The Infinite Message Limit of Interactive Source Coding

Prakash Ishwar

Joint work with Nan Ma

Electrical and Computer Engineering

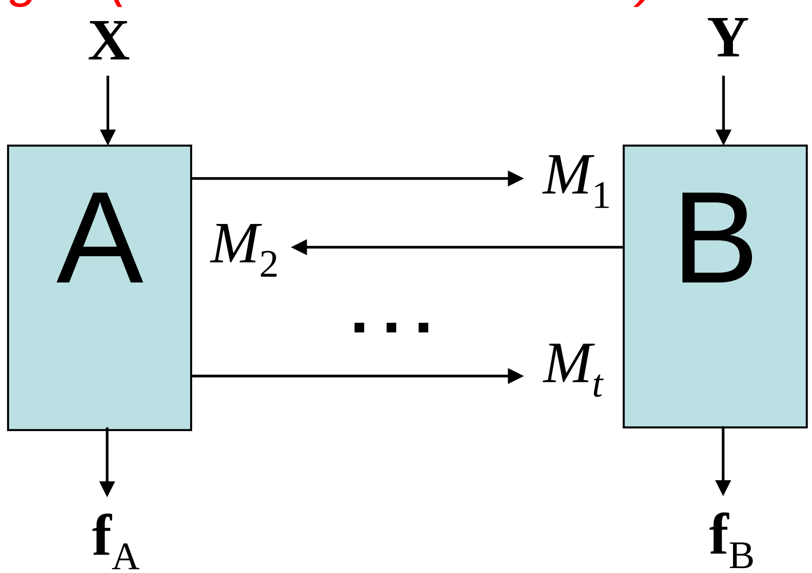
BOSTON UNIVERSITY

Research supported by

NSF

Motivation

- Degrees of freedom "resources" in distributed source coding and computation:
  - Blocklength (delay)
  - Rate (speed)
  - Quantizer step-size (quality)
  - Network size
  - Alphabet size …
  - *Number of messages (Multiround interaction)*

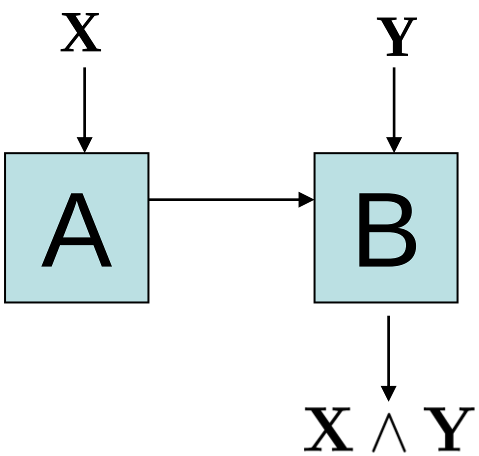


- Main goal: Characterize the *ultimate* benefit of interaction

Example: compute AND at terminal B

- Sources: $X \perp\!\!\!\perp Y, X \sim \mathrm{Ber}(p), Y \sim \mathrm{Ber}(q)$, Compute: $X \wedge Y$ at B

- One message: min rate = ?



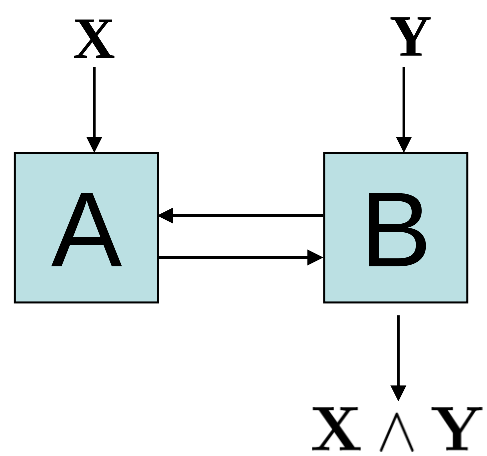
- 1-msg rate = $h(p)$

  [Yamamoto, IT'82]

- 3-msgs.? 4-msgs.? …

- What about ∞-msgs.?

- Two messages min rate = ?

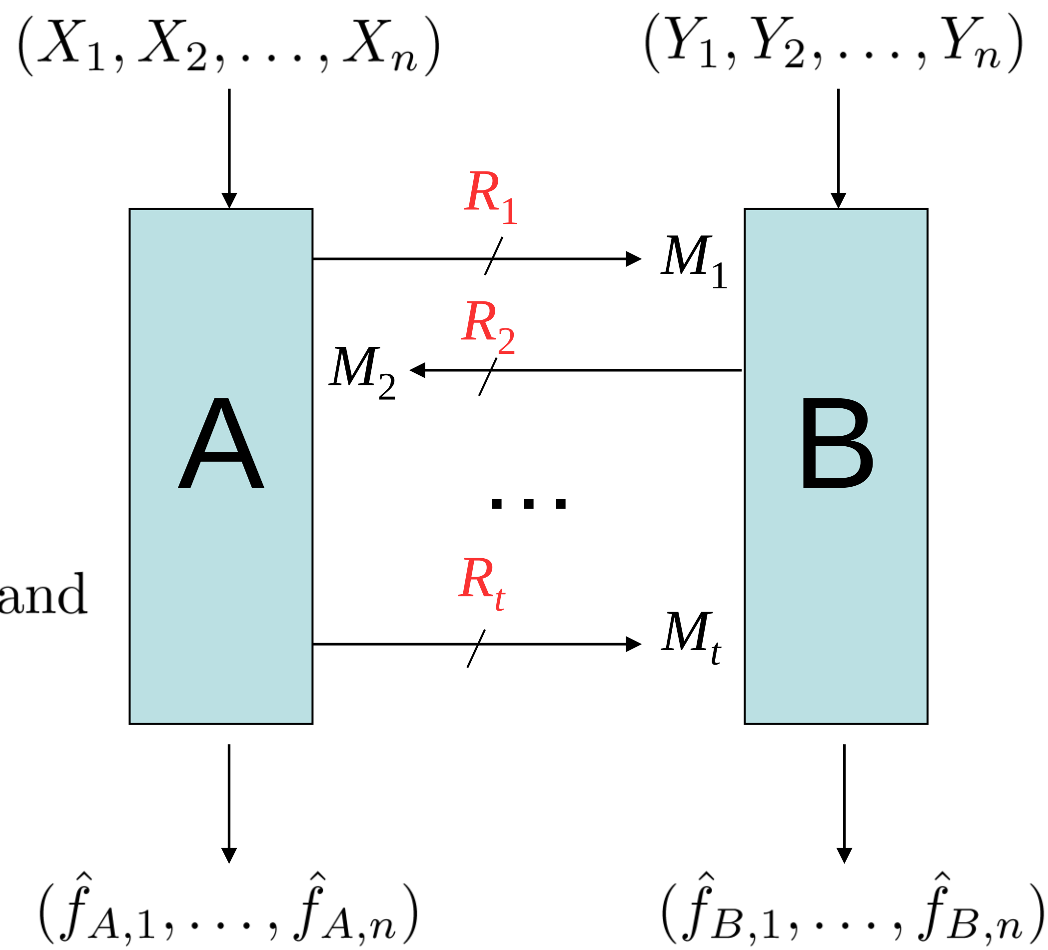


- 2-msg sum-rate ≤ $h(q) + q$

General 2-terminal interactive function computation

- n samples $(X_i, Y_i) \sim \text{iid } p_{XY}$
- Samplewise function computation at A and B
- t alternating messages
- $(R_1,\ldots,R_t)$ is admissible if there exists a sequence of codes: as $n \to \infty, \quad (\#\text{ bits msg } j)/n \to R_j$ and $\Pr[\text{comp. error}] \to 0$ (lossless, can extend to lossy)
- Minimum sum-rate:

$$R^A_{sum,t} = \min \sum_{i=1}^{t} R_t$$



need $(f_A(X_1,Y_1),\ldots,f_A(X_n,Y_n)), \quad (f_B(X_1,Y_1),\ldots,f_B(X_n,Y_n))$

Goal:

Characterize and compute

$$R_{sum,\infty} := \lim_{t \to \infty} R^A_{sum,t}$$

- Understand ultimate benefit of cooperative interaction
- "Unexplored" dimension for asymptotic analysis:
  - (possibly) **infinite messages with infinitesimal rate**
- "Untapped" resource: Multi-round interaction

Related work

- Communication complexity (Yao, Ahlswede, Cai, Orlitsky, Kushilevitz, Nisan, …, Braverman, A. Rao, …)
  - Usual focus on worst-case and comp. error $= 0$
  - Usually about bits not rate
- Two-way source coding [Kaspi IT'86]
  - Source reproduction
- Coding for computing [Orlitsky & Roche IT'00]
  - Two messages

$R_{sum,t}$ for finite t : Solved

Single-letter characterization [Ma, Ishwar: ISIT'08, IT'11]

$$R^A_{sum,t} = \min_{U^t}[I(X;U^t|Y) + I(Y;U^t|X)]$$

constraints

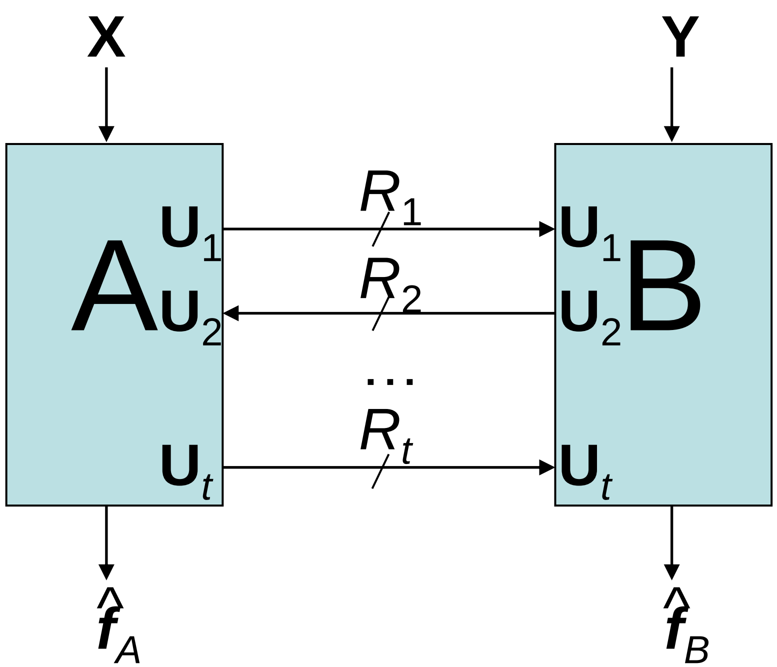
Markov chains	Entropy constraints	Cardinality bounds
$U_i - (X, U^{i-1}) - Y,\ i$ odd	$H(f_A(X,Y)\|X,U^t) = 0$	$\|\mathcal{U}_j\| \leq \begin{cases} \|\mathcal{X}\| \left(\prod_{i=1}^{j-1} \|\mathcal{U}_i\|\right) + t - j + 3, & j \text{ odd}, \\ \|\mathcal{Y}\| \left(\prod_{i=1}^{j-1} \|\mathcal{U}_i\|\right) + t - j + 3, & j \text{ even}. \end{cases}$
$U_i - (Y, U^{i-1}) - X,\ i$ even	$H(f_B(X,Y)\|Y,U^t) = 0$	

Achievability: (sequence of Wyner-Ziv codes)

- 1st msg: Quantizes **X** to $\mathbf{U}_1$ with side info **Y**
  $R_1 = I(X;U_1|Y)$, $U_1 - X - Y$
- 2nd msg: Quantizes ($\mathbf{Y}$,$\mathbf{U}_1$) to $\mathbf{U}_2$ with side info ($\mathbf{X}$,$\mathbf{U}_1$)
  $R_2 = I(Y;U_2|X,U_1)$, $U_2 - (Y,U_1) - X$

....

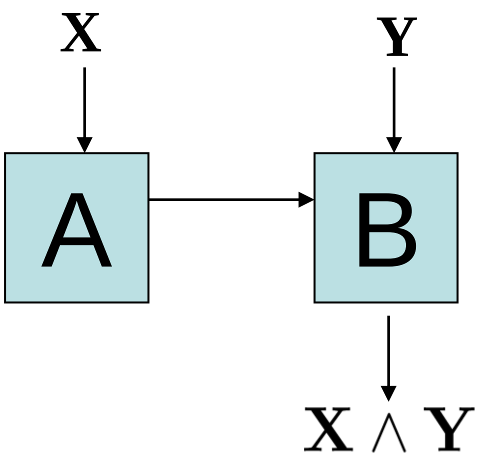
- Recover $\boldsymbol{f}_A$ based on ($\mathbf{X}$, $\mathbf{U}_1,\ldots,\mathbf{U}_t$): $H(f_A|X,U_1,\ldots,U_t)=0$
- Recover $\boldsymbol{f}_B$ based on ($\mathbf{Y}$, $\mathbf{U}_1,\ldots,\mathbf{U}_t$): $H(f_B|Y,U_1,\ldots,U_t)=0$

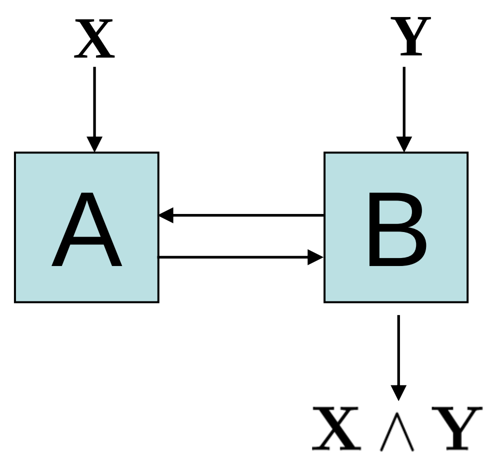


Example: compute AND at terminal B

- Sources: $X \perp\!\!\!\perp Y, X \sim \mathrm{Ber}(p), Y \sim \mathrm{Ber}(q)$, Compute: $X \wedge Y$ at B
- One message: min rate = ?

Two messages: min rate = ?





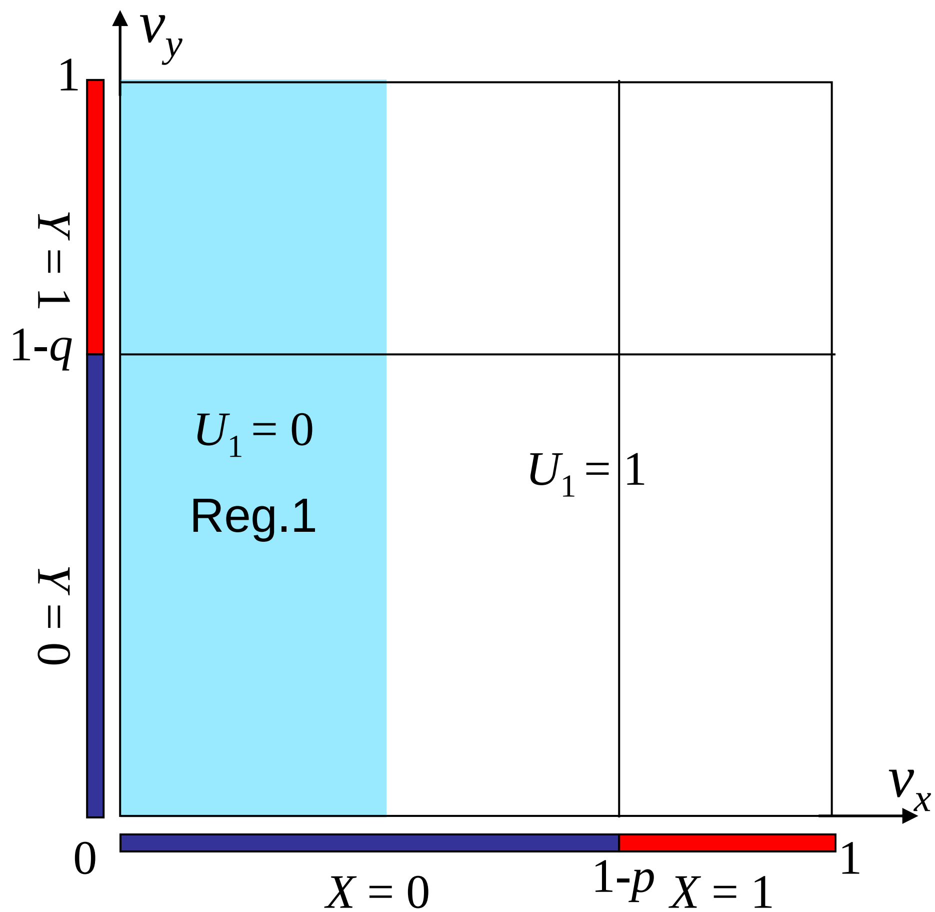
- 1-msg rate = $h(p)$

2-msg sum-rate ≤ $h(q) + q$

- 3-msgs.? 4-msgs.? …
- What about ∞-msgs.?

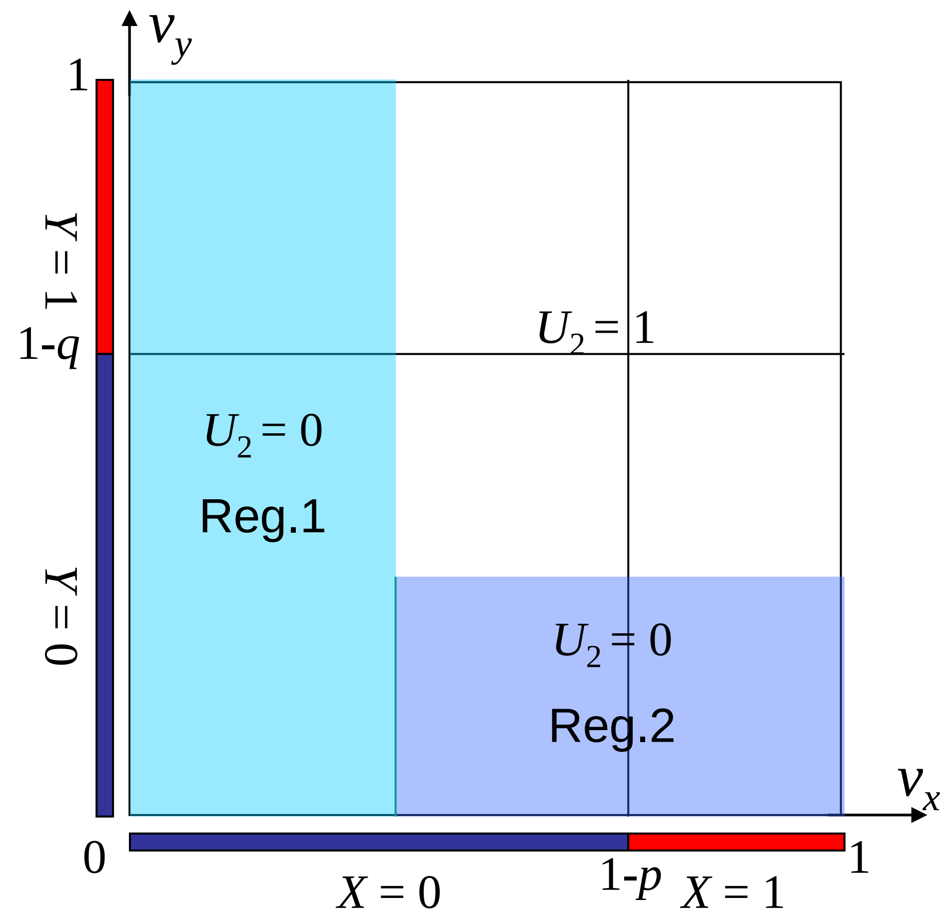
∞-msg interaction

- Auxiliary random variables
  - Hidden variables: (V_x, V_y) ~ Uniform(unit-sq.)
  - X, Y as functions of (V_x, V_y).
  - $U_1 = 0$ in Reg.1; $= 1$ otherwise



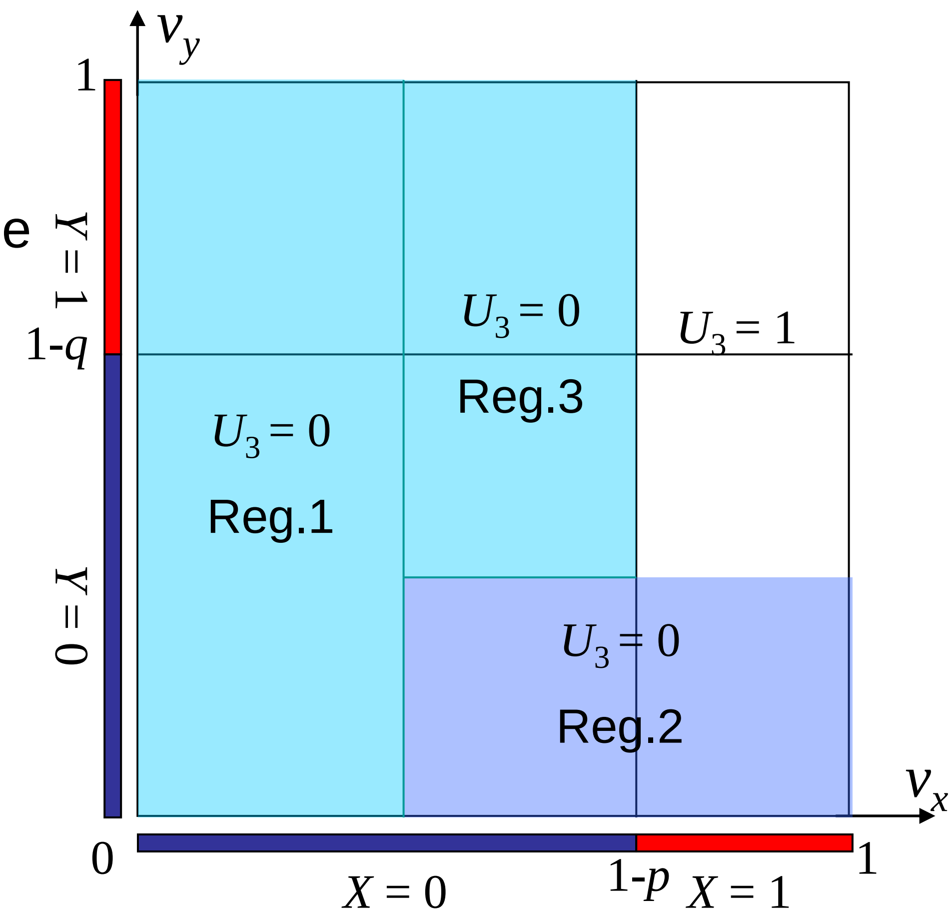
∞-msg interaction

- Auxiliary random variables
  - Hidden variables: $(V_x, V_y) \sim$ Uniform(unit-sq.)
  - X, Y as functions of (V_x, V_y).
  - $U_1 = 0$ in Reg.1; $= 1$ otherwise
  - $U_2 = 0$ in Reg.1 U 2; $= 1$ otherwise



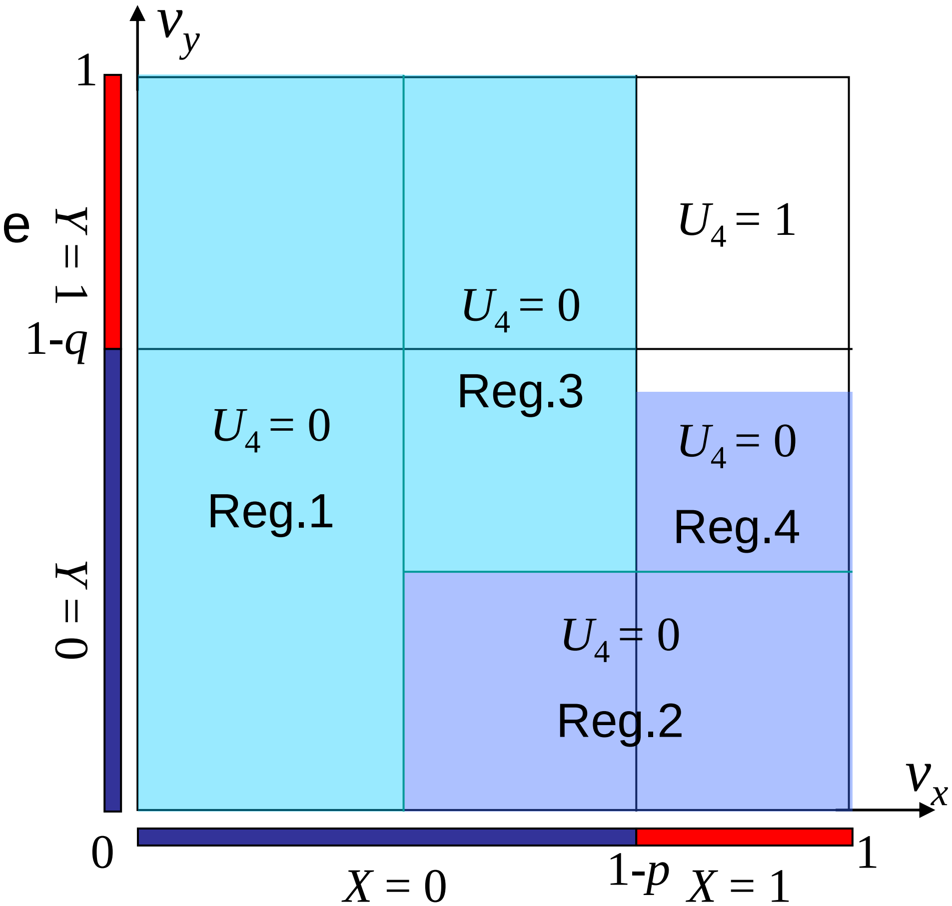
∞-msg interaction

- Auxiliary random variables
  - Hidden variables: $(V_x, V_y) \sim$ Uniform(unit-sq.)
  - X, Y as functions of (V_x, V_y).
  - $U_1 = 0$ in Reg.1; $= 1$ otherwise
  - $U_2 = 0$ in Reg.1 U 2; $= 1$ otherwise
  - $U_3 = 0$ in Reg.1 U 2 U 3; $= 1$ otherwise



∞-msg interaction

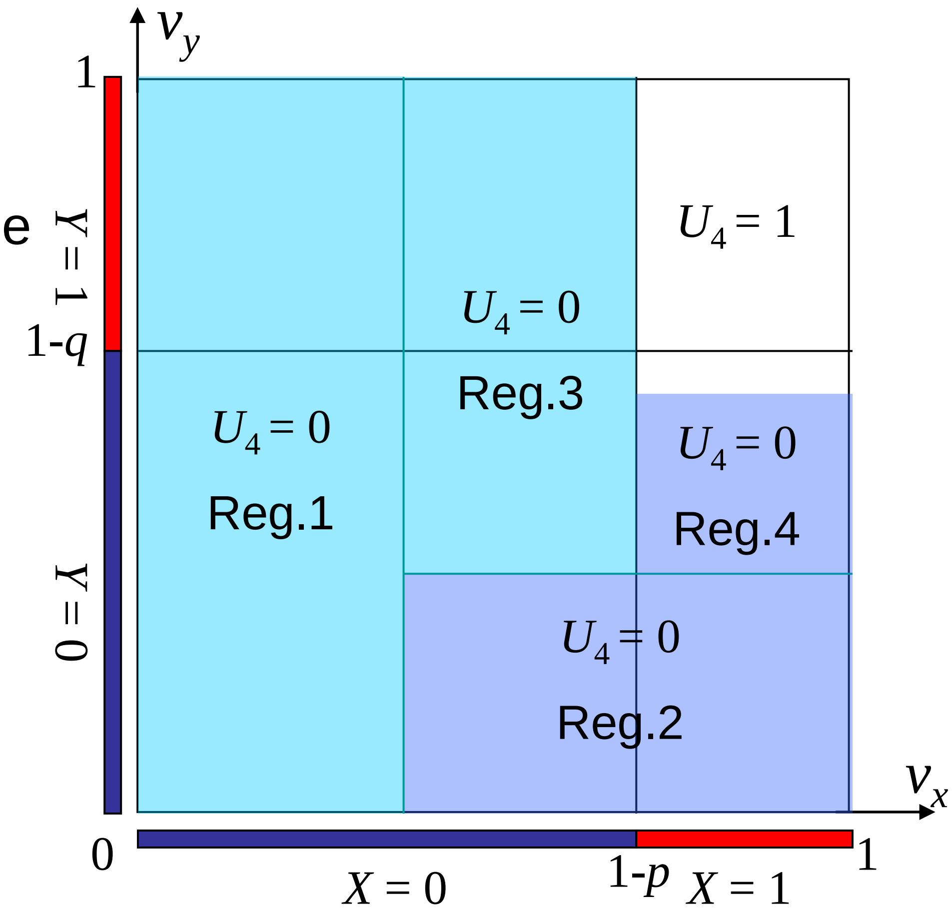
- Auxiliary random variables
  - Hidden variables: (V_x, V_y) ~ Uniform(unit-sq.)
  - X, Y as functions of (V_x, V_y).
  - $U_1 = 0$ in Reg.1; $= 1$ otherwise
  - $U_2 = 0$ in Reg.1 U 2; $= 1$ otherwise
  - $U_3 = 0$ in Reg.1 U 2 U 3; $= 1$ otherwise
  - $U_4 = 0$ in Reg.1 U 2 U 3 U 4; $= 1$ otherwise ($X \wedge Y = U_4 \wedge Y$)



∞-msg interaction

- Auxiliary random variables
  - Hidden variables: (V_x, V_y) ~ Uniform(unit-sq.)
  - X, Y as functions of (V_x, V_y).
  - $U_1 = 0$ in Reg.1; $= 1$ otherwise
  - $U_2 = 0$ in Reg.1 U 2; $= 1$ otherwise
  - $U_3 = 0$ in Reg.1 U 2 U 3; $= 1$ otherwise
  - $U_4 = 0$ in Reg.1 U 2 U 3 U 4;
    $= 1$ otherwise $(X \wedge Y = U_4 \wedge Y)$
  - Satisfy constraints

$$U_i - (X, U^{i-1}) - Y, \; i \text{ odd}$$
$$U_i - (Y, U^{i-1}) - X, \; i \text{ even}$$
$$H(f_A(X,Y)|X,U^t) = 0$$
$$H(f_B(X,Y)|Y,U^t) = 0$$

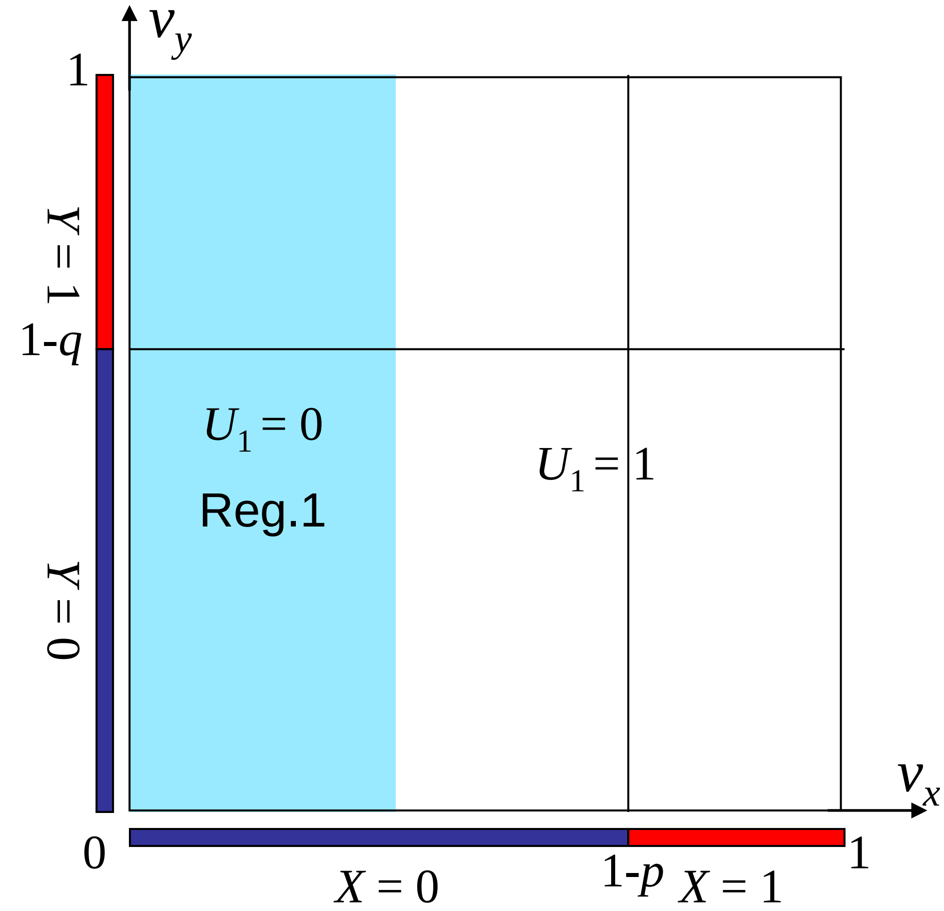


∞-msg interaction

- Admissible rates
  - 1st msg:

$$w_x(v_x, p) = \log \frac{1-x}{1-p-x}$$

$$I(X;U_1|Y) = \int_{\text{Reg.1}} w_x(v_x,p)\,dv_x dv_y$$



∞-msg interaction

- Admissible rates
  - 1st msg:

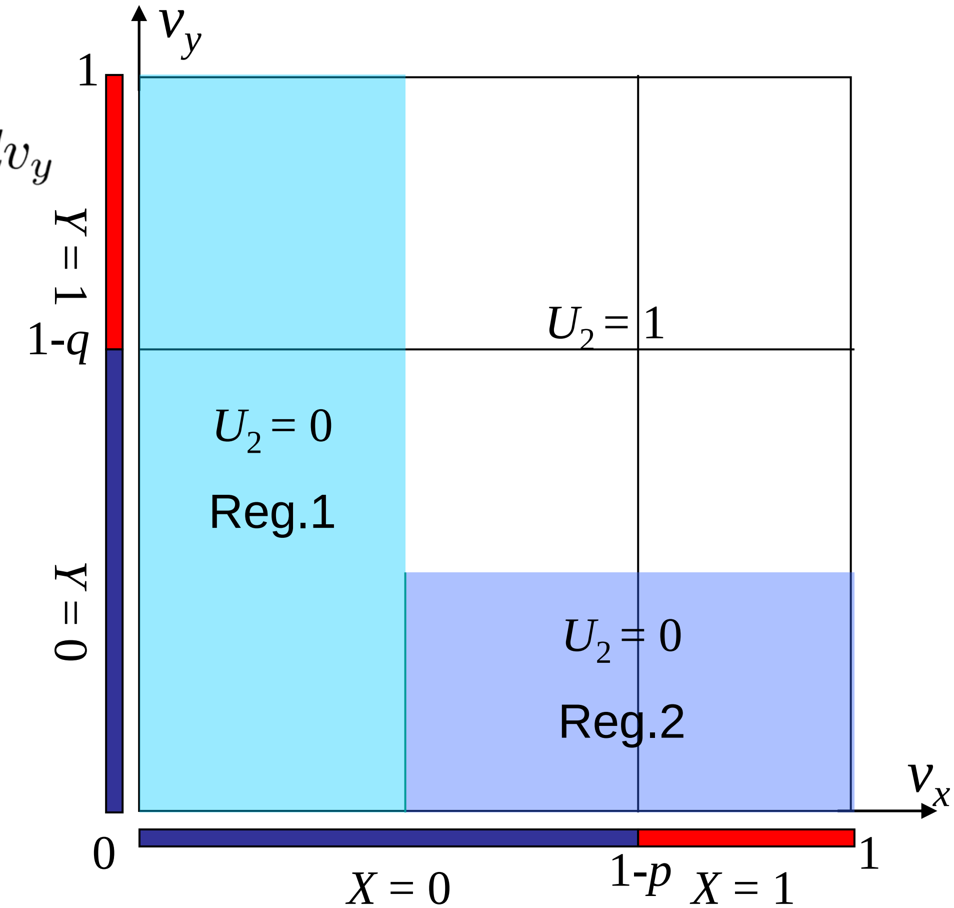
$$I(X;U_1|Y) = \int_{\text{Reg.1}} w_x(v_x,p)\,dv_x dv_y$$

  - 2nd msg:

$$I(Y;U_2|XU_1) = \int_{\text{Reg.2}} w_y(v_y,q)\,dv_x dv_y$$

$$w_x(v_x,p) = \log \frac{1-x}{1-p-x}$$

$$w_y(v_y,q) = \log \frac{1-y}{1-q-y}$$



∞-msg interaction

- Admissible rates
  - 1st msg:

$$I(X;U_1|Y) = \int_{\text{Reg.1}} w_x(v_x,p)\,dv_x dv_y$$

  - 2nd msg:

$$I(Y;U_2|XU_1) = \int_{\text{Reg.2}} w_y(v_y,q)\,dv_x dv_y$$

……...

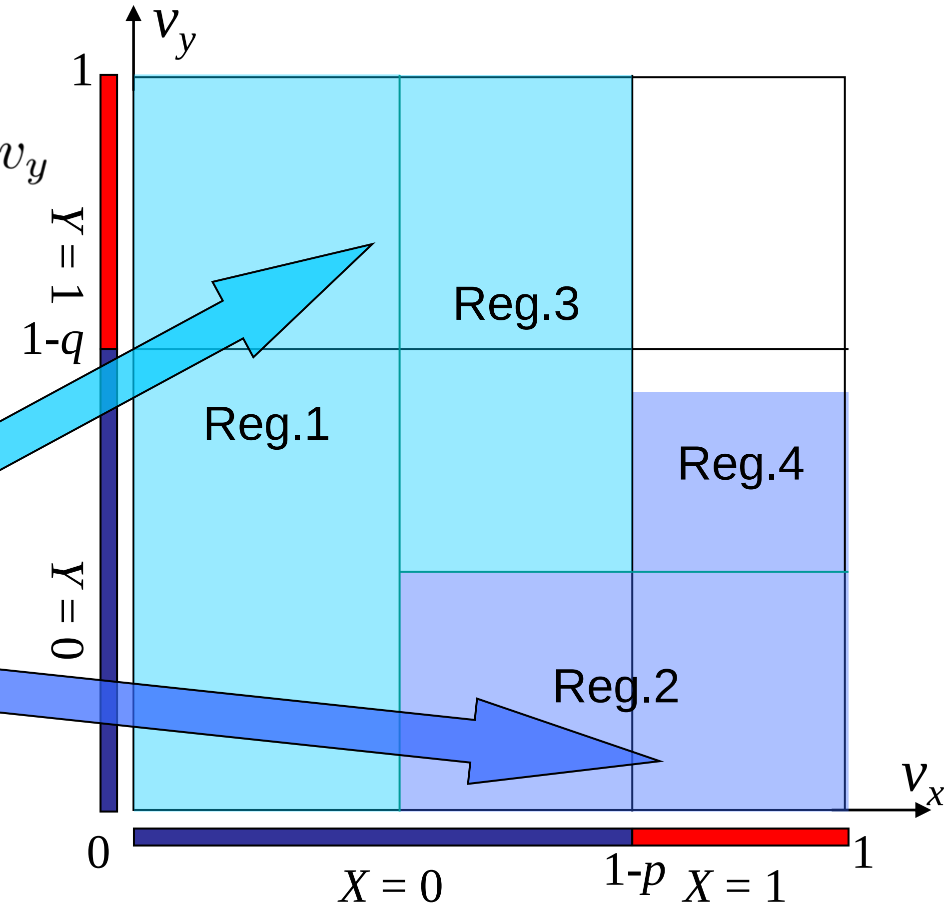
  - Sum-rate:

$$\int_{\text{vertical bars}} w_x(v_x,p)\,dv_x dv_y$$

$$+ \int_{\text{horizontal bars}} w_y(v_y,q)\,dv_x dv_y$$

$$w_x(v_x,p) = \log \frac{1-x}{1-p-x}$$

$$w_y(v_y,q) = \log \frac{1-y}{1-q-y}$$



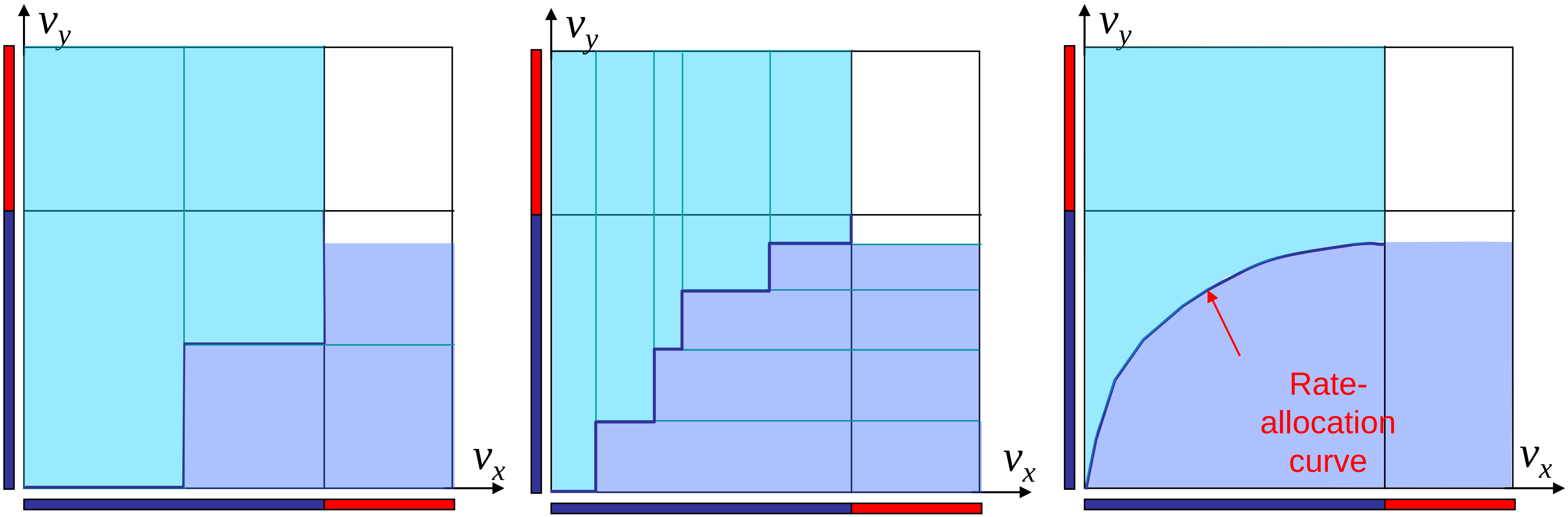
∞-msg interaction

- ***Admissible sum-rate:***

$$\int_{\text{vertical bars}} w_x(v_x, p)\,dv_x dv_y + \int_{\text{horizontal bars}} w_y(v_y, q)\,dv_x dv_y$$

- $t \to \infty$: $$\int_{\text{upper region}} w_x(v_x, p)\,dv_x dv_y + \int_{\text{lower region}} w_y(v_y, q)\,dv_x dv_y$$

- ***Rate-allocation curve:*** design parameter

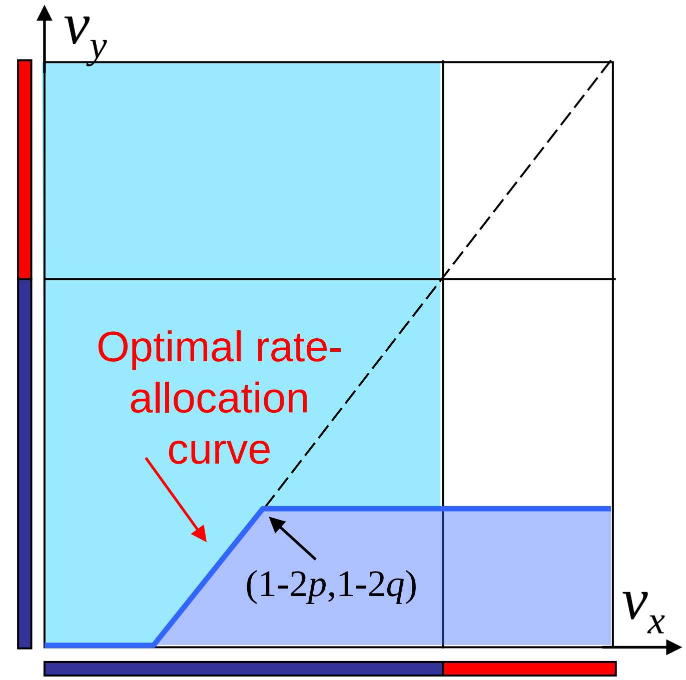


t = 4 msgs

t = 10 msgs

$t = \infty$ msgs

∞-msg interaction



- Optimize the rate-allocation curve:
- Cannot be achieved by finite msgs using this family of aux.r.v.'s (for general p, q)
- Admissible sum-rate in closed form:

$$R^*(p,q) = \begin{cases} h_2(p) + ph_2(q) + p\log_2 q + p(1-2q)\log_2 e, & \text{if } 0 \le p \le q \le 1/2, \\ R^*(q,p), & \text{if } 0 \le q \le p \le 1/2, \\ R^*(1-p,q), & \text{if } 0 \le q \le 1/2 \le p \le 1, \\ h_2(p), & \text{if } 1/2 \le q \le 1. \end{cases}$$

$R_{sum,t}$ for finite t : Solved

Single-letter characterization [Ma, Ishwar: ISIT'08, IT'11]

$$R^A_{sum,t} = \min_{U^t}[I(X;U^t|Y) + I(Y;U^t|X)]$$

constraints

Markov chains	Entropy constraints	Cardinality bounds
$U_i - (X, U^{i-1}) - Y,\ i$ odd $U_i - (Y, U^{i-1}) - X,\ i$ even	$H(f_A(X,Y)\|X,U^t) = 0$ $H(f_B(X,Y)\|Y,U^t) = 0$	$\|\mathcal{U}_j\| \leq \begin{cases} \|\mathcal{X}\| \left(\prod_{i=1}^{j-1} \|\mathcal{U}_i\|\right) + t - j + 3, & j \text{ odd}, \\ \|\mathcal{Y}\| \left(\prod_{i=1}^{j-1} \|\mathcal{U}_i\|\right) + t - j + 3, & j \text{ even}. \end{cases}$

- Achievability: sequence of Wyner-Ziv-like codes
- Finite dimensional optimization problem

How to compute $R_{sum,\infty}$?

- Idea 1:
  - Pick a large t, compute $R^A_{sum,t}$, pray this is a good approximation
    - ☺ Finite-dimensional optimization problem
    - ☹ How large t ?
    - ☹ Dimension explodes exponentially with t
- Idea 2:
  - Compute $R^A_{sum,t}$ for $t = 1, 2, \ldots$ till change is “negligible”
    - ☺ Finite-dimensional optimization problem (for each t)
    - ☹ Multiple problems, solve from scratch
    - ☹ Dimension explodes exponentially with t
- All this effort only for one p_{XY}
- Any hope?

A new approach

- View $R_{sum,\infty}(p_{XY})$ as a functional of p_{XY}

- New convex-geometric "limit-free" characterization of $R_{sum,\infty}$
  - For entire functional $R_{sum,\infty}(\bullet)$ (not for only one fixed p_{XY})
  - Provides optimality test for admissible sum-rate functionals
  - Family of lower bounds for $R_{sum,\infty}(\bullet)$

- Alternating "concavification" algorithm for $R_{sum,\infty}$
  - Each iteration uses "same amount" of computation
  - Reuses results from previous steps
  - Works with the entire $R_{sum,\infty}(p_{XY})$ "surface"

A new approach

Rest of this talk:

- Illustrate new approach through one simple example for 2-terminal lossless function computation

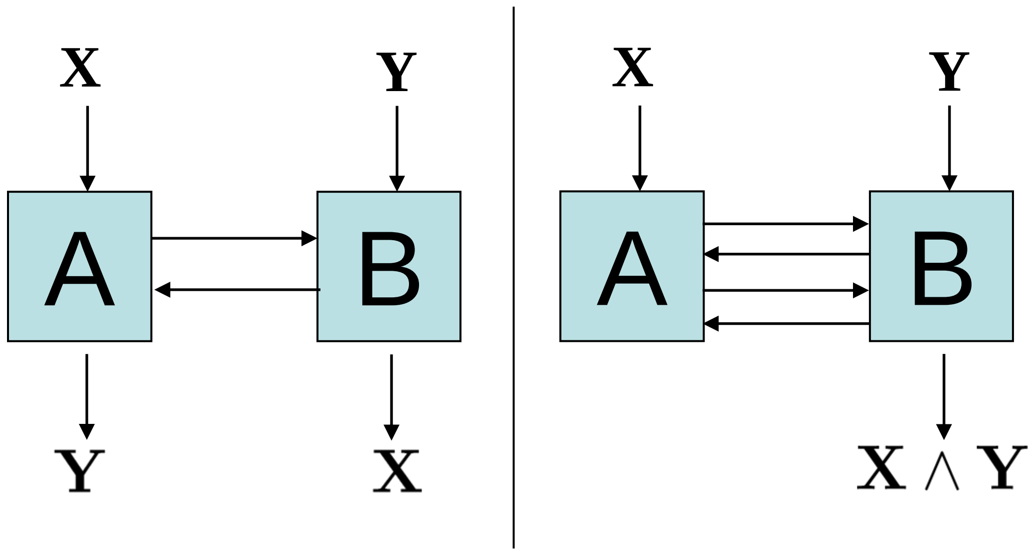
- Extension to general lossy two-terminal problem (brief)

- Extension to multi-terminal problems (very brief)

Example: compute AND at terminal B

- Sources: $X \perp\!\!\!\perp Y, X \sim \mathrm{Ber}(p), Y \sim \mathrm{Ber}(q)$
- Only B computes: $f_A(X,Y) = 0, \quad f_B(X,Y) = X \wedge Y$ (AND)
- Goal: Characterize $R_{sum,\infty}(p, q)$ as a function of (p, q)
- Rate reduction functional:

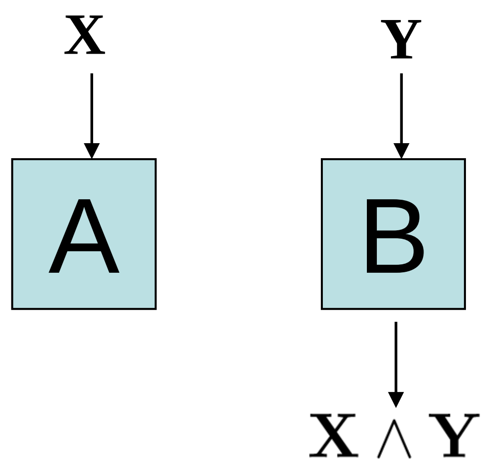
$$\rho_t^A := H(X|Y) + H(Y|X) \quad - \quad R_{sum,t}^A = h(p) + h(q) \quad - \quad R_{sum,t}^A$$



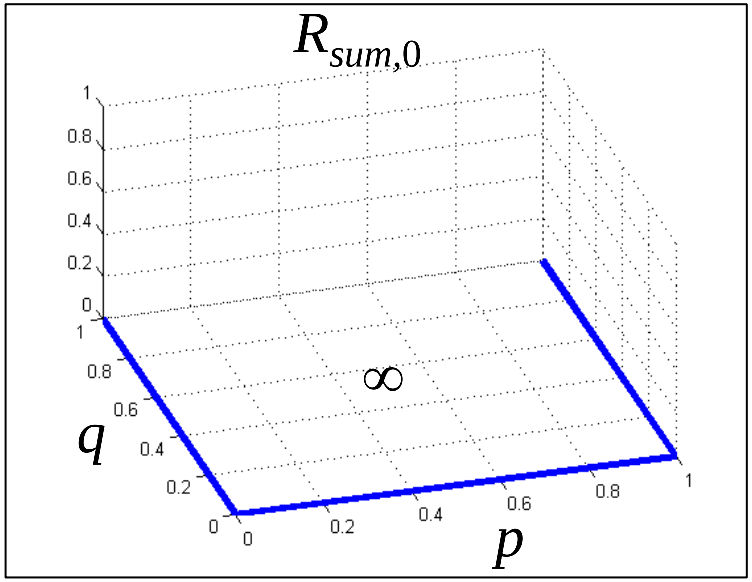
- Find $R_{sum,\infty} \iff$ find $\rho_\infty := H(X|Y) + H(Y|X) - R_{sum,\infty}$

Example: compute AND at terminal B

- Consider $t = 0$:
  - Feasible for special boundary distributions
  - If infeasible, define rate := ∞

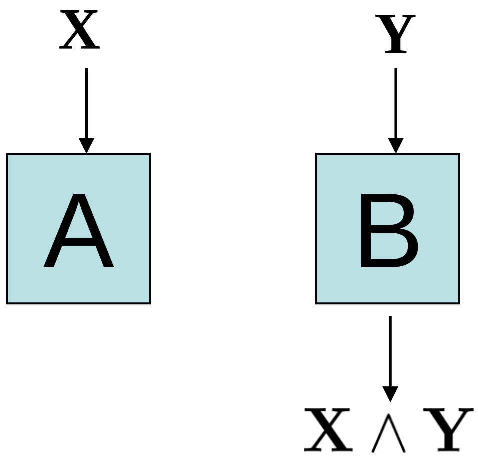


$$R_{sum,0}(p,q) = \begin{cases} 0, & \text{if } p = 0 \text{ or } q = 0 \quad (X \wedge Y = 0) \\ & \text{or } p = 1 \quad (X \wedge Y = Y) \\ \infty, & \text{otherwise} \quad (X \wedge Y \text{ not determined}) \end{cases}$$



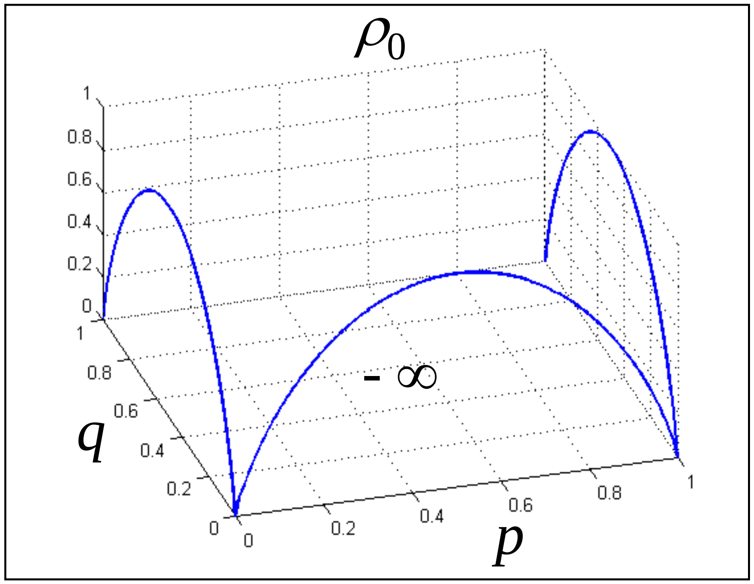
Example: compute AND at terminal B

- Consider $t = 0$:
  - Feasible for special boundary distributions
  - If infeasible, define rate := ∞



$$R_{sum,0}(p,q) = \begin{cases} 0, & \text{if } p = 0 \text{ or } q = 0 \quad (X \wedge Y = 0) \\ & \text{or } p = 1 \quad (X \wedge Y = Y) \\ \infty, & \text{otherwise} \quad (X \wedge Y \text{ not determined}) \end{cases}$$

$$\rho_0(p,q) = \begin{cases} h(p) + h(q), & \text{if } p = 0 \text{ or } q = 0 \\ & \text{or } p = 1 \\ -\infty, & \text{otherwise} \end{cases}$$



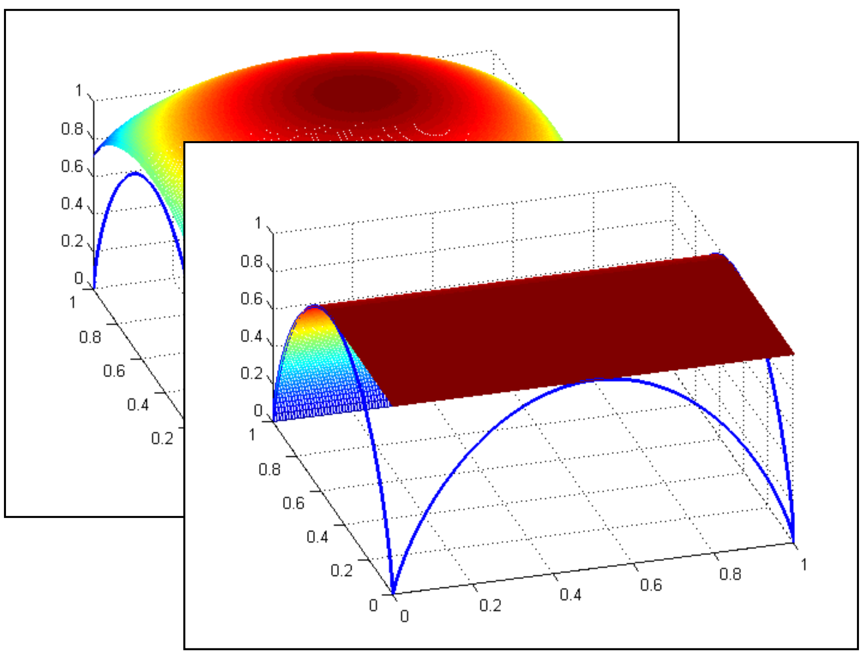
New characterization of ρ_∞

"Limit-free" characterization of ρ_∞ [Ma, Ishwar: Allerton 09]

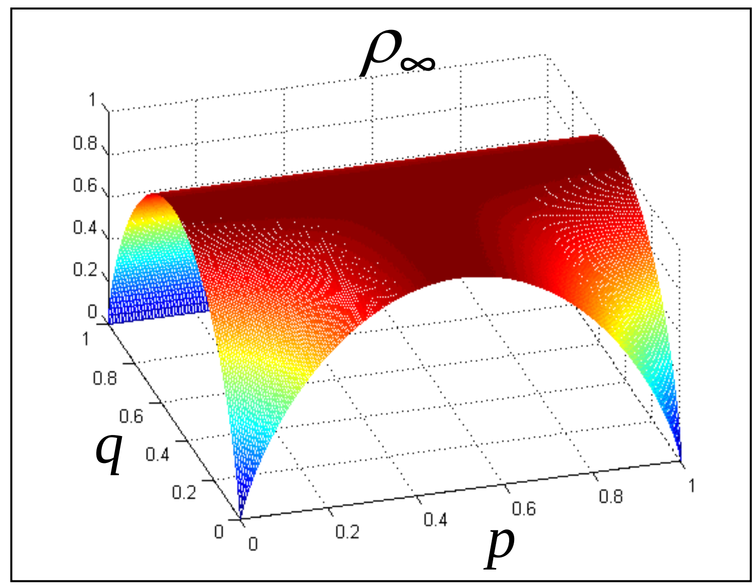
ρ_∞ is the least element of $\mathcal{F}$, where

$$\mathcal{F} := \left\{ \rho(p,q) \;\middle|\; \begin{array}{ll} 1. & \text{For all } (p,q),\ \rho(p,q) \geq \rho_0(p,q) \\ 2. & \text{For all } q,\ \rho(p,q) \text{ is concave w.r.t. } p \\ 3. & \text{For all } p,\ \rho(p,q) \text{ is concave w.r.t. } q \end{array} \right\}$$

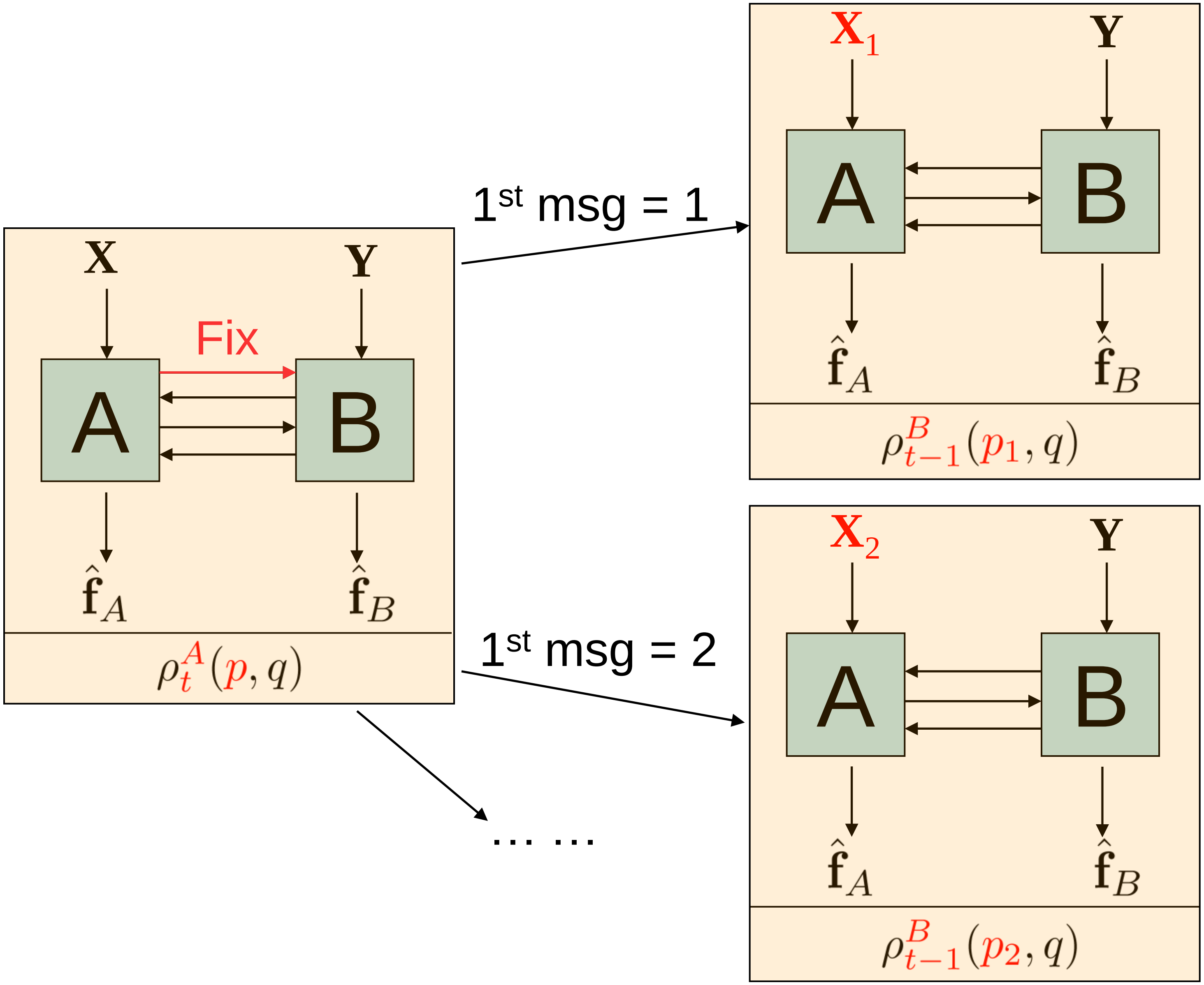
$\mathcal{F} = \{$



$\ldots \}$



Key insight: the subproblem viewpoint



Connection between ρ_t^A and ρ_{t-1}^B ?

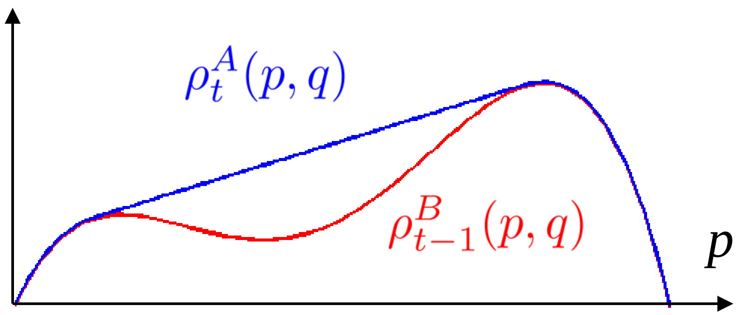
Subproblem viewpoint (continued)

$$R^A_{sum,t} = \min_{U^t}[I(X;U^t|Y) + I(Y;U^t|X)]$$

$$\Rightarrow \rho^A_t = \max_{U^t}[H(X|Y,U^t_2,U_1) + H(Y|X,U^t_2,U_1)]$$

$$= \max_{U_1}\left[\sum_{u_1} p_{U_1}(u_1)\rho^B_{t-1}(p_{u_1},q)\right]$$

- $\rho^A_t(p,q) = \max[$ convex combination of $\rho^B_{t-1}(p_1,q), \rho^B_{t-1}(p_2,q), \ldots]$

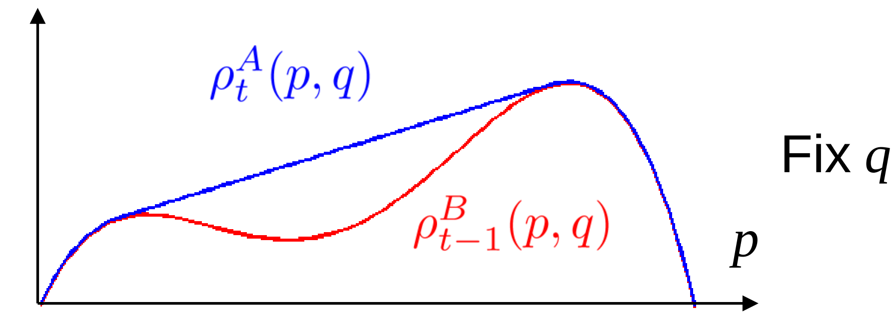


Fix q

"Concavify"

- $\rho^A_t =$ the smallest concave function above ρ^B_{t-1}
- $\mathrm{hypo}(\rho^A_t) =$ convex hull of $\mathrm{hypo}(\rho^B_{t-1})$

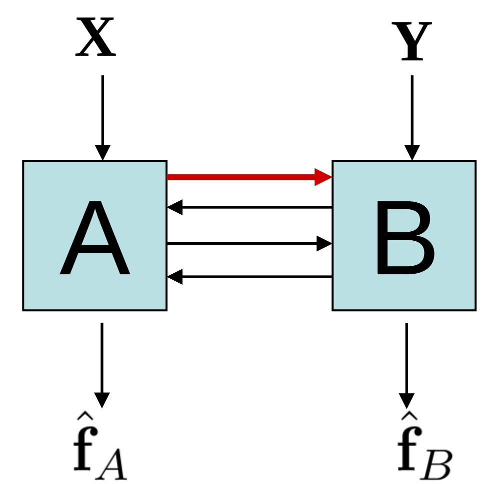
Concavity properties



- ρ_t^A concave in p, ρ_t^B concave in q

- ρ_{t-1}^B not concave in p $\Leftrightarrow$ $\rho_t^A > \rho_{t-1}^B$

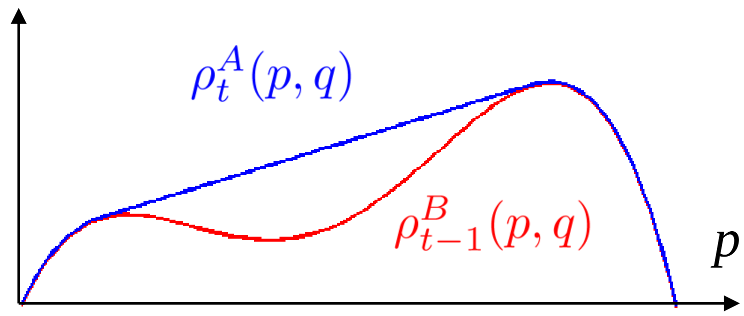
  $\Leftrightarrow$ beneficial to add a message A $\rightarrow$ B



- ρ_∞ : not beneficial to add any message

  $\Leftrightarrow$ concave in p and q, respectively

Optimality Test



- $\rho_t^B = \rho_\infty \Leftrightarrow$
- ρ_t^B is concave in p (for all q) $\Leftrightarrow$
- $\rho_t^B = \rho_{t+1}^A$ (for all p and q)
- Similar results by interchanging (A,p) and (B,q)
- Optimality Test: If ρ is an achievable rate-reduction function and:
  1. For all (p,q), $\rho(p,q) \geq \rho_0(p,q)$
  2. For all q, $\rho(p,q)$ is concave w.r.t. p
  3. For all p, $\rho(p,q)$ is concave w.r.t. q

  then $\rho = \rho_\infty$

Closed form expression of ρ_∞

- Optimality Test: If ρ is an achievable rate-reduction function and:
  1. For all (p, q), $\rho(p,q) \geq \rho_0(p,q)$
  2. For all q, $\rho(p,q)$ is concave w.r.t. p
  3. For all p, $\rho(p,q)$ is concave w.r.t. q

  then $\rho = \rho_\infty$

- Admissible sum-rate R^* (from beginning of talk):

$$R^*(p,q) = \begin{cases} h_2(p) + ph_2(q) + p\log_2 q + p(1-2q)\log_2 e, & \text{if } 0 \leq p \leq q \leq 1/2, \\ R^*(q,p), & \text{if } 0 \leq q \leq p \leq 1/2, \\ R^*(1-p,q), & \text{if } 0 \leq q \leq 1/2 \leq p \leq 1, \\ h_2(p), & \text{if } 1/2 \leq q \leq 1. \end{cases}$$

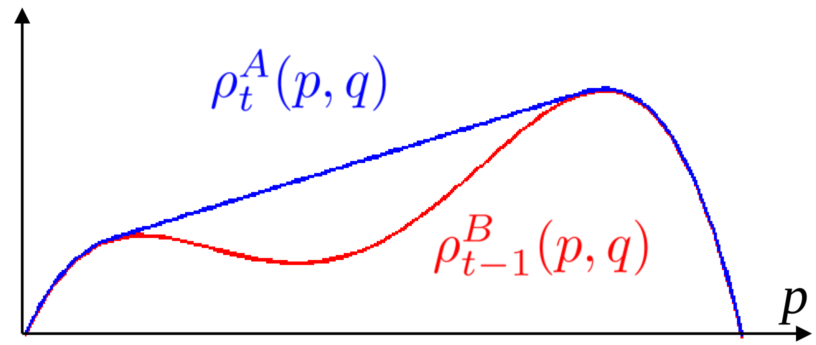
- Passes the optimality test!

- Therefore $R^* = R_{sum,\infty}$

Alternating concavification algorithm

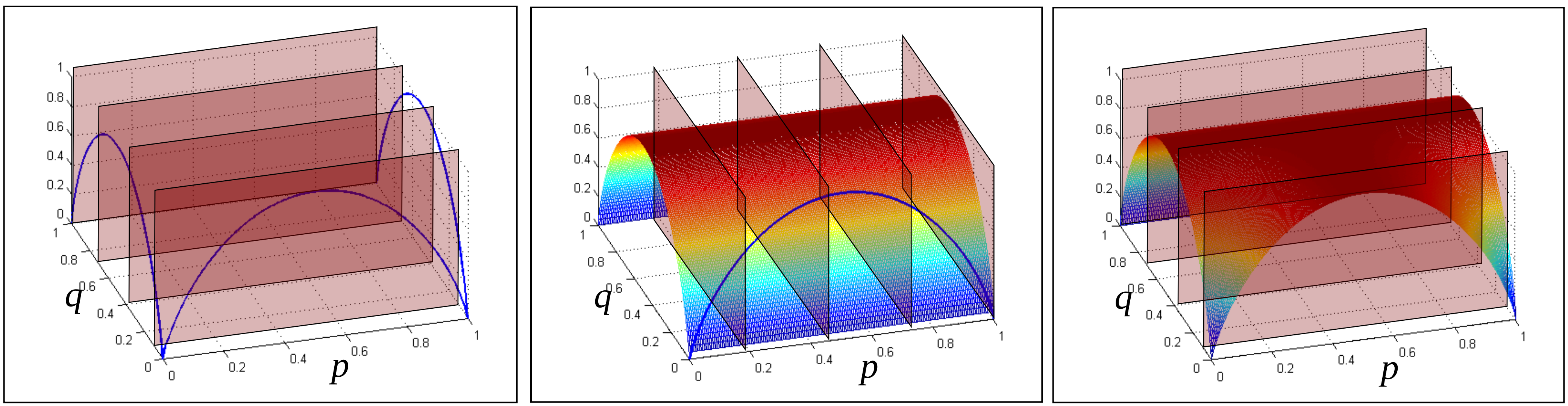
- Recall

$$\rho_{t-1}^{B}(p,q) \xrightarrow[\text{Fix } q]{\text{Concavify wrt } p} \rho_t^{A}(p,q)$$



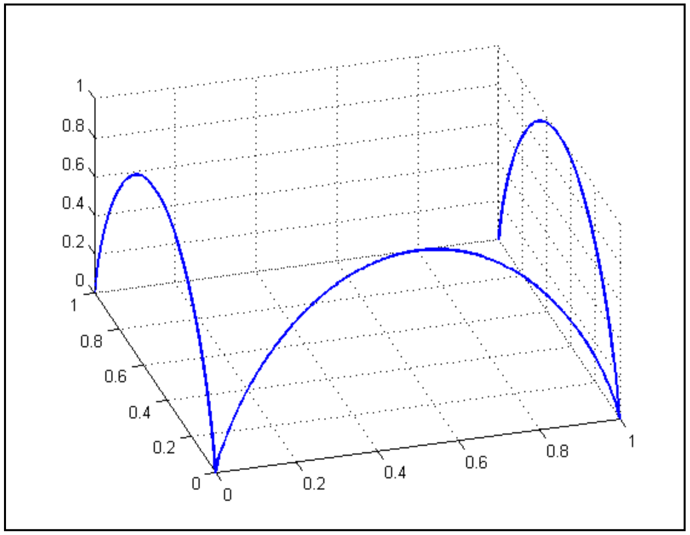
- Iterative algorithm

$$\rho_0(p,q) \xrightarrow[\text{Fix } q]{\text{Concavify wrt } p} \rho_1^{A}(p,q) \xrightarrow[\text{Fix } p]{\text{Concavify wrt } q} \rho_2^{B}(p,q) \longrightarrow \cdots$$

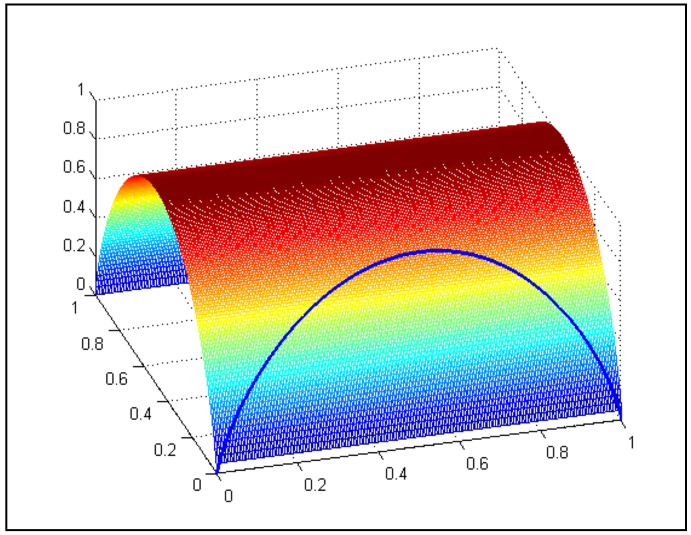


- Each iteration: “same amount” of computation
- Obtain the whole surface $\rho_t(p,q)$ for all (p,q)

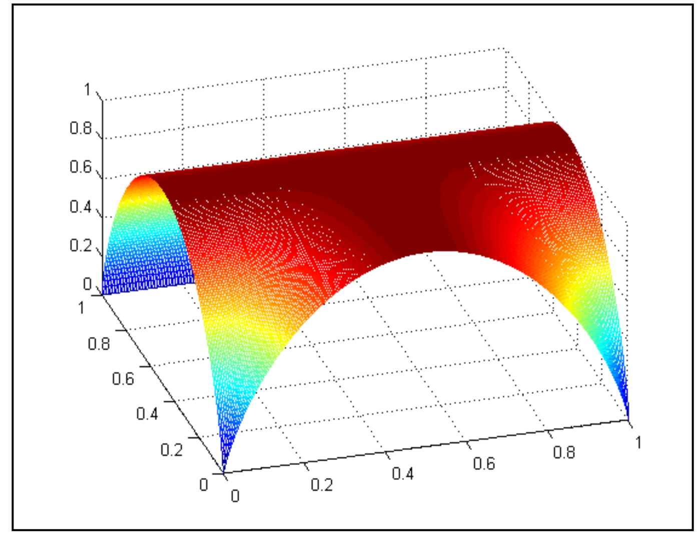
Alternating concavification algorithm



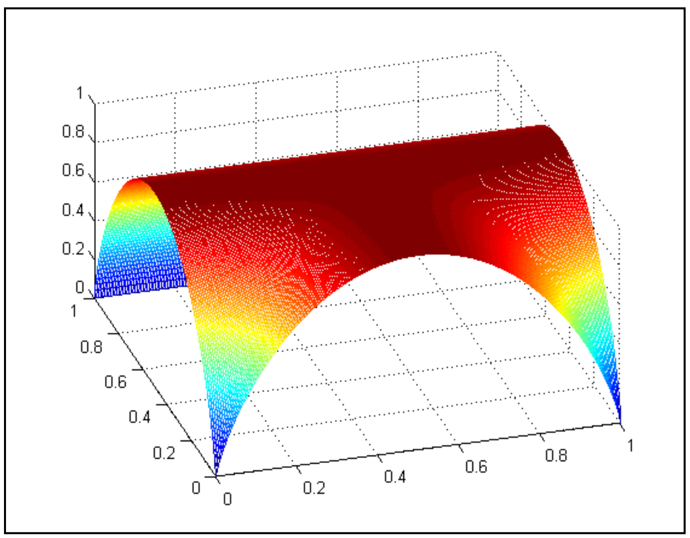
$t = 0$



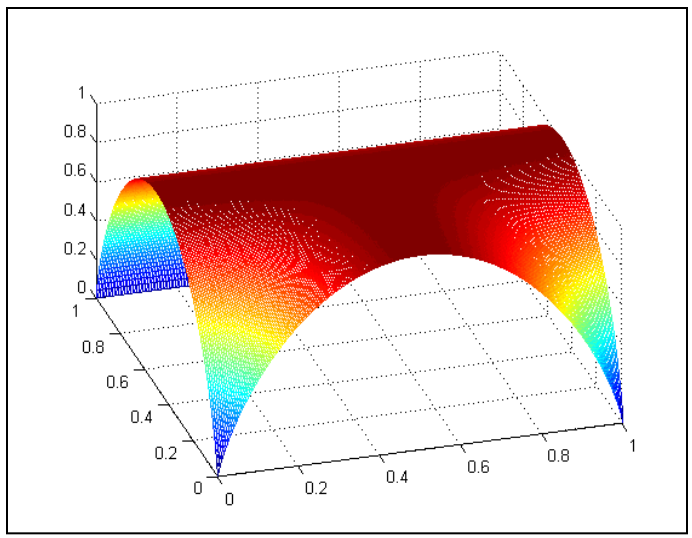
$t = 1$



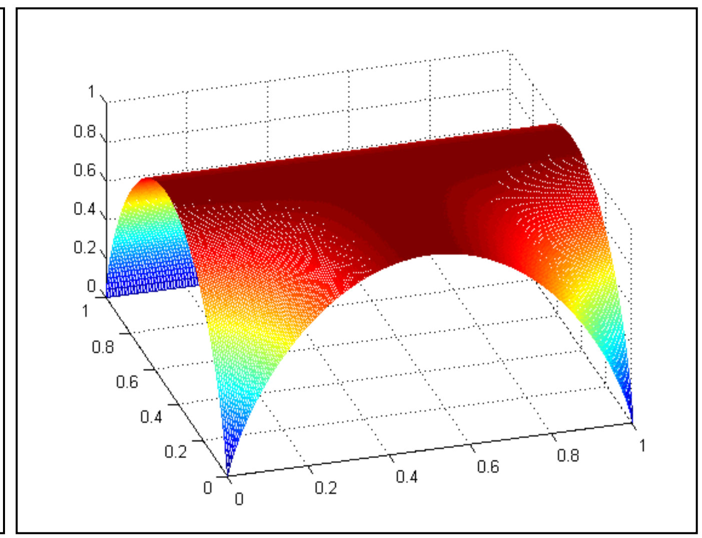
$t = 2$



$t = 3$

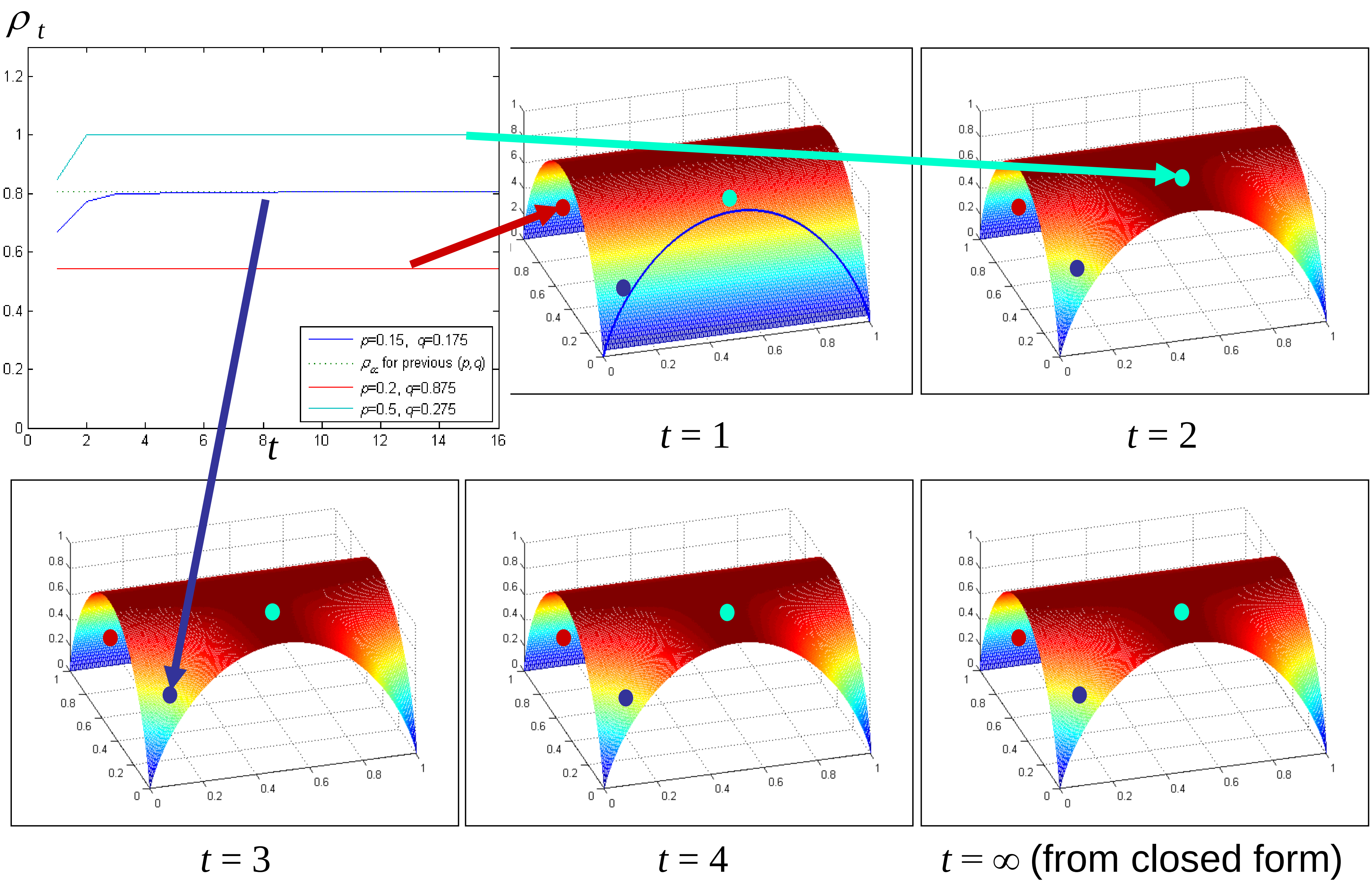


$t = 4$



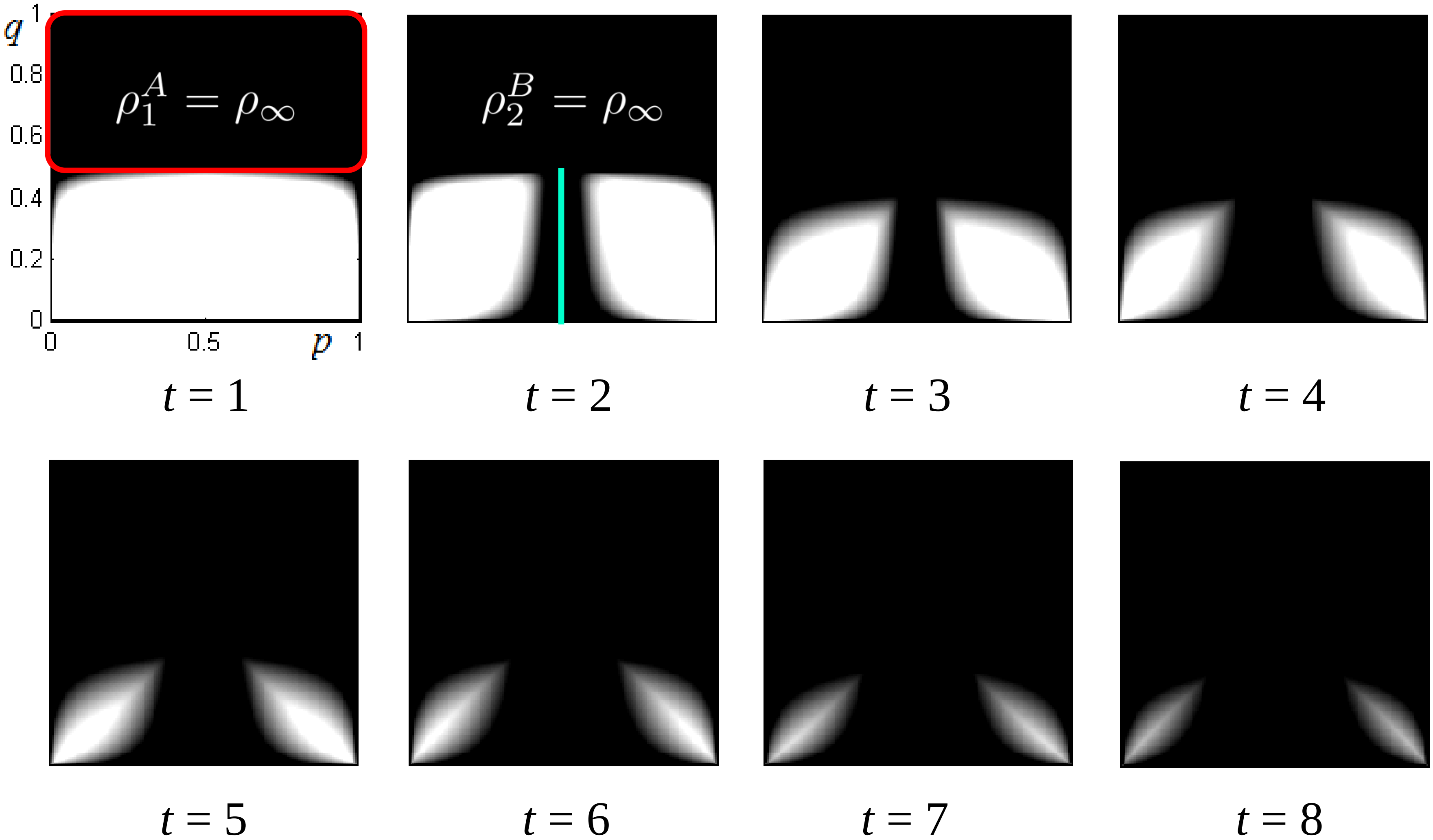
$t = \infty$ (from closed form)

How the surfaces evolve



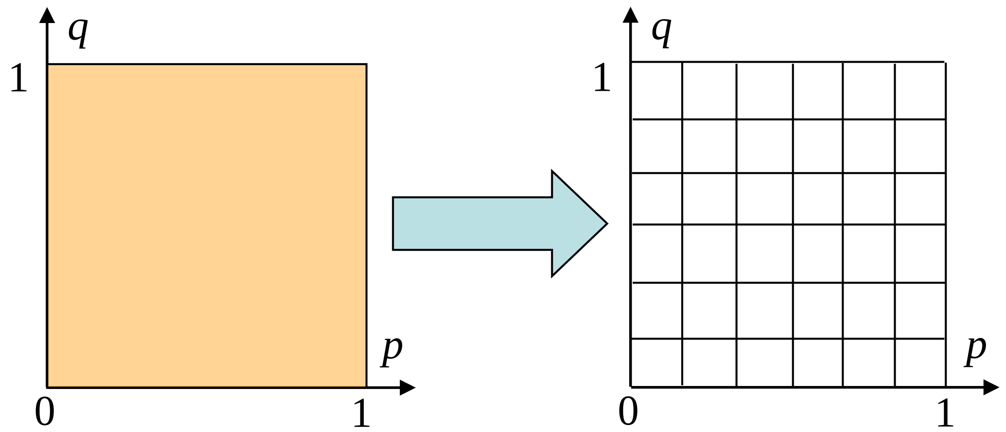
How the surfaces evolve

Brightness: $|\rho_t(p,q) - \rho_\infty(p,q)|$, black: $< 10^{-4}$

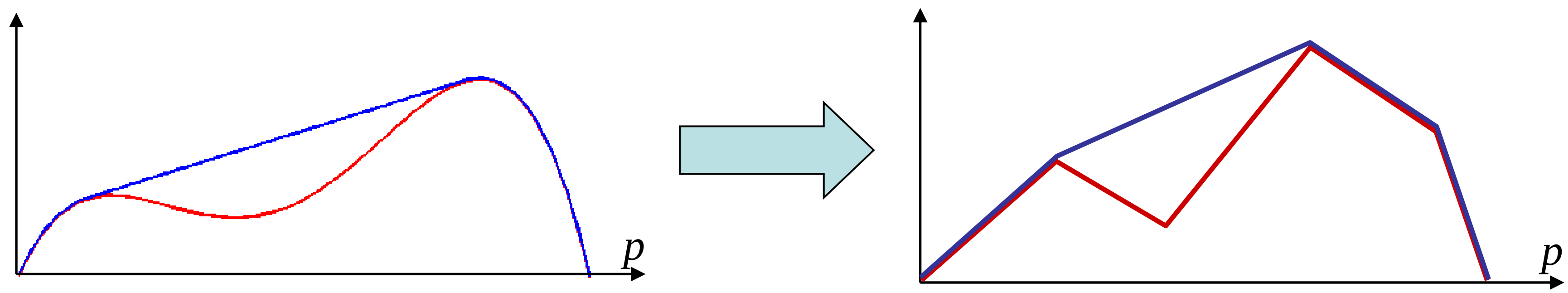


Discretization and convergence speed

- Numerical computation: discretize $[0,1]^2$ → N x N grid



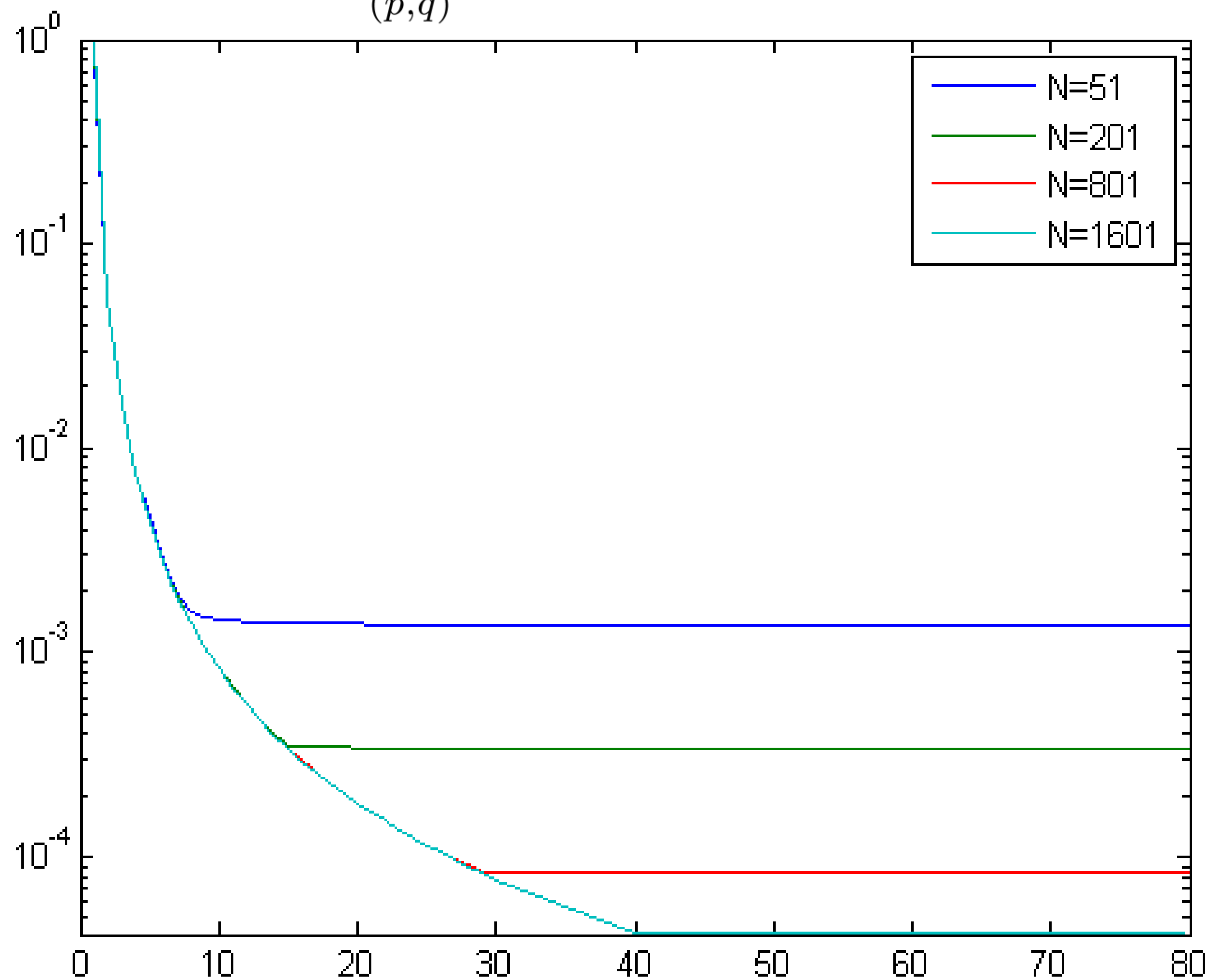
- Function "concavification" → convex hull of N points



Discretization and convergence speed

- Maximum error v.s. t v.s. N

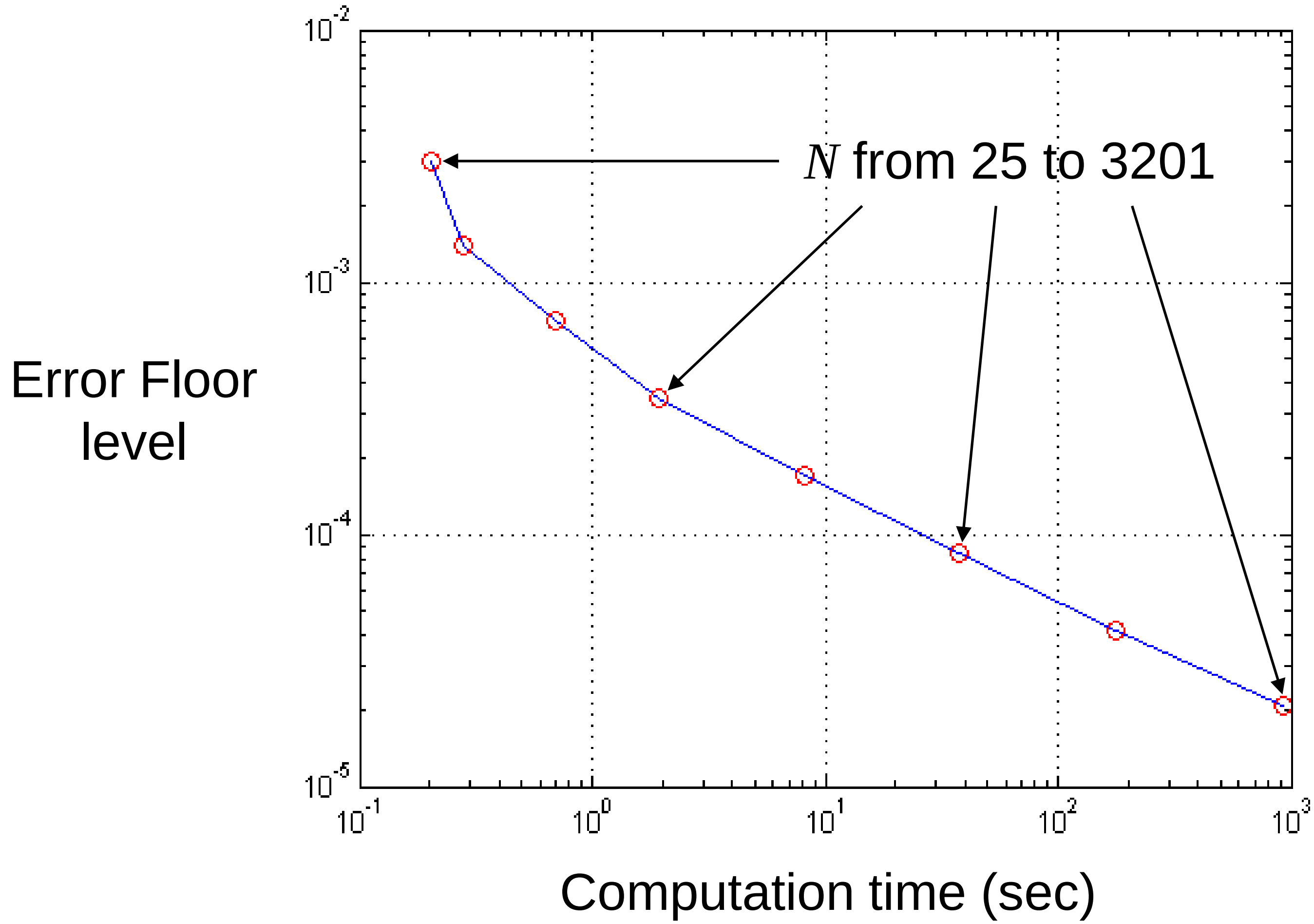
$$\max_{(p,q)} |\rho_t(p,q) - \rho_\infty(p,q)|$$



t : # iterations

Discretization and convergence speed

- Error floor level v.s. computation time
  - Experimentally, as N doubled, error floor halved, comp. time x4



General p_{XY} and f_A, f_B with distortions D_A, D_B

"Limit-free" characterization of $\rho_\infty(p_{XY}, D_A, D_B)$ [arXiv Oct'09]

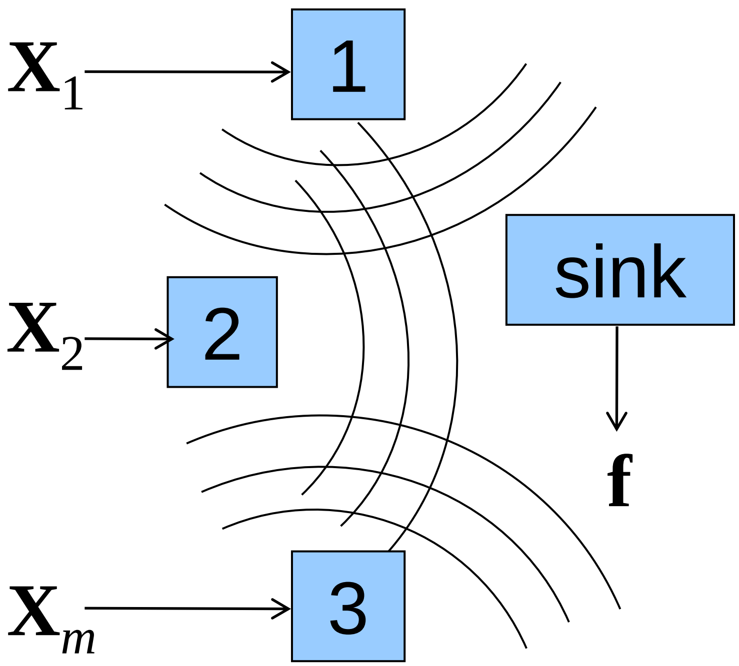
ρ_∞ is the least element of $\mathcal{F}$, where

$$\mathcal{F} := \left\{ \rho(p_{XY}, D_A, D_B) \;\middle|\; \begin{array}{l} 1.\ \rho \geq \rho_0 \\ 2.\ \text{For all } p_{Y|X},\ \rho(p_X p_{Y|X}, D_A, D_B) \text{ is concave w.r.t. } (p_X, D_A, D_B) \\ 3.\ \text{For all } p_{X|Y},\ \rho(p_Y p_{X|Y}, D_A, D_B) \text{ is concave w.r.t. } (p_Y, D_A, D_B) \end{array} \right\}$$

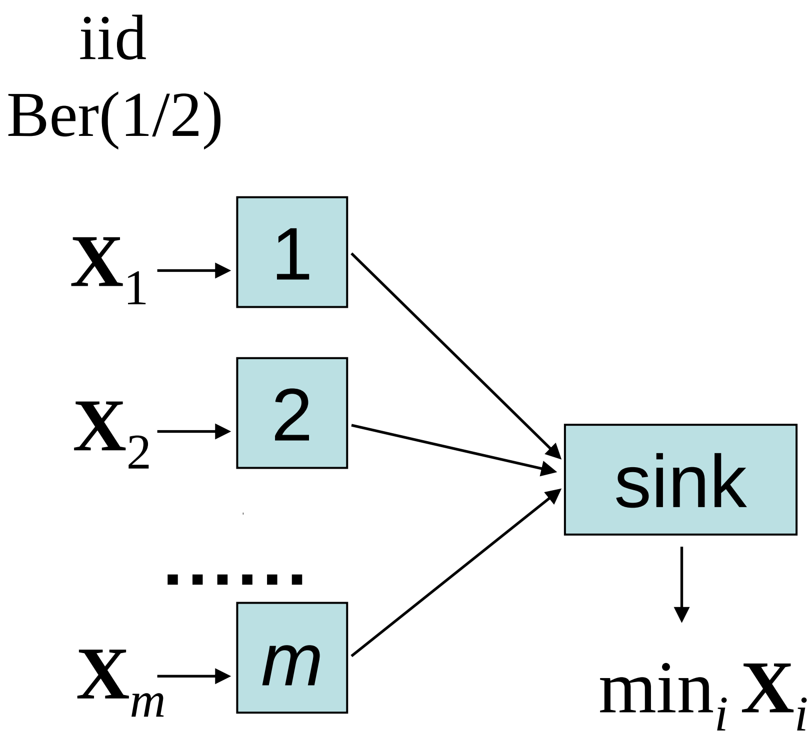
- 2 messages can strictly improve the Wyner-Ziv R-D function [ISIT'10] + next talk by Nan Ma
  - Resolves a question in Kaspi's 1985 two-way source coding paper

Collocated (broadcast) networks

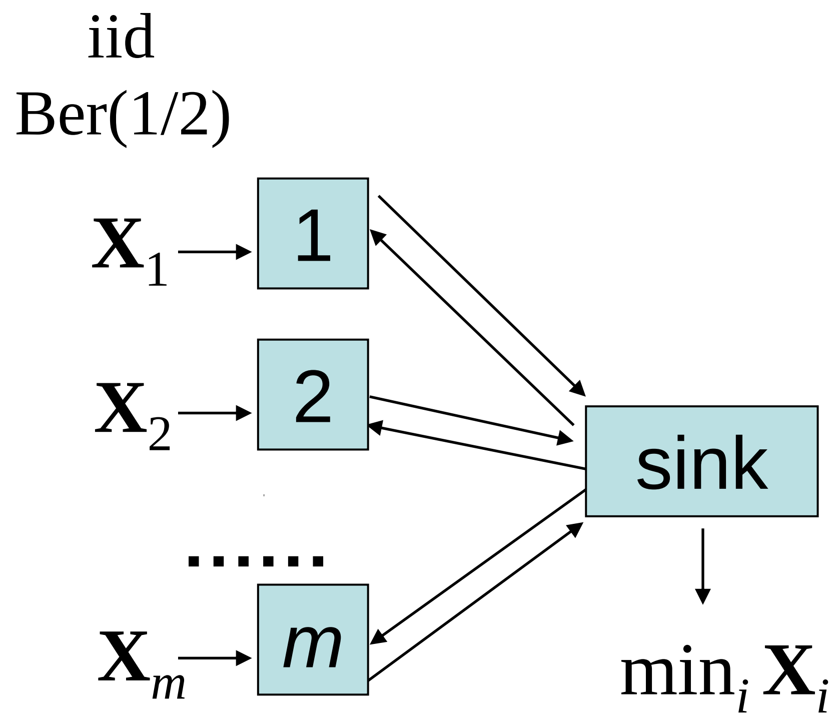
- General independent sources, general function:
  - Single-letter characterization for $R_{sum,t}$ [ISIT’09]
  - “Limit-free” convex-geometric characterization for $R_{sum,\infty}$ [ISIT’10]



Interaction changes the scaling law in star networks



$$R_{sum}(m) = m$$



$$1 \leq R_{sum}(m) < 6$$

Concluding remarks

- Characterizing the ultimate limits of interaction:
  - New type of functional single-letter characterization of $R_{sum,\infty}$
  - Alternating "concavification" algorithm for computing $R_{sum,\infty}$
- A new way to construct an infinite sequence of auxiliary random variables that is also optimal
- Infinite messages with infinitesimal rate => "Calculus" for source coding?

Open Problems

- Determine the computational complexity as a function of grid (discretization) size
- Characterize the rate of convergence
- Provide a simple procedure to directly determine for each given joint pmf, the smallest number of messages needed to reach the infinite-message limit
- Develop the functional characterization from first principles without using the single-letter characterization (functional equations and inequalities may be useful here)
- Find channel coding counterparts