Broadcast Channels

Chandra Nair

The Chinese University of Hong Kong

20 July, 2021

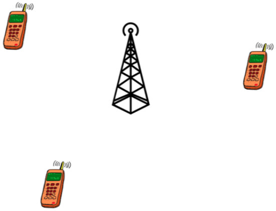
Dedication

Tom Cover

Katalin Marton

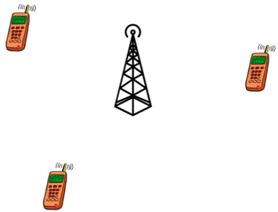
Broadcast Channel

Downlink Communication: From antenna to users in a cell

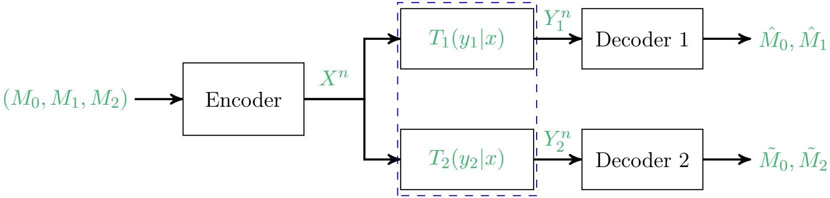


Broadcast Channel

Downlink Communication: From antenna to users in a cell

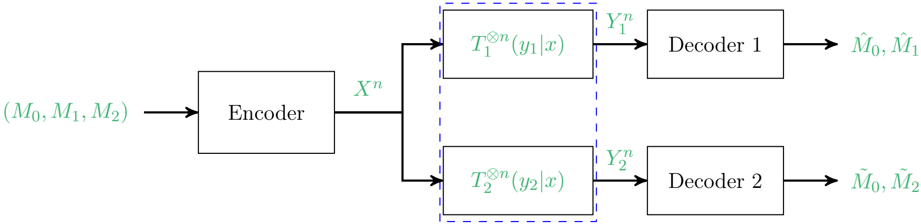


Mathematical Abstraction [Cover 1972]



Two-receiver broadcast channel

Memoryless Broadcast Channel

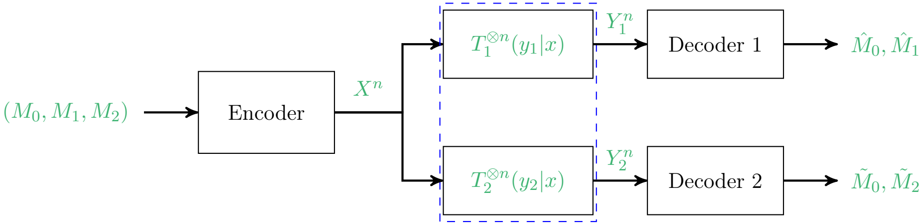


(R_0, R_1, R_2) is **achievable**: $\exists$ a sequence of **encoding maps** and **decoding maps** such that, as $n \to \infty$,

$$\mathrm{P}\Big(\{(\hat{M}_0, \hat{M}_1) \neq (M_0, M_1)\} \cup \{(\tilde{M}_0, \tilde{M}_1) \neq (M_0, M_1)\}\Big) \to 0,$$

when

$$(M_0, M_1, M_2) \sim \mathrm{Uni}\big([1 : \lfloor 2^{nR_0} \rfloor] \times [1 : \lfloor 2^{nR_1} \rfloor] \times [1 : \lfloor 2^{nR_2} \rfloor]\big).$$



(R_0, R_1, R_2) is **achievable**: $\exists$ a sequence of **encoding maps** and **decoding maps** such that, as $n \to \infty$,

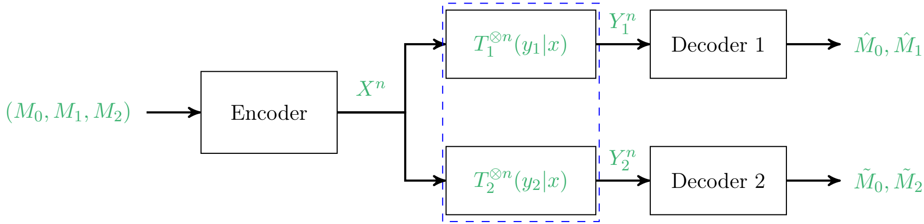
$$\mathrm{P}\Big(\{(\hat{M}_0, \hat{M}_1) \neq (M_0, M_1)\} \cup \{(\tilde{M}_0, \tilde{M}_1) \neq (M_0, M_1)\}\Big) \to 0,$$

when

$$(M_0, M_1, M_2) \sim \mathrm{Uni}\big([1 : \lfloor 2^{nR_0} \rfloor] \times [1 : \lfloor 2^{nR_1} \rfloor] \times [1 : \lfloor 2^{nR_2} \rfloor]\big).$$

Capacity Region, $\mathcal{C}(T_1, T_2)$: the closure of the set of all achievable (R_0, R_1, R_2).

Memoryless Broadcast Channel



(R_0, R_1, R_2) is **achievable**: $\exists$ a sequence of **encoding maps** and **decoding maps** such that, as $n \to \infty$,

$$\mathrm{P}\Big(\{(\hat{M}_0, \hat{M}_1) \neq (M_0, M_1)\} \cup \{(\tilde{M}_0, \tilde{M}_1) \neq (M_0, M_1)\}\Big) \to 0,$$

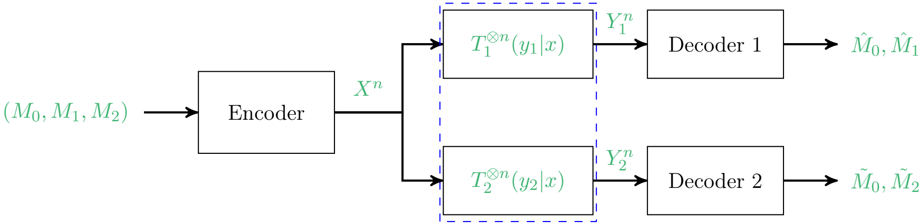
when

$$(M_0, M_1, M_2) \sim \mathrm{Uni}\big([1 : \lfloor 2^{nR_0} \rfloor] \times [1 : \lfloor 2^{nR_1} \rfloor] \times [1 : \lfloor 2^{nR_2} \rfloor]\big).$$

Capacity Region, $\mathcal{C}(T_1, T_2)$: the closure of the set of all achievable (R_0, R_1, R_2).

Goal: A **computable** characterization of the capacity region. **(open)**

Memoryless Broadcast Channel

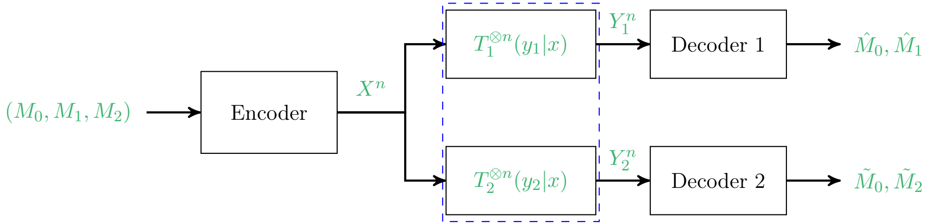


Computable characterization

$\max_{(R_0,R_1,R_2)\in\mathcal{C}(T_1,T_2)} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2$: expressed as a maximum of a continuous function over a compact set

Capacity Region, $\mathcal{C}(T_1, T_2)$: the closure of the set of all achievable (R_0, R_1, R_2).

Goal: A **computable** characterization of the capacity region. **(open)**



Computable characterization

$\max_{(R_0,R_1,R_2)\in\mathcal{C}(T_1,T_2)} \lambda_0 R_0 + \lambda_1 R_1 + \lambda_2 R_2$: expressed as a maximum of a continuous function over a compact set

This implies that $\exists$ Turing machine that can solve the *weak membership* problem [Corollary 6.2.5 in K. Weihrauch. Computable Analysis: An Introduction. Berlin, Heidelberg: Springer-Verlag, 2000. ISBN: 3540668179]

Capacity Region, $\mathcal{C}(T_1, T_2)$: the closure of the set of all achievable (R_0, R_1, R_2).

Goal: A **computable** characterization of the capacity region. **(open)**

This talk

- **Review**: Results from the classical period (1972-1982)
- **Main**: Results from the recent era (2004 -)
  - ◇ Capacity regions for new classes of channels
  - ◇ Optimality/Sub-optimality of certain coding strategies

This talk

- **Review**: Results from the classical period (1972-1982)
- **Main**: Results from the recent era (2004 -)
  - ◇ Capacity regions for new classes of channels
  - ◇ Optimality/Sub-optimality of certain coding strategies

My take: Recent results are mainly due to a change of perspective

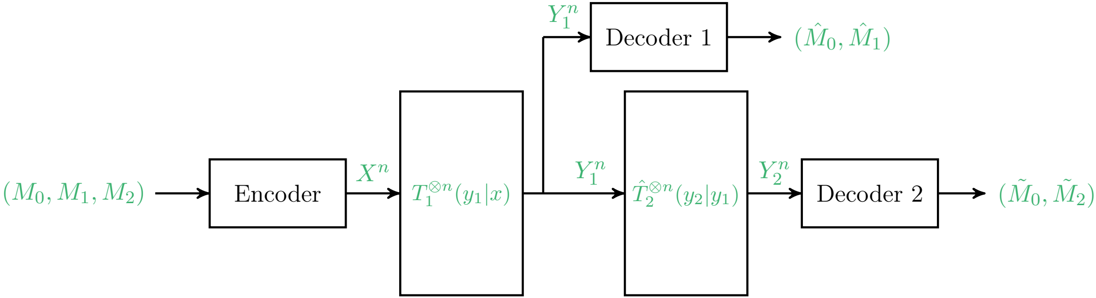
from *obtaining converses for coding theorems*

to *evaluation of inner and outer bounds* (non-convex optimization problems)

Review: Superposition coding idea for the broadcast channel

Superposition coding was developed as an achievable coding strategy for the **degraded** broadcast channel [Cov72]

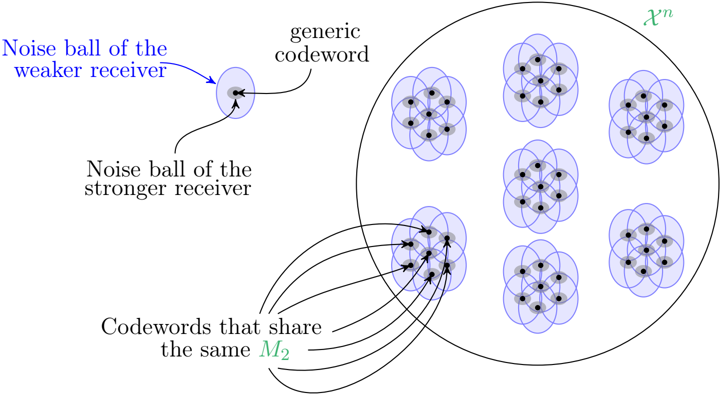
Cover



Review: Superposition coding idea for the broadcast channel

Superposition coding was developed as an achievable coding strategy for the **degraded** broadcast channel [Cov72]

Cover



Superposition coding was developed as an achievable coding strategy for the **degraded** broadcast channel [Cov72]

Cover

Theorem: Superposition coding achievable region

The set of rate triples (R_0, R_1, R_2) satisfying

$$R_0 + R_2 \leq I(V; Y_2)$$
$$R_0 + R_1 + R_2 \leq I(X; Y_1|V) + I(V; Y_2)$$
$$R_0 + R_1 + R_2 \leq I(X; Y_1)$$

for some $p(v, x)$ is achievable. Here $V \multimap X \multimap (Y_1, Y_2)$ is Markov. W.l.o.g. $|\mathcal{V}| \leq |\mathcal{X}| + 1$.

Review: Superposition coding idea for the broadcast channel

Superposition coding was developed as an achievable coding strategy for the **degraded** broadcast channel [Cov72]

Cover

Optimality of Superposition Coding Region

Degraded Gaussian broadcast channel, [Ber73]

- Use of Entropy Power Inequality to deduce Gaussian Optimality

Bergmans

Degraded discrete memoryless broadcast channel, [Gal74]

- Explicit identification of auxiliaries in the converse from distributions induced by codebooks

Gallager

Both arguments extend to k receivers

Review: Superposition coding idea for the broadcast channel

Superposition coding was developed as an achievable coding strategy for the **degraded** broadcast channel [Cov72]

Cover

Optimality of Superposition Coding Region

Less noisy broadcast channel, [KM75]

- $\forall p_{U|X}$ we have $I(U;Y_1) \geq I(U;Y_2)$

Projection of capacity region on $R_2 = 0$, [KM77a]
(Degraded message sets)

- Images of a set under two noisy channels [KM77b]
- First use of the identity
  $H(Y_1^n) - H(Y_2^n) = \sum_{i=1} \left(H(Y_{1i}|Y_1^{i-1}, Y_{2i+1}^n) - H(Y_{2i}|Y_1^{i-1}, Y_{2i+1}^n) \right)$
  - ◇ Staple equality for many converses or outer bounds

Körner

Marton

Review: Superposition coding idea for the broadcast channel

Superposition coding was developed as an achievable coding strategy for the **degraded** broadcast channel [Cov72]

Cover

Optimality of Superposition Coding Region

Less noisy broadcast channel, [KM75]

- $\forall p_{U|X}$ we have $I(U;Y_1) \geq I(U;Y_2)$

Körner

Projection of capacity region on $R_2 = 0$, [KM77a]
(Degraded message sets)

- Images of a set under two noisy channels [KM77b]
- First use of the identity $H(Y_1^n) - H(Y_2^n) = \sum_{i=1} \left(H(Y_{1i}|Y_1^{i-1}, Y_{2i+1}^n) - H(Y_{2i}|Y_1^{i-1}, Y_{2i+1}^n) \right)$
  - ◇ Staple equality for many converses or outer bounds

Marton

Both results were established only for 2 receivers

Open: Extension of *images of a set characterization to 3 receivers*

Review: Superposition coding idea for the broadcast channel

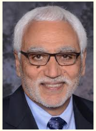
Superposition coding was developed as an achievable coding strategy for the **degraded** broadcast channel [Cov72]

Cover

Optimality of Superposition Coding Region

More capable broadcast channel, [El 79]

- $\forall p_X$ we have $I(X;Y_1) \geq I(X;Y_2)$.
- Equivalent: Any ϵ-error codebook for receiver Y_2 is "essentially" an ϵ-error codebook for receiver Y_1



El Gamal

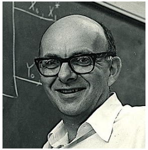
Remarks:

- Bypasses images of a set characterization (simpler)
- The proof contained the **UV outer bound** for two receiver broadcast channel
  - ◇ Focus was on converses, not outer bounds
- Result was established only for 2 receivers

Review: Random binning based achievable region

Random binning idea

- Compression of correlated sources, [SW73]



Slepian

Wolf

Optimality of random binning based achievable region

Deterministic Broadcast (1977-1978), [Gel77; Mar77; Pin78]

- $Y_1 = f(X), Y_2 = g(X)$

Semi-deterministic Broadcast (1978-1980), [Mar79; GP80]

- $Y_1 = f(X)$

Gelfand

Marton

Pinsker

Review: Product broadcast channels

Two **independent channel components**

- $X = (X_a, X_b)$, $Y_1 = (Y_{1a}, Y_{1b})$, $Y_2 = (Y_{2a}, Y_{2b})$
- $T_1(Y_1|X) = T_{1a}(Y_{1a}|X_a) \otimes T_{1b}(Y_{1b}|X_b)$
- $T_2(Y_2|X) = T_{2a}(Y_{2a}|X_a) \otimes T_{2b}(Y_{2b}|X_b)$

Capacity region for reversely degraded broadcast channel

Gaussian setting, [Hug75]

Hughes-Hartogs

Projection on $R_0 = 0$, [Pol77]

Poltyrev

Review: Product broadcast channels

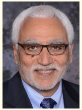
Two **independent channel components**

- $X = (X_a, X_b)$, $Y_1 = (Y_{1a}, Y_{1b})$, $Y_2 = (Y_{2a}, Y_{2b})$
- $T_1(Y_1|X) = T_{1a}(Y_{1a}|X_a) \otimes T_{1b}(Y_{1b}|X_b)$
- $T_2(Y_2|X) = T_{2a}(Y_{2a}|X_a) \otimes T_{2b}(Y_{2b}|X_b)$

Capacity region for reversely degraded broadcast channel

Full capacity region, [El 80]

- Optimality of the Minkowski sum of the two individual capacity regions



El Gamal

- The product channel setting has had a surprisingly large impact on recent results

Review: Achievable Region

By combining superposition coding and random binning the following region is achievable for any broadcast channel

Marton

Marton's achievable region $\mathcal{M}(T_1, T_2)$, [Mar79]

Any rate tuple (R_0, R_1, R_2) satisfying

$$\begin{aligned} R_0 &\leq \min\{I(W;Y_1), I(W;Y_2)\} \\ R_0 + R_1 &\leq I(UW;Y_1) \\ R_0 + R_2 &\leq I(VW;Y_2) \\ R_0 + R_1 + R_2 &\leq \{I(W;Y_1), I(W;Y_2)\} + I(U;Y_1|W) + I(V;Y_2|W) - I(U;V|W) \end{aligned}$$

for any $p_{UVWX} : (UVW) \multimap X \multimap (Y_1, Y_2)$ is achievable, i.e. $\mathcal{M}(T_1, T_2) \subseteq \mathcal{C}(T_1, T_2)$.

Caveat: This region was not computable

Review: Evaluation of achievable regions

Two results that are in the spirit of modern results in broadcast channel

The capacity region of the degraded binary symmetric (BSC) broadcast channel, [WZ73]

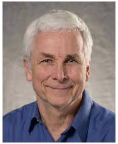
- Mrs. Gerber's lemma to evaluate superposition coding region

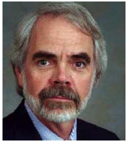

Wyner


Ziv

Evaluation of an achievable rate region for the broadcast channel, [HP79]

- Functional representation lemma
- Reduction of $\mathcal{M}(T_1, T_2)$, when X is binary and U, V, W are independent, to **randomized time-division** region
  - Makes the region computable


Hajek


Pursley

Outer bound to the capacity region (mostly classical)

There was some interest in outer bounds to the capacity region 2006 - 2010

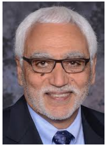
- Couple of outer bounds by


Kramer


Liang


Shamai

- Outer bound [NE07]
  - ◇ Improves on the bound by Körner-Marton
  - ◇ Employs an "XOR trick" to show equivalence between regions


El Gamal

UVW outer bound, $\mathcal{O}(T_1, T_2)$

The union of rate tuples (R_0, R_1, R_2) satisfying

$$\begin{aligned} R_0 &\leq \min\{I(W;Y_1), I(W;Y_2)\} \\ R_0 + R_1 &\leq \min\{I(W;Y_1), I(W;Y_2)\} + I(U;Y_1|W) \\ R_0 + R_2 &\leq \min\{I(W;Y_1), I(W;Y_2)\} + I(V;Y_2|W) \\ R_0 + R_1 + R_2 &\leq \{I(W;Y_1), I(W;Y_2)\} + I(U;Y_1|W) + I(X;Y_2|UW) \\ R_0 + R_1 + R_2 &\leq \{I(W;Y_1), I(W;Y_2)\} + I(V;Y_2|W) + I(X;Y_1|VW) \end{aligned}$$

for any $p_{UVWX} : (UVW) \multimap X \multimap (Y_1, Y_2)$ forms an outer bound, i.e. $\mathcal{C}(W_1, W_2) \subseteq \mathcal{O}(T_1, T_2)$. It suffices to consider $|\mathcal{W}| \leq |\mathcal{X}| + 5,\ |\mathcal{U}| \leq |\mathcal{X}| + 1,\ |\mathcal{V}| \leq |\mathcal{X}| + 1$. [Nai10a]

Remarks:

- This outer bound follows from classical arguments
- The projection of this region on $R_0 = 0$ is same as setting W to be constant
  - ◇ **UV outer bound** for private messages

Advances over the classical results - Main part

Investigate **optimality/sub-optimality** of **superposition coding region** (2007-2018)

- More instances where it matches capacity region
  - Evaluation of inner and outer bounds (ideas involved)
- Settings where it is sub-optimal

Advances over the classical results - Main part

Investigate **optimality/sub-optimality** of **superposition coding region** (2007-2018)

- More instances where it matches capacity region
  - Evaluation of inner and outer bounds (ideas involved)
- Settings where it is sub-optimal

Investigate **optimality/sub-optimality** of **Marton's inner bound** and **UVW outer bound** (2008-2015)

- More instances where the bounds coincide
  - Evaluation of inner and outer bounds (ideas involved)
- Settings where there is a gap between the bounds

Advances over the classical results - Main part

Investigate **optimality/sub-optimality** of **superposition coding region** (2007-2018)

- More instances where it matches capacity region
  - Evaluation of inner and outer bounds (ideas involved)
- Settings where it is sub-optimal

Investigate **optimality/sub-optimality** of **Marton's inner bound** and **UVW outer bound** (2008-2015)

- More instances where the bounds coincide
  - Evaluation of inner and outer bounds (ideas involved)
- Settings where there is a gap between the bounds

Recent advances (2019 -) and remarks

New Results: Optimality of Superposition Coding Region

Idea: **Evaluate** the **UVW Outer bound** and show that it is contained inside **Superposition Coding Region**

Difficulty: **Evaluation** of the regions are **non-convex** optimization problems

- Therefore, some optimization related insights are needed

New Results: Optimality of Superposition Coding Region

Idea: **Evaluate** the **UVW Outer bound** and show that it is contained inside **Superposition Coding Region**

Difficulty: **Evaluation** of the regions are **non-convex** optimization problems

- Therefore, some optimization related insights are needed

Example: Consider $T_1(y_1|x) \sim BSC(p)$ and $T_2(y_2|x) \sim BEC(\epsilon)$ [Nai10b]

Note: If $1 > \epsilon > H_2(p)$, then neither are more-capable than the other

- If $H_2(p) \leq \epsilon$ then Y_2 is more capable than Y_1

To show optimality of superposition coding region:

Step 1: Show that one can restrict to $p(x)$ to be the $\mathbf{Ber}(\frac{1}{2})$ to evaluate the UVW outer bound

- Employs a **symmetrization** argument

New Results: Optimality of Superposition Coding Region

Idea: **Evaluate** the **UVW Outer bound** and show that it is contained inside **Superposition Coding Region**

Difficulty: **Evaluation** of the regions are **non-convex** optimization problems

- Therefore, some optimization related insights are needed

Example: Consider $T_1(y_1|x) \sim BSC(p)$ and $T_2(y_2|x) \sim BEC(\epsilon)$ [Nai10b]

Note: If $1 > \epsilon > H_2(p)$, then neither are more-capable than the other

- If $H_2(p) \leq \epsilon$ then Y_2 is more capable than Y_1

To show optimality of superposition coding region:

Step 1: Show that one can restrict to $p(x)$ to be the **Ber**$(\frac{1}{2})$ to evaluate the UVW outer bound

- Employs a **symmetrization** argument

Step 2: When $p(x) \sim \text{Ber}(\frac{1}{2})$ show that for all $U : U -\!\!\circ\!\!- X -\!\!\circ\!\!- (Y_1, Y_2)$, we have $I(U;Y_1) \geq I(U;Y_2)$.

- The maximum of the **upper concave envelope** coincides with the maximum of the function.

New Results: Optimality of Superposition Coding Region

Idea: **Evaluate** the **UVW Outer bound** and show that it is contained inside **Superposition Coding Region**

Difficulty: **Evaluation** of the regions are **non-convex** optimization problems

- Therefore, some optimization related insights are needed

Generalization of this idea [Nai10b]:

- Essentially Less Noisy
- Essentially More Capable

Idea: **Evaluate** the **UVW Outer bound** and show that it is contained inside **Superposition Coding Region**

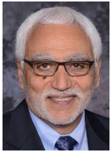
Difficulty: **Evaluation** of the regions are **non-convex** optimization problems

- Therefore, some optimization related insights are needed

Generalization of this idea [Nai10b]:

- Essentially Less Noisy
- Essentially More Capable

Effectively Less Noisy [KNE15]



El Gamal

H. Kim

Nachman

Main Issue: Image size characterization for 3 or more receivers

2. Image Size Characterization Problem.

The η-image size $g_W(A, \eta)$ of a set $A \subset X^n$ over a discrete memoryless channel (DMC) $\{W : X \to Y\}$ is the minimum cardinality of $B \subset Y^n$ such that $W^n(B \mid \mathbf{x}) \geq \eta$ for each $\mathbf{x} \in A$. The problem is to find, for a distribution P on X and DMCs $\{W_i : X \to Y_i\}$, $i = 1, \ldots, k$, a single-letter characterization of the limit of the sets of all $(k+1)$-dimensional vectors

-29-

$$\left[\frac{1}{n}\log|A|, \frac{1}{n}\log g_{W_1}(A,\eta), \ldots, \frac{1}{n}\log g_{W_k}(A,\eta)\right].$$

Here $A \subset X^n$ is any set of P-typical sequences, and $0 < \eta < 1$ is fixed (the result is independent of η).

Both problems are solved for $k = 2$ (cf. [1]) but not for $k \geq 3$. An interesting (unsolved) special case of Problem 2 for $k = 3$ is the follow-

Open problem in [CG87]

Csiszar

Note: Less noisy, more capable, degraded message sets used image sizes

For insiders: Past/future issue in identification of auxiliaries

Superposition Coding Region - 3 or more receivers

Less Noisy

- Superposition coding region is **optimal** for $k = 3$ [NW11]
  - ◇ **Novel ingredient**: Information inequality (specialized for less noisy) to aid single-letterization
- Optimality is **open** for $k \geq 4$

V. Wang

Superposition Coding Region - 3 or more receivers

V. Wang

Less Noisy

- Superposition coding region is **optimal** for $k = 3$ [NW11]
  - ◇ **Novel ingredient**: Information inequality (specialized for less noisy) to aid single-letterization
- Optimality is **open** for $k \geq 4$

More capable

Xia

- Superposition coding region is sub-optimal for $k = 3$ [NX12]
  - ◇ **"Counter Example"**: $T_1(y_1|x) \sim BEC(\epsilon_1)$, $T_2(y_2|x) \sim BEC(\epsilon_2)$, $T_3(y_3|x) \sim BSC(p)$
    - ⋆ $0 < \epsilon_1 < \epsilon_2 = H_2(p)$
  - ◇ Consider the weighted sum-rate $\frac{R_1}{1-\epsilon_1} + \frac{R_2+R_3}{1-\epsilon_2}$
    - ⋆ Superposition coding region yields maximum value 1
    - ⋆ Can be improved by ignoring receiver 2 ($R_2 = 0$) and considering the capacity region of Y_1, Y_3.

Superposition Coding Region - 3 or more receivers

Less Noisy

- Superposition coding region is **optimal** for $k = 3$ [NW11]
  - ◇ **Novel ingredient**: Information inequality (specialized for less noisy) to aid single-letterization
- Optimality is **open** for $k \geq 4$

V. Wang

More capable

- Superposition coding region is sub-optimal for $k = 3$ [NX12]
  - ◇ **"Counter Example"**: $T_1(y_1|x) \sim BEC(\epsilon_1)$, $T_2(y_2|x) \sim BEC(\epsilon_2)$, $T_3(y_3|x) \sim BSC(p)$
    - ⋆ $0 < \epsilon_1 < \epsilon_2 = H_2(p)$

Xia

Conjecture: Superposition coding region is optimal for the following setting:

$$Y_1 \overset{m.c.}{\gg} Y_2, Y_1 \overset{m.c.}{\gg} Y_3, Y_2 \overset{l.n.}{\succeq} Y_3.$$

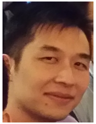
Less Noisy vs More Capable

Replacing a receiver by a **less noisy** receiver **cannot decrease** the capacity region of a two-receiver broadcast channel

On the other hand, replacing a receiver by a **more capable** receiver can **strictly decrease** the capacity region of a two-receiver broadcast channel [Gen+13b].

Geng

Shamai



V. Wang

Superposition Coding Region - 3 or more receivers

Degraded message sets

Focus on two message case only (simplest setting)

- "Common" message M_c to be decoded by Y_1, Y_2, Y_3

There are two possible scenarios:

- **Case A**: "Refined" message M_r to be decoded by Y_1
- **Case B**: "Refined" message M_r to be decoded by Y_1 and Y_2

Case A: 3-receiver degraded message sets

Case A: "Refined" message M_r to be decoded by Y_1

Superposition coding region: The set of rate pairs satisfying

$$
\begin{aligned}
R_c &\leq \min\{I(U;Y_2), I(U;Y_3)\} \\
R_c + R_r &\leq \min\{I(U;Y_2), I(U;Y_3)\} + I(X;Y_1|U) \\
R_c + R_r &\leq I(X;Y_1)
\end{aligned}
$$

for any $(U,X) : U \multimap X \multimap (Y_1, Y_2, Y_3)$ is achievable.

Case A: 3-receiver degraded message sets

Case A: "Refined" message M_r to be decoded by Y_1

Consider an **augmented** setting

- M_{123} to be decoded by Y_1, Y_2, Y_3
- M_{12} to be decoded by Y_1, Y_2
- M_{13} to be decoded by Y_1, Y_3
- M_1 to be decoded by Y_1

Observe that the collection of decoding sets is **upward closed**

Case A: 3-receiver degraded message sets

Case A: "Refined" message M_r to be decoded by Y_1

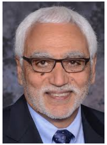
Consider an **augmented** setting

- M_{123} to be decoded by Y_1, Y_2, Y_3
- M_{12} to be decoded by Y_1, Y_2
- M_{13} to be decoded by Y_1, Y_3
- M_1 to be decoded by Y_1

Observe that the collection of decoding sets is **upward closed**

There is a natural extension of Marton's achievable region (combining superposition coding and mutual covering) to this setting

- The projection of the achievable region on the plane $R_{12} = R_{13} = 0$ yields an achievable rate for **Case A**
- There are instances when the projection is strictly larger than the superposition coding region [NE09]
- One instance is a **product** degraded erasure channel
  - ◇ The above projection matches the capacity region



El Gamal

Case B: 3-receiver degraded message sets

Case B: "Refined" message M_r to be decoded by Y_1 and Y_2

- M_c to be decoded by Y_1, Y_2, Y_3
- M_r to be decoded by Y_1, Y_2

Observe that the collection of decoding sets is **upward closed**

- A natural guess was that superposition coding was optimal

Superposition coding region: The set of rate pairs satisfying

$$\begin{aligned} R_c &\leq I(U;Y_3) \\ R_c + R_r &\leq I(U;Y_3) + \min\{I(X;Y_1|U), I(X;Y_2|U)\} \\ R_c + R_r &\leq \min\{I(X;Y_1), I(X;Y_2)\} \end{aligned}$$

for any $(U,X) : U \multimap X \multimap (Y_1,Y_2,Y_3)$ is achievable.

Case B: 3-receiver degraded message sets

Case B: "Refined" message M_r to be decoded by Y_1 and Y_2

Superposition coding region: The set of rate pairs satisfying

$$
\begin{aligned}
R_c &\leq I(U;Y_3) \\
R_c + R_r &\leq I(U;Y_3) + \min\{I(X;Y_1|U), I(X;Y_2|U)\} \\
R_c + R_r &\leq \min\{I(X;Y_1), I(X;Y_2)\}
\end{aligned}
$$

for any $(U,X) : U \multimap X \multimap (Y_1,Y_2,Y_3)$ is achievable.

After some years another intuition/conjecture

global tensorization *if and only if* **local tensorization**

suggested superposition coding may not be optimal.

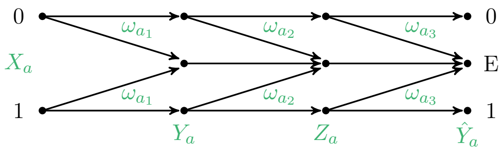
This intuition also helped find possible specific counterexamples (here and in other settings)

Product Degraded Erasure Channel

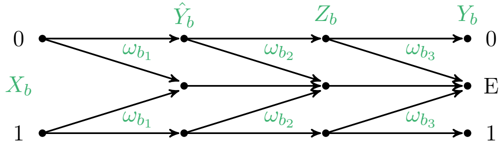
$$X_a \to Y_a : BEC(e_a),\ X_b \to Y_b : BEC(e_b)$$

$$X_a \to \hat{Y}_a : BEC(\hat{e}_a),\ X_b \to \hat{Y}_b : BEC(\hat{e}_b)$$

$$X_a \to Z_a : BEC(f_a),\ X_b \to Z_b : BEC(f_b)$$



$$\hat{e}_a \geq f_a \geq e_a \quad \& \quad e_b \geq f_b \geq \hat{e}_b$$



$$C_Z = (1 - f_a) + (1 - f_b)$$

Theorem [NY17]

For $e_a = 1/2$ $\quad \hat{e}_a = 1$ $\quad f_a = 17/22$

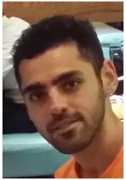
$e_b = 1/2$ $\quad \hat{e}_b = 0$ $\quad f_b = 9/34$

1-letter SC :

$$R_0 + R_1 \leq 1 \quad \text{and} \quad \lambda_1 R_0 + R_1 \leq \lambda_1 C_Z, \quad \lambda_1 = \frac{11}{10}$$

2-letter SC :

$$R_0 + R_1 \leq 1 \quad \text{and} \quad \lambda_2 R_0 + R_1 \leq \lambda_2 C_Z, \quad \lambda_2 = \frac{484}{435}$$



Yazdanpanah

Case B: 3-receiver degraded message sets

Evaluation of superposition coding regions entailed

- Symmetrization
- Representation using concave envelopes
- Slope of region at axis points
- Shannon-type inequalities (linear programming)
- Min-max interchange

Case B: 3-receiver degraded message sets

Körner had proposed a region for the image sizes over three channels [Kör84]

Körner

Theorem: For every RV's T, U, and V such that

$$TUV \to S \to XYZ,$$

nonnegative numbers t, t', and t'', the point (r_x, r_y, r_z) with coordinates

$$\begin{aligned} r_x &\triangleq \min\left[H(X), H(X|T)+t, H(X|TU)+t',\right. \\ &\qquad \left. H(X|TUV)+t''\right], \\ r_y &\triangleq \min\left[H(Y), H(Y|T)+t, H(Y|TU)+t',\right. \\ &\qquad \left. H(Y|TUV)+t''\right], \\ r_z &\triangleq \min\left[H(Z), H(Z|T)+t, H(Z|TU)+t',\right. \\ &\qquad \left. H(Z|TUV)+t''\right] \end{aligned} \tag{30}$$

is an element of $\mathscr{H}(X;Y;Z|S)$. □

The same example shows that such points do no exhaust $\mathscr{H}(X;Y;Z|S)$.

A (very) specific open problem

Let X_a^n, X_b^n be two n-bit binary random variables.

For $k, l \in \{0, 1, 2, \ldots, n\}$ define

$$\mathcal{H}_n(k,l) = \frac{1}{\binom{n}{k}\binom{n}{l}} \sum_{S,T \subseteq [n] : |S|=k, |T|=l} H(X_{aS}, X_{bT})$$

to be an averaged entropy function over sets of same size.

Consider the following mapping:

$$p_{X_a^n, X_b^n} \mapsto \left(\frac{1}{n}\mathcal{H}_n\left(\left\lfloor \frac{n}{2} \right\rfloor, \left\lfloor \frac{n}{2} \right\rfloor\right), \frac{1}{n}\mathcal{H}_n\left(0, n\right), \frac{1}{n}\mathcal{H}_n\left(\left\lfloor \frac{5n}{22} \right\rfloor, \left\lfloor \frac{25n}{34} \right\rfloor\right)\right)$$

Let $\mathcal{G}_n$ be the range of this mapping and $\mathcal{G} = \cup_n \mathcal{G}_n$.

A (very) specific open problem

Let X_a^n, X_b^n be two n-bit binary random variables.

For $k, l \in \{0, 1, 2, \ldots, n\}$ define

$$\mathcal{H}_n(k,l) = \frac{1}{\binom{n}{k}\binom{n}{l}} \sum_{S,T \subseteq [n]: |S|=k, |T|=l} H(X_{aS}, X_{bT})$$

to be an averaged entropy function over sets of same size.

Consider the following mapping:

$$p_{X_a^n, X_b^n} \mapsto \left(\frac{1}{n}\mathcal{H}_n\left(\left\lfloor \frac{n}{2} \right\rfloor, \left\lfloor \frac{n}{2} \right\rfloor\right), \frac{1}{n}\mathcal{H}_n\left(0, n\right), \frac{1}{n}\mathcal{H}_n\left(\left\lfloor \frac{5n}{22} \right\rfloor, \left\lfloor \frac{25n}{34} \right\rfloor\right)\right)$$

Let $\mathcal{G}_n$ be the range of this mapping and $\mathcal{G} = \cup_n \mathcal{G}_n$.

Question: Determine a computable characterization of $\mathcal{G}$

Remarks

- Identical to asking for the capacity region of the previous example
- A simple (yet non-trivial) instance of Csiszar's **open** question of image-size characterization over three channels
- The answer is known if you only had $\left(\frac{1}{n}\mathcal{H}_n\left(n\alpha_1, n\beta_1\right), \frac{1}{n}\mathcal{H}_n\left(n\alpha_2, n\beta_2\right)\right)$

A (very) specific open problem

Lemma

- $\mathcal{H}_n(k,l) \le \mathcal{H}_n(k+k_0, l+l_0) \le \frac{k+k_0}{k}\frac{l+l_0}{l}\mathcal{H}_n(k,l)$ *for* $0 \le k_0 \le n-k, 0 \le l_0 \le n-l$.
- $\frac{k-k_0}{k}\frac{l-l_0}{l}\mathcal{H}_n(k,l) \le \mathcal{H}_n(k-k_0, l-l_0) \le \mathcal{H}_n(k,l)$ *for* $0 \le k_0 \le k, 0 \le l_0 \le l$.
- *(Concavity)* $\frac{m}{n}\mathcal{H}_n(k_1,l) + \frac{n-m}{n}\mathcal{H}_n(k_2,l) \le \mathcal{H}_n\left(\frac{mk_1+(n-m)k_2}{n}, l\right)$ *for* $0 \le m, k_1, k_2, l \le n$.

With these properties we can obtain the following inequalities

$$\mathcal{H}_n(\frac{n}{2},\frac{n}{2}) \le \mathcal{H}_n(\frac{n}{2},\frac{25n}{34}),$$
$$\frac{5}{11}\mathcal{H}_n(\frac{n}{2},\frac{25n}{34}) + \frac{6}{11}\mathcal{H}_n(0,\frac{25n}{34}) \le \mathcal{H}(\frac{5n}{22},\frac{25n}{34}),$$
$$\frac{8}{17}\mathcal{H}_n(0,n) + \frac{9}{17}\mathcal{H}_n(0,\frac{n}{2}) \le \mathcal{H}_n(0,\frac{25n}{34}),$$
$$\frac{17}{25}\mathcal{H}_n(0,\frac{25n}{34}) \le \mathcal{H}_n(0,\frac{n}{2}).$$

A (very) specific open problem

Lemma

- $\mathcal{H}_n(k,l) \le \mathcal{H}_n(k+k_0, l+l_0) \le \frac{k+k_0}{k}\frac{l+l_0}{l}\mathcal{H}_n(k,l)$
  for $0 \le k_0 \le n-k, 0 \le l_0 \le n-l$.
- $\frac{k-k_0}{k}\frac{l-l_0}{l}\mathcal{H}_n(k,l) \le \mathcal{H}_n(k-k_0, l-l_0) \le \mathcal{H}_n(k,l)$ *for* $0 \le k_0 \le k, 0 \le l_0 \le l$.
- *(Concavity)* $\frac{m}{n}\mathcal{H}_n(k_1,l) + \frac{n-m}{n}\mathcal{H}_n(k_2,l) \le \mathcal{H}_n\left(\frac{mk_1+(n-m)k_2}{n}, l\right)$
  for $0 \le m, k_1, k_2, l \le n$.

A linear combination of the inequalities shows that for all $p_{X_a^n,X_b^n}$

$$\frac{1}{n}\left(\frac{85}{160}\mathcal{H}(\frac{n}{2},\frac{n}{2}) + \frac{75}{160}\mathcal{H}(0,n) - \frac{187}{160}\mathcal{H}(\frac{5n}{22},\frac{25n}{34})\right) \le 0$$

A (very) specific open problem

Lemma

- $\mathcal{H}_n(k,l) \le \mathcal{H}_n(k+k_0, l+l_0) \le \frac{k+k_0}{k}\frac{l+l_0}{l}\mathcal{H}_n(k,l)$ *for* $0 \le k_0 \le n-k, 0 \le l_0 \le n-l$.
- $\frac{k-k_0}{k}\frac{l-l_0}{l}\mathcal{H}_n(k,l) \le \mathcal{H}_n(k-k_0, l-l_0) \le \mathcal{H}_n(k,l)$ *for* $0 \le k_0 \le k, 0 \le l_0 \le l$.
- *(Concavity)* $\frac{m}{n}\mathcal{H}_n(k_1,l) + \frac{n-m}{n}\mathcal{H}_n(k_2,l) \le \mathcal{H}_n\left(\frac{mk_1+(n-m)k_2}{n}, l\right)$ *for* $0 \le m, k_1, k_2, l \le n$.

An outer bound for the example

Any achievable (R_0, R_1) satisfies

$$R_0 + R_1 \le 1 \quad \text{and} \quad \frac{187}{160}R_0 + R_1 \le \frac{18}{16}.$$

An **explicit** outer bound in a non-traditional way!

Advances over the classical results - Main part

Investigate optimality/sub-optimality of superposition coding region (2007-2018)

- More instances where it matches capacity region
  - ◇ Evaluation of inner and outer bounds (ideas involved)
- Settings where it is sub-optimal

Investigate **optimality/sub-optimality** of **Marton's inner bound** and **UVW outer bound** (2008-2015)

- More instances where the bounds coincide
  - ◇ Evaluation of inner and outer bounds (ideas involved)
- Settings where there is a gap between the bounds

Advances over the classical results - Main part

Investigate optimality/sub-optimality of superposition coding region (2007-2018)

- More instances where it matches capacity region
  - Evaluation of inner and outer bounds (ideas involved)
- Settings where it is sub-optimal

Investigate **optimality/sub-optimality** of **Marton's inner bound** and **UVW outer bound** (2008-2015)

- More instances where the bounds coincide
  - Evaluation of inner and outer bounds (ideas involved)
- Settings where there is a gap between the bounds

Very recent advances (2019 -)

Vector Gaussian Broadcast Channel

Consider the broadcast channel (of interest in multi-antenna wireless communication)

$$\mathbf{Y}_1 = A\mathbf{X} + \mathbf{Z}$$
$$\mathbf{Y}_2 = B\mathbf{X} + \mathbf{Z}$$

where $Z \sim \mathcal{N}(0, I)$, and A, B are matrices. Assume $\mathrm{tr}(\mathrm{E}(\mathbf{X}\mathbf{X}^T)) \leq P$.

On the projection of the region on $R_0 = 0$

- Marton's Inner Bound and UVW outer bound coincide [WSS06]

Steinberg-Weingarten-Shamai

Vector Gaussian Broadcast Channel

Consider the broadcast channel (of interest in multi-antenna wireless communication)

$$\mathbf{Y}_1 = A\mathbf{X} + \mathbf{Z}$$
$$\mathbf{Y}_2 = B\mathbf{X} + \mathbf{Z}$$

where $Z \sim \mathcal{N}(0, I)$, and A, B are matrices. Assume $\mathrm{tr}(\mathrm{E}(\mathbf{X}\mathbf{X}^T)) \leq P$.

On the projection of the region on $R_0 = 0$

- Marton's Inner Bound and UVW outer bound coincide [WSS06]

Idea:

- Identified a parameterized family of **channel enhancements** that each had a degraded structure
- Used Entropy Power Inequality to evaluate the outer bound for the enhanced channel

Steinberg-Weingarten-Shamai

Vector Gaussian Broadcast Channel

Consider the broadcast channel (of interest in multi-antenna wireless communication)

$$\mathbf{Y}_1 = A\mathbf{X} + \mathbf{Z}$$
$$\mathbf{Y}_2 = B\mathbf{X} + \mathbf{Z}$$

where $Z \sim \mathcal{N}(0, I)$, and A, B are matrices. Assume $\mathrm{tr}(\mathrm{E}(\mathbf{X}\mathbf{X}^T)) \leq P$.

On the projection of the region on $R_0 = 0$

- Marton's Inner Bound and UVW outer bound coincide [WSS06]

Idea:

- Identified a parameterized family of **channel enhancements** that each had a degraded structure
- Used Entropy Power Inequality to evaluate the outer bound for the enhanced channel
- Used the Dirty Paper Coding [Cos83] inspired auxiliaries to evaluate Marton's Inner Bound
- Showed that these two "relaxed versions" of inner and outer bounds coincide and sandwich the true capacity region.

Steinberg-Weingarten-Shamai

Costa

Vector Gaussian Broadcast Channel

Consider the broadcast channel (of interest in multi-antenna wireless communication)

$$\mathbf{Y}_1 = A\mathbf{X} + \mathbf{Z}$$
$$\mathbf{Y}_2 = B\mathbf{X} + \mathbf{Z}$$

where $Z \sim \mathcal{N}(0, I)$, and A, B are matrices. Assume $\mathrm{tr}(\mathrm{E}(\mathbf{X}\mathbf{X}^T)) \leq P$.

For the entire capacity region

- Marton's Inner Bound and UVW outer bound coincide [GN14]

Idea:

- Used the Dirty Paper Coding [Cos83] inspired auxiliaries to evaluate Marton's Inner Bound
- Evaluated the outer bound directly

Geng

Vector Gaussian Broadcast Channel

Consider the broadcast channel (of interest in multi-antenna wireless communication)

$$\mathbf{Y}_1 = A\mathbf{X} + \mathbf{Z}$$
$$\mathbf{Y}_2 = B\mathbf{X} + \mathbf{Z}$$

where $Z \sim \mathcal{N}(0, I)$, and A, B are matrices. Assume $\mathrm{tr}(\mathrm{E}(\mathbf{X}\mathbf{X}^T)) \le P$.

For the entire capacity region

- Marton's Inner Bound and UVW outer bound coincide [GN14]

Idea:

- Used the Dirty Paper Coding [Cos83] inspired auxiliaries to evaluate Marton's Inner Bound
- Evaluated the outer bound directly
  - ◇ Step 1: Identify **sub-additive functionals** using the arguments in the outer bound (with additivity only under some independence constraints)
  - ◇ Step 2: Used rotated version of two independent copies of the maximizers and use the above argument to deduce **independence of the rotated versions**
  - ◇ 3: Use Bernstein's characterization theorem to conclude that optimizers must be Gaussian

Geng

Illustration of the Idea

Korner-Marton functional - extremal distribution

Maximize, for $\lambda > 1$, the value of the functional

$$\mathcal{C}_{\mu_X}[h(AX+Z) - \lambda h(BX+Z)]$$

over $X : \mathrm{E}(XX^T) \preceq K$, where A, B are invertible matrices and $Z \sim \mathcal{N}(0, I)$.

The **upper concave envelope** $\mathcal{C}_x[f(x)]$ of f is equivalently characterized by :

- Smallest concave upper bound: $\mathcal{C}_x[f(x)] := \inf_{g \geq f \text{ concave}} g(x)$.
- **Largest convex combination**: $\mathcal{C}_x[f(x)] := \sup_{\substack{p(\tilde{x}) \\ \mathrm{E}[\tilde{X}]=x}} \mathrm{E}[f(\tilde{X})]$.
- Fenchel dual characterization: $\mathcal{C}_x[f(x)] := \inf_\alpha \left(\sup_{\hat{x}} \left(f(\hat{x}) - \langle \alpha, \hat{x} \rangle \right) + \langle \alpha, x \rangle \right)$, where the infimum is over continuous linear functional α.

Illustration of the Idea

Korner-Marton functional - extremal distribution

Maximize, for $\lambda > 1$, the value of the functional

$$\mathcal{C}_{\mu_X}[h(AX + Z) - \lambda h(BX + Z)]$$

over $X : \mathrm{E}(XX^T) \preceq K$, where A, B are invertible matrices and $Z \sim \mathcal{N}(0, I)$.

We will see that the maximum value is

$$h(AX_* + Z) - \lambda h(BX_* + Z),$$

where $X_* \sim \mathcal{N}(0, K')$ for some $K' \preceq K$.

Illustration of the Idea

Korner-Marton functional - extremal distribution

Maximize, for $\lambda > 1$, the value of the functional

$$\mathcal{C}_{\mu_X}[h(AX+Z) - \lambda h(BX+Z)]$$

over $X : \mathrm{E}(XX^T) \preceq K$, where A, B are invertible matrices and $Z \sim \mathcal{N}(0, I)$.

We will see that the maximum value is

$$h(AX_* + Z) - \lambda h(BX_* + Z),$$

where $X_* \sim \mathcal{N}(0, K')$ for some $K' \preceq K$.

Lemma: $\mathcal{C}_{\mu_X}[h(AX+Z) - \lambda h(BX+Z)]$ is sub-additive.

Proof: For any μ_{X_1,X_2} and $\mu_{U|X_1,X_2}$

$$\begin{aligned} & h(AX_1+Z_1, AX_2+Z_2|U) - \lambda h(BX_1+Z_1, BX_2+Z_2|U) \\ & = h(AX_1+Z_1|U_1) - \lambda h(BX_1+Z_1|U_1) + h(AX_2+Z_2|U_2) - \lambda h(BX_2+Z_2|U_2) \\ & \quad -(\lambda-1)I(AX_2+Z_2; BX_1+Z_1|U), \end{aligned}$$

where $U_1 = (U, AX_2+Z_2)$ and $U_2 = (U, BX_1+Z_1)$

Illustration of the Idea

Korner-Marton functional - extremal distribution

Maximize, for $\lambda > 1$, the value of the functional

$$\mathcal{C}_{\mu_X}[h(AX+Z) - \lambda h(BX+Z)]$$

over $X : \mathrm{E}(XX^T) \preceq K$, where A, B are invertible matrices and $Z \sim \mathcal{N}(0, I)$.

We will see that the maximum value is

$$h(AX_* + Z) - \lambda h(BX_* + Z),$$

where $X_* \sim \mathcal{N}(0, K')$ for some $K' \preceq K$.

Lemma: $\mathcal{C}_{\mu_X}[h(AX+Z) - \lambda h(BX+Z)]$ is sub-additive.

Proof: For any μ_{X_1,X_2}

$$\begin{aligned} &\mathcal{C}_{\mu_{X_1,X_2}}[h(AX_1+Z_1, AX_2+Z_2) - \lambda h(BX_1+Z_1, BX_2+Z_2)] \\ &\leq \mathcal{C}_{\mu_{X_1}}[h(AX_1+Z_1) - \lambda h(BX_1+Z_1)] \\ &\quad + \mathcal{C}_{\mu_{X_2}}[h(AX_2+Z_2) - \lambda h(BX_2+Z_2)] \end{aligned}$$

Gaussian optimality: ctd..

Let $(U_\dagger, X_\dagger)$ be a maximizer, i.e.

$$V = \max_{\mu_X} \mathcal{C}_{\mu_X}[h(AX+Z) - \lambda h(BX+Z)] = h(AX_\dagger + Z|U_\dagger) - \lambda h(BX_\dagger + Z|U_\dagger).$$

Let (X_a, U_a) and (X_b, U_b) be i.i.d. according to $(U_\dagger, X_\dagger)$.

Gaussian optimality: ctd..

Let $(U_\dagger, X_\dagger)$ be a maximizer, i.e.

$$V = \max_{\mu_X} \mathcal{C}_{\mu_X}[h(AX+Z) - \lambda h(BX+Z)] = h(AX_\dagger + Z|U_\dagger) - \lambda h(BX_\dagger + Z|U_\dagger).$$

Let (X_a, U_a) and (X_b, U_b) be i.i.d. according to $(U_\dagger, X_\dagger)$.

Setting $U = (U_a, U_b)$, $X_+ = \frac{X_a+X_b}{\sqrt{2}}$ and $X_- = \frac{X_a-X_b}{\sqrt{2}}$ the proof of sub-additivity yields

$$\begin{aligned} 2V &= \mathcal{C}_{\mu_{X_1,X_2}}[h(AX_1+Z_1, AX_2+Z_2) - \lambda h(BX_1+Z_1, BX_2+Z_2)]\Big|_{(\mu_{X_+,X_-})} \\ &\leq \mathcal{C}_{\mu_{X_1}}[h(AX_1+Z_1) - \lambda h(BX_1+Z_1)]\Big|_{\mu_{X_+}} \\ &\quad + \mathcal{C}_{\mu_{X_2}}[h(AX_2+Z_2) - \lambda h(BX_2+Z_2)]\Big|_{\mu_{X_-}} \\ &\quad\quad -(\lambda-1)I(AX_- + Z_2; BX_+ + Z_1|U_a, U_b) \\ &\leq V + V \end{aligned}$$

Gaussian optimality: ctd..

Let $(U_\dagger, X_\dagger)$ be a maximizer, i.e.

$$V = \max_{\mu_X} \mathcal{C}_{\mu_X}[h(AX+Z) - \lambda h(BX+Z)] = h(AX_\dagger + Z|U_\dagger) - \lambda h(BX_\dagger + Z|U_\dagger).$$

Let (X_a, U_a) and (X_b, U_b) be i.i.d. according to $(U_\dagger, X_\dagger)$.

Setting $U = (U_a, U_b)$, $X_+ = \frac{X_a+X_b}{\sqrt{2}}$ and $X_- = \frac{X_a-X_b}{\sqrt{2}}$ the proof of sub-additivity yields

$$\begin{aligned}
2V &= \mathcal{C}_{\mu_{X_1,X_2}}[h(AX_1+Z_1, AX_2+Z_2) - \lambda h(BX_1+Z_1, BX_2+Z_2)]\Big|_{(\mu_{X_+,X_-})} \\
&\le \mathcal{C}_{\mu_{X_1}}[h(AX_1+Z_1) - \lambda h(BX_1+Z_1)]\Big|_{\mu_{X_+}} \\
&\quad + \mathcal{C}_{\mu_{X_2}}[h(AX_2+Z_2) - \lambda h(BX_2+Z_2)]\Big|_{\mu_{X_-}} \\
&\quad\quad -(\lambda-1)I(AX_- + Z_2; BX_+ + Z_1|U_a, U_b) \\
&\le V + V
\end{aligned}$$

Therefore: we get that conditioned on (U_a, U_b): $X_+ \perp X_-$.

Gaussian optimality: ctd..

Let $(U_\dagger, X_\dagger)$ be a maximizer, i.e.

$$V = \max_{\mu_X} \mathcal{C}_{\mu_X}[h(AX+Z) - \lambda h(BX+Z)] = h(AX_\dagger + Z|U_\dagger) - \lambda h(BX_\dagger + Z|U_\dagger).$$

Let (X_a, U_a) and (X_b, U_b) be i.i.d. according to $(U_\dagger, X_\dagger)$.

Note: Thus, conditioned on (U_a, U_b):

- $X_a \perp X_b$ (from construction)
- $(X_a + X_b) \perp (X_a - X_b)$ (from proof of sub-additivity)

Gaussian optimality: ctd..

Let $(U_\dagger, X_\dagger)$ be a maximizer, i.e.

$$V = \max_{\mu_X} \mathcal{C}_{\mu_X}[h(AX+Z) - \lambda h(BX+Z)] = h(AX_\dagger + Z|U_\dagger) - \lambda h(BX_\dagger + Z|U_\dagger).$$

Let (X_a, U_a) and (X_b, U_b) be i.i.d. according to $(U_\dagger, X_\dagger)$.

Note: Thus, conditioned on (U_a, U_b):

- $X_a \perp X_b$ (from construction)
- $(X_a + X_b) \perp (X_a - X_b)$ (from proof of sub-additivity)
- Implies that conditioned on (U_a, U_b): X_a, X_b are Gaussian
  - ◇ Characterization of Gaussians (Bernstein '40s)
  - ◇ Proof: Using characteristic functions (Fourier transforms)

Let $(U_\dagger, X_\dagger)$ be a maximizer, i.e.

$$V = \max_{\mu_X} \mathcal{C}_{\mu_X}[h(AX+Z) - \lambda h(BX+Z)] = h(AX_\dagger + Z|U_\dagger) - \lambda h(BX_\dagger + Z|U_\dagger).$$

Let (X_a, U_a) and (X_b, U_b) be i.i.d. according to $(U_\dagger, X_\dagger)$.

Note: Thus, conditioned on (U_a, U_b):

- $X_a \perp X_b$ (from construction)
- $(X_a + X_b) \perp (X_a - X_b)$ (from proof of sub-additivity)
- Implies that conditioned on (U_a, U_b): X_a, X_b are Gaussian
  - ◇ Characterization of Gaussians (Bernstein '40s)
  - ◇ Proof: Using characteristic functions (Fourier transforms)

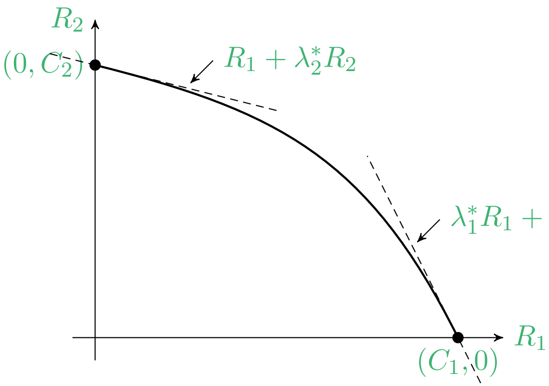
Note: There are some similarities with work of Lieb and Barthe (90s)

They also use rotations (but not information measures and its algebra)

Slope at the axis points of the capacity region

The capacity region of a **generic broadcast channel** has the following shape.

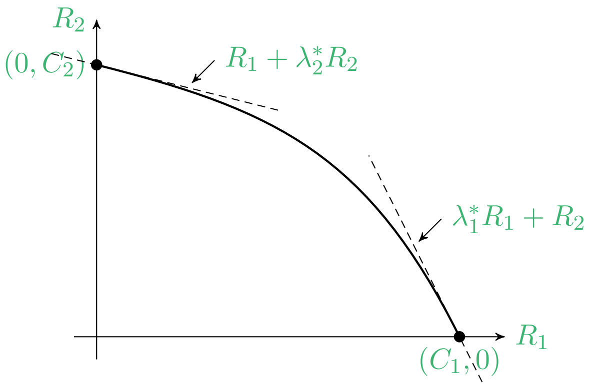
Question: what is the slope of the capacity region at the points $(C_1, 0)$ and $(0, C_2)$.



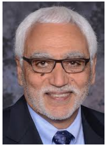
Slope at the axis points of the capacity region

The capacity region of a **generic broadcast channel** has the following shape.

Question: what is the slope of the capacity region at the points $(C_1, 0)$ and $(0, C_2)$.



Theorem: The slope of Marton's achievable region ($\mathcal{M}(T_1, T_2)$) **matches** the slope of UVW outer bound ($\mathcal{O}(T_1, T_2)$) at the points $(C_1, 0)$ and $(0, C_2)$. [NKG16]



El Gamal

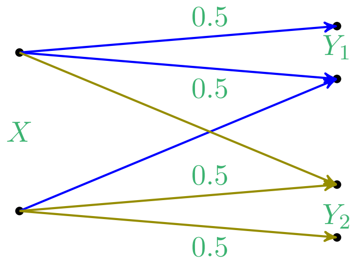
H. Kim

Comparing $\mathcal{M}(T_1, T_2)$ with $\mathcal{O}(T_1, T_2)$

Could it be possible that $\mathcal{M}(T_1, T_2) = \mathcal{O}(T_1, T_2)$ (and hence $\mathcal{C}(T_1, T_2)$)?

- Both regions are "hard" to evaluate
- $\mathcal{M}(T_1, T_2)$ was not computable
- They coincided for lots of classes of channels, some with very limited structure such as vector Gaussian
- Both bounds gave the same slope at axis points

Comparing $\mathcal{M}(T_1, T_2)$ with $\mathcal{O}(T_1, T_2)$



Binary skew-symmetric broadcast channel (BSSC)

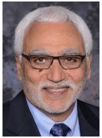
Conjectured [NW08] that for the above channel,

$$I(U;Y_1) + I(V;Y_2) - I(U;V) \le \max\{I(X;Y_1), I(X;Y_2)\}$$

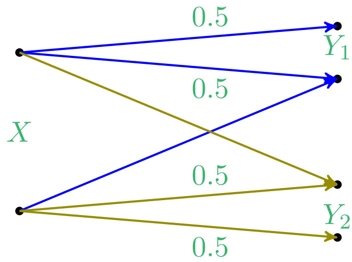
for any $(U,V) : (U,V) \multimap X \multimap (Y_1,Y_2)$

- If the inequality is true then $\max_{(R_1,R_2)\in\mathcal{M}(T_1,T_2)} R_1 + R_2 \in (0.3615, 0.3617)$ [HP79]
- On the other hand [NE07] $\max_{(R_1,R_2)\in\mathcal{O}(T_1,T_2)} R_1 + R_2 \in (0.3725, 0.3726)$

V. Wang



El Gamal



Binary skew-symmetric broadcast channel (BSSC)

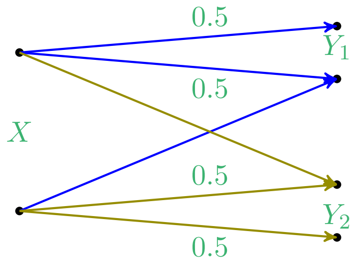
Marton's region was shown to be computable [GA09]

- Suffices to consider $|\mathcal{U}| \leq |\mathcal{X}|$, $|\mathcal{V}| \leq |\mathcal{X}|$ and $|\mathcal{W}| \leq |\mathcal{X}| + 4$ to evaluate $\mathcal{M}(T_1, T_2)$
- **Pertubation** approach to bound cardinality of **extremal** auxiliaries

Anantharam

Gohari

Comparing $\mathcal{M}(T_1, T_2)$ with $\mathcal{O}(T_1, T_2)$



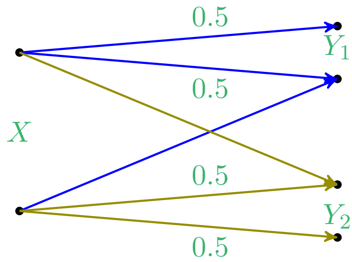
Binary skew-symmetric broadcast channel (BSSC)

Marton's region was shown to be computable [GA09]

- Suffices to consider $|\mathcal{U}| \leq |\mathcal{X}|$, $|\mathcal{V}| \leq |\mathcal{X}|$ and $|\mathcal{W}| \leq |\mathcal{X}| + 4$ to evaluate $\mathcal{M}(T_1, T_2)$
- **Pertubation** approach to bound cardinality of **extremal** auxiliaries
- Numerically verified the previous information inequality
- **Proved** that $\mathcal{M}(T_1, T_2)$ cannot match $\mathcal{O}(T_1, T_2)$ for BSSC

Anantharam

Gohari



Binary skew-symmetric broadcast channel (BSSC)

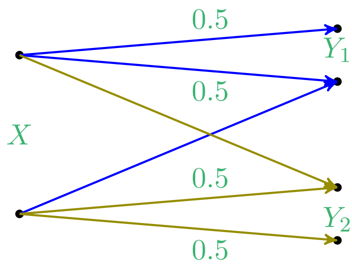
- Proved [JN10] that for the above channel,

$$I(U;Y_1)+I(V;Y_2)-I(U;V) \leq \max\{I(X;Y_1), I(X;Y_2)\}$$

for any $(U,V) : (U,V) - X - (Y_1, Y_2)$
  - ◇ Extending the **perturbation approach**

Jog

Comparing $\mathcal{M}(T_1, T_2)$ with $\mathcal{O}(T_1, T_2)$



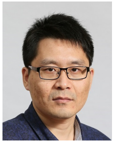
Binary skew-symmetric broadcast channel (BSSC)

Extended the same inequality to **any binary input** broadcast channel [Gen+13a], i.e.

$$I(U;Y_1)+I(V;Y_2)-I(U;V) \leq \max\{I(X;Y_1), I(X;Y_2)\}$$

for any $(U,V) : (U,V) \multimap X \multimap (Y_1, Y_2)$

- $\max_{(R_1,R_2)\in\mathcal{M}(T_1,T_2)} R_1 + R_2$ is given by randomized-time-division
- Immediate to evaluate the sum-rate for any binary input broadcast channel



Geng

Jog

V. Wang

Optimality/sub-optimality of $\mathcal{M}(T_1, T_2)$ and/or $\mathcal{O}(T_1, T_2)$

Known: $\mathcal{M}(T_1, T_2)$ is optimal *if and only if*

$$\mathcal{M}(T_1 \otimes T_1, T_2 \otimes T_2) = \mathcal{M}(T_1, T_2) \oplus \mathcal{M}(T_1, T_2) \quad \forall (T_1, T_2).$$

Optimality/sub-optimality of $\mathcal{M}(T_1, T_2)$ and/or $\mathcal{O}(T_1, T_2)$

Known: $\mathcal{M}(T_1, T_2)$ is optimal *if and only if*

$$\mathcal{M}(T_1 \otimes T_1, T_2 \otimes T_2) = \mathcal{M}(T_1, T_2) \oplus \mathcal{M}(T_1, T_2) \quad \forall (T_1, T_2).$$

Investigated product broadcast channels [Gen+14]

- Demonstrated a product broadcast channel where

$$\begin{aligned}\mathcal{C}(T_1 \otimes \hat{T}_1, T_2 \otimes \hat{T}_2) &= \mathcal{M}(T_1 \otimes \hat{T}_1, T_2 \otimes \hat{T}_2) \\ &\supsetneq \mathcal{C}(T_1, T_2) \oplus \mathcal{C}(\hat{T}_1, \hat{T}_2) = \mathcal{M}(T_1, T_2) \oplus \mathcal{M}(\hat{T}_1, \hat{T}_2)\end{aligned}$$

- Developed sufficient conditions for $\mathcal{M}(T_1, T_2)$ to be optimal
  - ◇ **Key idea**: Min-max interchange
  - ◇ Establish capacity regions of new classes of **product** broadcast channels
  - ◇ Developed a new outer bound for product broadcast channels

Geng

Gohari

Yu

Optimality/sub-optimality of $\mathcal{M}(T_1, T_2)$ and/or $\mathcal{O}(T_1, T_2)$

Known: $\mathcal{M}(T_1, T_2)$ is optimal *if and only if*

$$\mathcal{M}(T_1 \otimes T_1, T_2 \otimes T_2) = \mathcal{M}(T_1, T_2) \oplus \mathcal{M}(T_1, T_2) \quad \forall (T_1, T_2).$$

Investigated product broadcast channels [Gen+14]

- Demonstrated a product broadcast channel where

$$\begin{aligned}\mathcal{C}(T_1 \otimes \hat{T}_1, T_2 \otimes \hat{T}_2) &= \mathcal{M}(T_1 \otimes \hat{T}_1, T_2 \otimes \hat{T}_2) \\ &\supsetneq \mathcal{C}(T_1, T_2) \oplus \mathcal{C}(\hat{T}_1, \hat{T}_2) = \mathcal{M}(T_1, T_2) \oplus \mathcal{M}(\hat{T}_1, \hat{T}_2)\end{aligned}$$

Geng

- Developed sufficient conditions for $\mathcal{M}(T_1, T_2)$ to be optimal
  - **Key idea**: Min-max interchange
  - Establish capacity regions of new classes of **product** broadcast channels
  - Developed a new outer bound for product broadcast channels

Gohari

- Demonstrated a product broadcast channel where

$$\mathcal{M}(T_1 \otimes \hat{T}_1, T_2 \otimes \hat{T}_2) = \mathcal{C}(T_1 \otimes \hat{T}_1, T_2 \otimes \hat{T}_2) \subsetneq \mathcal{O}(T_1 \otimes \hat{T}_1, T_2 \otimes \hat{T}_2).$$

Yu

Hence, sub-optimality of $\mathcal{O}(T_1 \otimes \hat{T}_1, T_2 \otimes \hat{T}_2)$.

On optimality of $\mathcal{M}(T_1, T_2)$

Let $T_1(y_1|x)$ and $T_2(y_2|x)$ be given channels, $\alpha \in [0,1]$, and $\lambda \geq 1$. Define

$$F_{\lambda,\alpha}^{T_1,T_2}(\mu_X) := \mathcal{C}_{\mu_X}\left[(\lambda-\alpha)H(Y_1) - \alpha H(Y_2) + \max_{p(u,v|x)}\{\lambda I(U;Y_1) + I(V;Y_2) - I(U;V)\}\right]$$

On optimality of $\mathcal{M}(T_1, T_2)$

Let $T_1(y_1|x)$ and $T_2(y_2|x)$ be given channels, $\alpha \in [0,1]$, and $\lambda \geq 1$. Define

$$F_{\lambda,\alpha}^{T_1,T_2}(\mu_X) := \mathcal{C}_{\mu_X}\left[(\lambda-\alpha)H(Y_1) - \alpha H(Y_2) + \max_{p(u,v|x)}\{\lambda I(U;Y_1) + I(V;Y_2) - I(U;V)\}\right]$$

Results [AGN13; AGN19]

- To evaluate $F_{\lambda,\alpha}^{T_1,T_2}(\mu_X)$ it suffices to consider (U,V): $|\mathcal{U}| + |\mathcal{V}| \leq |\mathcal{X}| + 1$
  - ◇ Employs the perturbation approach to obtain cardinality bounds
  - ◇ Feasible to get very good approximations by fine grid search for small $|\mathcal{X}|$ (say, $|\mathcal{X}| \leq 4$)

Anantharam

Gohari

On optimality of $\mathcal{M}(T_1, T_2)$

Let $T_1(y_1|x)$ and $T_2(y_2|x)$ be given channels, $\alpha \in [0,1]$, and $\lambda \geq 1$. Define

$$F_{\lambda,\alpha}^{T_1,T_2}(\mu_X) := \mathcal{C}_{\mu_X}\left[(\lambda-\alpha)H(Y_1) - \alpha H(Y_2) + \max_{p(u,v|x)}\{\lambda I(U;Y_1) + I(V;Y_2) - I(U;V)\}\right]$$

Results [AGN13; AGN19]

Anantharam

- To evaluate $F_{\lambda,\alpha}^{T_1,T_2}(\mu_X)$ it suffices to consider (U,V): $|\mathcal{U}| + |\mathcal{V}| \leq |\mathcal{X}| + 1$
  - ◇ Employs the perturbation approach to obtain cardinality bounds
  - ◇ Feasible to get very good approximations by fine grid search for small $|\mathcal{X}|$ (say, $|\mathcal{X}| \leq 4$)
- If the following **sub-additivity** (over product channels) holds: $F_{\lambda,\alpha}^{T_1\otimes\hat{T}_1,T_2\otimes\hat{T}_2}(\mu_{X_1,\hat{X}_1}) \leq F_{\lambda,\alpha}^{T_1,T_2}(\mu_{X_1}) + F_{\lambda,\alpha}^{\hat{T}_1,\hat{T}_2}(\mu_{\hat{X}_1})$ then $\mathcal{M}(T_1,T_2)$ is optimal.
- Numerical simulations have not yet yielded counterexamples
- Can prove the sub-additivity when $\alpha = 0$ or $\alpha = 1$.

Gohari

On optimality of $\mathcal{M}(T_1, T_2)$

Let $T_1(y_1|x)$ and $T_2(y_2|x)$ be given channels, $\alpha \in [0,1]$, and $\lambda \geq 1$. Define

$$F_{\lambda,\alpha}^{T_1,T_2}(\mu_X) := \mathcal{C}_{\mu_X}\left[(\lambda-\alpha)H(Y_1) - \alpha H(Y_2) + \max_{p(u,v|x)}\{\lambda I(U;Y_1) + I(V;Y_2) - I(U;V)\}\right]$$

Results [AGN13; AGN19]

- To evaluate $F_{\lambda,\alpha}^{T_1,T_2}(\mu_X)$ it suffices to consider (U,V): $|\mathcal{U}| + |\mathcal{V}| \leq |\mathcal{X}| + 1$
  - Employs the perturbation approach to obtain cardinality bounds
  - Feasible to get very good approximations by fine grid search for small $|\mathcal{X}|$ (say, $|\mathcal{X}| \leq 4$)

Anantharam

- If the following **sub-additivity** (over product channels) holds:
  $F_{\lambda,\alpha}^{T_1\otimes\hat{T}_1,T_2\otimes\hat{T}_2}(\mu_{X_1,\hat{X}_1}) \leq F_{\lambda,\alpha}^{T_1,T_2}(\mu_{X_1}) + F_{\lambda,\alpha}^{\hat{T}_1,\hat{T}_2}(\mu_{\hat{X}_1})$
  then $\mathcal{M}(T_1, T_2)$ is optimal.

Gohari

- Numerical simulations have not yet yielded counterexamples
  - **Request**: Can others also try numerical experiments.

- Can prove the sub-additivity when $\alpha = 0$ or $\alpha = 1$.

On optimality of $\mathcal{M}(T_1, T_2)$

Let $T_1(y_1|x)$ and $T_2(y_2|x)$ be given channels, $\alpha \in [0,1]$, and $\lambda \geq 1$. Define

$$F_{\lambda,\alpha}^{T_1,T_2}(\mu_X) := \mathcal{C}_{\mu_X}\left[(\lambda-\alpha)H(Y_1) - \alpha H(Y_2) + \max_{p(u,v|x)}\{\lambda I(U;Y_1) + I(V;Y_2) - I(U;V)\}\right]$$

Results [AGN13; AGN19]

- To evaluate $F_{\lambda,\alpha}^{T_1,T_2}(\mu_X)$ it suffices to consider (U,V): $|\mathcal{U}| + |\mathcal{V}| \leq |\mathcal{X}| + 1$
  - ◇ Employs the perturbation approach to obtain cardinality bounds
  - ◇ Feasible to get very good approximations by fine grid search for small $|\mathcal{X}|$ (say, $|\mathcal{X}| \leq 4$)

Anantharam

- If the following **sub-additivity** (over product channels) holds:
  $F_{\lambda,\alpha}^{T_1\otimes\hat{T}_1,T_2\otimes\hat{T}_2}(\mu_{X_1,\hat{X}_1}) \leq F_{\lambda,\alpha}^{T_1,T_2}(\mu_{X_1}) + F_{\lambda,\alpha}^{\hat{T}_1,\hat{T}_2}(\mu_{\hat{X}_1})$
  then $\mathcal{M}(T_1,T_2)$ is optimal.

Gohari

- Numerical simulations have not yet yielded counterexamples
- Can prove the sub-additivity when $\alpha = 0$ or $\alpha = 1$.

Therefore **current evidence** points to a **potential optimality** of $\mathcal{M}(T_1,T_2)$.

On optimality of $\mathcal{M}(T_1, T_2)$

Let $T_1(y_1|x)$ and $T_2(y_2|x)$ be given channels, $\alpha \in [0,1]$, and $\lambda \geq 1$. Define

$$F_{\lambda,\alpha}^{T_1,T_2}(\mu_X) := \mathcal{C}_{\mu_X}\left[(\lambda-\alpha)H(Y_1) - \alpha H(Y_2) + \max_{p(u,v|x)}\{\lambda I(U;Y_1) + I(V;Y_2) - I(U;V)\}\right]$$

A natural extension (projection of **upward closed** decoding sets) of Martons region to **three receivers** is shown to be strictly sub-optimal. [PP18]

- Used algebraic structured codes

Padakandla

Pradhan

On optimality of $\mathcal{M}(T_1, T_2)$

Let $T_1(y_1|x)$ and $T_2(y_2|x)$ be given channels, $\alpha \in [0,1]$, and $\lambda \geq 1$. Define

$$F_{\lambda,\alpha}^{T_1,T_2}(\mu_X) := \mathcal{C}_{\mu_X}\left[(\lambda-\alpha)H(Y_1) - \alpha H(Y_2) + \max_{p(u,v|x)}\{\lambda I(U;Y_1) + I(V;Y_2) - I(U;V)\}\right]$$

A natural extension (projection of **upward closed** decoding sets) of Martons region to **three receivers** is shown to be strictly sub-optimal. [PP18]

- Used algebraic structured codes

Padakandla

Pradhan

Structured codes traces origins in network information theory to

- Modulo-two-sum problem [KM79]

Körner

Marton

Advances over the classical results - Main part

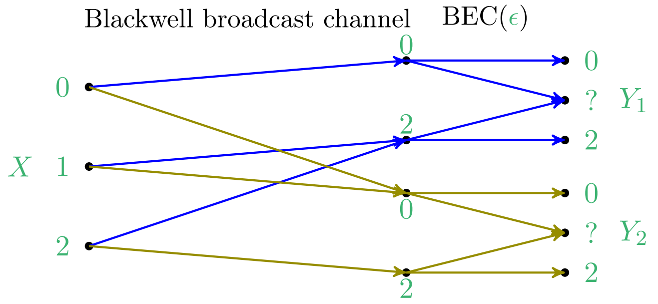
Investigate optimality/sub-optimality of superposition coding region (2007-2018)

- More instances where it matches capacity region
  - ◇ Evaluation of inner and outer bounds (ideas involved)
- Settings where it is sub-optimal

Investigate optimality/sub-optimality of Marton's inner bound and UVW outer bound (2008-2015)

- More instances where the bounds coincide
  - ◇ Evaluation of inner and outer bounds (ideas involved)
- Settings where there is a gap between the bounds

Recent advances (2019 -) and remarks

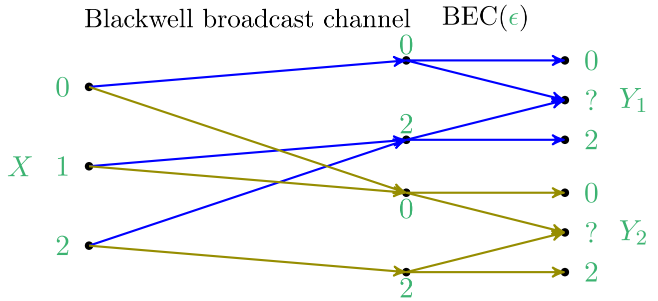


Erased Blackwell Broadcast Channel

- What is the capacity region of the above channel?
- Is it $(1-\epsilon) \times \mathcal{C}(Blackwell)$?
  - ◇ Outer bounds seem to suggest so (each mutual information term of the form $I(U;Y_i|V)$ scales by $(1-\epsilon)$)
- On the other hand, $\mathcal{M}(Blackwell)$ has a term $I(U;V)$ that does not scale by $(1-\epsilon)$

Gohari

Recent Advances



Erased Blackwell Broadcast Channel

- The capacity region $\subsetneq (1-\epsilon)\mathcal{C}(Blackwell)$ [GN20]
  - Notion of "auxiliary receiver" to develop two new outer bounds for broadcast channel (both strictly improve on $\mathcal{O}(T_1, T_2)$)
  - The capacity region is still **open** for this setting
  - Even determining the **corner-point** of the form $(1-\epsilon, R_2^*)$ is **open**.

Gohari

Concluding Remarks

The ideas involved in developing these results have proved useful in other settings as well

- Sub-optimality of Han-Kobayashi achievable region for the Interference channel
  - ◇ Ideas involved in evaluation of the regions (non-covex optimization problems)
- New outer bounds in Interference, Relay etc
  - ◇ Auxiliary receiver approach
- (Re)-discovered connections to hypercontractivity

Concluding Remarks

The ideas involved in developing these results have proved useful in other settings as well

- Sub-optimality of Han-Kobayashi achievable region for the Interference channel
  - ◇ Ideas involved in evaluation of the regions (non-covex optimization problems)
- New outer bounds in Interference, Relay etc
  - ◇ Auxiliary receiver approach
- (Re)-discovered connections to hypercontractivity

Working on these settings also gave rise to a **meta-conjecture**

- *Global tensorization if and only if local tensorization* for a family of functionals (that frequently arise in multiuser settings) [Nai20]
- Some counterexamples were inspired by this meta-conjecture

Concluding Remarks

The ideas involved in developing these results have proved useful in other settings as well

- Sub-optimality of Han-Kobayashi achievable region for the Interference channel
  - ◇ Ideas involved in evaluation of the regions (non-covex optimization problems)
- New outer bounds in Interference, Relay etc
  - ◇ Auxiliary receiver approach
- (Re)-discovered connections to hypercontractivity

Working on these settings also gave rise to a **meta-conjecture**

- *Global tensorization if and only if local tensorization* for a family of functionals (that frequently arise in multiuser settings) [Nai20]
- Some counterexamples were inspired by this meta-conjecture

Most of the progress have been achieved by considering the oxymoronic **simple yet hard** instances.

- Listed a few **open** problems of the above flavor in this talk as well

Thanks

Thanks

- My collaborators for this wonderful journey
- The organizers for the opportunity
- The virtual audience for sparing your valuable time

References I

[Cov72] T. Cover. “Broadcast Channels”. In: IEEE Trans. Info. Theory IT-18 (Jan. 1972), pp. 2–14.

[Ber73] P. F. Bergmans. “Coding theorem for broadcast channels with degraded components”. In: IEEE Trans. Info. Theory IT-15 (Mar. 1973), pp. 197–207.

[SW73] D. Slepian and J. Wolf. “Noiseless coding of correlated information sources”. In: IEEE Transactions on Information Theory 19.4 (1973), pp. 471–480. DOI: 10.1109/TIT.1973.1055037.

[WZ73] A. Wyner and J. Ziv. “A Theorem on the Entropy of Certain Binary Sequences and Applications: Part I”. In: IEEE Trans. Inform. Theory IT-19.6 (Nov. 1973), pp. 769–772.

[Gal74] R. G. Gallager. “Capacity and coding for degraded broadcast channels”. In: Probl. Peredac. Inform. 10(3) (1974), pp. 3–14.

[Hug75] D. Hughes-Hartogs. “The capacity of the degraded spectral Gaussian broadcast channel”. PhD thesis. Stanford University, 1975.

References II

[KM75] J. Körner and K. Marton. "Comparison of two noisy channels". In: Topics in Inform. Theory(ed. by I. Csiszar and P.Elias), Keszthely, Hungary (Aug. 1975), pp. 411–423.

[Gel77] S. I. Gel'fand. "Capacity region of a broadcast channel". Russian. In: Probl. Peredachi Inf. 13.3 (1977), pp. 106–108. ISSN: 0555-2923.

[KM77a] J. Körner and K. Marton. "General broadcast channels with degraded message sets". In: IEEE Trans. Info. Theory IT-23 (Nov. 1977), pp. 60–64.

[KM77b] J. Körner and K. Marton. "Images of a set via two channels and their role in Mult-User Communication". In: IEEE Trans. Info. Theory IT-23 (Jan. 1977), pp. 751–761.

[Mar77] K. Marton. "The capacity region of deterministic broadcast channels". In: Trans. Int. Symp. Inform. Theory (1977).

[Pol77] G. S. Poltyrev. "Capacity for parallel broadcast channels with degraded components". Russian. In: Probl. Peredachi Inf. 13.2 (1977), pp. 23–35. ISSN: 0555-2923.

[Pin78] M. S. Pinsker. "Capacity of noiseless broadcasting channels". Russian. In: Probl. Peredachi Inf. 14.2 (1978), pp. 28–34. ISSN: 0555-2923.

[El 79] A. El Gamal. "The Capacity of a Class of Broadcast Channels". In: IEEE Trans. Info. Theory IT-25 (Mar. 1979), pp. 166–169.

[HP79] B. Hajek and M. Pursley. "Evaluation of an Achievable Rate Region for the Broadcast Channel". In: IEEE Trans. Info. Theory IT-25 (Jan. 1979), pp. 36–46.

[KM79] J. Körner and K. Marton. "How to encode the modulo-two sum of binary sources (Corresp.)". In: IEEE Transactions on Information Theory 25.2 (1979), pp. 219–221. DOI: 10.1109/TIT.1979.1056022.

[Mar79] K. Marton. "A Coding theorem for the discrete memoryless broadcast channel". In: IEEE Trans. Info. Theory IT-25 (May 1979), pp. 306–311.

[El 80] A. El Gamal. "Capacity of the product and sum of two unmatched broadcast channels". English. In: Probl. Inf. Transm. 16 (1980), pp. 1–16. ISSN: 0032-9460; 1608-3253/e.

References IV

[GP80] S. I. Gel'fand and M. S. Pinsker. "Capacity of a broadcast channel with one deterministic component". English. In: Probl. Inf. Transm. 16 (1980), pp. 17–25. ISSN: 0032-9460; 1608-3253/e.

[Cos83] M. Costa. "Writing on dirty paper (Corresp.)". In: Information Theory, IEEE Transactions on 29.3 (May 1983), pp. 439–441. ISSN: 0018-9448. DOI: 10.1109/TIT.1983.1056659.

[Kör84] J. Körner. "OPEC or a Basic Problem in Source Networks". In: IEEE Trans. Inf. Theor. 30.1 (Sept. 1984), pp. 68–77. ISSN: 0018-9448. DOI: 10.1109/TIT.1984.1056841. URL: https://doi.org/10.1109/TIT.1984.1056841.

[CG87] T. Cover and B. Gopinath, eds. Open problems in communication and computation. eng. New York: Springer-Verlag, 1987. ISBN: 0387966218.

[Wei00] K. Weihrauch. Computable Analysis: An Introduction. Berlin, Heidelberg: Springer-Verlag, 2000. ISBN: 3540668179.

References V

[WSS06] H. Weingarten, Y. Steinberg, and S. Shamai. "The Capacity Region of the Gaussian Multiple-Input Multiple-Output Broadcast Channel". In: Information Theory, IEEE Transactions on 52.9 (Sept. 2006), pp. 3936–3964. ISSN: 0018-9448. DOI: 10.1109/TIT.2006.880064.

[NE07] C. Nair and A. El Gamal. "An outer bound to the Capacity Region of the Broadcast Channel". In: IEEE Trans. Info. Theory IT-53 (Jan. 2007), pp. 350–355.

[NW08] C. Nair and Z. V. Wang. "On the inner and outer bounds for 2-receiver discrete memoryless broadcast channels". In: Proceedings of the ITA Workshop (2008). eprint: cs.IT/0804.3825.

[GA09] A. Gohari and V. Anantharam. "Evaluation of Marton's Inner Bound for the General Broadcast Channel". In: International Symposium on Information Theory (2009), pp. 2462–2466.

[NE09] C. Nair and A. El Gamal. "The Capacity Region of a Class of Three-Receiver Broadcast Channels With Degraded Message Sets". In: Information Theory, IEEE Transactions on 55.10 (Oct. 2009), pp. 4479–4493. ISSN: 0018-9448. DOI: 10.1109/TIT.2009.2027512.

References VI

[JN10] V. Jog and C. Nair. “An information inequality for the BSSC channel”. In: Proceedings of the ITA Workshop (2010).

[Nai10a] C. Nair. “A note on outer bounds for broadcast channel”. In: Presented at International Zurich Seminar (2010). eprint: http://arXiv.org/abs/1101.0640.

[Nai10b] C. Nair. “Capacity Regions of Two New Classes of Two-Receiver Broadcast Channels”. In: Information Theory, IEEE Transactions on 56.9 (Sept. 2010), pp. 4207–4214. ISSN: 0018-9448. DOI: 10.1109/TIT.2010.2054310.

[NW11] C. Nair and Z.V. Wang. “The Capacity Region of the Three Receiver Less Noisy Broadcast Channel”. In: Information Theory, IEEE Transactions on 57.7 (July 2011), pp. 4058–4062. ISSN: 0018-9448. DOI: 10.1109/TIT.2011.2145970.

[NX12] C. Nair and L. Xia. “On Three-Receiver More Capable Channels”. In: 2012 IEEE International Symposium on Information Theory (ISIT’2012). Cambridge, Massachusetts, USA, July 2012, pp. 383–387.

[AGN13] V. Anantharam, A. Gohari, and C. Nair. “Improved cardinality bounds on the auxiliary random variables in Marton’s inner bound”. In: International Symposium on Information Theory. Istanbul, Turkey, July 2013, pp. 1272–1276.

[Gen+13a] Y. Geng et al. “An Information Inequality and Evaluation of Marton’s Inner Bound for Binary Input Broadcast Channels”. In: IEEE Transactions on Information Theory 59.7 (2013), pp. 4095–4105.

[Gen+13b] Y. Geng et al. “On Broadcast Channels With Binary Inputs and Symmetric Outputs”. In: Information Theory, IEEE Transactions on 59.11 (2013), pp. 6980–6989. ISSN: 0018-9448. DOI: 10.1109/TIT.2013.2279392.

[GN14] Y. Geng and C. Nair. “The Capacity Region of the Two-Receiver Gaussian Vector Broadcast Channel With Private and Common Messages”. In: Information Theory, IEEE Transactions on 60.4 (Apr. 2014), pp. 2087–2104. ISSN: 0018-9448. DOI: 10.1109/TIT.2014.2304457.

References VIII

[Gen+14] Y. Geng et al. "On Marton's Inner Bound and Its Optimality for Classes of Product Broadcast Channels". In: Information Theory, IEEE Transactions on 60.1 (2014), pp. 22–41. ISSN: 0018-9448. DOI: 10.1109/TIT.2013.2285925.

[KNE15] H. Kim, B. Nachman, and A. El Gamal. "Superposition coding is almost always optimal for the Poisson broadcast channel". In: 2015 IEEE International Symposium on Information Theory (ISIT). 2015, pp. 834–838. DOI: 10.1109/ISIT.2015.7282572.

[NKG16] C. Nair, H. Kim, and A. El Gamal. "On the optimality of randomized time division and superposition coding for the broadcast channel". In: 2016 IEEE Information Theory Workshop (ITW). Sept. 2016, pp. 131–135. DOI: 10.1109/ITW.2016.7606810.

[NY17] C. Nair and B. Yazdanpanah. "Sub-optimality of superposition coding region for three receiver broadcast channel with two degraded message sets". In: 2017 IEEE International Symposium on Information Theory (ISIT). 2017, pp. 1038–1042. DOI: 10.1109/ISIT.2017.8006686.

[PP18] A. Padakandla and S. Pradhan. "Achievable Rate Region for Three User Discrete Broadcast Channel Based on Coset Codes". In: IEEE Transactions on Information Theory 64.4 (2018), pp. 2267–2297. DOI: 10.1109/TIT.2018.2798669.

[AGN19] V. Anantharam, A. Gohari, and C. Nair. "On the Evaluation of Marton's Inner Bound for Two-Receiver Broadcast Channels". In: IEEE Transactions on Information Theory 65.3 (2019), pp. 1361–1371. DOI: 10.1109/TIT.2018.2880241.

[GN20] A. Gohari and C. Nair. "New Outer Bounds for the Two-Receiver Broadcast Channel". In: 2020 IEEE International Symposium on Information Theory (ISIT). 2020, pp. 1492–1497. DOI: 10.1109/ISIT44484.2020.9174003.

[Nai20] C. Nair. "On Marton's Achievable Region: Local Tensorization for Product Channels with a Binary Component". In: 2020 Information Theory and Applications Workshop (ITA). 2020, pp. 1–7. DOI: 10.1109/ITA50056.2020.9244997.

An Observation

Note: Family of functionals that showed up in network information theory

$$\sum_{S \subseteq [n]} \alpha_S H(X_S), \ \alpha_S \in \mathbb{R}.$$

Usually, one is interested in testing the sub-additivity of

$$\mathcal{C}_{\mu_X}[\alpha_S H(X_S)].$$

This is equivalent to testing a global tensorization property.

Definition

A functional $\sum_{S \subseteq [n]} \alpha_S H(X_S)$ is said to satisfy **global tensorization** if a `product distribution maximizes` $G_{12}^{\mu}(\gamma_1, \gamma_2)$ for all γ_1, γ_2, where

$$G_{12}^{\mu}(\gamma_1, \gamma_2) := \sum_{S \subseteq [n]} \alpha_S H(X_{1S}, X_{2S}) - \mathrm{E}(\gamma_1(\mathbf{X}_1)) - \mathrm{E}(\gamma_2(\mathbf{X}_2))$$

An Observation

Definition

A functional $\sum_{S\subseteq[n]} \alpha_S H(X_S)$ is said to satisfy **local tensorization** if `the product of local maximizers` of $G^{\mu_1}(\gamma_1)$ and $G^{\mu_2}(\gamma_2)$ `is a local maximizer` of $G^{\mu}_{12}(\gamma_1, \gamma_2)$ for all γ_1, γ_2, where

$$G^{\mu}_1(\gamma_1) := \sum_{S\subseteq[n]} \alpha_S H(X_{1S}) - \mathrm{E}(\gamma_1(\mathbf{X}_1))$$

$$G^{\mu}_2(\gamma_2) := \sum_{S\subseteq[n]} \alpha_S H(X_{2S}) - \mathrm{E}(\gamma_2(\mathbf{X}_2))$$

$$G^{\mu}_{12}(\gamma_1, \gamma_2) := \sum_{S\subseteq[n]} \alpha_S H(X_{1S}, X_{2S}) - \mathrm{E}(\gamma_1(\mathbf{X}_1)) - \mathrm{E}(\gamma_2(\mathbf{X}_2))$$

An Observation

Definition

A functional $\sum_{S\subseteq[n]} \alpha_S H(X_S)$ is said to satisfy **local tensorization** if `the product of local maximizers` of $G^{\mu_1}(\gamma_1)$ and $G^{\mu_2}(\gamma_2)$ `is a local maximizer` of $G^{\mu}_{12}(\gamma_1,\gamma_2)$ for all γ_1,γ_2, where

$$G^{\mu}_1(\gamma_1) := \sum_{S\subseteq[n]} \alpha_S H(X_{1S}) - \mathrm{E}(\gamma_1(\mathbf{X}_1))$$

$$G^{\mu}_2(\gamma_2) := \sum_{S\subseteq[n]} \alpha_S H(X_{2S}) - \mathrm{E}(\gamma_2(\mathbf{X}_2))$$

$$G^{\mu}_{12}(\gamma_1,\gamma_2) := \sum_{S\subseteq[n]} \alpha_S H(X_{1S},X_{2S}) - \mathrm{E}(\gamma_1(\mathbf{X}_1)) - \mathrm{E}(\gamma_2(\mathbf{X}_2))$$

Observation (Conjecture)

For functionals in this family global tensorization holds if and only if local tensorization holds

Note: Similarity to testing concavity using a local (second derivative) condition