Improving on the Cutset Bound via a Geometric Analysis of Typical Sets

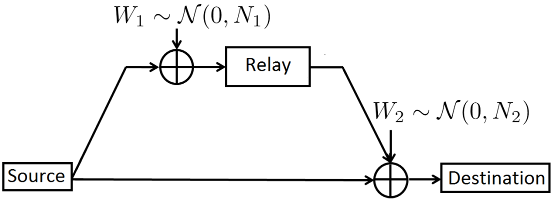
Ayfer Özgür

Stanford University

CUHK, May 28, 2016

Joint work with Xiugang Wu (Stanford).

Gaussian Relay Channel



$$Z = X + W_1 \qquad\qquad Y = X + X_r + W_2$$

Capacity is the largest achievable end-to-end reliable communication rate between the source and the destination.

Early work

One of the central problems in information theory:

- Introduced by van der Meulen in 1971.
- Seminal work by Cover and El Gamal in 1979.
  - Two achievable Schemes: Decode-Forward, Compress-Forward.
  - A general upper bound: Cutset Bound.

Over the Following 35 Years

- Major research effort over the last two decades however the problem remains open.

Over the Following 35 Years

- Major research effort over the last two decades however the problem remains open.
- Decode-Forward and Compress-Forward have been extended to multi-relay networks (Xie-Kumar'05,Kramer-Gastpar-Gupta'05) and unified (Wu-Xie'14).

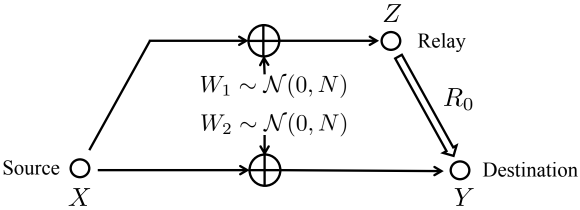
Over the Following 35 Years

- Major research effort over the last two decades however the problem remains open.
- Decode-Forward and Compress-Forward have been extended to multi-relay networks (Xie-Kumar'05,Kramer-Gastpar-Gupta'05) and unified (Wu-Xie'14).
- Many new relaying strategies have been discovered:
  - e.g., Amplify-Forward (Schein-Gallager'00), Hash-Forward (Kim'08), Compute-Forward (Nazer-Gastpar'11), Quantize-Map-Forward (Avestimehr-Diggavi-Tse'11), Noisy Network Coding (Lim-Kim-El Gamal-Chung'11).

Over the Following 35 Years

- Major research effort over the last two decades however the problem remains open.
- Decode-Forward and Compress-Forward have been extended to multi-relay networks (Xie-Kumar'05,Kramer-Gastpar-Gupta'05) and unified (Wu-Xie'14).
- Many new relaying strategies have been discovered:
  - e.g., Amplify-Forward (Schein-Gallager'00), Hash-Forward (Kim'08), Compute-Forward (Nazer-Gastpar'11), Quantize-Map-Forward (Avestimehr-Diggavi-Tse'11), Noisy Network Coding (Lim-Kim-El Gamal-Chung'11).
- Cutset bound remains as the only upper bound on capacity of the Gaussian Relay Channel.
  - Consistently used as a benchmark for performance.
  - It is not known if this bound is indeed achievable or not.

Gaussian Primitive Relay Channel



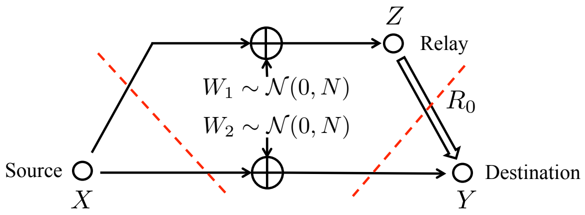
Cutset Bound: (Cover-El Gamal'79)

If R achievable, $\exists$ some X with $\mathbb{E}[X^2] \leq P$ such that:

$$R \leq I(X; Y, Z) \qquad \text{Broadcast Bound}$$

$$R \leq I(X; Y) + R_0 \qquad \text{Multiple Access Bound}$$

Gaussian Primitive Relay Channel



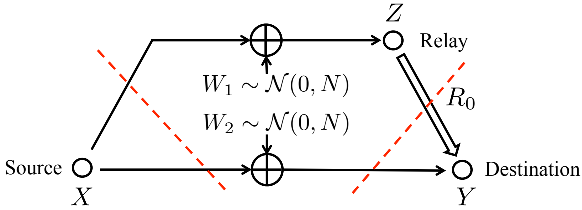
Cutset Bound: (Cover-El Gamal'79)

If R achievable, $\exists$ some X with $\mathbb{E}[X^2] \leq P$ such that:

$R \leq I(X; Y, Z)$ Broadcast Bound

$R \leq I(X; Y) + R_0$ Multiple Access Bound

Gaussian Primitive Relay Channel



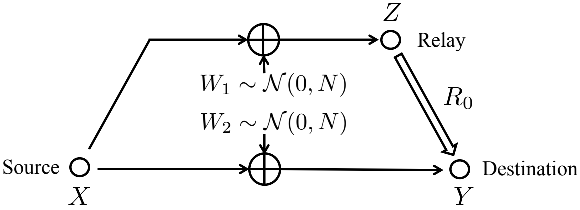
Cutset Bound: (Cover-El Gamal'79)

If R achievable, then

$$R \leq \frac{1}{2} \log \left(1 + \frac{2P}{N}\right)$$

$$R \leq \frac{1}{2} \log \left(1 + \frac{P}{N}\right) + R_0$$

Cutset Bound is not Tight

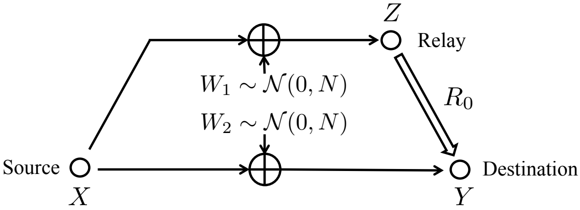


Theorem: (Wu-Ozgur'15)

If R achievable, then $\exists$ some X with $\mathbb{E}[X^2] \leq P$ and $a \geq 0$ such that

$R \leq I(X; Y, Z)$	Broadcast Bound
$R \leq I(X; Y) + R_0 - a$	Modified Multiple Access Bound
$R \leq I(X; Y) + a + \sqrt{2a \ln 2} \log e$	New constraint involving a

Cutset Bound is not Tight



Theorem: (Wu-Ozgur'15)

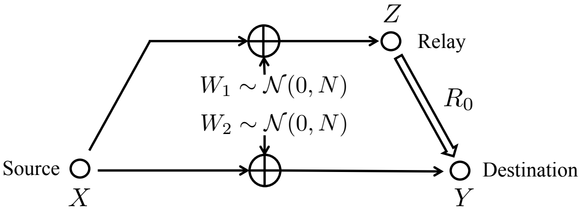
If R achievable, then $\exists$ some $a \geq 0$ such that

$$R \leq \frac{1}{2}\log\left(1+\frac{2P}{N}\right)$$

$$R \leq \frac{1}{2}\log\left(1+\frac{P}{N}\right) + R_0 - a$$

$$R \leq \frac{1}{2}\log\left(1+\frac{P}{N}\right) + a + \sqrt{2a\ln 2}\log e$$

Cutset Bound is not Tight



Theorem: (Wu-Ozgur'15)

If R achievable, then

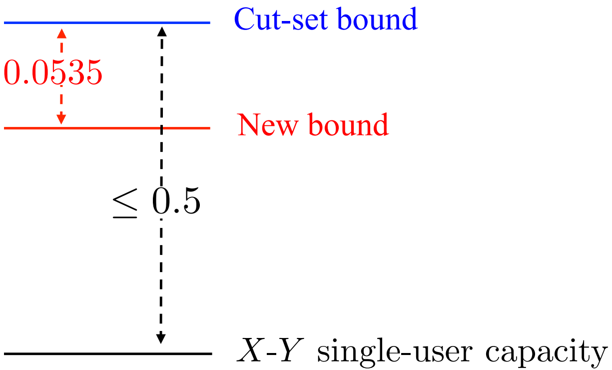
$$R \leq \frac{1}{2} \log \left(1 + \frac{2P}{N}\right)$$

$$R \leq \frac{1}{2} \log \left(1 + \frac{P}{N}\right) + R_0 - a^*$$

where a^* is such that $R_0 = 2a^* + \sqrt{2a^* \ln 2} \log e$

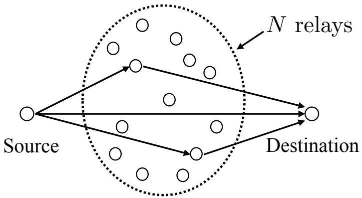
New bound is strictly tighter than the cutset bound.

Gap to the Cutset Bound



Largest gap among all $\left(\frac{P}{N}, R_0\right)$: 0.0535.

Capacity Approximation for Gaussian Relay Networks



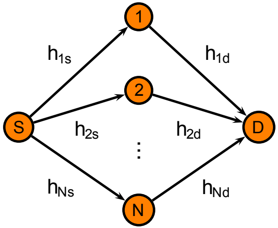
Theorem: (Avestimehr-Diggavi-Tse'11)

The capacity of any Gaussian relay network can be approximated with the cutset bound within a gap that is independent of the channel configurations and depends on the network topology only through the number of nodes.

Approximation gap:

Gap $\leq 7.5N$	(Avestimehr-Diggavi-Tse'11)
$\leq 1.5N$	(Ozgur-Diggavi'13)
$\leq 0.6N$	(Lim-Kim-El Gamal-Chung'13)
$\leq 0.5N$	(Lim-Kim-Kim'15)

The Gaussian N-Relay Diamond Network



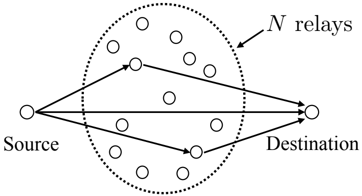
Chern and Ozgur'12

We can approximate the capacity of any Gaussian N-Relay Diamond network with its cutset bound within a

$$\text{Gap} \leq \log N$$

independent of the channel coefficients and the SNR's.

Sublinear gap to the cut-set bound?



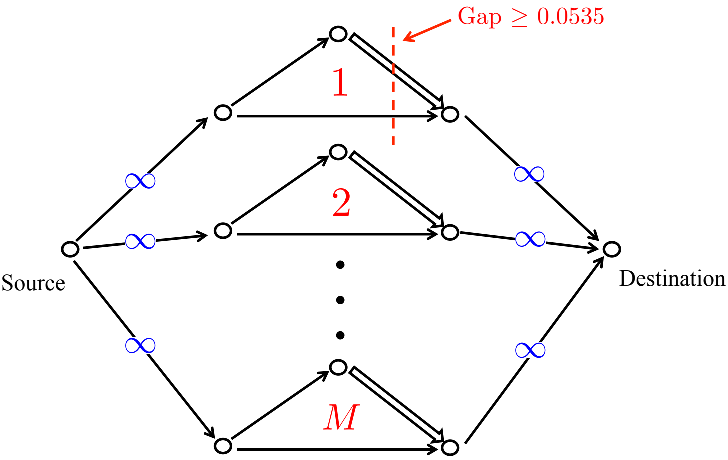
Theorem:(Courtade-Ozgur'15)

Sublinear gap is achievable iff cutset bound is tight for all Gaussian relay networks.

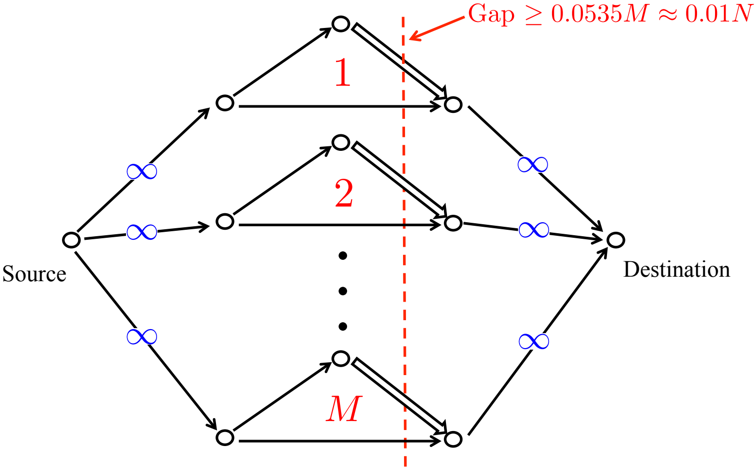
- Linear gap to the cutset bound is order-optimal:

$$0.5N \geq \text{Gap} \geq 0.01N.$$

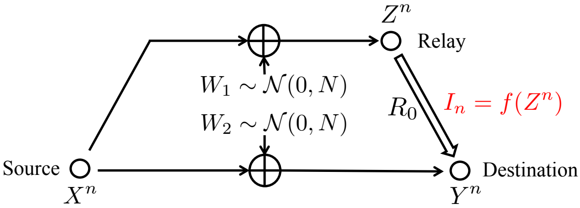
A Tensorization Argument



A Tensorization Argument



Derivation of the Cutset Bound



- Step 1) Apply Fano's inequality:

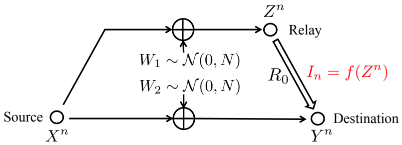
$$nR \leq I(X^n; Y^n, I_n) + n\epsilon_n$$

- Step 2) Bound with single-letter expressions.

Derivation of the Cutset Bound

- BC Constraint:

$$\begin{aligned} nR &\le I(X^n; Y^n, I_n) + n\epsilon_n \\ &\le I(X^n; Y^n, Z^n) + n\epsilon_n \\ &\le n(I(X; Y, Z) + \epsilon_n) \end{aligned}$$



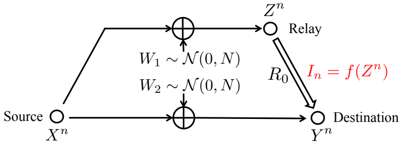
- MAC Constraint:

$$\begin{aligned} nR &\le I(X^n; Y^n, I_n) + n\epsilon_n \\ &\le I(X^n; Y^n) + I(X^n; I_n|Y^n) + n\epsilon_n \\ &\le I(X^n; Y^n) + \underbrace{H(I_n|Y^n)}_{\le nR_0} - \underbrace{H(I_n|Y^n, X^n)}_{\ge 0} + n\epsilon_n \\ &\le n(I(X; Y) + R_0 + \epsilon_n) \end{aligned}$$

Derivation of the New Bound

- BC Constraint:

$$nR \leq I(X^n; Y^n, I_n) + n\epsilon_n$$
$$\leq I(X^n; Y^n, Z^n) + n\epsilon_n$$
$$\leq n(I(X; Y, Z) + \epsilon_n)$$

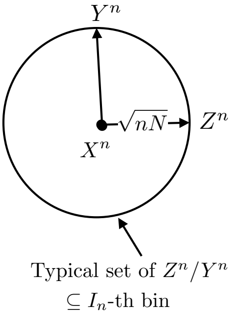


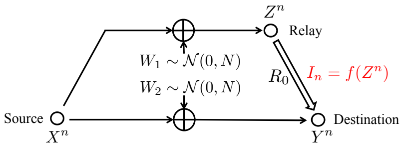
- MAC Constraint:

$$nR \leq I(X^n; Y^n, I_n) + n\epsilon_n$$
$$\leq I(X^n; Y^n) + I(X^n; I_n|Y^n) + n\epsilon_n$$
$$\leq I(X^n; Y^n) + \underbrace{H(I_n|Y^n)}_{\leq nR_0} - \underbrace{H(I_n|Y^n, X^n)}_{=H(I_n|X^n)=na} + n\epsilon_n$$
$$\leq n(I(X; Y) + R_0 - a + \epsilon_n)$$

Need a lower bound on a.

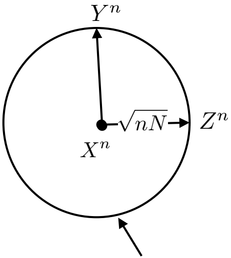
a cannot be arbitrarily small

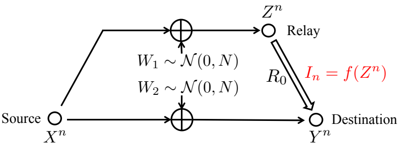




If $H(I_n|X^n) = na = 0$,

a cannot be arbitrarily small



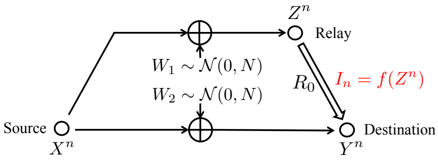


If $H(I_n|X^n) = na = 0$, then $H(I_n|Y^n) = 0$

$$R \leq I(X^n; Y^n) + \underbrace{H(I_n|Y^n)}_{=0} - \underbrace{H(I_n|X^n)}_{=0} + n\epsilon_n$$

$$\leq n(I(X;Y) + \epsilon_n)$$

In general

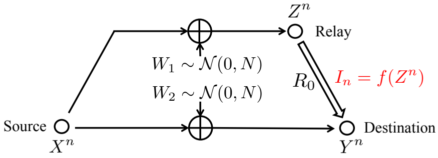


$I_n = f(Z^n) - Z^n - X^n - Y^n$, and Y^n and Z^n are i.i.d. given X^n.

$$R \leq I(X^n; Y^n) + \underbrace{H(I_n|Y^n)}_{\leq ?} - \underbrace{H(I_n|X^n)}_{= na} + n\epsilon_n$$

$$\leq I(X^n; Y^n) + \underbrace{I(I_n; X^n)}_{= n(R_0 - a)} - \underbrace{I(I_n; Y^n)}_{\geq ?} + n\epsilon_n$$

In general



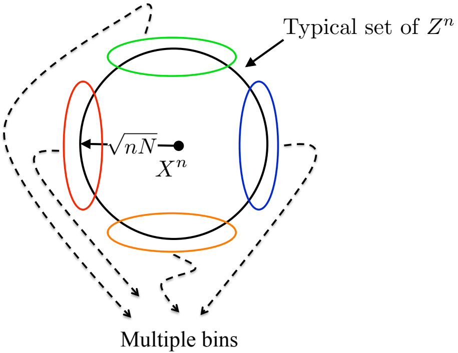
$I_n = f(Z^n) - Z^n - X^n - Y^n$, and Y^n and Z^n are i.i.d. given X^n.

$$R \le I(X^n;Y^n) + \underbrace{H(I_n|Y^n)}_{\le ?} - \underbrace{H(I_n|X^n)}_{=na} + n\epsilon_n$$

$$\le I(X^n;Y^n) + \underbrace{I(I_n;X^n)}_{=n(R_0-a)} - \underbrace{I(I_n;Y^n)}_{\ge ?} + n\epsilon_n$$

Strong Data Processing: Let $U - X - Y$. Fix $I(U;X)$ and bound $I(U;Y)$.

$H(I_n|X^n) = na$ and $a > 0$

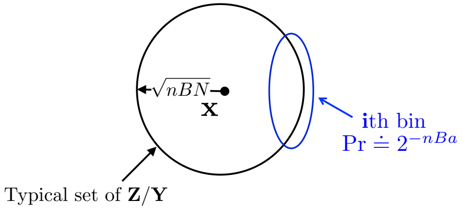


$\#$ of bins $=?$ $\quad$ $\mathbb{P}$(each bin) $=?$

From n- to nB- Dimensional Space

- B-length i.i.d. extension $\{(X^n(b), Y^n(b), Z^n(b), I_n(b))\}_{b=1}^{B}$.
- Signals in nB-dimensional space: $\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \mathbf{I}$.
- Law of Large Numbers:
  - n dimensional space: $H(I_n|X^n) = na$.
  - nB dimensional space:

$$\text{For any typical } (\mathbf{x}, \mathbf{i}), \qquad \mathbb{P}(\mathbf{Z} \in \mathbf{i}\text{'th bin}|\mathbf{x}) \doteq 2^{-nBa}.$$



Gaussian Measure Concentration

Theorem (Functional)

If $X \sim \mathcal{N}(0, I_k)$, *and* $Z = f(X)$ *such that* $f : \mathbb{R}^k \to \mathbb{R}$ *is L-Lipschitz, i.e.* $|f(x) - f(y)| \leq L||x - y||_2$, $\forall x, y$ *then*

$$\mathbb{P}(Z - \mathbb{E}[Z] > t) \leq e^{-t^2/(2L^2)}, \qquad \forall t > 0.$$

Gaussian Measure Concentration

Theorem (Functional)

If $X \sim \mathcal{N}(0, I_k)$*, and* $Z = f(X)$ *such that* $f : \mathbb{R}^k \to \mathbb{R}$ *is L-Lipschitz, i.e.* $|f(x) - f(y)| \le L||x - y||_2, \forall x, y$ *then*

$$\mathbb{P}(Z - \mathbb{E}[Z] > t) \le e^{-t^2/(2L^2)}, \qquad \forall t > 0.$$

Theorem (Geometric)

If $X \sim \mathcal{N}(0, I_k)$*, and let* $B_t = \{x \in R^k : \exists y \in B \, s.t. \, ||y - x||_2 \le t\}$ *for* $B \subseteq \mathbb{R}^k$*, then*

$$\mathbb{P}(X \notin B_t) \ge e^{-\frac{1}{2}\left(t - \sqrt{-2\log \mathbb{P}(B)}\right)^2}.$$

Gaussian Measure Concentration

Theorem (Functional)

If $X \sim \mathcal{N}(0, I_k)$*, and* $Z = f(X)$ *such that* $f : \mathbb{R}^k \to \mathbb{R}$ *is L-Lipschitz, i.e.* $|f(x) - f(y)| \leq L||x - y||_2, \forall x, y$ *then*

$$\mathbb{P}(Z - \mathbb{E}[Z] > t) \leq e^{-t^2/(2L^2)}, \qquad \forall t > 0.$$

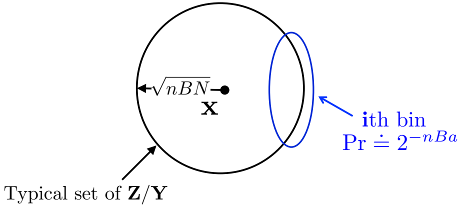
Theorem (Geometric)

If $X \sim \mathcal{N}(0, I_k)$*, and let* $B_t = \{x \in R^k : \exists y \in B \, s.t. \, ||y - x||_2 \leq t\}$ *for* $B \subseteq \mathbb{R}^k$*, then*

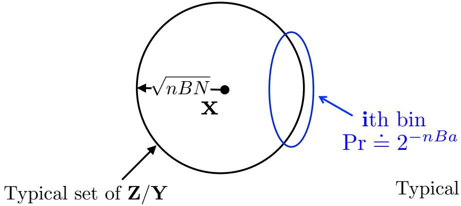
$$\mathbb{P}(X \notin B_t) \geq e^{-\frac{1}{2}\left(t - \sqrt{-2 \log \mathbb{P}(B)}\right)^2}.$$

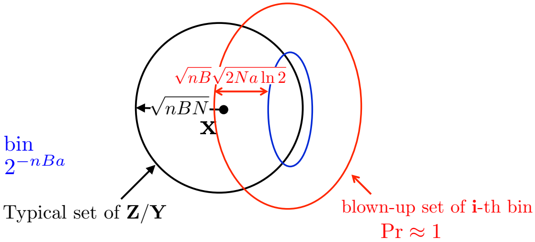
If $P(B) = e^{-kb}$, blow up by a radius $t = \sqrt{k}(\sqrt{2b} + \epsilon)$, then $P(X \notin B_t) \leq e^{-k\epsilon^2}$.

Gaussian Measure Concentration

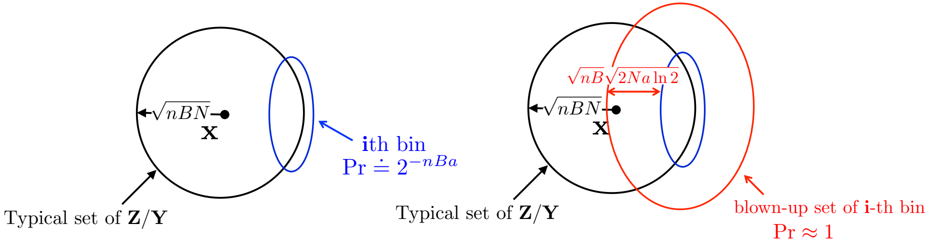


Gaussian Measure Concentration





Gaussian Measure Concentration



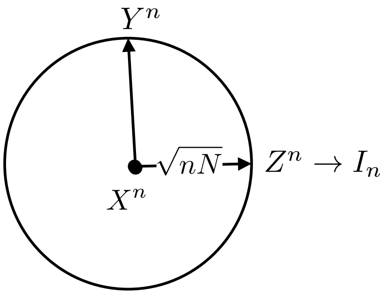
$$\mathbb{P}(\mathbf{Z} \in \text{blown-up set of } \mathbf{i}\text{'th bin}|\mathbf{x}) \approx 1.$$

$$\Downarrow$$

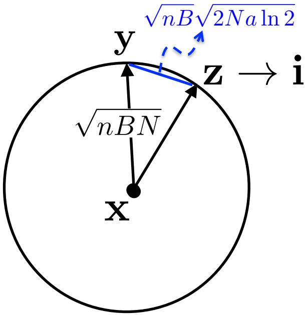
$$\mathbb{P}(\mathbf{Y} \in \text{blown-up set of } \mathbf{i}\text{'th bin}|\mathbf{x}) \approx 1.$$

Geometry of Typical Sets

- n-dimensional space:
  Typical (X^n, Y^n, Z^n, I_n)

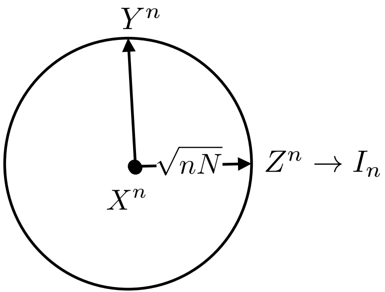


- nB-dimensional space:
  Typical $(\mathbf{x}, \mathbf{y}, \mathbf{i})$

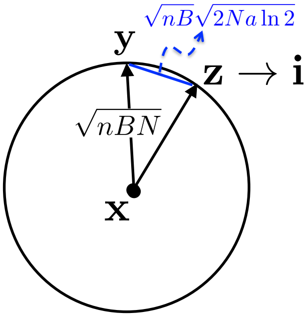


Geometry of Typical Sets

- n-dimensional space:
  Typical (X^n, Y^n, Z^n, I_n)



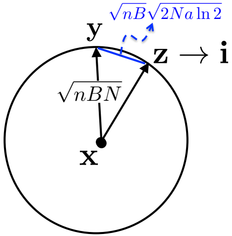
- nB-dimensional space:
  Typical $(\mathbf{x}, \mathbf{y}, \mathbf{i})$



Translating Geometry to Information Inequalities

Assume $\mathbf{y}, \mathbf{z}$ were discrete:

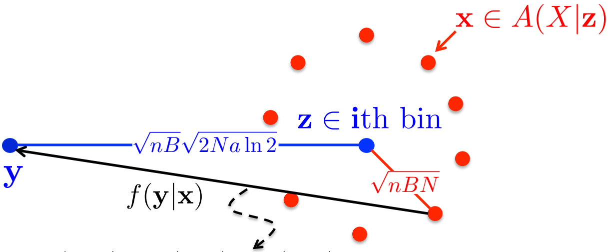
- $B(r)$: Ball of radius r around $\mathbf{y}$.
- $|B(r)|$: number of $\mathbf{z}$ sequences.
- $H(\mathbf{I}|\mathbf{Y}) \leq \log |B(\sqrt{2nBNa \log 2})|$.



If $I(I_n; X^n) = n(R_0 - a)$,

$$I(I_n; Y^n) \geq nR_0 - \frac{1}{B} \log |B(\sqrt{2nBNa \log 2})|.$$

Bounding $f(\mathbf{y}|\mathbf{i})$ for a typical $(\mathbf{y}, \mathbf{i})$



$$d(\mathbf{x},\mathbf{y}) \leq d(\mathbf{x},\mathbf{z}) + d(\mathbf{z},\mathbf{y})$$
$$\leq \sqrt{nB}(\sqrt{N} + \sqrt{2Na\ln 2})$$

$$f(\mathbf{y}|\mathbf{i}) \geq \sum_{\mathbf{x}\in A(X^n|\mathbf{z})} p(\mathbf{x}|\mathbf{i}) f(\mathbf{y}|\mathbf{x})$$

$$|A(X^n|\mathbf{z})| \doteq 2^{BH(X^n|Z^n)}$$
$$p(\mathbf{x}|\mathbf{i}) \doteq 2^{-BH(X^n|I_n)}$$
$$f(\mathbf{y}|\mathbf{x}) \geq \frac{1}{(2\pi N)^{nB/2}} e^{-\frac{nB(\sqrt{N}+\sqrt{2Na\ln 2})^2}{2N}}$$

Bounding $f(\mathbf{y}|\mathbf{i})$

For a typical $(\mathbf{y}, \mathbf{i})$:

1. $f(\mathbf{y}|\mathbf{i}) \geq 2^{-B(H(X^n|I_n)-H(X^n|Z^n)+h(Y^n|X^n)+n(a+\sqrt{2a\ln 2}\log e))}$
2. $f(\mathbf{y}|\mathbf{i}) \doteq 2^{-Bh(Y^n|I_n)}$

A new entropy inequality:

$$h(Y^n|I_n) \leq H(X^n|I_n) - H(X^n|Z^n) + h(Y^n|X^n) + n(a + \sqrt{2a\ln 2}\log e).$$

Bounding $f(\mathbf{y}|\mathbf{i})$

For a typical $(\mathbf{y}, \mathbf{i})$:

1. $f(\mathbf{y}|\mathbf{i}) \geq 2^{-B(H(X^n|I_n)-H(X^n|Z^n)+h(Y^n|X^n)+n(a+\sqrt{2a\ln 2}\log e))}$
2. $f(\mathbf{y}|\mathbf{i}) \doteq 2^{-Bh(Y^n|I_n)}$

A new entropy inequality:

$$h(Y^n|I_n) \leq H(X^n|I_n) - H(X^n|Z^n) + h(Y^n|X^n) + n(a + \sqrt{2a\ln 2}\log e).$$

Equivalently:

$$I(I_n; X^n) - I(I_n; Y^n) \leq n(a + \sqrt{2a\ln 2}\log e).$$

Bounding $f(\mathbf{y}|\mathbf{i})$

For a typical $(\mathbf{y}, \mathbf{i})$:

1. $f(\mathbf{y}|\mathbf{i}) \geq 2^{-B(H(X^n|I_n)-H(X^n|Z^n)+h(Y^n|X^n)+n(a+\sqrt{2a\ln 2}\log e))}$
2. $f(\mathbf{y}|\mathbf{i}) \doteq 2^{-Bh(Y^n|I_n)}$

A new entropy inequality:

$$h(Y^n|I_n) \leq H(X^n|I_n) - H(X^n|Z^n) + h(Y^n|X^n) + n(a + \sqrt{2a\ln 2}\log e).$$

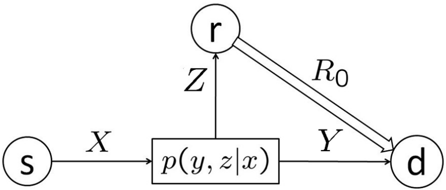
Equivalently:

$$I(I_n; X^n) - I(I_n; Y^n) \leq n(a + \sqrt{2a\ln 2}\log e).$$

A new constraint:

$$\begin{aligned} R &\leq I(X;Y) + I(I_n; X^n) - I(I_n; Y^n) \\ &\leq I(X;Y) + a + \sqrt{2a\ln 2}\log e. \end{aligned}$$

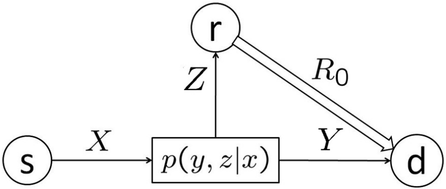
Discrete Memoryless Primitive Relay Channel



Y and Z are conditionally I.I.D. given X.

Joint work with Liang-Liang Xie.

An Open Problem

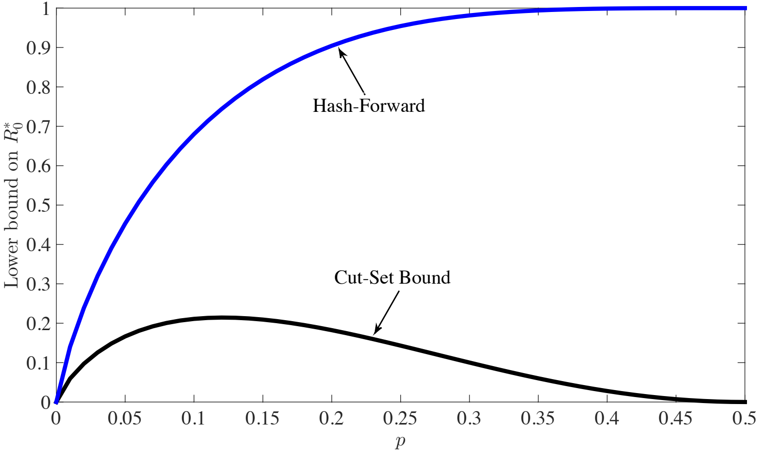


Cover, Open problems in Communication and Computation, 1987

What is the minimum R_0 (denoted by R_0^*) needed to achieve $C_{XYZ} = \max_{p(x)} I(X; Y, Z)$?

For example, when the channel from X to Y and Z is BSC(p), $R_0^* \leq 1$.

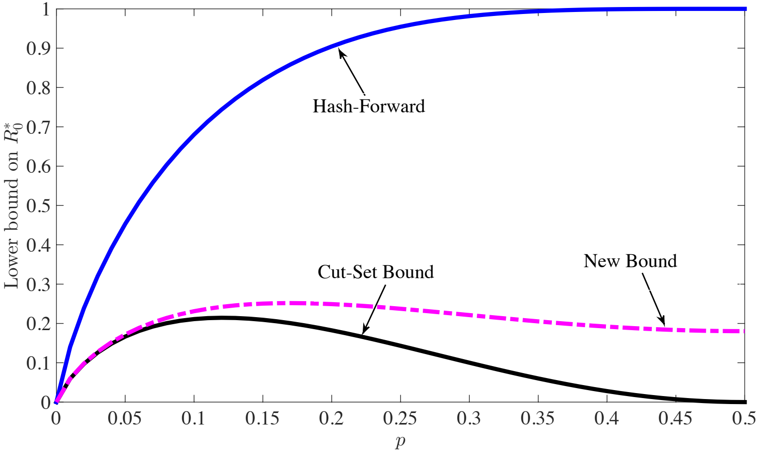
Bounds on $R_0^*(BSC)$



A striking dichotomy when $p \to 0.5 \;\Rightarrow\; C_{XYZ} \to 0$:

- Achievable schemes require $R_0 \to 1$.
- Upper bounds allow for $R_0 \to 0$.

Bounds on $R_0^*(BSC)$



Strictly positive R_0 needed to achieve $C_{XYZ} \to 0$!

Conclusion

- We developed new upper bounds on the capacity of the relay channel that are tighter than the cutset bound.
- Our proof used ideas from typicality and measure concentration.
- It would be interesting to see if the approach we develop in this paper, i.e. deriving information inequalities by studying the geometry of typical sets, in particular using measure concentration, can be used to make progress on other long-standing open problems in network information theory.