-round Scheme of Two-Way Source Coding
- Sum-Rate-Distortion Function
- Achievability and Converse Proof
- Some questions remain open

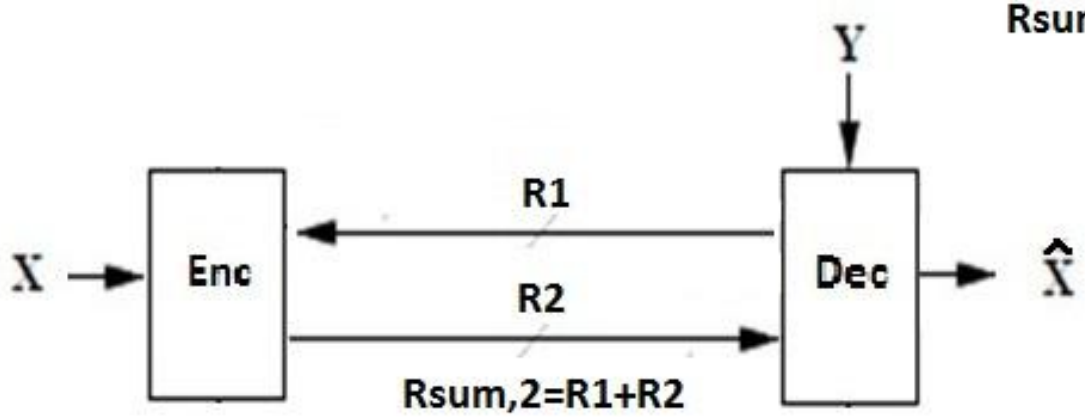
Problems Remain Open

1. Does interaction strictly improves rate-distortion function?
2. Does an unbounded K helps?
3. The existence of an optimal K* s.t. $K^* < \infty$
4. Zero-error worst-length case
5. Probability of block error for lossless reproduction
6. How many bits we can save?
7. Interaction in function computation case

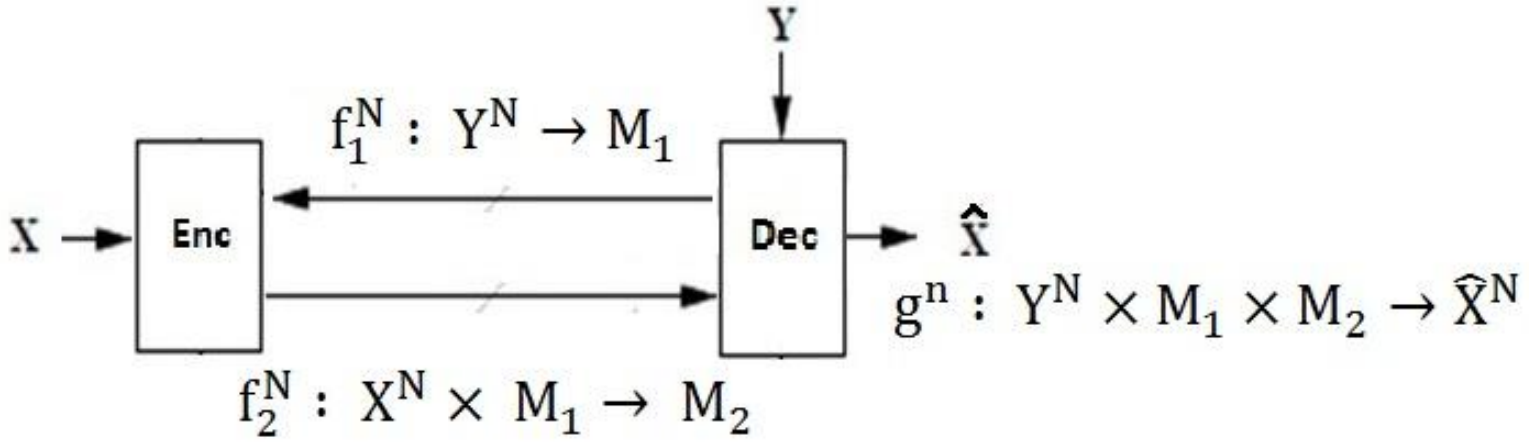
Interaction Improves Rate-Distortion Function



Rsum2 Strictly less than Rsum1 ?

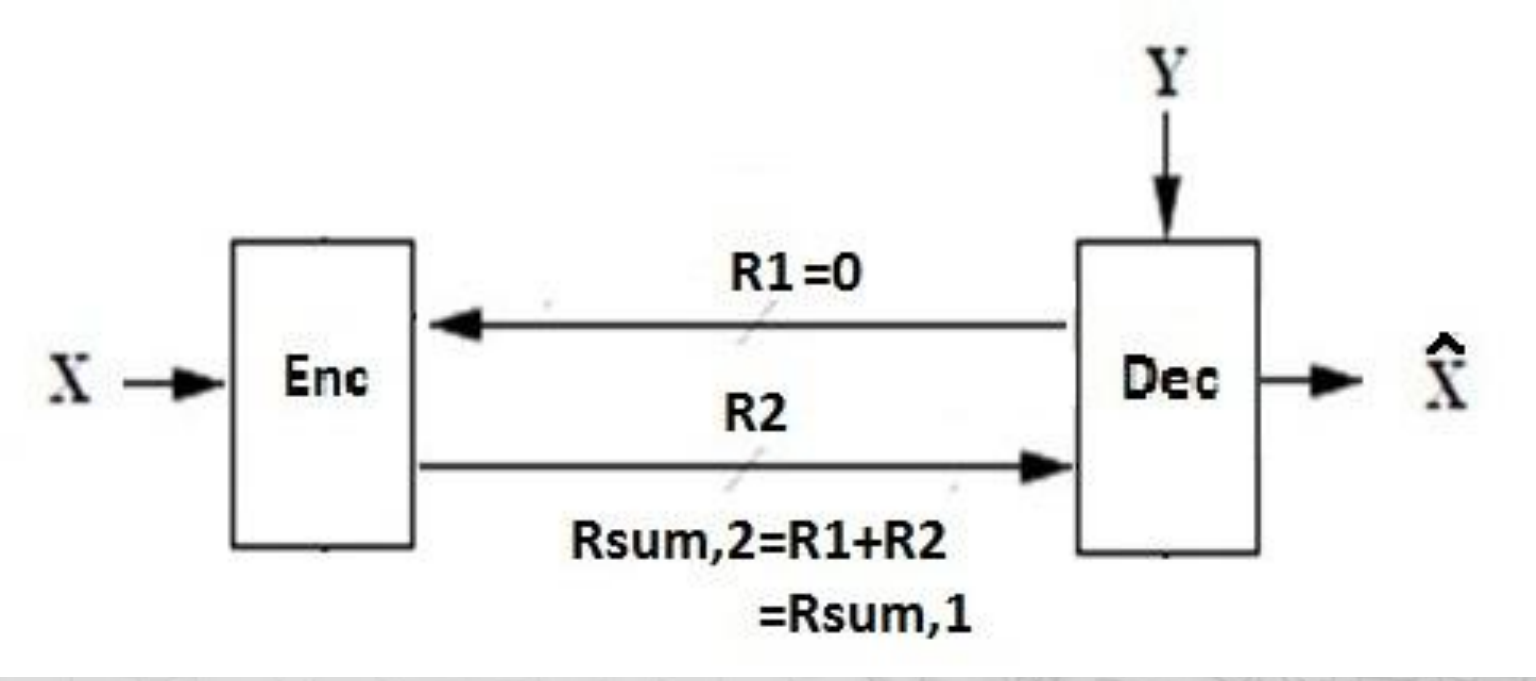


Interaction Improves Rate-Distortion Function



Interaction Improves Rate-Distortion Function

$$R_{sum,2}(D) = \min_{E[d(X,g(V_1,V_2,Y))\leq D]} \{I(Y;V_1|X) + I(X;V_2|Y,V_1)\}$$



Interaction Improves Rate-Distortion Function

We have

$$R_{sum,1}(D) \geq R_{sum,2}(D)$$

Question:

Is the inequality strict?

Interaction Improves Rate-Distortion Function

Key tool : rate reduction functionals

Definition:

$$\rho_i(p_{XY}, D) := H(X|Y) + H(Y|X) - R_{sum,i}(p_{XY}, D)$$

Interaction Improves Rate-Distortion Function

$$R_{sum,1}(D) > R_{sum,2}(D) \leftrightarrow \rho_1(p_{XY}, D) < \rho_2(p_{XY,}D)$$

Lemma 1:

The following two conditions are equivalent

(1) For all p_{XY} and $D, \rho_1(p_{XY}, D) = \rho_2(p_{XY,}D)$

(2) For all $p_{X|Y}$, $\rho_1(p_{X|Y}p_Y, D)$ is concave w.r.t. (p_Y, D)

Interaction Improves Rate-Distortion Function

Lemma 2: Let f(p) be a function differentiable around p=0 such that f(0)=0 and f'(0)>0. Then

$$\lim_{p \to 0} \frac{h(f(p))}{h(p)} = f'(0)$$

Can be proved by the l'Hopital rule

Interaction Improves Rate-Distortion Function

Theorem 1: There exists a distortion function d, a joint distribution p_{xy}, and a distortion level D for which

$$R_{sum,1}(D) > R_{sum,2}(D)$$

Lemma 1,2 will be used in the proof of Theorem1

Interaction Improves Rate-Distortion Function

Let $X = Y = \{0,1\}$

$\hat{X} = \{0,1,e\}$

Let d be the binary erasure distortion function

$$d: \{0,1\} \times \{0,1\} \rightarrow \{0,1,e\}$$

	0	**1**	**e**
0	0	INF	1
1	INF	0	1

Interaction Improves Rate-Distortion Function

Let (X,Y) ~ DSBS(p)

$$p_{XY}(x,y) = 0.5(1-p)\delta_{xy} + 0.5p(1-\delta_{xy})$$

Where a Kronecker function is used

$$\delta_{xy} = \begin{cases} 0 & x \neq y \\ 1 & x = y \end{cases}$$

Marginal distribution

X~Ber(1/2) Y~Ber(1/2)

00	**0.5(1-p)**
11	**0.5(1-p)**
01	**0.5p**
10	**0.5p**

Interaction Improves Rate-Distortion Function

By Lemma 1, it is sufficient to prove there exist $p_{Y,1}$ and $p_{Y,2}$ such that

$$\rho_1\left(p_{X|Y}\frac{p_{Y,1}+p_{Y,2}}{2},D\right)<\frac{\rho_1\left(p_{X|Y}p_{Y,1},D\right)+\rho_1\left(p_{X|Y}p_{Y,2},D\right)}{2}$$

This can be proved by the following 5 propositions

Interaction Improves Rate-Distortion Function

$$\rho_1(p_{XY}, D) = \max_{p_{U|X}, g: E[d(X,g(U,Y))] \le D} H(X|Y,U) + H(Y|X)$$

Proposition 1

$$p_{U|X}(u|x) = \begin{cases} a_{0e} & \text{if } x = 0, u = e, \\ 1 - a_{0e} & \text{if } x = 0, u = 0, \\ a_{1e} & \text{if } x = 1, u = e, \\ 1 - a_{1e} & \text{if } x = 1, u = 1, \\ 0 & \text{otherwise,} \end{cases}$$

	0	**1**	**e**
0	0	INF	1
1	INF	0	1

$$E[d(X,U)] = p_X(0)a_{0e} + p_X(1)a_{1e} \le D$$

Interaction Improves Rate-Distortion Function

Proposition 2

$$\rho_1(p_{XY}, D) = \max_{\substack{\alpha_{0e},\alpha_{1e}\in[0,1]:\\ \Phi(p_{XY},\alpha_{0e},\alpha_{1e})\leq D}} \Psi(p_{XY}, \alpha_{0e}, \alpha_{1e})$$

Where,

$$\Psi(p_{XY}, \alpha_{0e}, \alpha_{1e}) := (p_{00}\alpha_{0e} + p_{10}\alpha_{1e})h\left(\frac{p_{00}\alpha_{0e}}{p_{00}\alpha_{0e} + p_{10}\alpha_{1e}}\right)$$
$$+(p_{01}\alpha_{0e} + p_{11}\alpha_{1e})h\left(\frac{p_{01}\alpha_{0e}}{p_{01}\alpha_{0e} + p_{11}\alpha_{1e}}\right)$$
$$+(p_{00} + p_{01})h\left(\frac{p_{00}}{p_{00} + p_{01}}\right) + (p_{11} + p_{10})h\left(\frac{p_{11}}{p_{11} + p_{10}}\right)$$

$$\Phi(p_{XY}, \alpha_{0e}, \alpha_{1e}) := p_X(0)\alpha_{0e} + p_X(1)\alpha_{1e}$$

Interaction Improves Rate-Distortion Function

Proposition 3 The rate reduction funtionals can be reduced to the compact expression for binary erasure distortion and DSBS source

$$\rho_1(p_{XY}, D) = (1 + D)h(p)$$

Interaction Improves Rate-Distortion Function

Proposition 4

$$\frac{\rho_1\left(p_{X|Y}p_{Y,1}, D\right) + \rho_1\left(p_{X|Y}p_{Y,2}, D\right)}{2} \geq C(p, q, \alpha_{0e}, 1)$$

Holds for

$$D = \eta\,(p, q, \alpha_{0e}, 1)$$

Where $C(p, q, \alpha_{0e}, \alpha_{1e}) := \Psi\left(p_{X|Y}p_{Y,1}, \alpha_{0e}, \alpha_{1e}\right)$

$$\eta\,(p, q, \alpha_{0e}, \alpha_{1e}) := \Phi\left(p_{X|Y}p_{Y,1}, \alpha_{0e}, \alpha_{1e}\right)$$

Interaction Improves Rate-Distortion Function

Proposition 5 For all $q \in (0, 1/2)$ and all $\alpha_{0e} \in (0, 1)$, there exists $p \in (0, 1)$ such that the strict inequality

$$\rho_1(p_{XY}, D) < C(p, q, \alpha_{0e}, 1)$$

holds for

$$D = \eta(p, q, \alpha_{0e}, 1)$$

Then theorem 1 has been proved.

Interaction Improves Rate-Distortion Function

Theorem 2: If d is the binary erasure distortion and p_{XY} the joint pmf of a DSBS with parameter p, then for all L>0 there exists an admissible two-message rate-distortion tuple (R_1, R_2, D) such that

$$\frac{R_{sum,1}(p_{XY}, D)}{R_1 + R_2} > L \; and \; \frac{R_1}{R_2} < \frac{1}{L}$$

In which case Interaction improves the rate?

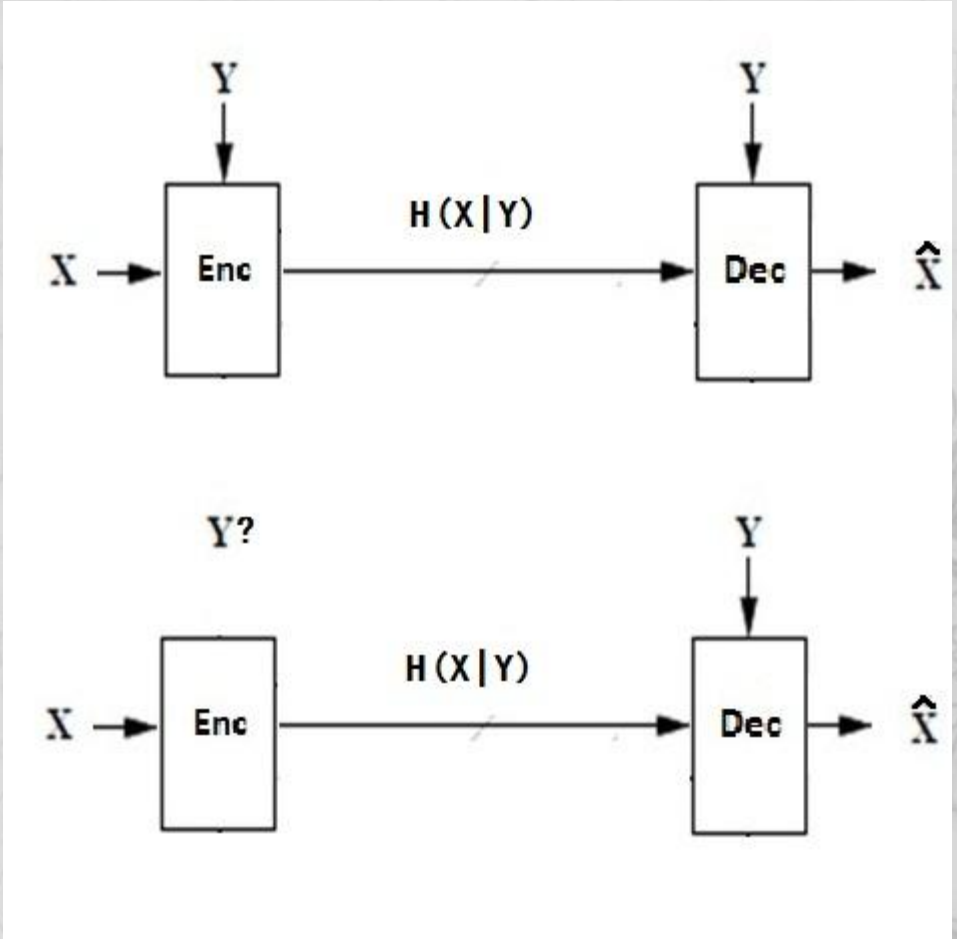
	LOSSY	LOSSLESS	ZERO-ERROR
Source Reconstruction	**Yes**	**No**	**Yes**
Function Computation	**Yes**	**Yes**	**Yes**

Outline

- Problem Setup (Two-Way Source Coding)
- Interaction in Lossless Expected Length Case
- Interaction in Zero-error Worst Length Case
- Interaction in Function Computation

Lossless Expected Length Case

- Known Y at the encoder does not improve the rate



Outline

- Problem Setup (Two-Way Source Coding)
- Interaction in Lossless Expected Length Case
- Interaction in Zero-error Worst Length Case
- Interaction in Function Computation

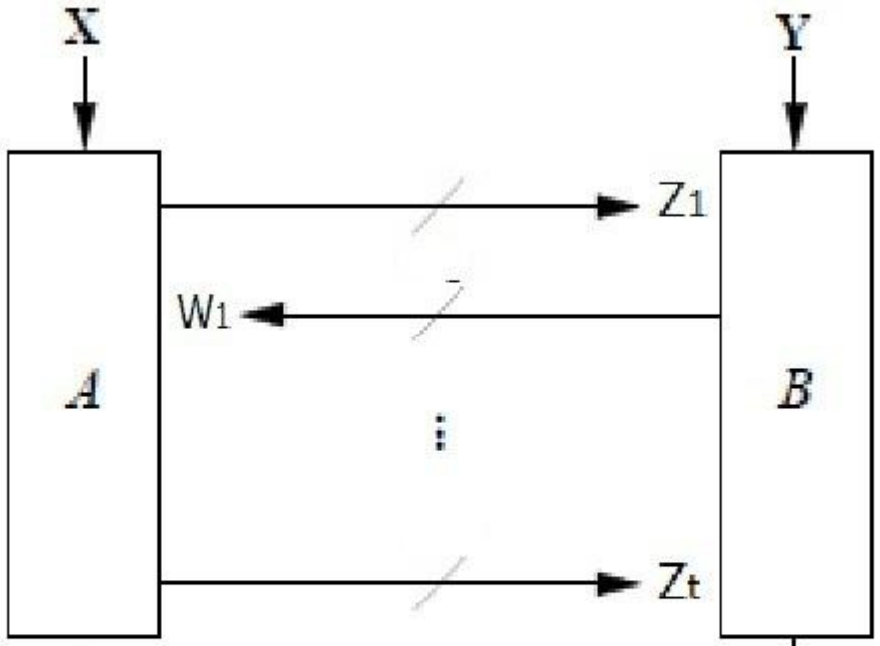
Outline

- Interaction in Zero-error Worst-Length Case
  - Problem Setup
  - Definitions and Properties
  - Results in Zero-error Worst-Length Case
  - Proof of the Results
  - Conclusion

Outline

- Interaction in Zero-error Worst-Length Case
  - Problem Setup
  - Definitions and Properties
  - Results in Zero-error Worst-Length Case
  - Proof of the Results
  - Conclusion

Problem Setup



$$\min_{\varphi} \max_{(x,y)} \sum_{i=1}^{m_\varphi} |\sigma_i(x,y)| \qquad \hat{x} \;\; P(x \neq \hat{x}) = 0$$

Outline

- Interaction in Zero-error Worst-Length Case
  - Problem Setup
  - Definitions and Properties
  - Results in Zero-error Worst-Length Case
  - Proof of the Results
  - Conclusion

Definitions and Properties

Support set of (X,Y)

$$S_{X,Y} \triangleq \{(x, y): p(x, y) > 0\}$$

Transmission length of the input (x,y)

$$l_{\varphi}(x, y) \triangleq \sum_{i=1}^{m_{\varphi}} |\sigma_i(x, y)|$$

Definitions and Properties

Worst-case complexity of an protocol

$$\hat{L}_{\varphi} \triangleq \sup \{l_{\varphi}(x, y): (x, y) \in S_{X,Y}\}$$

M-message complexity of (X,Y)

$$\hat{C}_{m}(X|Y) \triangleq \inf\{\hat{L}_{\varphi}: \varphi \text{ is an M protocol for} (X, Y)\}$$

Definitions and Properties

$C_m(X|Y)$ is a decreasing function of m since empty message works

$$C_1=\{C_1,0,0\}$$

Can define $C_\infty(X|Y)$

Also

$$\hat{C}_1(X|Y) \geq \hat{C}_2(X|Y) \geq \cdots \geq \hat{C}_m(X|Y)$$

Definitions and Properties

Define

$$S_X \triangleq \{x: \text{for some } y, (x,y) \in S_{X,Y}\}$$

Y's ambiguity set

$$S_{X|Y}(y) \triangleq \{x: (x,y) \in S_{X,Y}\}$$

Definition and Properties

Ambiguity

$$\hat{\mu}_{X|Y}(y) \triangleq \left\lceil S_{X|Y}(y) \right\rceil$$

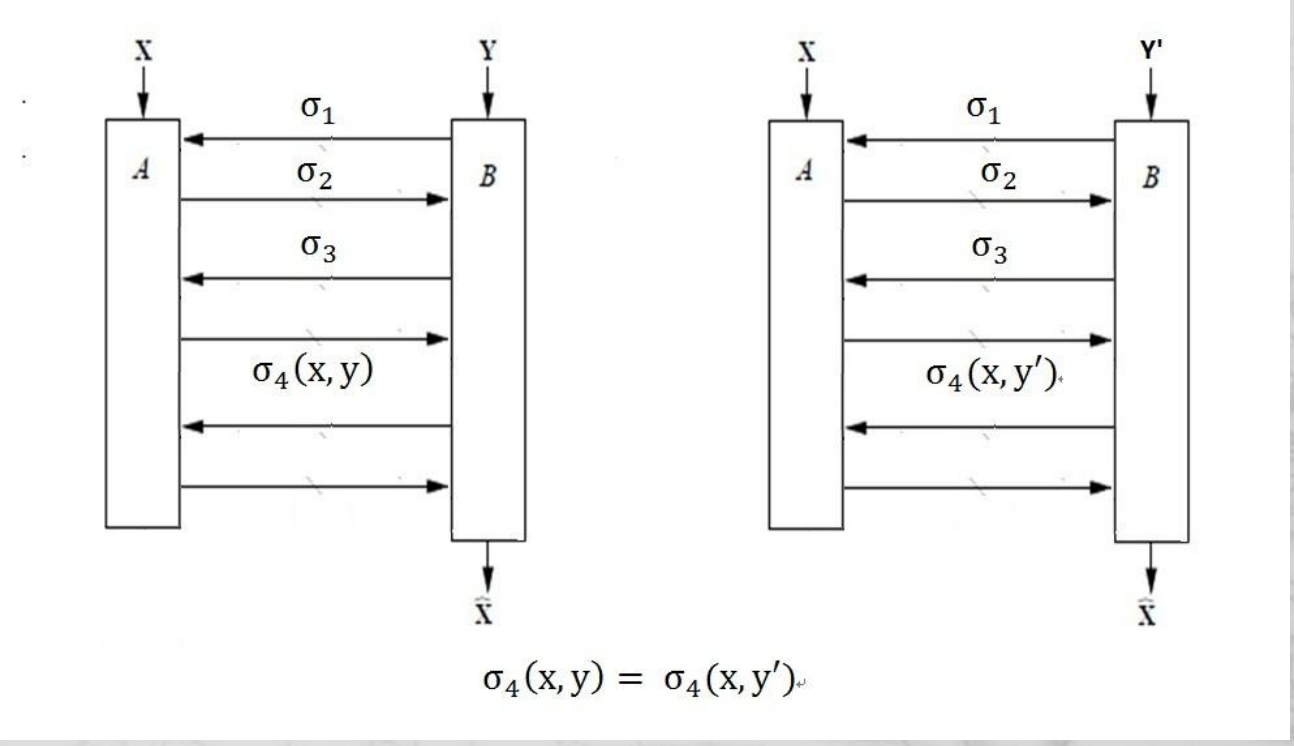
Maximum ambiguity

$$\hat{\mu}_{X|Y} \triangleq \sup \{\hat{\mu}_{X|Y}(y) : y \in S_Y\}$$

Definitions and Properties

Separate-transmissions property

$$< \sigma_j(x,y) >_{j=1}^{i-1} = < \sigma_j(x,y') >_{j=1}^{i-1} \rightarrow \sigma_i(x,y) = \sigma_i(x,y')$$

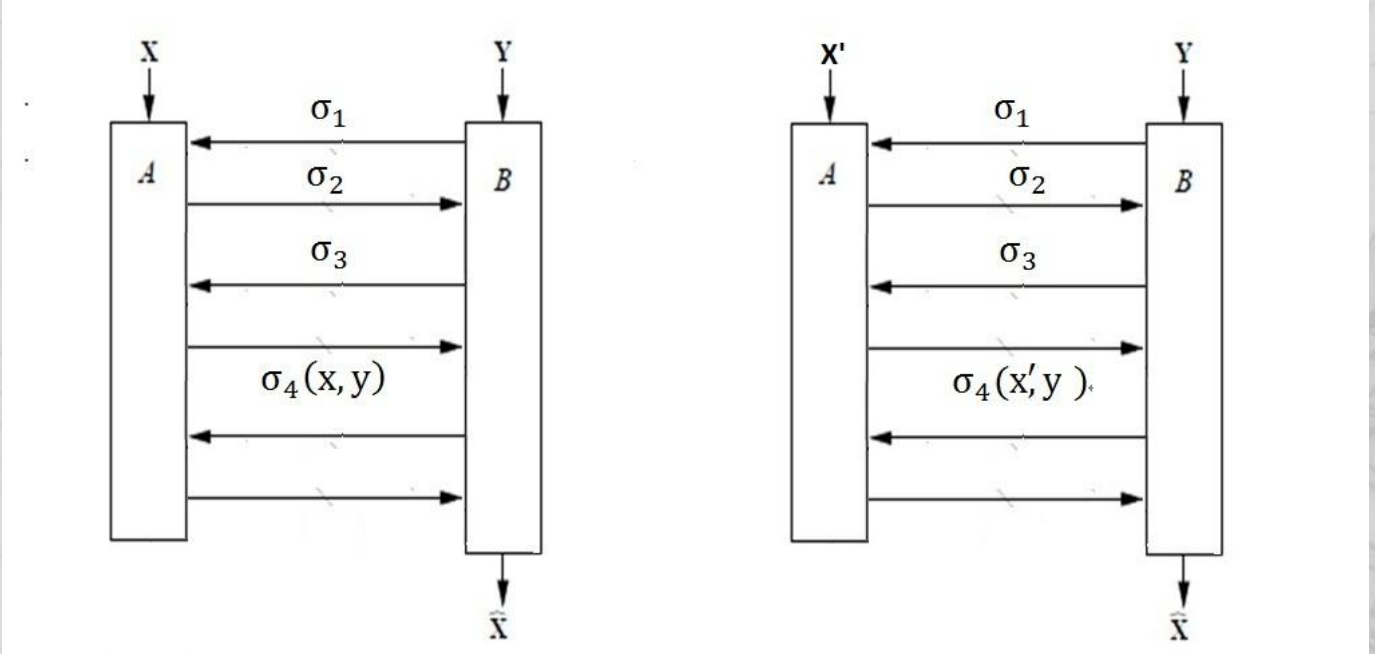


Definitions and Properties

Implicit-termination property

$$< \sigma_j(x,y) >_{j=1}^{i-1} = < \sigma_j(x',y) >_{j=1}^{i-1}$$

$$\rightarrow \sigma_i(x,y) \text{ and } \sigma_i(x',y) \text{ are not prefix of each other}$$

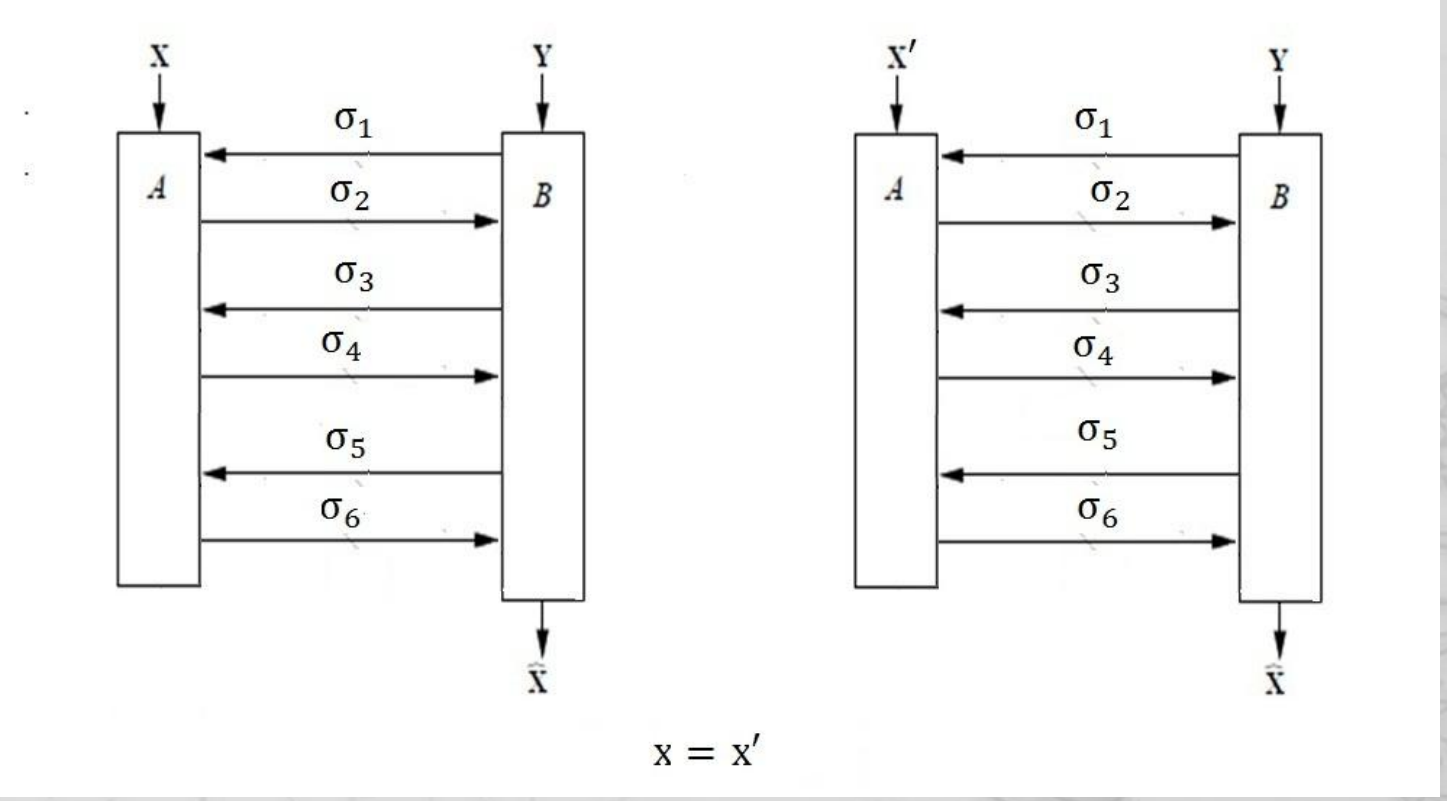


$\sigma_4(x,y)$ $\sigma_4(x',y)$can not be a prefix of each other.

Definitions and Properties

Correct-decision property

$$< \sigma_j(x,y) >_{j=1}^{\infty} = < \sigma_j(x',y) >_{j=1}^{\infty} \rightarrow x = x'$$



Hypergraph G(V,E)

Ordered pair (V,E)

Adjacent

Coloring of the hypergraph

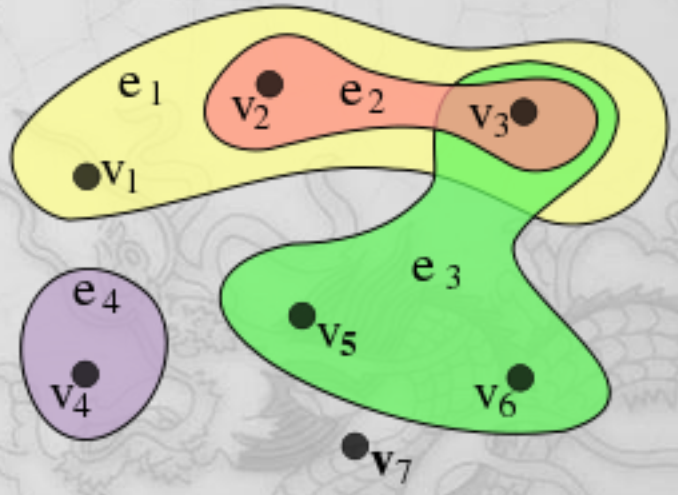
$C(V_1) \neq C(V_2)$

if V_1 and V_2 are adjacent

K-colorable

Chromatic number

$\omega\big(G(X|Y)\big)$

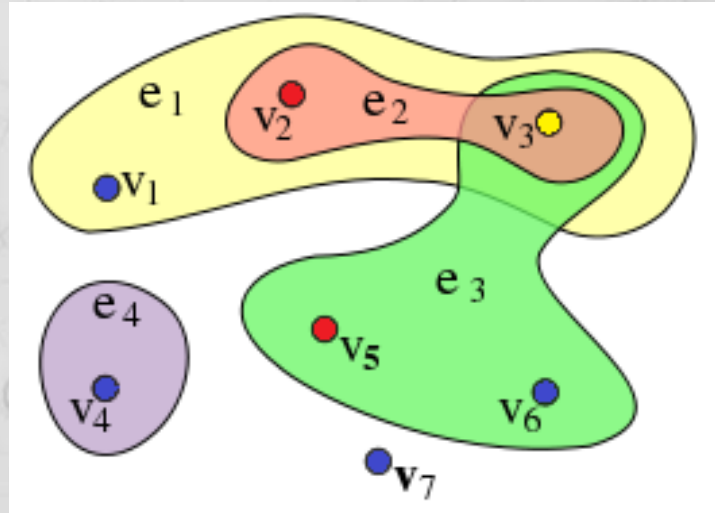


Hypergraph G(V,E)

K-colorable (K=3,4,5…)

Chromatic number

$$\omega\big(G(X|Y)\big) = 3$$



Outline

- Interaction in Zero-error Worst-Length Case
  - Problem Setup
  - Definitions and Properties
  - Results in Zero-error Worst-Length Case
  - Proof of the Results
  - Conclusion

Results

One-Way Complexity

The Limits of Interaction

Two Messages are Optimal

Two Messages are Almost Optimal

Two Messages are Not Optimal

Results

One-Way Complexity

The Limits of Interaction

Two Messages are Optimal

Two Messages are Almost Optimal

Two Messages are Not Optimal

Results

One-Way Complexity

$$\hat{C}_1(X|Y) = \lceil \log\omega(G(X|Y)) \rceil$$

The one-way complexity is the chromatic number of the characteristic hypergraph of (X,Y)

Results

One-Way Complexity

The Limits of Interaction

Two Messages are Optimal

Two Messages are Almost Optimal

Two Messages are Not Optimal

Results

The Limits of Interaction

$$\hat{C}_1(X|Y) \geq \hat{C}_2(X|Y) \geq \cdots \geq \hat{C}_m(X|Y)$$

$$\hat{C}_\infty(X|Y) \geq \lceil \log \hat{C}_1(X|Y) \rceil + 1$$

The minimum number of bits we need to reconstruct X with zero-error

Results

One-Way Complexity

The Limits of Interaction

Two Messages are Optimal

Two Messages are Almost Optimal

Two Messages are Not Optimal

Results

Two Messages are Optimal

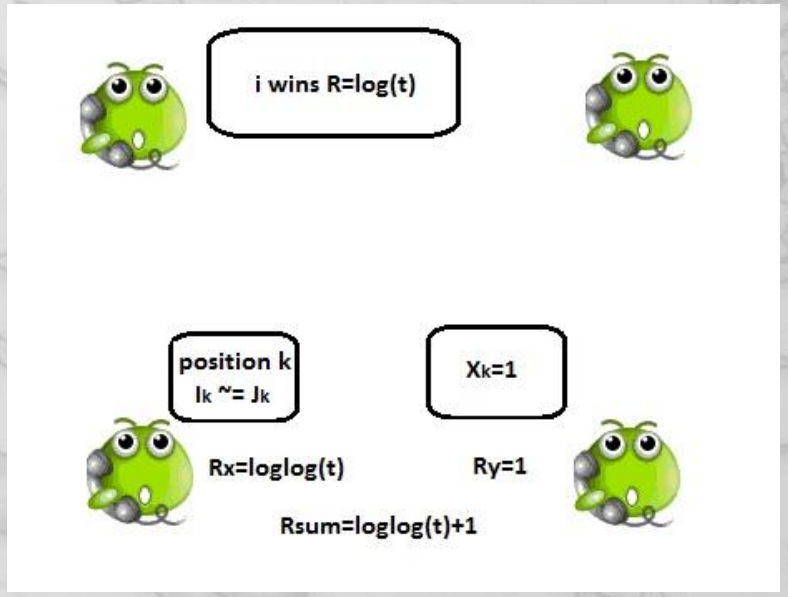
$$\hat{C}_2(X|Y) = \cdots = \hat{C}_m(X|Y) = \left\lceil \log \hat{C}_1(X|Y) \right\rceil + 1$$

In some cases, two message are enough to achieve the bound.

See example 1 english league

$$\hat{C}_2(X|Y) = \left\lceil \log \hat{C}_1(X|Y) \right\rceil + 1$$

$$= \log\log(t) + 1$$



Results

One-Way Complexity

The Limits of Interaction

Two Messages are Optimal

Two Messages are Almost Optimal

Two Messages are Not Optimal

Results

Two Messages are Almost Optimal

In general case, we can prove

$$\hat{C}_2(X|Y) \leq 4\hat{C}_\infty(X|Y) + 3$$

Two messages: log-reduction

More than two messages: linear-reduction

Results

One-Way Complexity

The Limits of Interaction

Two Messages are Optimal

Two Messages are Almost Optimal

Two Messages are Not Optimal

Results

Two Messages are Not Optimal

$$\hat{C}_2(X|Y) \geq (2-\epsilon)\hat{C}_3(X|Y) \geq (2-\epsilon)\hat{C}_\infty(X|Y)$$

In some cases, two messages are not optimal.

See example 2 Playoffs

Outline

- Interaction in Zero-error Worst-Length Case
  - Problem Setup
  - Definitions and Properties
  - Results in Zero-error Worst-Length Case
  - Proof of the Results
  - Conclusion

Proof of the Results

One-Way Complexity

The Limits of Interaction

Two Messages are Optimal

Two Messages are Almost Optimal

Two Messages are Not Optimal

Proof of the Results

One-Way Complexity

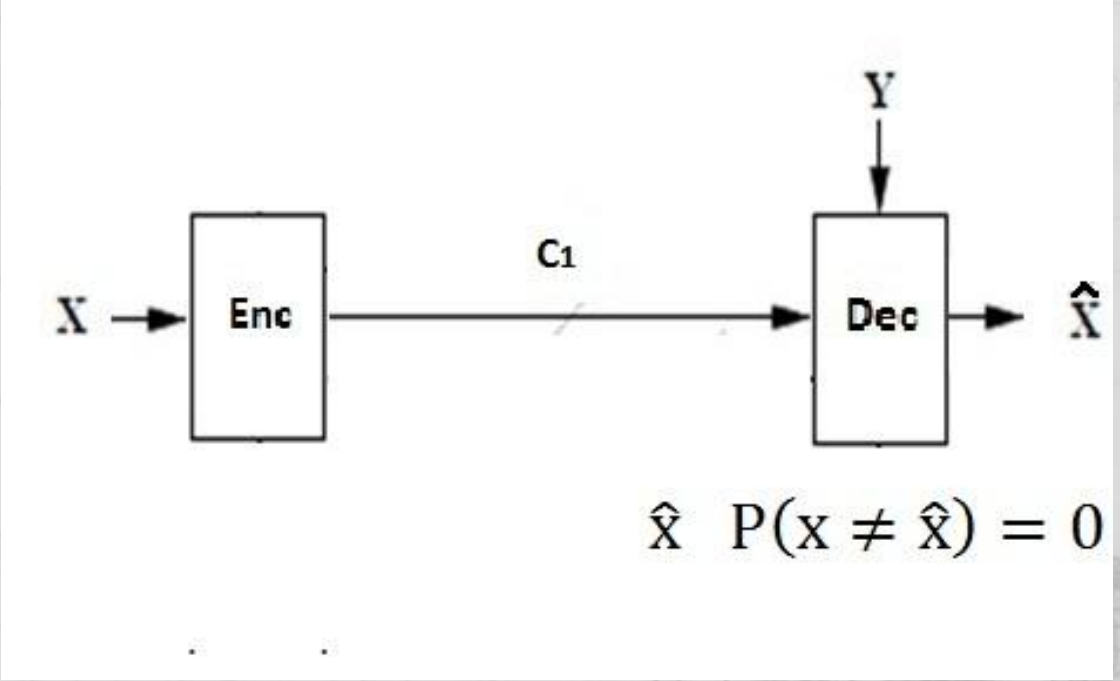
The Limits of Interaction

Two Messages are Optimal

Two Messages are Almost Optimal

Two Messages are Not Optimal

One-Way Complexity



One-Way Complexity

Define ω(G(X|Y)) as the chromatic number of G

Then,

$$\hat{C}_1(X|Y) = \lceil \log \omega(G(X|Y)) \rceil$$

Proof of the Results

One-Way Complexity

The Limits of Interaction

Two Messages are Optimal

Two Messages are Almost Optimal

Two Messages are Not Optimal

The Limits of Interaction

For all nontrivial (X,Y) pairs

$$\hat{C}_{\infty}(X|Y) \geq \lceil \log \hat{C}_1(X|Y) \rceil + 1$$

Here we prove an one-bit weaker result first

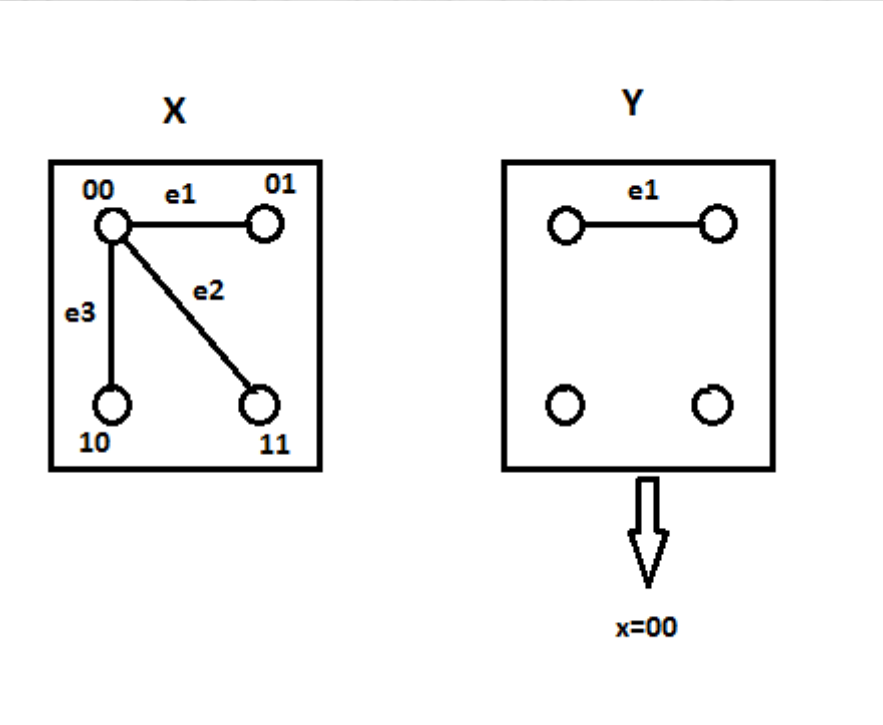
$$\hat{C}_{\infty}(X|Y) \geq \lceil \log \hat{C}_1(X|Y) \rceil$$

The Limits of Interaction

High-level Idea of the proof

X sends a sub-graph with edges contain vertex x

Y decodes x based on the edge Y=y



Proof of the Results

One-Way Complexity

The Limits of Interaction

Two Messages are Optimal

Two Messages are Almost Optimal

Two Messages are Not Optimal

Two Messages are Optimal

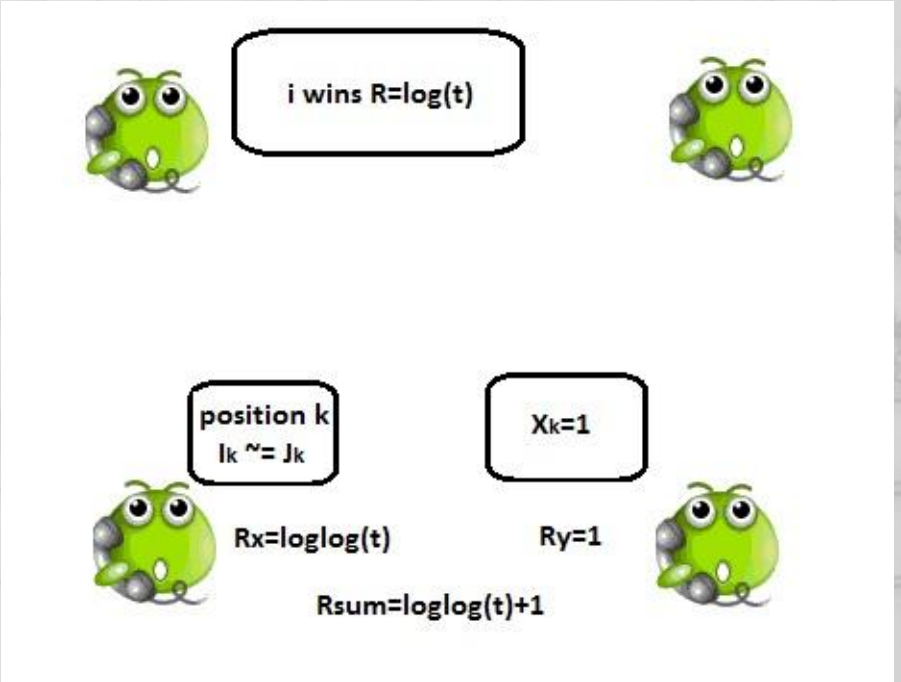
Two messages are optimal when

Hypergragh degenerates to graph (example 1)

For given Y=y

(team i vs team j)

$$\hat{\mu}_{X|Y} = 2$$



Example 1 English League

t clubs in english league(t=16), two random teams play versus each other.

Source X:

Jayant knows Chelsea won

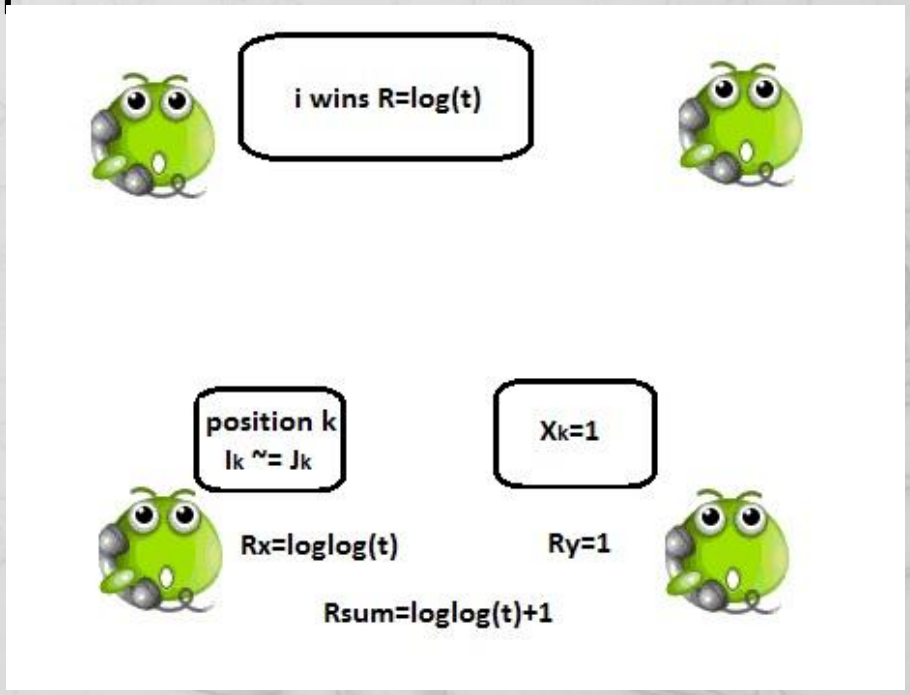
Source Y:

I know Chelsea Vs MU

Aim:

I know Chelsea won

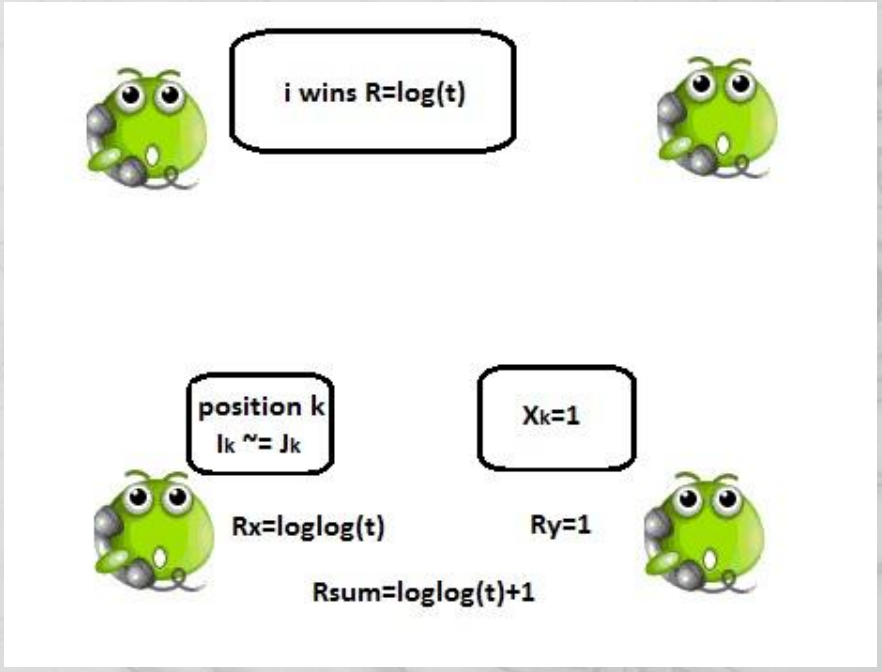
(Reconstruct X on Y side)



Two Messages are Optimal

High level idea of the proof

Construct a communication scheme as shown in Example 1



Two Messages are Optimal

For all (X,Y) pairs with $\hat{\mu}_{X|Y} = 2$

$$\hat{C}_{\infty}(X|Y) = \left\lceil \log \hat{C}_1(X|Y) \right\rceil + 1$$

Only need to show

$$\hat{C}_{\infty}(X|Y) \leq \left\lceil \log \hat{C}_1(X|Y) \right\rceil + 1$$

By construct a protocol

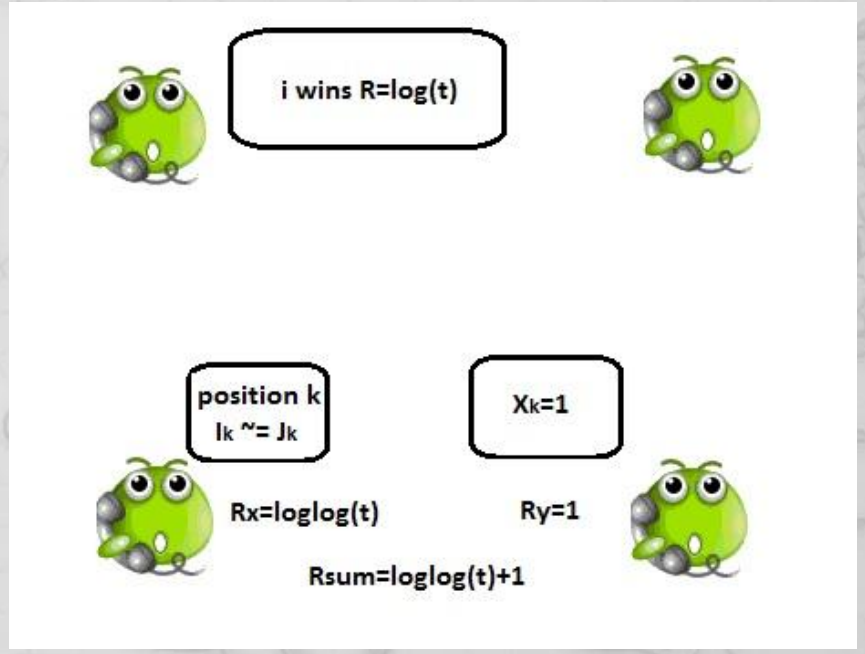
$$\hat{C}_{\infty}(X|Y) \leq \hat{C}_2(X|Y) = \left\lceil \log \hat{C}_1(X|Y) \right\rceil + 1$$

Two Messages are Optimal

X and Y agree on a $\omega(G(X|Y))$ and on a $\log(\omega(G(X|Y)))$ bit encoding of color

Y transmits the location of the two color differs

X transmits that value

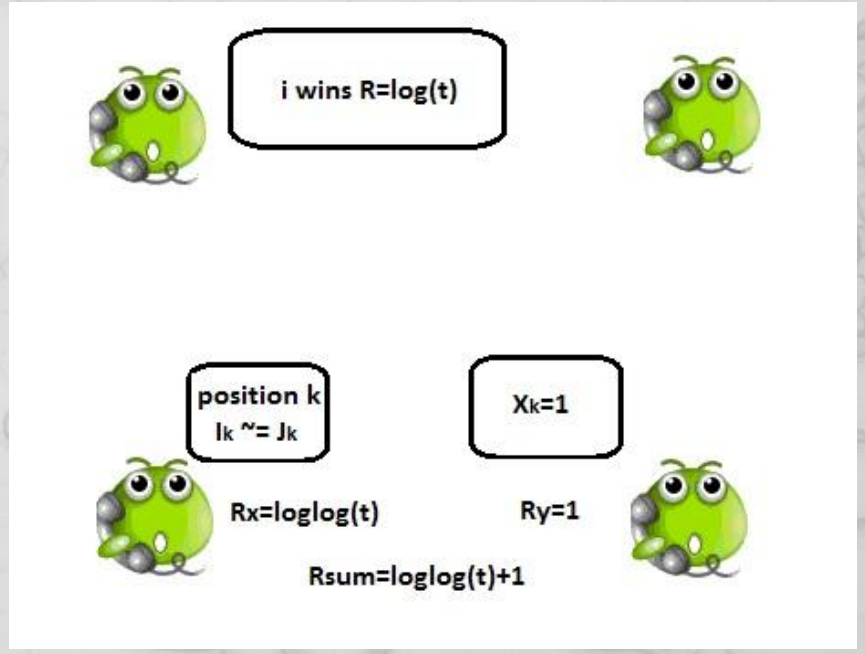


Two Messages are Optimal

General Scheme

Y transmits a sub graph that only need 2 color

X gives the color



Proof of the Results

One-Way Complexity

The Limits of Interaction

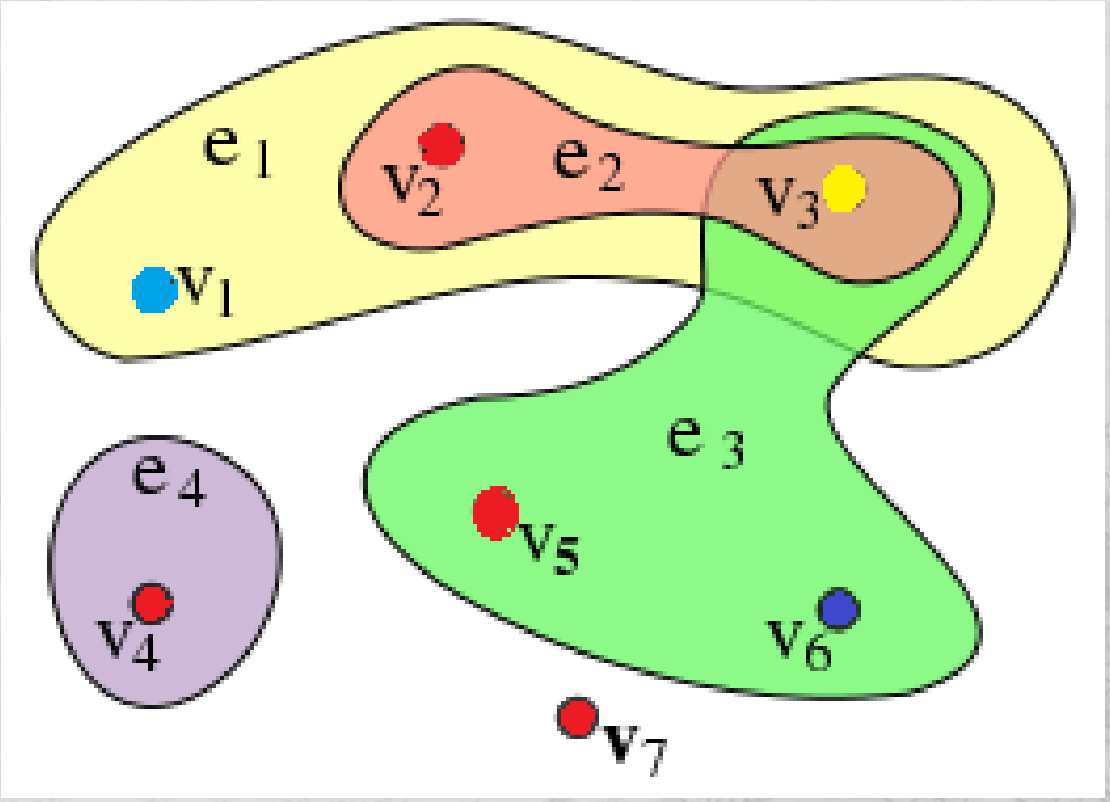
Two Messages are Optimal

Two Messages are Almost Optimal

Two Messages are Not Optimal

Two Message are Almost Optimal

$\hat{\mu}_{X|Y} > 2$



Two Message are Almost Optimal

For all nontrivial (X,Y) pairs with $\hat{\mu}_{X|Y} > 2$

We have

$$\hat{C}_2(X|Y) \leq \log\log\omega\left(G(X|Y)\right) + 3\log\hat{\mu}_{X|Y} + 4$$

Then we can show

$$\hat{C}_2(X|Y) \leq 4\hat{C}_\infty(X|Y) + 3$$

Two Message are Almost Optimal

High level idea of the proof

Y transmits a sub-hyper-graph using $\lceil \log |F| \rceil$ bits

The chromatic number of each sub-graph is b

(b>2)

X gives back the color

The idea of perfect hash functions are used here

Proof of the Results

One-Way Complexity

The Limits of Interaction

Two Messages are Optimal

Two Messages are Almost Optimal

Two Messages are Not Optimal

Two Message are Not Optimal

$$\hat{C}_2(X|Y) \geq (2-\epsilon)\hat{C}_3(X|Y) \geq (2-\epsilon)\hat{C}_\infty(X|Y)$$

High level idea of the proof

Chromatic Decomposition number $D(G(X_{l,t}|Y_{l,t}))$

Prove in some cases

$$\hat{C}_2(X|Y) \geq D(G(X_{l,t}|Y_{l,t})) \geq (2-\epsilon)\hat{C}_3(X|Y)$$

Two Message are Not Optimal

Chromatic-decomposition number $D(G(X_{l,t}|Y_{l,t}))$

Define edge cover:

$$E \subseteq \bigcup_{i=1}^{k} E_i$$

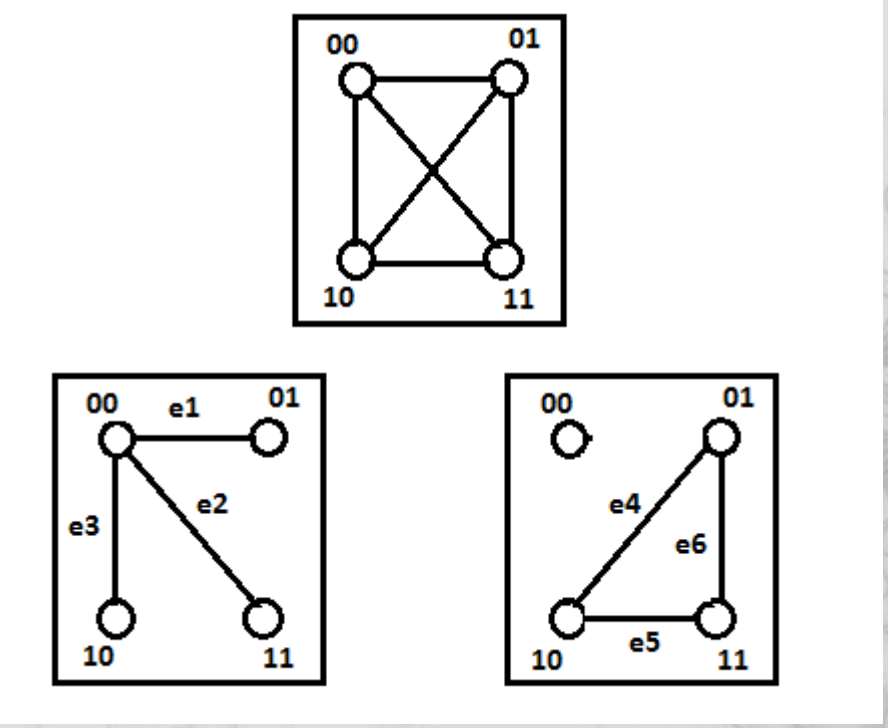
Define chromatic-decomposition number:

$$D(G) = \min_{\bigcup_{i=1}^{k} E_i} \{ \sum_{i=1}^{k} \omega(G_i) : E \subseteq \bigcup_{i=1}^{k} E_i \}$$

Two Message are Not Optimal

Edge Cover: $E = E_1 \cup E_2$

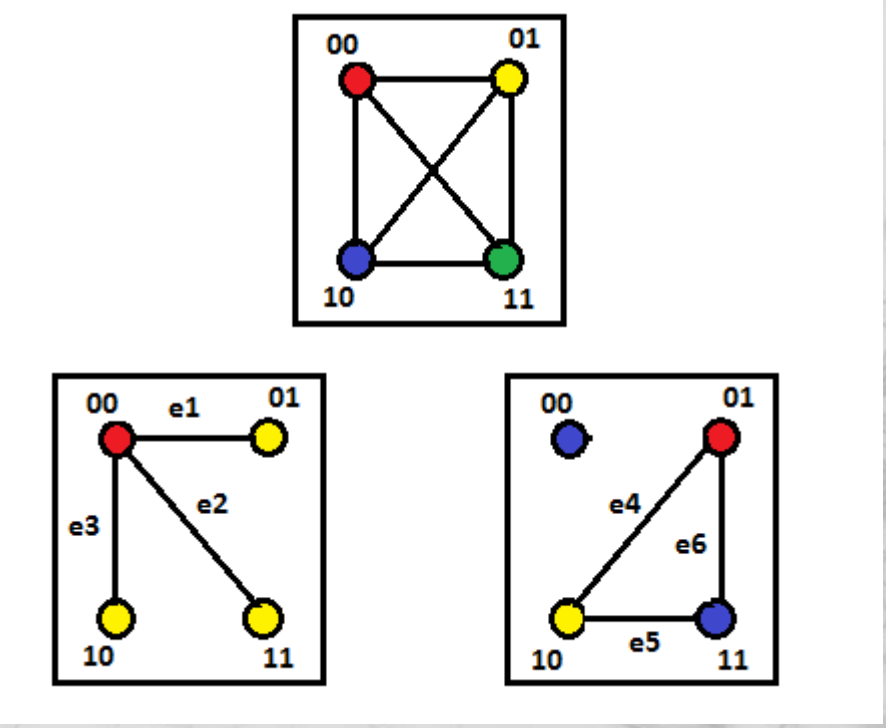
E_1={e1 e2 e3} E_2={e4 e5 e6}



Two Message are Not Optimal

Chromatic-decomposition

$\omega(E_1)=2$ $\omega(E_2)=3$



Two Message are Not Optimal

Will show, in Example 2 Playoffs

$$\hat{C}_2(X|Y) \geq D(G) \geq (2-\epsilon)\hat{C}_3(X|Y)$$

Example 2 Playoffs

L sub-leagues, t teams for each sub-league.

Totally l*t teams in the great association

First 2 teams of each sub-league come into playoffs

Source X:

Jayant know the result (champion/canceled)

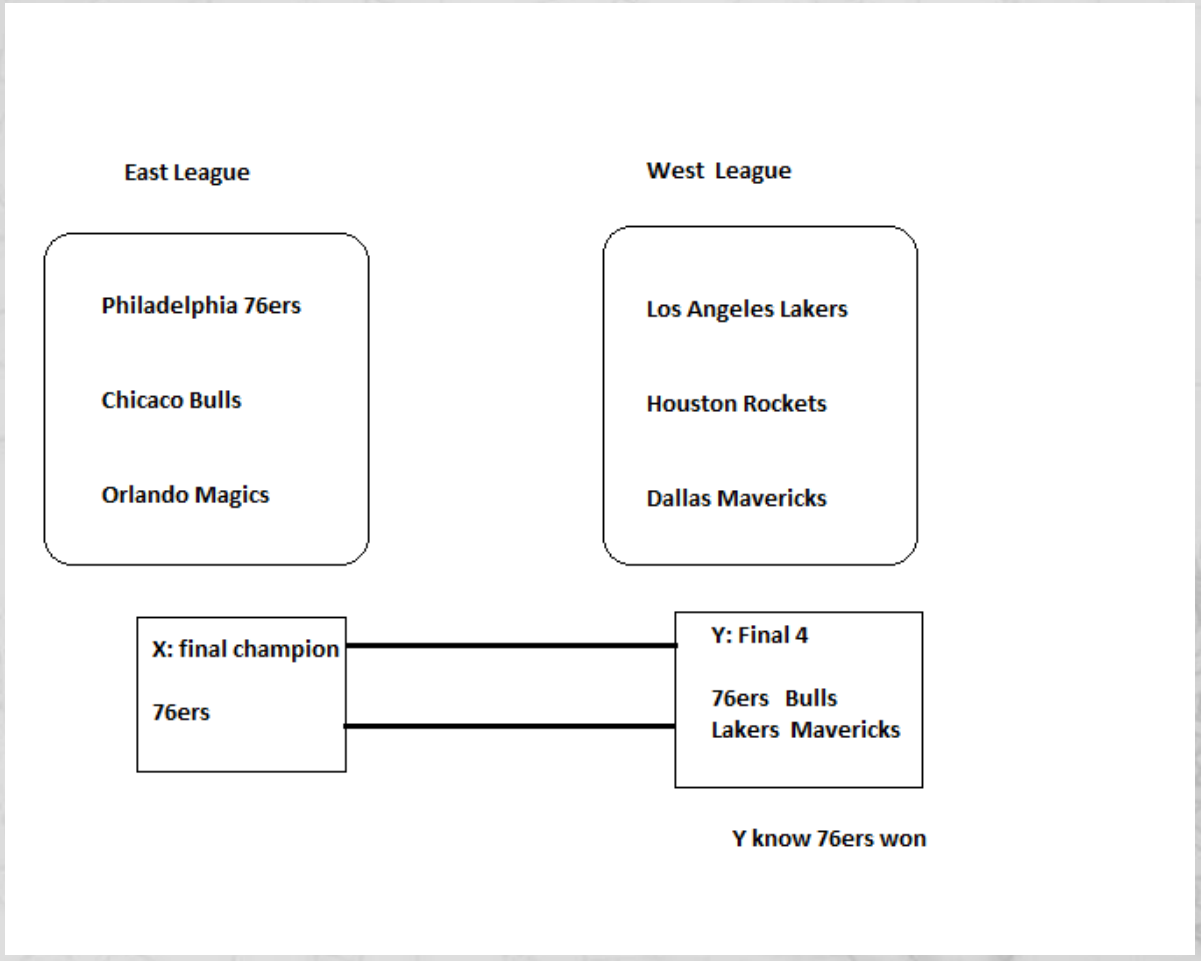
Source Y:

I know the 2l teams in the playoffs

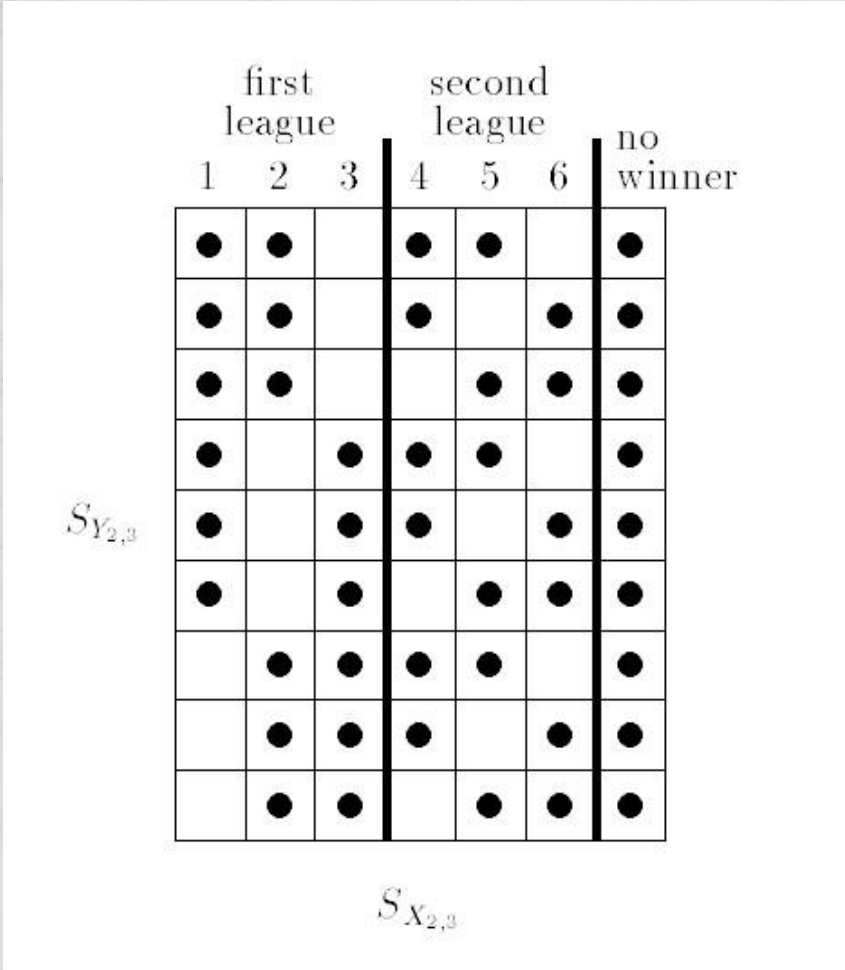
Aim: Reconstruct X on Y side (I know the result)

Example 2 Playoffs

i.e. t=3 l=2



Characteristic Table



Example 2 Playoffs

Chromatic number in example 2

$$\log\omega\big(G(X|Y)\big) = l * t + 1 = 7$$

Any two teams belong to a common edge.

(no teams can share the color, l*t color needed)

“Cancel” belong to all edges

(additional 1 color is needed for “cancel”)

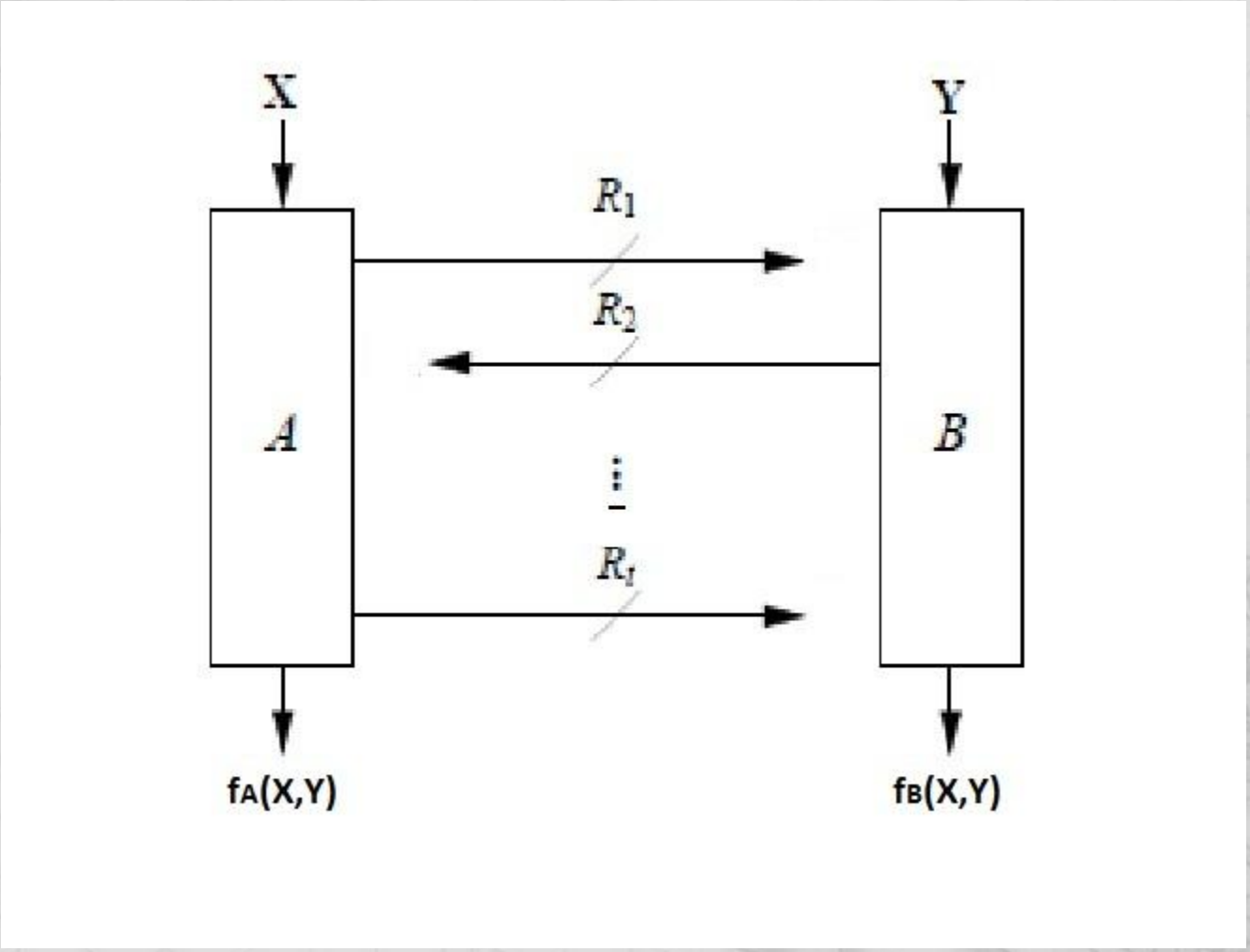
Outline

- Interaction in Zero-error Worst-Length Case
  - Problem Setup
  - Definitions and Properties
  - Results in Zero-error Worst-Length Case
  - Proof of the Results
  - Conclusion

Outline

- Problem Setup (Two-Way Source Coding)
- Interaction in Lossless Expected Length Case
- Interaction in Zero-error Worst Length Case
- Interaction in Function Computation

Interaction in Function Computation



Interaction in Function Computation

- Graph entropy
- Example 1

  Benefit can be arbitrary large
- Example 2

  Achievable Infinite-message sum-rate
- Conclusion

Interaction in Function Computation

- Graph entropy
- Example 1

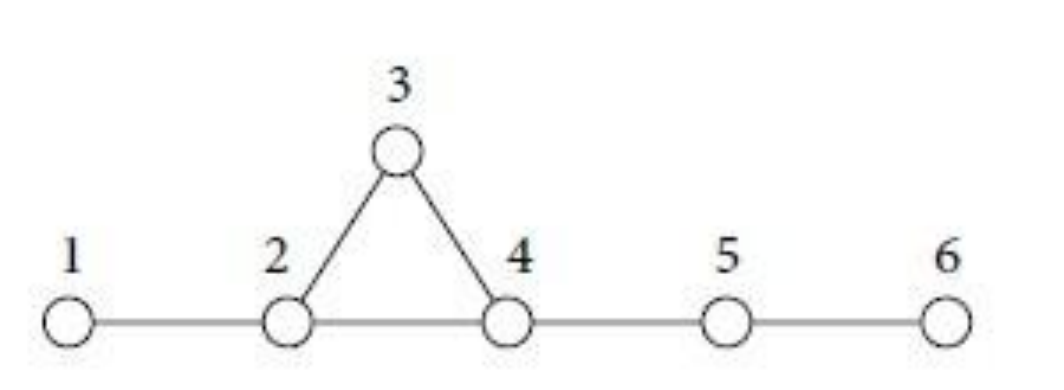
  Benefit can be arbitrary large
- Example 2

  Achievable Infinite-message sum-rate
- Conclusion

Graph Entropy

Maximum independent sets of G(V,E)

$$\Gamma(G) = \{\{1,3,5\},\{1,3,6\},\{1,4,6\},\{2,5\},\{2,6\}\}$$

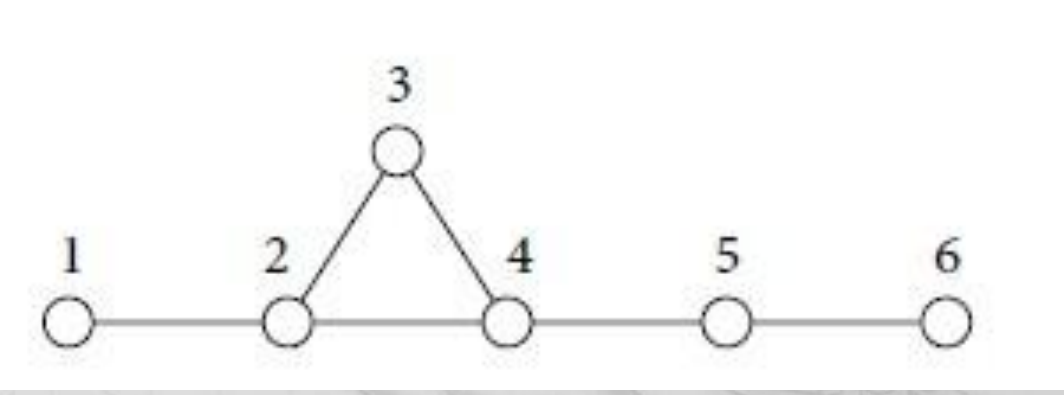


Graph Entropy

Define random viable W, $W \in \Gamma(G)$

$$p(w|x) = 0 \text{ if } x \text{ is not in } w$$

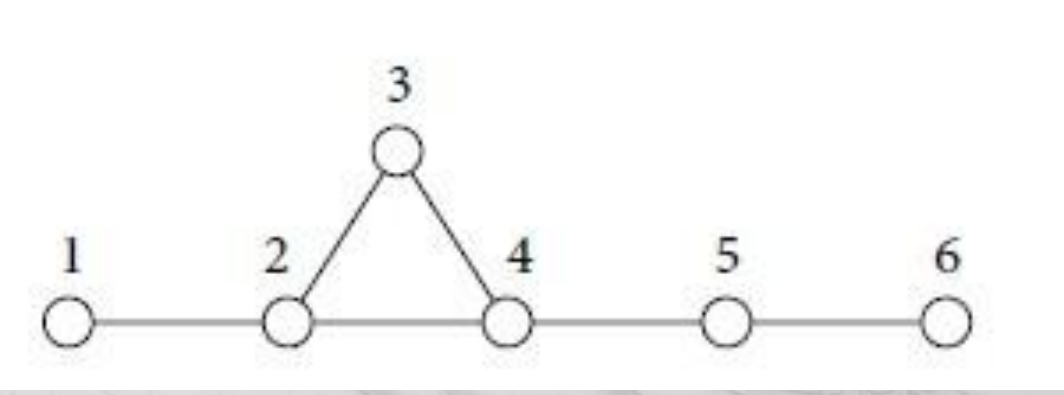
$$\sum_{w:x \in w} p(w|x) = 1$$



Graph Entropy

Graph Entropy:

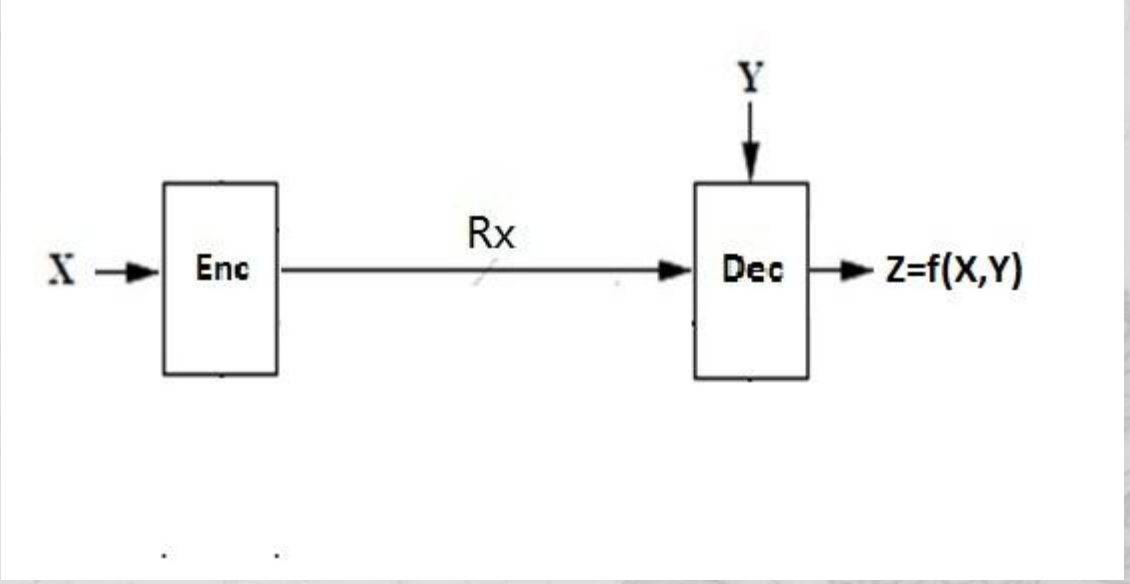
$$H_G(X) = \min_{p(w|x)} I(X;W)$$



Graph Entropy

Optimal rate for function computation satisfies:

$$H(f(X,Y)|Y) \leq R \leq H(X|Y)$$



Graph Entropy

By the definition of Characteristic graph G, we have

$$R = H_G(X|Y)$$

Compare with our chromatic number result

Interaction in Function Computation

- Graph entropy
- Example 1

  Benefit can be arbitrary large
- Example 2

  Achievable Infinite-message sum-rate
- Conclusion

Example 1

The “buffer” example

X want to send message to a buffer Y

Buffer will output the message if it’s not full

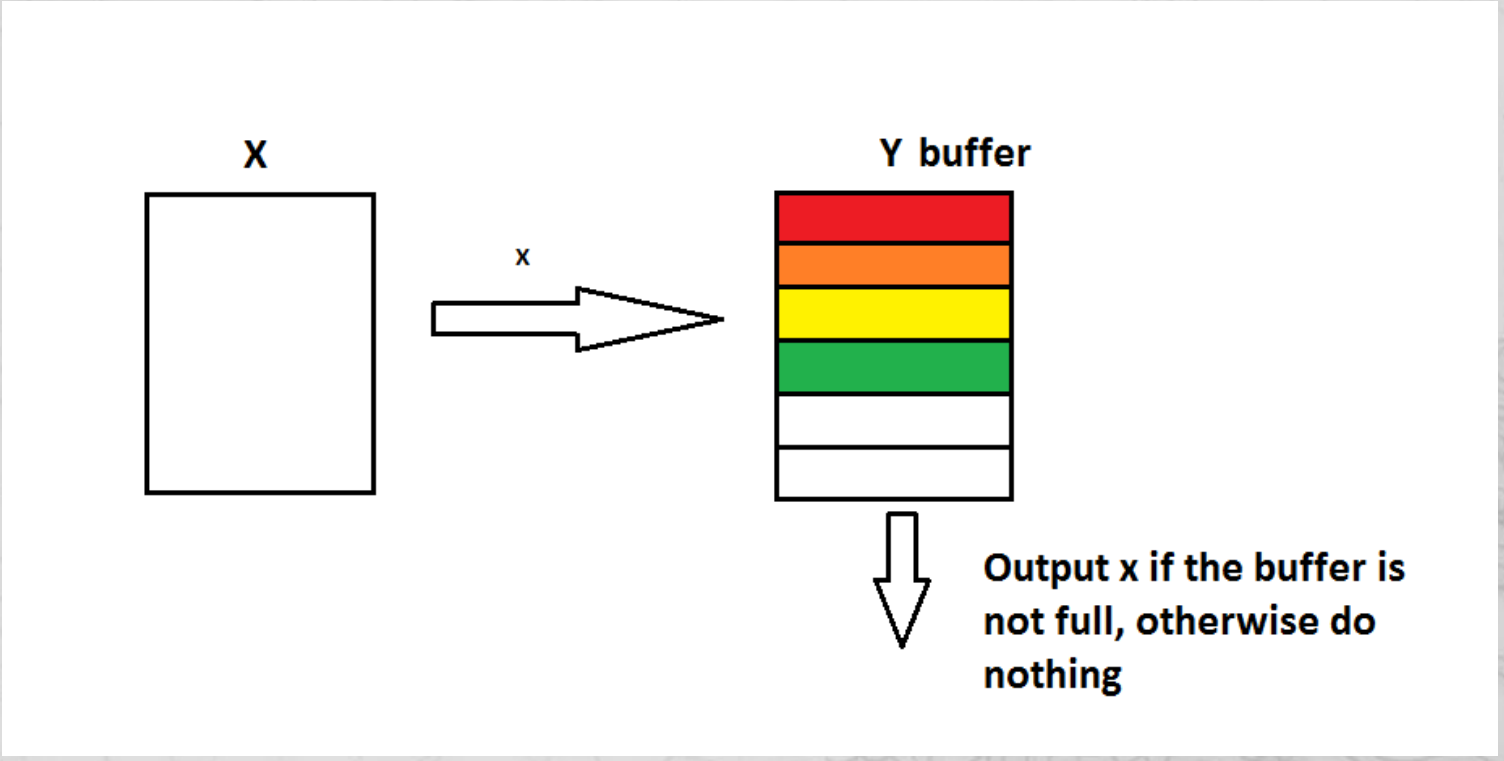
But, will throw away any new coming message if it’s full

X~Uniform{1…L} Y~Ber(p)

$f_a(x,y):=0$ $f_b(x,y):=xy$

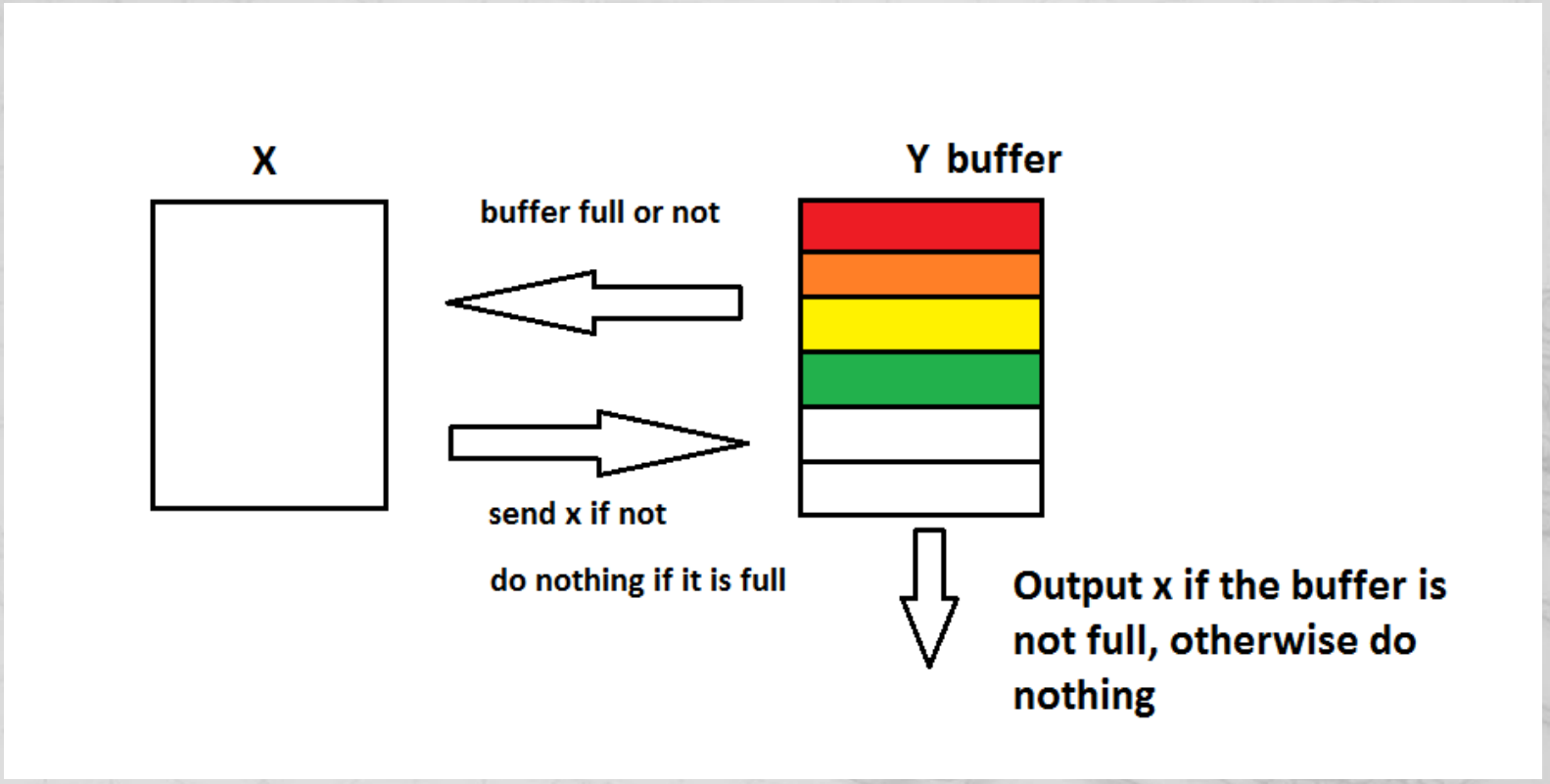
Example 1

Scheme 1: X directly sends message to Y



Example 1

Scheme 2: Y tells X if it's full or not first



Example 1

Scheme 1: X directly sends message to Y

$$R_{sum,1} = H(X|Y) = H(X)$$

Scheme 2: Y tells X if it's full or not first

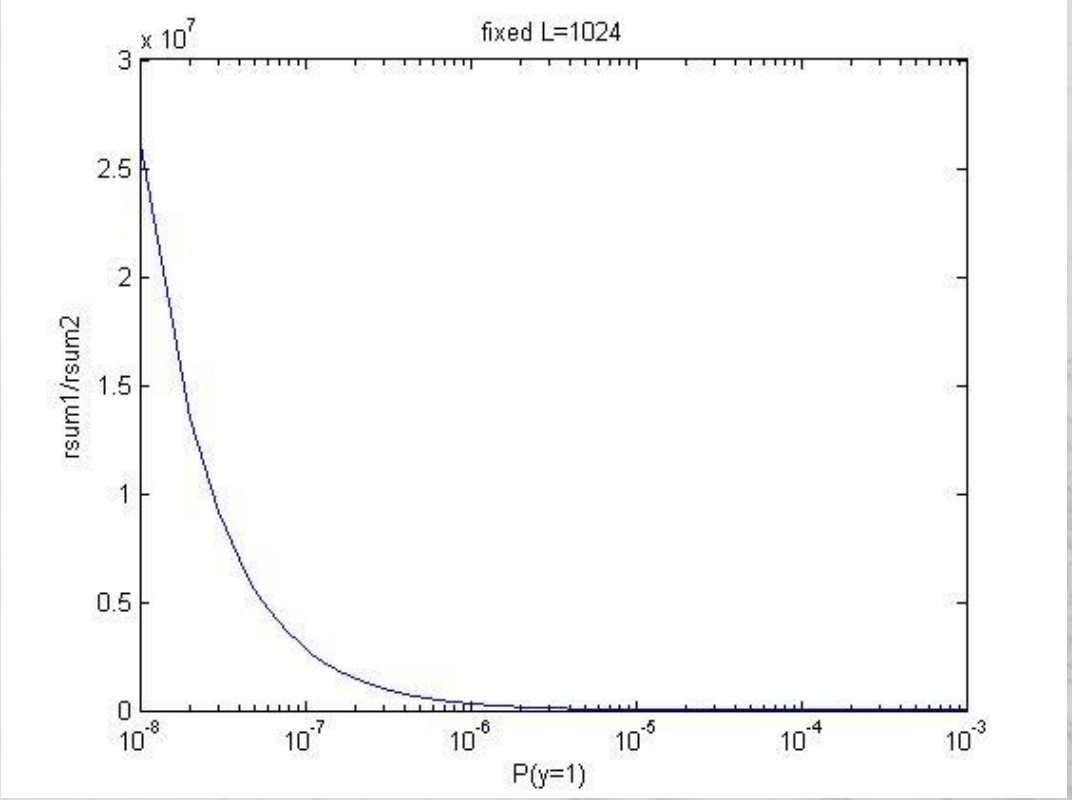
$$R_1 = H(Y) = h(p)$$

$$R_2 = H(f_b(X,Y)|Y) = p\log_2 L$$

Example 1

Fixed L, $R_{sum,1}/R_{sum,2}$ can be arbitrarily large

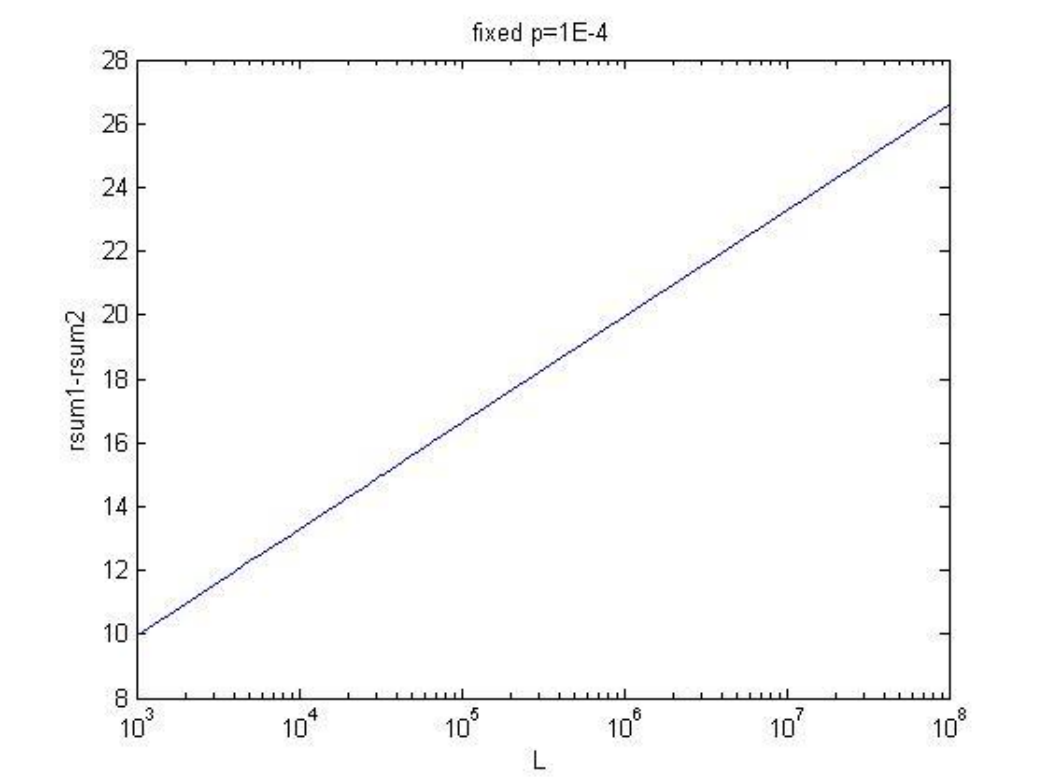
i.e. L=1024



Example 1

Fixed p, $R_{sum,1}$-$R_{sum,2}$ can be arbitrarily large

i.e. p=1E-4



Interaction in Function Computation

- Graph entropy
- Example 1

  Benefit can be arbitrary large
- Example 2

  Achievable Infinite-message sum-rate
- Conclusion

Example 2

An achievable infinite-message sum-rate as a definite integral with inginitesimal-rate messages

X~Ber(p) Y~Ber(q)

X,Y independent

$f_A(x,y)=f_B(x,y)=x^y$

Example 2

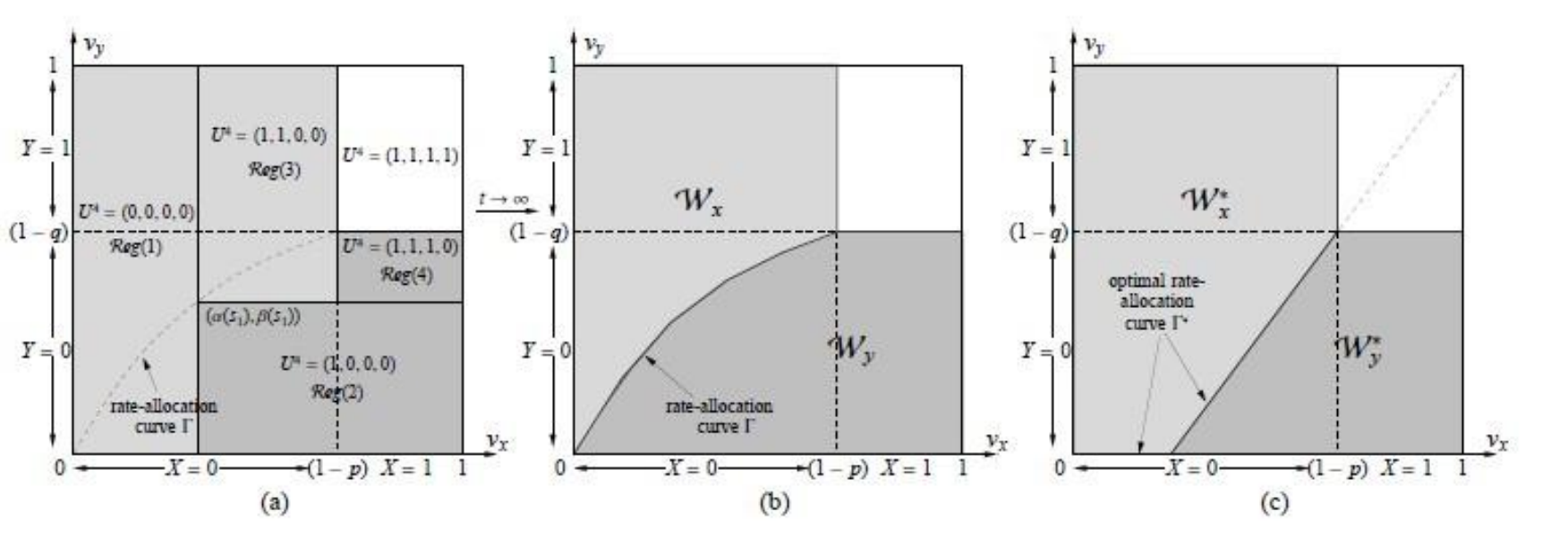
High level idea of the design:

- Define real auxiliary random variable pair
- Use real multiplication instead of AND
- Sum-rate changes to a definite integral
- Define a rate allocation curve to minimize the sum-rate

Example 2

Sum-rate changes to a definite integral

Define a rate allocation curve to minimize the sum-rate



Example 2

$$R_{sum,\infty} = \min\left\{\iint_{w_x} w_x(v_x, p)\, dv_x dv_y + \iint_{w_y} w_y(v_y, p)\, dv_x dv_y\right\}$$

Optimize by the rate allocation curve

Can have

$$R_{sum,\infty} = h(p) + p * h(q) + p * \log(q) + p(1-q)\log(e)$$

Compare with

$$R_{sum,2} = H(X|Y) + h(X^Y|X) = h(p) + p * h(q)$$

Example 2

i.e. p=0.5 q=0.5

$$R_{sum,\infty} \approx 1.361$$

$$R_{sum,2} = 1.5$$

Interaction in Function Computation

- Graph entropy
- Example 1

  Benefit can be arbitrary large
- Example 2

  Achievable Infinite-message sum-rate
- Conclusion

Reference

[1]Amiram H. Kaspi “Two-Way Source Coding with a Fidelity Critertion”

[2]Abbas El Gamal, Yound-Han Kim “Network Information Theory” Chapter 20,21

[3]Alon Orlitsky “Worst-Case Interactive Communication I: Two Messages are Almost Optimal”

[4]Alon Orlitsky “Worst-Case Interactive Communication II: Two Messages are Not Optimal”

[5]Nan Ma, Prakash Ishwar “Interaction Strictly Improves the Wyner-Ziv Rate-distortion Function”

[6]Nan Ma, Prakash Ishwar “Distributed Source Coding for Interactive Function Computation”

[7]Nan Ma, Prakash Ishwar “Infinite-message Distributed Source Coding for Two-terminal Interactive Computing”