The Benefit of Interaction in Lossy Source Coding

Nan Ma

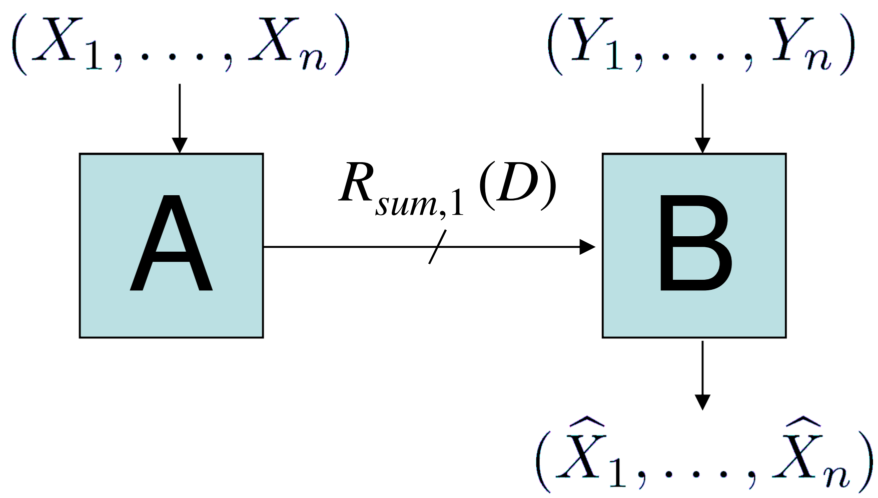
UC Berkeley

Joint work with Prakash Ishwar, Boston University

Acknowledgment:
Prof. Ram Zamir, Dr. Vinod Prabhakaran

Research supported by:

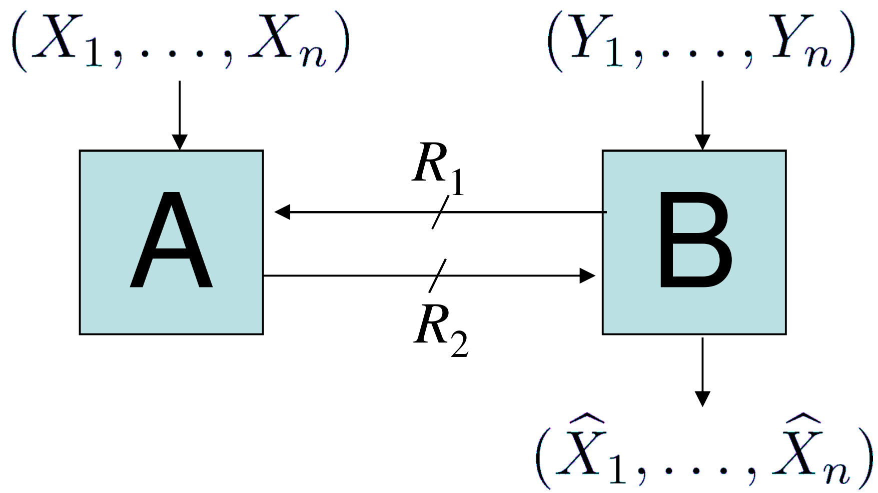
Wyner-Ziv problem



- n samples $(X_i, Y_i) \sim \text{iid } p_{XY}$
- Per-sample distortion measure $d(x, \hat{x})$
- Wyner-Ziv rate-distortion function [Wyner & Ziv IT'76]:

$$R_{sum,1}(D) = \min_{\substack{U-X-Y \\ \widehat{X}=g(U,Y) \\ E[d(X,\widehat{X})] \leq D}} I(X;U|Y)$$

Kaspi's 2-way src coding problem (simplified version)

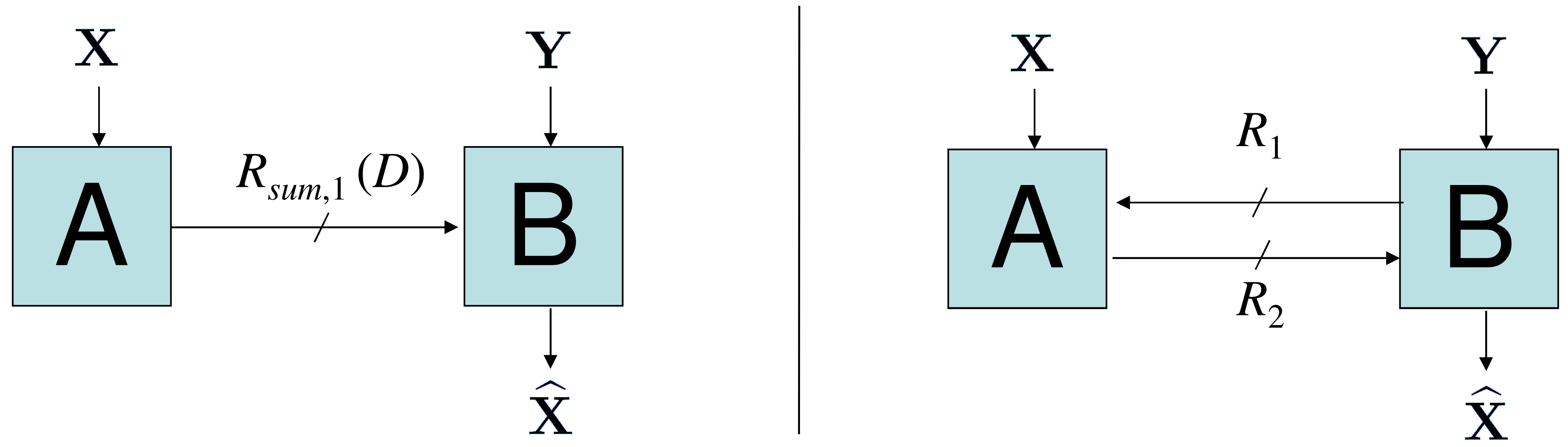


- Same objective: lossy source reproduction
- Two-message interaction
  - Sum-rate: $R_1 + R_2$, minimum sum-rate for distortion D is $R_{sum,2}(D)$
- Sum-rate-distortion function [Kaspi IT'85] :

$$R_{sum,2}(D) = \min_{\substack{V_1 - Y - X \\ V_2 - (XV_1) - Y \\ \widehat{X} = g(V_1, V_2, Y) \\ E[d(X,\widehat{X})] \le D}} \{I(Y;V_1|X) + I(X;V_2|YV_1)\}$$

Main question

- One message v.s. two messages

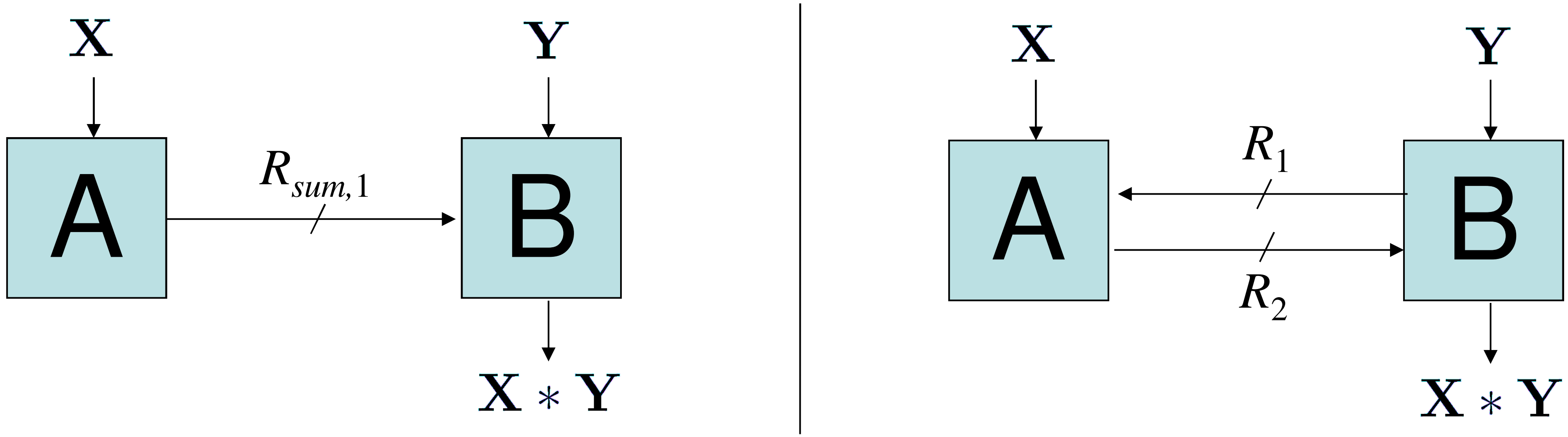


- $R_{sum,1} \geq R_{sum,2}$ always holds
- Question [Kaspi IT'85]: Is interaction useful?

$R_{sum,1} = R_{sum,2}$ or $R_{sum,1} > R_{sum,2}$ for some D?

Related results

- Lossless function computation [Orlitsky & Roche IT'01], [Ma & Ishwar ISIT'08]
  - Independent sources: $X \sim$ Uniform$\{1, 2, \ldots, L\}$, $Y \sim$ Bernoulli (p)
  - B computes $X * Y$

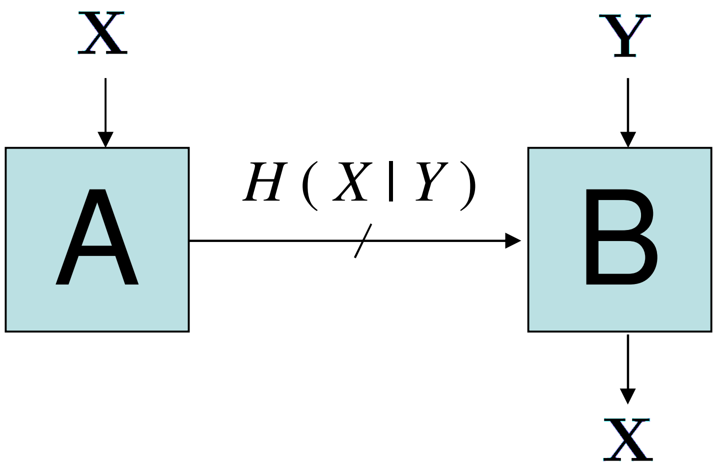


$R_{sum,1} = \log_2 L$ *strictly* > $R_1 + R_2 = h_2(p) + p \log_2 L$

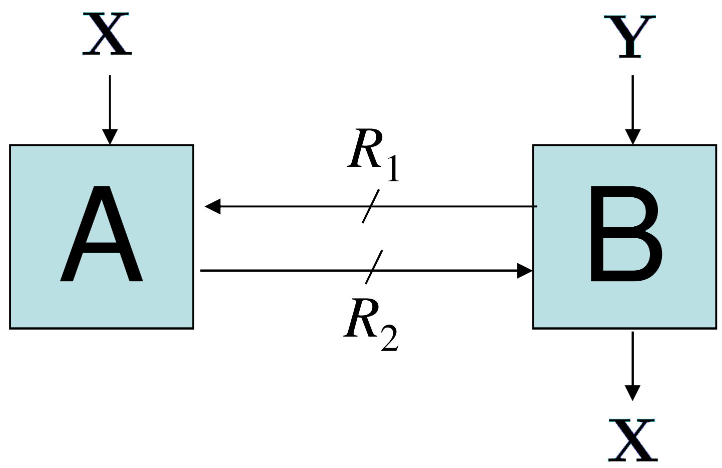
- $R_{sum,1} / (R_1 + R_2)$ can be arbitrarily large

Related results

- Lossless source reproduction [Slepian & Wolf IT'73] and cutset bound



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



even if $R_1 = H(Y)$, $R_2 \geq H(X|Y)$

  - No benefit in using two messages
  - Caveat: interaction may help for nonergodic sources [Yang & He, ISIT'08]

Contribution

- Main question:
  - Lossy source reproduction: $R_{sum,1} = R_{sum,2}$ or $R_{sum,1} > R_{sum,2}$?

- Answer: $R_{sum,1} > R_{sum,2}$ ---- *interaction is useful* [Ma & Ishwar ISIT'10]
  - Will first show this <u>without explicitly constructing a 2-msg scheme</u>
  - Will then show explicit construction in which

    1) Gain of interaction is arbitrarily large and simultaneously,

    2) Feedback rate is arbitrarily small, compared to the forward rate

- Key tool: rate reduction functional

Rate reduction functional [Ma & Ishwar Allerton'09, arXiv Oct'09]

- Functional viewpoint: $R_{sum,i}$ is a functional of (p_{XY}, D)
- Rate reduction functional

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

Rate to exchange
sources losslessly

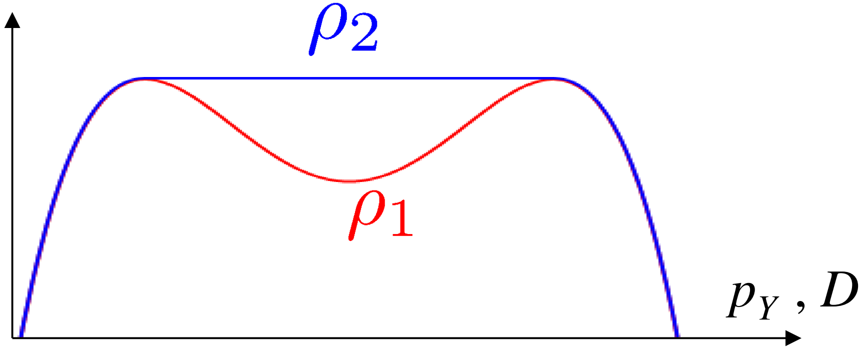
$$\rho_1(p_{XY}, D) = \max_{\substack{U-X-Y \\ \widehat{X}=g(U,Y) \\ E[d(X,\widehat{X})]\leq D}} [H(Y|X) + H(Y|U,X)]$$

- Since $R_{sum,1} \geq R_{sum,2}$ always holds, $\rho_1 \leq \rho_2$ always holds
- Thus, $R_{sum,1} > R_{sum,2}$ iff $\rho_1 < \rho_2$ iff $\rho_1 \neq \rho_2$

Rate reduction functional [Ma & Ishwar Allerton'09, arXiv Oct'09]

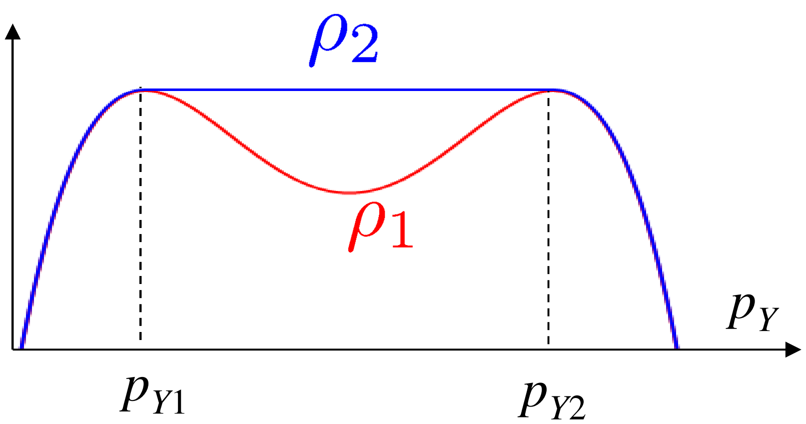
- Key corollary of result from previous talk:

$\rho_1 = \rho_2$ iff $\rho_1(p_{X|Y}\, p_Y, D)$ is concave w.r.t. (p_Y, D)



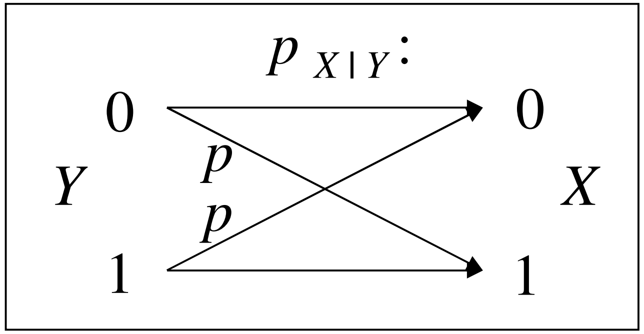
- If ρ_1 is not concave, "concavification" improves performance
- Will pick $p_{X|Y}$ and distortion function and show that:
  - $\rho_1(p_{X|Y}\, p_Y, D)$ is not concave w.r.t. p_Y which implies
  - $\rho_1 \neq \rho_2 \Rightarrow \rho_1 < \rho_2 \Rightarrow R_{sum,1} > R_{sum,2}$

$\rho_1(p_{X|Y}\, p_Y, D)$ is not concave w.r.t. p_Y



- Let $p_{Y_1} \sim Ber(q)$ and $p_{Y_2} \sim Ber(\bar{q})$, where, $\bar{q} := (1-q)$

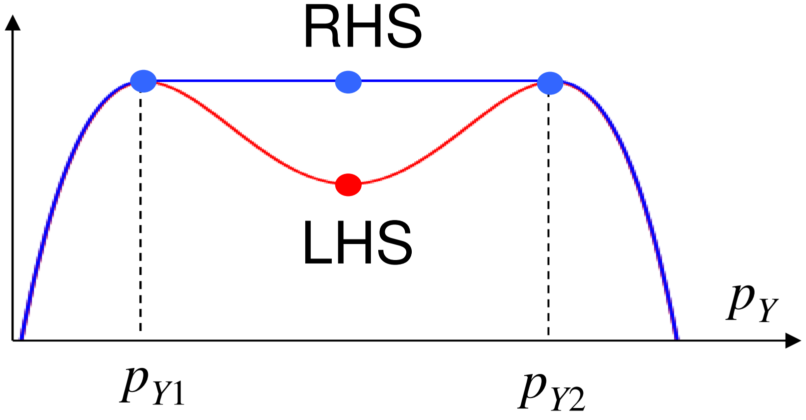
- Let $p_{X|Y} \sim BSC(p) =$



- Let d = binary erasure distortion =

$d(x,\hat{x})$			
$x \backslash \hat{x}$	0	e	1
0	0	1	∞
1	∞	1	0

$\rho_1(p_{X|Y}\, p_Y, D)$ is not concave w.r.t. p_Y



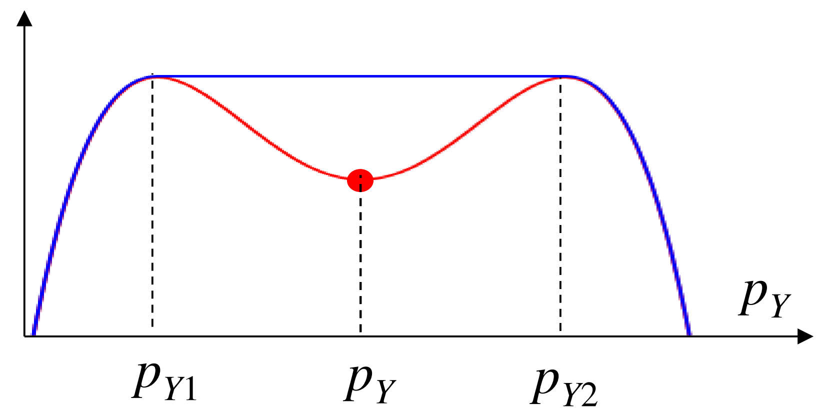
- Will prove that there exist (p, q, D) such that

$$\rho_1\left(p_{X|Y}\frac{p_{Y_1}+p_{Y_2}}{2}, D\right) < \frac{1}{2}\rho_1(p_{X|Y}p_{Y_1}, D) + \frac{1}{2}\rho_1(p_{X|Y}p_{Y_2}, D)$$

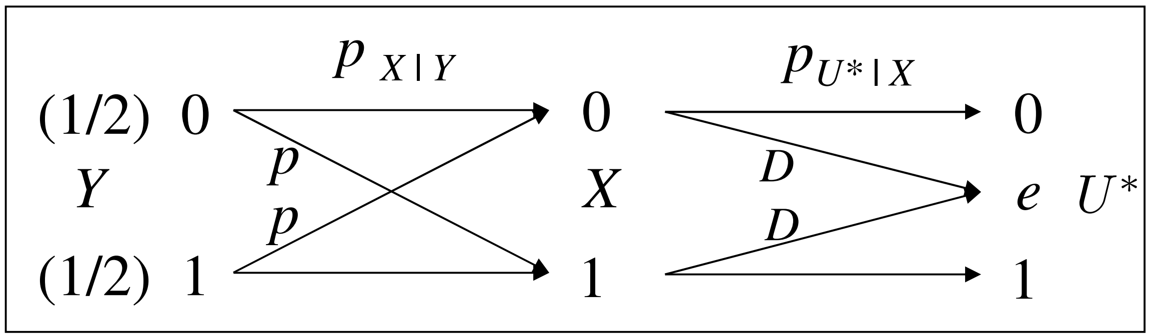
(would have been $\geq$ if concave)

Rate reduction for DSBS

- Objective: $\rho_1\left(p_{X|Y}\frac{p_{Y_1}+p_{Y_2}}{2}, D\right) < \frac{1}{2}\rho_1(p_{X|Y}p_{Y_1}, D) + \frac{1}{2}\rho_1(p_{X|Y}p_{Y_2}, D)$



- Left-side: $p_Y := \frac{p_{Y_1}+p_{Y_2}}{2} \sim Ber(1/2); \quad p_{X|Y} \sim BSC(p)$
  - For $D \in [0,1]$, the optimal auxiliary variables are

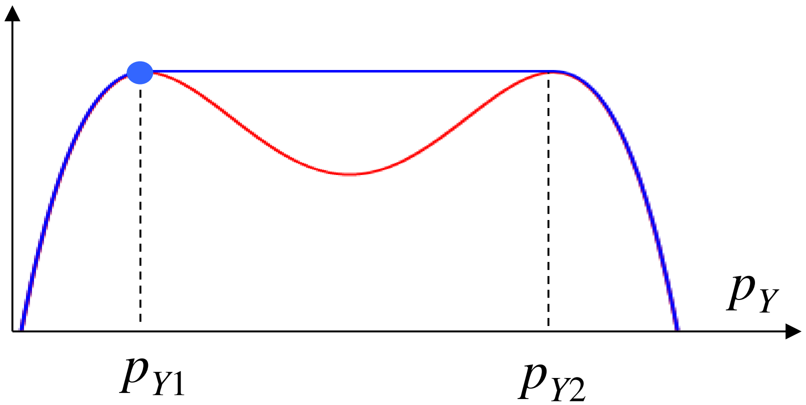


$$\widehat{X}^* = U^*$$

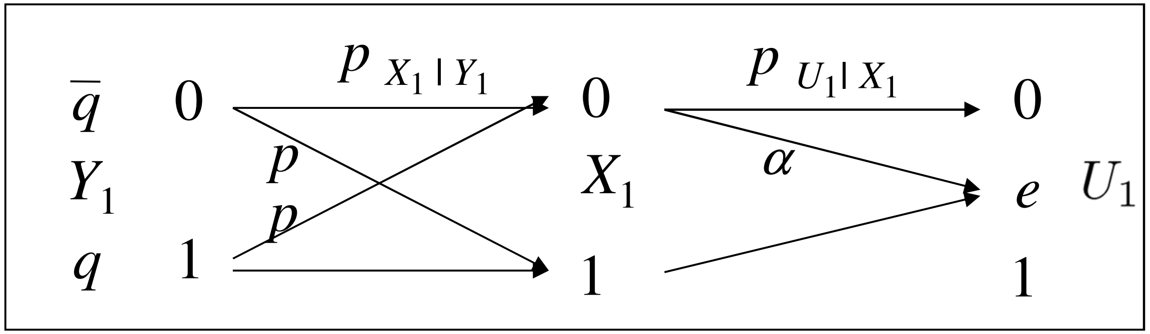
  - Rate reduction: $\rho_1(p_{X|Y}p_Y, D) = (1+D)h(p)$

Rate reduction for symmetrically correlated sources

- Objective: $\rho_1\left(p_{X|Y}\frac{p_{Y_1}+p_{Y_2}}{2}, D\right) < \frac{1}{2}\rho_1(p_{X|Y}p_{Y_1}, D) + \frac{1}{2}\rho_1(p_{X|Y}p_{Y_2}, D)$



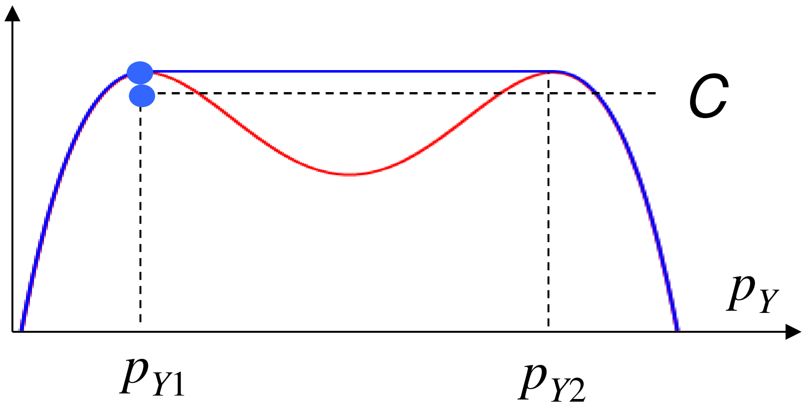
- 1st term on right-side: $p_{Y_1} \sim Ber(q);\ \ p_{X|Y} \sim BSC(p)$
  - A valid choice of (suboptimal) auxiliary variables is



$\hat{X}_1 = U_1$

Rate reduction for symmetrically correlated sources

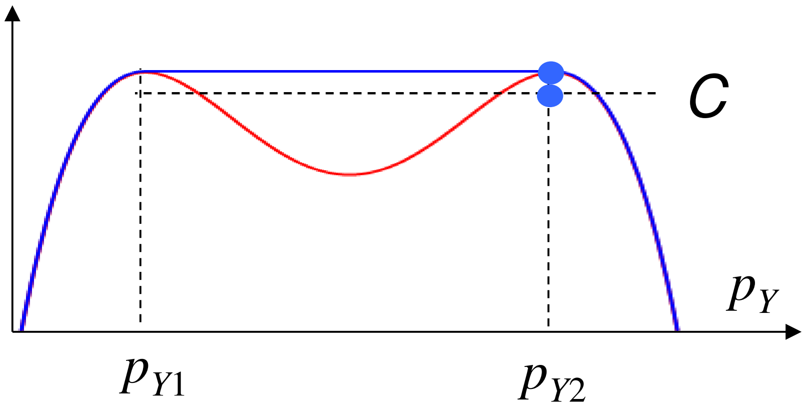
- Objective: $\rho_1\left(p_{X|Y}\frac{p_{Y_1}+p_{Y_2}}{2}, D\right) < \frac{1}{2}\rho_1(p_{X|Y}p_{Y_1}, D) + \frac{1}{2}\rho_1(p_{X|Y}p_{Y_2}, D)$



- 1st term on right-side: $p_{Y_1} \sim Ber(q)$; $p_{X|Y} \sim BSC(p)$
  - Distortion: $D'(p,q,\alpha) = (\bar{p}\bar{q} + pq)\alpha + (\bar{p}q + p\bar{q})$
  - Rate reduction: $\rho_1(p_{X|Y}p_{Y_1}, D') \geq H(Y_1|X_1) + H(X_1|Y_1,U_1) =: C(p,q,\alpha)$

Rate reduction for symmetrically correlated sources

- Objective: $\rho_1\left(p_{X|Y}\frac{p_{Y_1}+p_{Y_2}}{2}, D\right) < \frac{1}{2}\rho_1(p_{X|Y}p_{Y_1}, D) + \frac{1}{2}\rho_1(p_{X|Y}p_{Y_2}, D)$



- 2nd term on right-side: $p_{Y_2} \sim Ber(\bar{q})$; $p_{X|Y} \sim BSC(p)$
  - By symmetry:

$$\rho_1(p_{X|Y}p_{Y_2}, D') = \rho_1(p_{X|Y}p_{Y_1}, D') \geq C(p, q, \alpha)$$

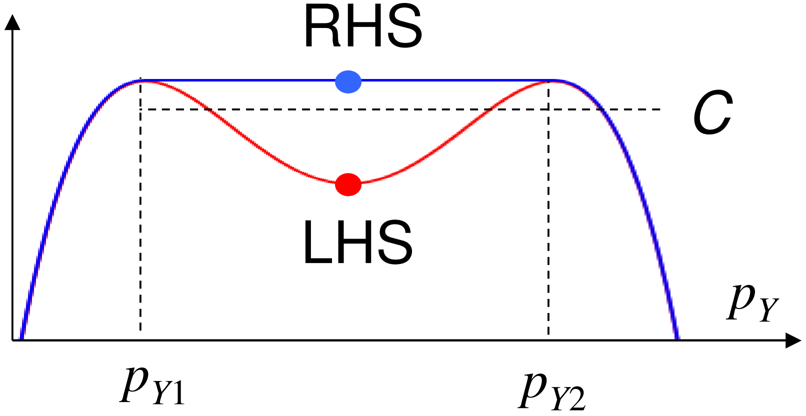
where

$$D'(p, q, \alpha) = (\bar{p}\bar{q} + pq)\alpha + (\bar{p}q + p\bar{q})$$

Comparing left-side with right-side

- Objective: $\rho_1\left(p_{X|Y}\frac{p_{Y_1}+p_{Y_2}}{2}, D\right) < \frac{1}{2}\rho_1(p_{X|Y}p_{Y_1}, D) + \frac{1}{2}\rho_1(p_{X|Y}p_{Y_2}, D)$

- Left-side $= (1+D)h(p)$



- Right-side $\geq C(p,q,\alpha)$, $D = D'(p,q,\alpha)$

- As $p \to 0$: $$\lim_{p\to 0}\frac{(1+D)h(p)}{h(p)} = 2-\bar{q}(1-\alpha)$$ $$\lim_{p\to 0}\frac{C(p,q,\alpha)}{h(p)} = 2-q(1-\alpha)$$

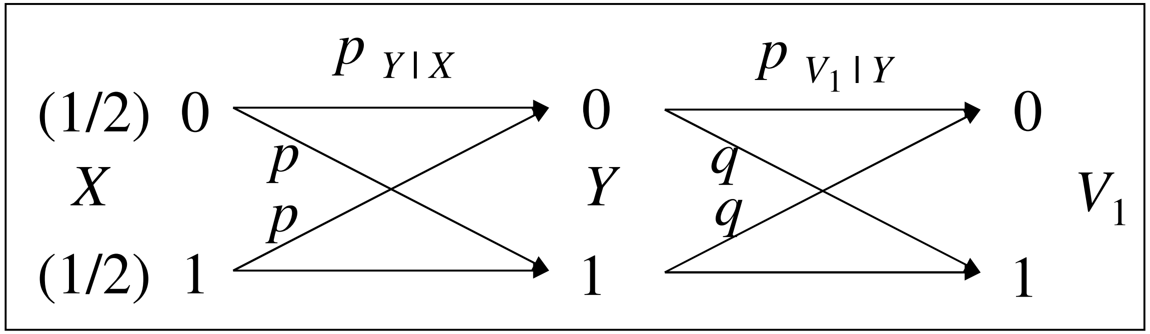
- For $q\in(0,1/2), 0<p\ll 1$:
  - left-side *strictly* < C ≤ right-side

- $\rho_1(p_{X|Y}\,p_Y, D)$ is not concave w.r.t. $p_Y \Rightarrow \rho_1 \neq \rho_2$

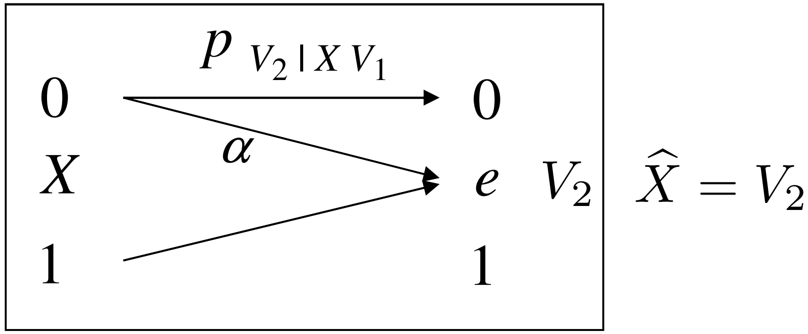
Explicit construction of 2-message aux.r.v.'s

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

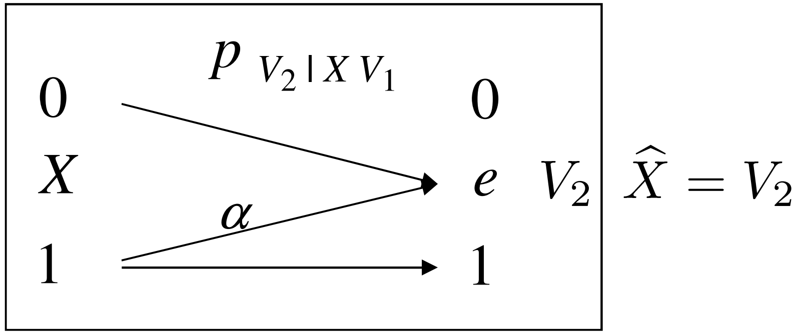
- Let $(X, Y) \sim$ DSBS (p) and
- d = binary erasure distortion
- Choose V_1 to be:



- Choose V_2 as follows:
  - Given $V_1 = 0$

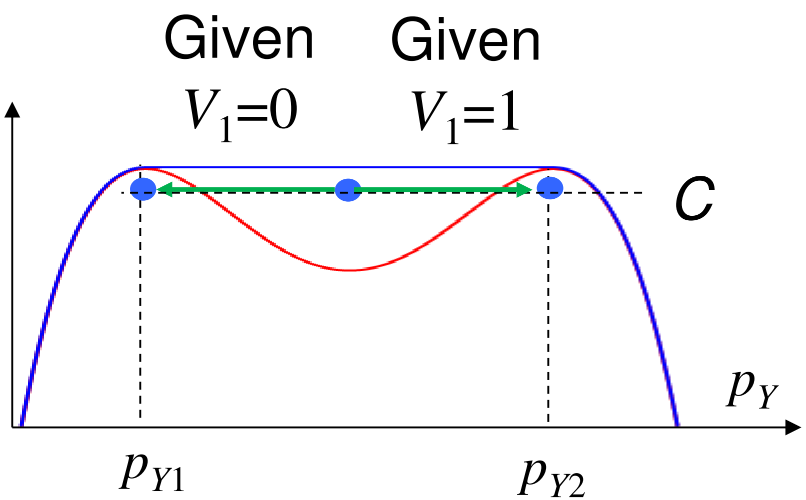


  - Given $V_1 = 1$



Explicit construction of 2-message aux.r.v.

- Connecting choice of V_1 , V_2 to previous discussion:
  - Conditioning on 1st msg, two-msg system back to one-msg system



  - Given $V_1 = 0$: $p_{YXV_2|V_1} = p_{Y_1X_1U_1}$, $\widehat{X} = V_2 = U_1$
  - Given $V_1 = 1$: symmetric case

Explicit construction of 2-message aux.r.v.

- Distortion: $D'(p, q, \alpha)$
- Rate reduction: $C(p, q, \alpha)$
- Take limits:

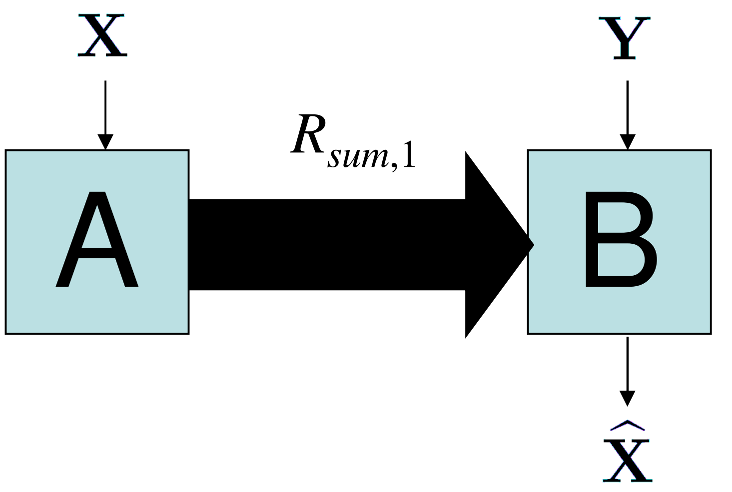
$$\lim_{p \to 0} \frac{R_1}{h(p)} = \lim_{p \to 0} \frac{I(Y; V_1 | X)}{h(p)} = 0$$

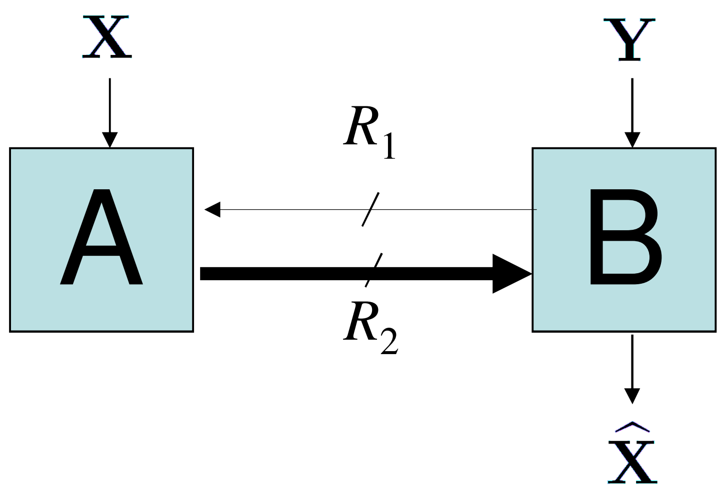
$$\lim_{p \to 0} \frac{R_2}{h(p)} = \lim_{p \to 0} \frac{I(X; V_2 | X, V_1)}{h(p)} = q(1 - \alpha)$$

$$\lim_{p \to 0} \frac{R_{sum,1}}{h(p)} = \bar{q}(1 - \alpha)$$

Explicit construction of 2-message aux.r.v.

- When $0 < q \ll 1, 0 < p \ll 1, \quad R_1 \ll R_2 \ll R_{sum,1} \ll 1$





  - $R_{sum,1} / (R_1 + R_2)$ can be arbitrarily large and simultaneously
  - R_1 / R_2 can be arbitrarily small

Another example: arbitrarily large additive gain

- Can these three conditions hold simultaneously?
  - $R_{sum,1}$ - ($R_1 + R_2$) arbitrarily large
  - $R_{sum,1}$ / ($R_1 + R_2$) arbitrarily large
  - R_1 / R_2 arbitrarily small

- Yes!

Another example: arbitrarily large additive gain

- Extension of a non-interactive example in [Cohen, Zamir '08]
  - Large alphabets
  - Planar difference set
- Size of alphabets: $\Theta(a^2)$
- When a grows,
  - $R_{sum,1}$ - ($R_1 + R_2$) ~ $\log a$ arbitrarily large
  - $R_{sum,1}$ / ($R_1 + R_2$) ~ $2 (\log a)^2$ arbitrarily large
  - R_1 / R_2 ~ $(\log\log a)$ / $(\log a)$ arbitrarily small

Concluding remarks

- Interaction strictly improves the Wyner-Ziv R-D function for lossy source reproduction
  - The benefit of a very small feedback rate can be huge
- Powerful tool: rate reduction functional
  - Connects one-message and two-message performance
  - *Concavity* of ρ_1 functional is equivalent to *optimality* of ρ_1
- There are interesting questions beyond the single-letter characterization