Information inequalities in network information theoretic settings

Chandra Nair

Department of Information Engineering

The Chinese University of Hong Kong
chandra.nair@gmail.com

April 16, 2013

Network Information Theory

Information theoretic study of multi-user communication scenarios

- Extension of point-to-point communication theory [Shannon ‘48]
- Started with Shannon’s two-way channel [‘62]
  - Model of a telephone conversation
  - A very hard model (unfortunately)
    - feedback
    - interference
    - message cognition
- Other simpler channels were proposed and studied
  - Multiple access channel [Shannon ’61, van-der-Muelen ’71]
  - Broadcast channel [Cover ‘72]
  - Interference channel [Ahlswede ‘74]
  - Relay channel [van-der-Muelen ’71]
- Most fundamental problems remain open

MULTIPLE ACCESS CHANNEL [VAN-DER-MUELEN '71]

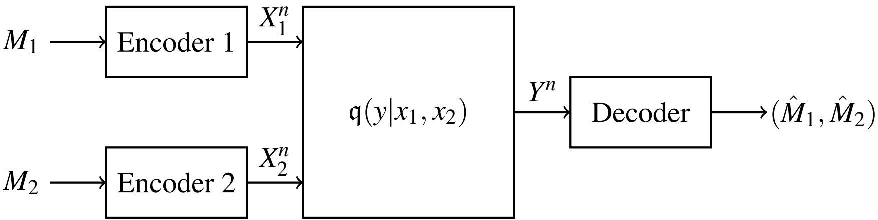


Figure : Discrete memoryless multiple access channel (MAC)

Goal: Compute *Capacity Region* or set of achievable rates?

One of the few settings where the *capacity region* is established

Capacity region of MAC [Ahlswede '71,'74], [Liao '72]

The capacity region is the set of rate pairs (R_1, R_2) such that they satisfy

$$\begin{aligned} R_1 &\leq I(X_1; Y|X_2, Q) \\ R_2 &\leq I(X_2; Y|X_1, Q) \\ R_1 + R_2 &\leq I(X_1, X_2; Y|Q) \end{aligned}$$

for some pmf $p(q)p(x_1|q)p(x_2|q)$, with $|Q| \leq 2$.

Remarks

- Q: convexification random variable (time-sharing)
- The region corresponds to a general cut-set outer bound

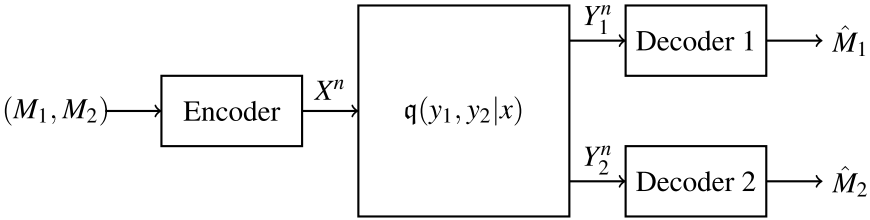


Figure : Discrete memoryless broadcast channel

Goal: Compute *Capacity Region* or set of achievable rates?

Capacity region: open; established in several classes of channels (including additive Gaussian noise model)

BEST KNOWN ACHIEVABLE REGION (M-IB) [MARTON '79]

The union of rate pairs (R_1, R_2) satisfying

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

over all $(U, V, W) \to X \to (Y_1, Y_2)$ is achievable.

Remarks

- U,V,W: *auxiliary* random variables
- Suffices: $|W| \leq |X| + 4, |U| \leq |X|, |V| \leq |X|$ [Gohari-Anantharam '11]
- New: $|W| \leq |X| + 4, |U| + |V| \leq |X| + 1$ [Gohari-Nair-Anantharam '13]
- Not known if it is optimal or not
- Factorization conjecture [Gohari-Nair-Anantharam '12]: If true, the above region is capacity

BEST OUTER BOUND [EL GAMAL '79, NAIR-EL GAMAL '07]

The union of rate pairs (R_1, R_2) satisfying

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

over all $(U,V) \to X \to (Y_1, Y_2)$ forms an outer bound.

Remarks

- Suffices: $|U| \leq |X| + 1, |V| \leq |X| + 1$
- Strictly improves on Körner-Marton outer bound [Nair-El Gamal '07]
- Differs from M-IB [Nair-Wang '08, Gohari-Anantharam '11]
- Known to be strictly suboptimal [Geng-Gohari-Nair-Yu '11]

INTERFERENCE CHANNEL (AHLSWEDE 1974)

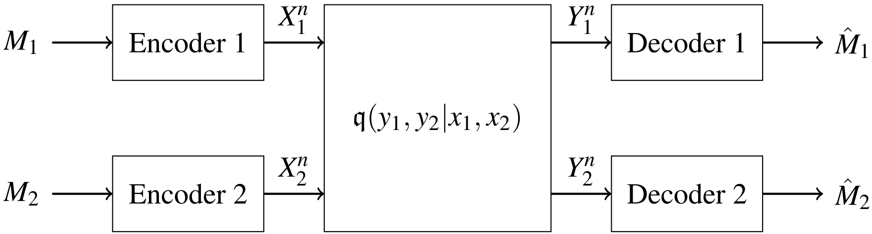


Figure : Discrete memoryless interference channel

Goal: Compute *Capacity Region* or set of achievable rates?

Capacity region: open problem; known for some (limited) classes

Not known even in the *scalar Gaussian* case

- Known in strong-interference regime [Sato '78, Han-Kobayashi '81]

Best Inner bound (HK-IB) [Han-Kobayashi '81]

A rate pair (R_1, R_2) is achievable if

$$
\begin{aligned}
R_1 &\leq I(X_1;Y_1|V,Q) \\
R_2 &\leq I(X_2;Y_2|U,Q) \\
R_1 + R_2 &\leq I(X_1,V;Y_1|Q) + I(X_2;Y_2|U,V,Q) \\
R_1 + R_2 &\leq I(X_2,U;Y_2|Q) + I(X_1;Y_1|U,V,Q) \\
R_1 + R_2 &\leq I(X_1,V;Y_1|U,Q) + I(X_2,U;Y_2|V,Q) \\
2R_1 + R_2 &\leq I(X_1,V;Y_1|Q) + I(X_1;Y_1|U,V,Q) + I(X_2,U;Y_2|V,Q) \\
R_1 + 2R_2 &\leq I(X_2,U;Y_2|Q) + I(X_2;Y_2|U,V,Q) + I(X_1,V;Y_1|U,Q)
\end{aligned}
$$

for some pmf $p(q)p(u_1,x_1|q)p(v,x_2|q)$, where $|U_1| \leq |X_1| + 4$, $|V| \leq |X_2| + 4$, and $|Q| \leq 6$.

Remarks

- Q: coded time-division (more than convexification)
- Not known if it is optimal or not
- Above representation is due to [Chong-Motani-Garg-El Gamal '08]

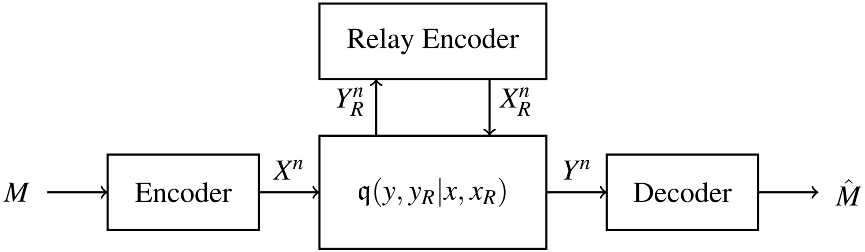


Figure : Point-to-point system with relay

Remarks

- Capacity is unknown; very few special cases known
- Cut set outer bound: $C \leq \max_{p(x_1,x_2)} \min\{I(X_1, X_2; Y_3), I(X_1; Y_2, Y_3|X_2)\}$
- Cut set outer bound is strictly sub-optimal [Aleksic,Razagi,Yu '09]

Comments on the problems

One approach [Avestimehr-Diggavi-Tse, ...]

- These problems have been open for long time
- May be simple expressions for capacity regions do not exist
- Why not settle for appromations

A counter argument [this talk]

- Before we give up, at the very least, we need to
  - Determine if Marton's inner bound is optimal/sub-optimal for the broadcast channel
  - Determine if Han-Kobayashi inner bound is optimal/sub-optimal for the interference channel

Information inequalities play a central role to answer these two questions

- Inequalities to prove optimality of inner bounds (*factorization inequalities*)
- Inequalities to compute extreme points of inner bounds (*extremal inequalities*)

Comments on the problems

One approach [Avestimehr-Diggavi-Tse, ...]

- These problems have been open for long time
- May be simple expressions for capacity regions do not exist
- Why not settle for appromations

A counter argument [this talk]

- Before we give up, at the very least, we need to
  - Determine if Marton's inner bound is optimal/sub-optimal for the broadcast channel
  - Determine if Han-Kobayashi inner bound is optimal/sub-optimal for the interference channel

Information inequalities play a central role to answer these two questions

- Inequalities to prove optimality of inner bounds (*factorization inequalities*)
- Inequalities to compute extreme points of inner bounds (*extremal inequalities*)

COMMENTS ON THE PROBLEMS

One approach [Avestimehr-Diggavi-Tse, ...]

- These problems have been open for long time
- May be simple expressions for capacity regions do not exist
- Why not settle for appromations

A counter argument [this talk]

- Before we give up, at the very least, we need to
  - Determine if Marton's inner bound is optimal/sub-optimal for the broadcast channel
  - Determine if Han-Kobayashi inner bound is optimal/sub-optimal for the interference channel

Information inequalities play a central role to answer these two questions

- Inequalities to prove optimality of inner bounds (*factorization inequalities*)
- Inequalities to compute extreme points of inner bounds (*extremal inequalities*)

Converses to coding theorems

Illustrative example: Degraded broadcast channel

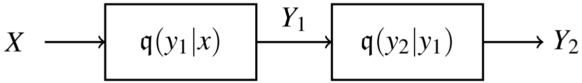


Figure : Degraded broadcast channel

The capacity region is the set of rate pairs (R_1, R_2) such that they satisfy

$$R_1 \leq I(X_1; Y_1|U)$$
$$R_2 \leq I(U; Y_2)$$

for some pmf $p(u)p(x|u)$, with $|U| \leq |X|$. [Cover '72, Gallager '74]

Given any sequence of codebooks with $P_e^{(n)} \to 0$, observe that

$$
\begin{aligned}
nR_2 &= H(M_2) && \text{(message is uniformly distributed)} \\
&= I(M_2; Y_2^n) + H(M_2|Y^n) \\
&\le I(M_2; Y_2^n) + nR_2 P_e^{(n)} && \text{(Fano's inequality)} \\
&= \sum_{i=1}^{n} I(M_2; Y_{2i}|Y_{21}^{i-1}) + nR_2 P_e^{(n)} && \text{(chain rule)} \\
&\le \sum_{i=1}^{n} I(M_2, Y_{21}^{i-1}; Y_{2i}) + nR_2 P_e^{(n)} \\
&= \sum_{i=1}^{n} I(U_i; Y_{2i}) + nR_2 P_e^{(n)} && \text{where } U_i := (M_2, Y_{21}^{i-1}) \\
&= n\Big(I(U_Q; Y_2|Q) + R_2 P_e^{(n)}\Big) && \text{where } Q \sim \text{uniform } \{1, \ldots, n\} \\
&\le n\Big(I(U^{(n)}; Y_2) + R_2 P_e^{(n)}\Big) && \text{where } U^{(n)} := (Q, U_Q)
\end{aligned}
$$

Similarly

$$
\begin{aligned}
nR_1 &= H(M_1) = I(M_1; Y_1^n | M_2) + H(M_1 | M_2, Y_1^n) \\
&\leq \sum_{i=1}^{n} I(M_1; Y_{1i} | M_2, Y_{11}^{i-1}) + nR_1 \mathrm{P}_e^{(n)} && \text{(Fano's inequality)} \\
&\leq \sum_{i=1}^{n} I(X_i; Y_{1i} | M_2, Y_{11}^{i-1}) + nR_1 \mathrm{P}_e^{(n)} && (M_1, Y_{11}^{i-1}) \to X_i \to Y_{1i} \\
&= \sum_{i=1}^{n} I(X_i; Y_{1i} | M_2, Y_{11}^{i-1}, Y_{21}^{i-1}) + nR_1 \mathrm{P}_e^{(n)} && Y_{21}^{i-1} \to Y_{11}^{i-1} \to (M_2, X_i) \\
&\leq \sum_{i=1}^{n} I(X_i; Y_{1i} | M_2, Y_{21}^{i-1}) + nR_1 \mathrm{P}_e^{(n)} && (M_1, Y_{11}^{i-1}, Y_{21}^{i-1}) \to X_i \to Y_{1i} \\
&= \sum_{i=1}^{n} I(X_i; Y_{1i} | U_i) + nR_1 \mathrm{P}_e^{(n)} && \text{Recall } U_i = (M_2, Y_{21}^{i-1})
\end{aligned}
$$

$$
\begin{aligned}
nR_1 &= \sum_{i=1}^{n} I(X_i; Y_{1i}|U_i) + nR_1 \mathrm{P}_e^{(n)} \\
&= n\Big(I(X; Y_1|Q, U_Q) + R_1 \mathrm{P}_e^{(n)}\Big) \qquad \text{Recall } Q \sim \text{uniform } \{1, \dots, n\} \\
&= n\Big(I(X; Y_1|U^{(n)}) + R_2 \mathrm{P}_e^{(n)}\Big)
\end{aligned}
$$

By Caratheodory's theorem there is $(\tilde{U}^{(n)}, X)$ with $|\tilde{U}^{(n)}| \leq |X|$ such that

$$
I(X; Y_1|U^{(n)}) = I(X; Y_1|\tilde{U}^{(n)}) \quad \text{and} \quad I(U^{(n)}; Y_2) = I(\tilde{U}^{(n)}; Y_2).
$$

Letting $n \to \infty$, we see that $\exists$ (U, X) such that

$$
\begin{aligned}
R_2 &\leq I(U; Y_2) \\
R_1 &\leq I(X; Y_1|U) \quad \square
\end{aligned}
$$

Comments on Gallagers proof and more generally

- Absolutely brilliant argument
  - Explicitly constructs an auxiliary U
  - Every converse and outer bound known today follows this approach
- However
  - It is doing more than it should
    - Another proof may exhibit existence of auxiliary U without identifying it
  - The auxiliary U constructed in the converse has no relation to the U *actually employed* in the coding phase
- Maybe
  - The problem is in the use of auxiliaries in the representation
  - Seek other representations
  - These might lead to different converse arguments
- This leads us to information inequalities [this talk]
  - Factorization inequalities - prove optimality directly
  - Extremal inequality - To simplify inner bounds

COMMENTS ON GALLAGERS PROOF AND MORE GENERALLY

- Absolutely brilliant argument
  - Explicitly constructs an auxiliary U
  - Every converse and outer bound known today follows this approach
- However
  - It is doing more than it should
    - Another proof may exhibit existence of auxiliary U without identifying it
  - The auxiliary U constructed in the converse has no relation to the U *actually employed* in the coding phase
- Maybe
  - The problem is in the use of auxiliaries in the representation
  - Seek other representations
  - These might lead to different converse arguments
- This leads us to information inequalities [this talk]
  - Factorization inequalities - prove optimality directly
  - Extremal inequality - To simplify inner bounds

Comments on Gallagers proof and more generally

- Absolutely brilliant argument
  - Explicitly constructs an auxiliary U
  - Every converse and outer bound known today follows this approach
- However
  - It is doing more than it should
    - Another proof may exhibit existence of auxiliary U without identifying it
  - The auxiliary U constructed in the converse has no relation to the U *actually employed* in the coding phase
- Maybe
  - The problem is in the use of auxiliaries in the representation
  - Seek other representations
  - These might lead to different converse arguments
- This leads us to information inequalities [this talk]
  - Factorization inequalities - prove optimality directly
  - Extremal inequality - To simplify inner bounds

Revisit the capacity region for degraded broadcast channel

The capacity region is the set of rate pairs (R_1, R_2) such that they satisfy

$$R_1 \leq I(X_1; Y_1|U)$$
$$R_2 \leq I(U; Y_2)$$

for some pmf $p(u)p(x|u)$, with $|U| \leq |X|$. [Cover '72, Gallager '74]

Seek: $C_\lambda := \max_{(R_1,R_2)\in\mathcal{C}} R_1 + \lambda R_2$, for $\lambda \geq 1$ (supporting hyperplanes)

Elementary manipulations

$$\begin{aligned} C_\lambda &= \max_{p(u,x)} \lambda I(U;Y_2) + I(X;Y_1|U) && \text{(corner point)} \\ &= \max_{p(u,x)} \lambda I(X;Y_2) + I(X;Y_1|U) - \lambda I(X;Y_2|U) && U \to X \to Y_2 \\ &= \max_{p(x)} \Big(\lambda I(X;Y_2) + \max_{p(u|x)} (I(X;Y_1|U) - \lambda I(X;Y_2|U)) \Big) \\ &= \max_{p(x)} \Big(\lambda I(X;Y_2) + \mathfrak{C}[I(X;Y_1) - \lambda I(X;Y_2)] \Big) \end{aligned}$$

UPPER CONCAVE ENVELOPES

Given a function $f(x)$, its upper concave envelope is

$$\mathfrak{C}[f] := \inf\{g(x) : g(x) \geq f(x), g(x) \text{ is concave}\}.$$

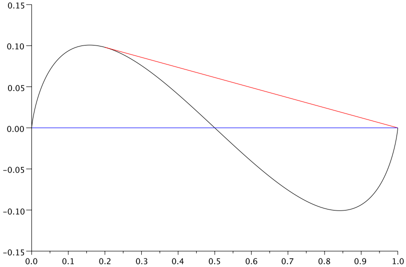


Figure : Illustration of concave envelope

An observation

Fix a broadcast channel $\mathfrak{q}(y_1, y_2|x)$.

Can treat $I(X;Y_1) - \lambda I(X;Y_2)$ as a function of $p(x)$.

For $\lambda \geq 1$ define

$$T_\lambda(X) := \mathfrak{C}[I(X;Y_1) - \lambda I(X;Y_2)]$$

It is easy to see that

$$\max_{p(u|x)} I(X;Y_1|U) - \lambda I(X;Y_2|U) = \mathfrak{C}[I(X;Y_1) - \lambda I(X;Y_2)]$$

Define $C_\lambda(\mathfrak{q}) = \max_{p(x)} \big(\lambda I(X;Y_2) + T_\lambda(X)\big)$

Capacity region of degraded broadcast channel can be expressed as

$$\bigcap_{\lambda \geq 1} \{(R_1, R_2) \subset \mathbb{R}_+^2 : R_1 + \lambda R_2 \leq C_\lambda(\mathfrak{q})\}$$

An observation

Fix a broadcast channel $\mathfrak{q}(y_1, y_2|x)$.

Can treat $I(X;Y_1) - \lambda I(X;Y_2)$ as a function of $p(x)$.

For $\lambda \geq 1$ define

$$T_\lambda(X) := \mathfrak{C}[I(X;Y_1) - \lambda I(X;Y_2)]$$

It is easy to see that

$$\max_{p(u|x)} I(X;Y_1|U) - \lambda I(X;Y_2|U) = \mathfrak{C}[I(X;Y_1) - \lambda I(X;Y_2)]$$

Define $C_\lambda(\mathfrak{q}) = \max_{p(x)} \big(\lambda I(X;Y_2) + T_\lambda(X)\big)$

Capacity region of degraded broadcast channel can be expressed as

$$\bigcap_{\lambda \geq 1} \{(R_1, R_2) \subset \mathbb{R}_+^2 : R_1 + \lambda R_2 \leq C_\lambda(\mathfrak{q})\}$$

A CONVERSATION

Alice: Hey, we have a new characterization for the superposition coding region without auxiliaries

$$\bigcap_{\lambda \geq 1} \{(R_1, R_2) \subset \mathbb{R}_+^2 : R_1 + \lambda R_2 \leq C_\lambda(\mathfrak{q})\}$$

Bob: But is it really different? Isn't your concave envelope hiding the auxiliaries

Alice: But the only auxliaries that show up are the ones that are "extremal". i.e. the only interesting ones are the ones that help compute the upper concave envelope

Bob: So what? Does this buy you anything: Can we get new results? Can we simplify old proofs?

Alice: Yes, when we apply this to other settings. Focusing on extremal auxiliaries has yielded new results. Greatly simplifies old proofs.

A conversation

Alice: Hey, we have a new characterization for the superposition coding region without auxiliaries

$$\bigcap_{\lambda \geq 1} \{(R_1, R_2) \subset \mathbb{R}_+^2 : R_1 + \lambda R_2 \leq C_\lambda(\mathfrak{q})\}$$

Bob: But is it really different? Isn't your concave envelope hiding the auxiliaries

Alice: But the only auxliaries that show up are the ones that are "extremal". i.e. the only interesting ones are the ones that help compute the upper concave envelope

Bob: So what? Does this buy you anything: Can we get new results? Can we simplify old proofs?

Alice: Yes, when we apply this to other settings. Focusing on extremal auxiliaries has yielded new results. Greatly simplifies old proofs.

A CONVERSATION

Alice: Hey, we have a new characterization for the superposition coding region without auxiliaries

$$\bigcap_{\lambda \geq 1} \{(R_1, R_2) \subset \mathbb{R}_+^2 : R_1 + \lambda R_2 \leq C_\lambda(\mathfrak{q})\}$$

Bob: But is it really different? Isn't your concave envelope hiding the auxiliaries

Alice: But the only auxliaries that show up are the ones that are "extremal". i.e. the only interesting ones are the ones that help compute the upper concave envelope

Bob: So what? Does this buy you anything: Can we get new results? Can we simplify old proofs?

Alice: Yes, when we apply this to other settings. Focusing on extremal auxiliaries has yielded new results. Greatly simplifies old proofs.

DEGRADED BINARY SYMMETRIC BROADCAST CHANNEL

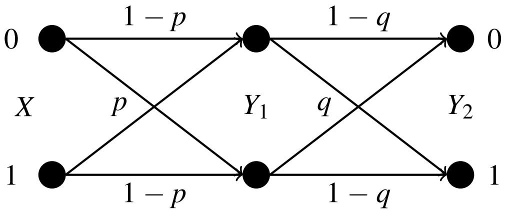


Figure : Degraded BSC broadcast channel

To compute the union of rate pairs (R_1, R_2) such that they satisfy

$$R_1 \leq I(X_1; Y_1|U)$$
$$R_2 \leq I(U; Y_2)$$

for some pmf $p(u)p(x|u)$, with $|U| \leq |X|$, it suffices to consider $U \mapsto X \sim BSC(s)$ [Cover '72, Wyner-Ziv '73]

Proof is non-trivial; uses Mrs. Gerber's lemma (convexity of $h(p * h^{-1}(x))$)

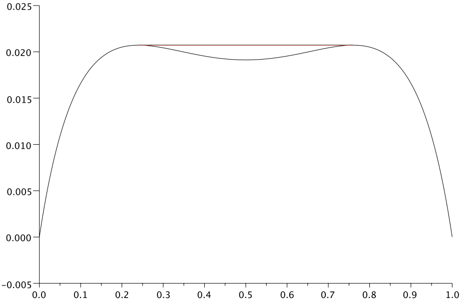


Figure : Illustration of $I(X;Y_1) - \lambda I(X;Y_2)$

Immediate that a global maximum exists when $\mathrm{P}(X=0) = \frac{1}{2}$ and

$\mathrm{P}(X=0|U=0) = s, \mathrm{P}(X=0|U=1) = 1-s$, i.e, $U \mapsto X \sim BSC(s)$

Optimality using concave envelope representation

Consider two degraded broadcast channels $\mathfrak{q}_1(y_{11}, y_{21}|x_1)$ and $\mathfrak{q}_2(y_{12}, y_{22}|x_2)$

Form a product broadcast channel: $\mathfrak{q}_1(y_{11}, y_{21}|x_1) \otimes \mathfrak{q}_2(y_{12}, y_{22}|x_2)$

$$T_\lambda(X_1, X_2) := \mathfrak{C}[I(X_1, X_2; Y_{11}, Y_{12}) - \lambda I(X_1; X_2; Y_{21}, Y_{22})]$$

$T_\lambda(X_1, X_2)$: function of $p(x_1, x_2)$

Claim: If the following *factorization inequality* holds

$$T_\lambda(X_1, X_2) \leq T_\lambda(X_1) + T_\lambda(X_2)$$

then one has optimality of the region

$$\bigcap_{\lambda \geq 1} \{(R_1, R_2) \subset \mathbb{R}_+^2 : R_1 + \lambda R_2 \leq C_\lambda(\mathfrak{q})\},$$

where $C_\lambda(\mathfrak{q}) = \max_{p(x)} \big(\lambda I(X; Y_2) + T_\lambda(X)\big)$

REMARKS

- The above claim is much stronger than what is needed for proving optimality
- Note that $I(X;Y)$ has a similar behaviour, i.e. for $\mathfrak{q}_1(y_1|x_1) \otimes \mathfrak{q}_2(y_2|x_2)$

$$
\begin{aligned}
& I(X_1,X_2;Y_1,Y_2) \\
&\quad = I(X_1,X_2;Y_1) + I(X_1,X_2;Y_2|Y_1) \\
&\quad = I(X_1;Y_1) + I(X_2;Y_2|Y_1) \qquad\qquad (Y_2 \to X_2 \to X_1 \to Y_1) \\
&\quad \le I(X_1;Y_1) + I(X_2;Y_2)
\end{aligned}
$$

Proof of Claim: If the factorization inequality holds

$$
\begin{aligned}
\mathcal{C}_\lambda(\mathfrak{q}\otimes\cdots\otimes\mathfrak{q}) &= \max_{p(x^n)} \lambda I(X_1^n;Y_{21}^n) + T_\lambda(X^n) \\
&\le \max_{p(x^n)} \sum_{i=1}^n \left(\lambda I(X_i;Y_{2i}) + T_\lambda(X_i)\right) \qquad \text{(By assumption)} \\
&\le n \max_{p(x)} \lambda I(X;Y_2) + T_\lambda(X) = nC_\lambda(\mathfrak{q})
\end{aligned}
$$

Remarks

- The above claim is much stronger than what is needed for proving optimality
- Note that $I(X;Y)$ has a similar behaviour, i.e. for $\mathfrak{q}_1(y_1|x_1) \otimes \mathfrak{q}_2(y_2|x_2)$

$$\begin{aligned} & I(X_1, X_2; Y_1, Y_2) \\ & \quad = I(X_1, X_2; Y_1) + I(X_1, X_2; Y_2|Y_1) \\ & \quad = I(X_1; Y_1) + I(X_2; Y_2|Y_1) && (Y_2 \to X_2 \to X_1 \to Y_1) \\ & \quad \leq I(X_1; Y_1) + I(X_2; Y_2) \end{aligned}$$

Proof of Claim: If the factorization inequality holds

$$\begin{aligned} \mathcal{C}_\lambda(\mathfrak{q} \otimes \cdots \otimes \mathfrak{q}) &= \max_{p(x^n)} \lambda I(X_1^n; Y_{21}^n) + T_\lambda(X^n) \\ &\leq \max_{p(x^n)} \sum_{i=1}^n (\lambda I(X_i; Y_{2i}) + T_\lambda(X_i)) && \text{(By assumption)} \\ &\leq n \max_{p(x)} \lambda I(X; Y_2) + T_\lambda(X) = n C_\lambda(\mathfrak{q}) \end{aligned}$$

PROOF - CTD

To complete the argument, observe

$$\begin{aligned} & nR_1 + n\lambda R_2 \\ & \quad \leq \lambda I(M_1, M_2; Y_{21}^n) + I(M_1; Y_{11}^n | M_2) - \lambda I(M_1; Y_{21}^n | M_2) + n\epsilon_n \quad \text{(Fano)} \\ & \quad \overset{(a)}{\leq} \lambda I(X^n; Y_{21}^n) + \mathfrak{C}[I(X^n; Y_{11}^n) - \lambda I(X_1^n; Y_{21}^n)] \\ & \leq \max_{p(x^n)} \lambda I(X_1^n; Y_{21}^n) + T_\lambda(X^n) \end{aligned}$$

where the (a) follows from $(M_1, M_2) \to X^n \to (Y_{11}^n, Y_{21}^n)$

Important: The proof I demonstrated is a generic proof

- If one can demonstrate an appropriate *factorization inequality* then
  - Optimality follows directly from Fano's inequality
  - Optimality of the n-letter form (known in many cases)
- † If a weaker *factorization inequality* does not hold then
  - Inner bound is strictly sub-optimal

To complete the argument, observe

$$\begin{aligned} & nR_1 + n\lambda R_2 \\ & \quad \leq \lambda I(M_1, M_2; Y_{21}^n) + I(M_1; Y_{11}^n | M_2) - \lambda I(M_1; Y_{21}^n | M_2) + n\epsilon_n \quad \text{(Fano)} \\ & \quad \stackrel{(a)}{\leq} \lambda I(X^n; Y_{21}^n) + \mathfrak{C}[I(X^n; Y_{11}^n) - \lambda I(X_1^n; Y_{21}^n)] \\ & \leq \max_{p(x^n)} \lambda I(X_1^n; Y_{21}^n) + T_\lambda(X^n) \end{aligned}$$

where the (a) follows from $(M_1, M_2) \to X^n \to (Y_{11}^n, Y_{21}^n)$

Important: The proof I demonstrated is a generic proof

- If one can demonstrate an appropriate *factorization inequality* then
  - Optimality follows directly from Fano's inequality
  - Optimality of the n-letter form (known in many cases)
- † If a weaker *factorization inequality* does not hold then
  - Inner bound is strictly sub-optimal

Factorization inequalities: introduction

Inequalities that *factor* over product channels

Triivial Example:

- $I(X_1, X_2; Y_1, Y_2) \leq I(X_1; Y_1) + I(X_2; Y_2)$ when $Y_1 \to X_1 \to X_2 \to Y_2$ is Markov.

Non-trivial example:

For any product broadcast channel $\mathfrak{q}_1(y_{11}, y_{21}|x_1) \otimes \mathfrak{q}_2(y_{12}, y_{22}|x_2)$ and $\lambda \geq 1$

$$T_\lambda(X_1, X_2) \leq T_\lambda(X_1) + T_\lambda(X_2),$$

where $T_\lambda(X_1, X_2) = \mathfrak{C}\left[I(X_1, X_2; Y_{11}, Y_{12}) - \lambda I(X_1, X_2; Y_{21}, Y_{22})\right]$

- Stronger than what I needed earlier (no degradedness assumption)
- Implies the (known) capacity region for *degraded message sets* (two receivers)

Inequalities that *factor* over product channels

Triivial Example:

- $I(X_1, X_2; Y_1, Y_2) \leq I(X_1; Y_1) + I(X_2; Y_2)$ when $Y_1 \to X_1 \to X_2 \to Y_2$ is Markov.

Non-trivial example:

For any product broadcast channel $\mathfrak{q}_1(y_{11}, y_{21}|x_1) \otimes \mathfrak{q}_2(y_{12}, y_{22}|x_2)$ and $\lambda \geq 1$

$$T_\lambda(X_1, X_2) \leq T_\lambda(X_1) + T_\lambda(X_2),$$

where $T_\lambda(X_1, X_2) = \mathfrak{C}\big[I(X_1, X_2; Y_{11}, Y_{12}) - \lambda I(X_1, X_2; Y_{21}, Y_{22})\big]$

- Stronger than what I needed earlier (no degradedness assumption)
- Implies the (known) capacity region for *degraded message sets* (two receivers)

OTHER FACTORIZATION INEQUALITIES

Define for a broadcast channel

$$S(X) := \mathfrak{C}\big[I(X;Y_1) - I(X;Y_2) + \mathfrak{C}[I(X;Y_2) - I(X;Y_1)]\big]$$

For any product broadcast channel $\mathfrak{q}_1(y_{11}, y_{21}|x_1) \otimes \mathfrak{q}_2(y_{12}, y_{22}|x_2)$

$$S(X_1, X_2) \le S(X_1) + S(X_2)$$

- This yields the secrecy capacity (known result)

For an interference channel, † define

$$R(X_1;X_2) := \mathfrak{C}\big[I(X_1;Y_1|X_2) - I(X_1;Y_2|X_2)\big]$$

For any product interference channel
$\mathfrak{q}_1(y_{11}, y_{21}|x_{11}, x_{21}) \otimes \mathfrak{q}_2(y_{12}, y_{22}|x_{12}, x_{22})$

$$R(X_{11}, X_{12}; X_{21}, X_{22}) \le R(X_{11}; X_{21}) + R(X_{12}; X_{22})$$

Other factorization inequalities

Define for a broadcast channel

$$S(X) := \mathfrak{C}\Big[I(X;Y_1) - I(X;Y_2) + \mathfrak{C}[I(X;Y_2) - I(X;Y_1)]\Big]$$

For any product broadcast channel $\mathfrak{q}_1(y_{11}, y_{21}|x_1) \otimes \mathfrak{q}_2(y_{12}, y_{22}|x_2)$

$$S(X_1, X_2) \leq S(X_1) + S(X_2)$$

- This yields the secrecy capacity (known result)

For an interference channel, † define

$$R(X_1; X_2) := \mathfrak{C}\Big[I(X_1; Y_1|X_2) - I(X_1; Y_2|X_2)\Big]$$

For any product interference channel
$\mathfrak{q}_1(y_{11}, y_{21}|x_{11}, x_{21}) \otimes \mathfrak{q}_2(y_{12}, y_{22}|x_{12}, x_{22})$

$$R(X_{11}, X_{12}; X_{21}, X_{22}) \leq R(X_{11}; X_{21}) + R(X_{12}; X_{22})$$

A factorization inequality - Conjecture

For a three-receiver broadcast channel $\mathfrak{q}(y_1, y_2, y_3|x)$, $\mu \in [0,1]$, $\lambda \geq 1$

$$T_{\mu,\lambda}(X) := \mathfrak{C}\big[\mu I(X;Y_1) + (1-\mu)I(X;Y_2) - \lambda I(X;Y_3)\big]$$

For any product broadcast channel $\mathfrak{q}_1(y_{11}, y_{21}, y_{31}|x_1) \otimes \mathfrak{q}_2(y_{12}, y_{22}, y_{32}|x_2)$

$$T_{\mu,\lambda}(X_1, X_2) \leq T_{\mu,\lambda}(X_1) + T_{\mu,\lambda}(X_2) \qquad \text{(Conjecture)}$$

Remarks

- Know this holds when $\mu \in \{0,1\}$
- If true, would imply the capacity region of three receiver broadcast channel with two degraded message sets
  - Problem has been open since the 70s
- Numerically verified for channels of small sizes

A FACTORIZATION INEQUALITY - CONJECTURE

For a three-receiver broadcast channel $\mathfrak{q}(y_1, y_2, y_3|x)$, $\mu \in [0,1]$, $\lambda \geq 1$

$$T_{\mu,\lambda}(X) := \mathfrak{C}\Big[\mu I(X;Y_1) + (1-\mu)I(X;Y_2) - \lambda I(X;Y_3)\Big]$$

For any product broadcast channel $\mathfrak{q}_1(y_{11}, y_{21}, y_{31}|x_1) \otimes \mathfrak{q}_2(y_{12}, y_{22}, y_{32}|x_2)$

$$T_{\mu,\lambda}(X_1, X_2) \leq T_{\mu,\lambda}(X_1) + T_{\mu,\lambda}(X_2) \qquad \text{(Conjecture)}$$

Remarks

- Know this holds when $\mu \in \{0, 1\}$
- If true, would imply the capacity region of three receiver broadcast channel with two degraded message sets
  - Problem has been open since the 70s
- Numerically verified for channels of small sizes

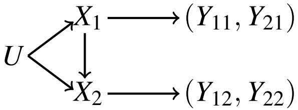
PROOF OF A FACTORIZATION INEQUALITY

Claim: For any product broadcast channel $\mathfrak{q}_1(y_{11}, y_{21}|x_1) \otimes \mathfrak{q}_2(y_{12}, y_{22}|x_2)$ and $\lambda \geq 1$

$$T_\lambda(X_1, X_2) \leq T_\lambda(X_1) + T_\lambda(X_2),$$

where $T_\lambda(X_1, X_2) = \mathfrak{C}\big[I(X_1, X_2; Y_{11}, Y_{12}) - \lambda I(X_1, X_2; Y_{21}, Y_{22})\big]$

Note: This is equivalent to showing that for any U such that



there exists $U_1 \to X_1 \to (Y_{11}, Y_{21})$ and $U_2 \to X_2 \to (Y_{12}, Y_{22})$ such that

$$\begin{aligned} & I(X_1, X_2; Y_{11}, Y_{12}|U) - \lambda I(X_1, X_2; Y_{21}, Y_{22}|U) \\ & \leq I(X_1; Y_{11}|U_1) - \lambda I(X_1; Y_{21}|U_1) + I(X_1; Y_{12}|U_2) - \lambda I(X_1; Y_{22}|U_2) \end{aligned}$$

PROOF CTD..

$$
\begin{aligned}
&I(X_1, X_2; Y_{11}, Y_{12}|U) - \lambda I(X_1, X_2; Y_{21}, Y_{22}|U) \\
&\quad = I(X_1; Y_{11}|U) - \lambda I(X_1; Y_{21}|U, Y_{22}) && \text{using the Markov} \\
&\qquad + I(X_2; Y_{12}|U, Y_{11}) - \lambda I(X_2; Y_{22}|U) && \text{structure} \\
&\quad = I(X_1; Y_{11}|U, Y_{22}) - \lambda I(X_1; Y_{21}|U, Y_{22}) && \text{subtract/add term} \\
&\qquad + I(X_2; Y_{12}|U, Y_{11}) - \lambda I(X_2; Y_{22}|U, Y_{11}) && I(Y_{11}; Y_{22}|U) \\
&\quad = I(X_1; Y_{11}|U_1) - \lambda I(X_1; Y_{21}|U_1) && U_2 := (U, Y_{22}) \\
&\qquad + I(X_2; Y_{12}|U_2) - \lambda I(X_2; Y_{22}|U_1) && U_1 := (U, Y_{11})
\end{aligned}
$$

Note that $(U, Y_{22}) \to X_1 \to (Y_{11}, Y_{21})$ and $(U, Y_{11}) \to X_2 \to (Y_{12}, Y_{22})$ hold as desired □

Remarks

- This proof is motivated by an argument of Csiszar
- This inequality has much deeper consequences
- Need: an alternate proof of this inequality

$$
\begin{aligned}
& I(X_1, X_2; Y_{11}, Y_{12}|U) - \lambda I(X_1, X_2; Y_{21}, Y_{22}|U) \\
&\quad = I(X_1; Y_{11}|U) - \lambda I(X_1; Y_{21}|U, Y_{22}) && \text{using the Markov} \\
&\qquad + I(X_2; Y_{12}|U, Y_{11}) - \lambda I(X_2; Y_{22}|U) && \text{structure} \\
&\quad = I(X_1; Y_{11}|U, Y_{22}) - \lambda I(X_1; Y_{21}|U, Y_{22}) && \text{subtract/add term} \\
&\qquad + I(X_2; Y_{12}|U, Y_{11}) - \lambda I(X_2; Y_{22}|U, Y_{11}) && I(Y_{11}; Y_{22}|U) \\
&\quad = I(X_1; Y_{11}|U_1) - \lambda I(X_1; Y_{21}|U_1) && U_2 := (U, Y_{22}) \\
&\qquad + I(X_2; Y_{12}|U_2) - \lambda I(X_2; Y_{22}|U_1) && U_1 := (U, Y_{11})
\end{aligned}
$$

Note that $(U, Y_{22}) \to X_1 \to (Y_{11}, Y_{21})$ and $(U, Y_{11}) \to X_2 \to (Y_{12}, Y_{22})$ hold as desired □

Remarks

- This proof is motivated by an argument of Csiszar
- This inequality has much deeper consequences
- Need: an alternate proof of this inequality

- Most proofs of factorization inequalities are motivated by converses/outer bounds
  - Though factorization implies optimality, not the other way
- One recent converse motivated by factorization
  - Capacity region of reversely semi-deterministic broadcast channel [Geng-Gohari-Nair-Yu '11]
  - Show strictly sub-optimality of outer bound for broadcast channels

For most problems where one has good inner bounds one can conjecture a factorization inequality

- Easily check (numerically) if these inequalities hold for small cardinalities

Need: A method for proving these factorization inequalities, an understanding of them, and a more formal and precise statement

Remarks about factorization inequalities

- Most proofs of factorization inequalities are motivated by converses/outer bounds
  - Though factorization implies optimality, not the other way
- One recent converse motivated by factorization
  - Capacity region of reversely semi-deterministic broadcast channel [Geng-Gohari-Nair-Yu '11]
  - Show strictly sub-optimality of outer bound for broadcast channels

For most problems where one has good inner bounds one can conjecture a factorization inequality

- Easily check (numerically) if these inequalities hold for small cardinalities

Need: A method for proving these factorization inequalities, an understanding of them, and a more formal and precise statement

Remarks about factorization inequalities

- Most proofs of factorization inequalities are motivated by converses/outer bounds
  - Though factorization implies optimality, not the other way
- One recent converse motivated by factorization
  - Capacity region of reversely semi-deterministic broadcast channel [Geng-Gohari-Nair-Yu '11]
  - Show strictly sub-optimality of outer bound for broadcast channels

For most problems where one has good inner bounds one can conjecture a factorization inequality

- Easily check (numerically) if these inequalities hold for small cardinalities

Need: A method for proving these factorization inequalities, an understanding of them, and a more formal and precise statement

Recall: Factorization inequalities are inequalities that factor over product channels

Extremal inequalities: Inequalities that compute the *extremal* auxiliary random variables

Usual example: Entropy power inequality (EPI) and its variants

- More on this later

Another example: Mrs. Gerber's Lemma ("discrete analogue" of EPI)

Begin with: A couple of discrete extremal inequalities

- Motivated by Marton's inner bound for broadcast channels

Extremal inequalities: Introduction

Recall: Factorization inequalities are inequalities that factor over product channels

Extremal inequalities: Inequalities that compute the *extremal* auxiliary random variables

Usual example: Entropy power inequality (EPI) and its variants

- More on this later

Another example: Mrs. Gerber's Lemma ("discrete analogue" of EPI)

Begin with: A couple of discrete extremal inequalities

- Motivated by Marton's inner bound for broadcast channels

The union of rate pairs (R_1, R_2) satisfying

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

over all $(U, V, W) \to X \to (Y_1, Y_2)$ is achievable.

$T(X) := \max_{p(uv|x)} I(U; Y_1) + I(V; Y_2) - I(U; V)$

Conjecture [Gohari-Nair-Anantharam'12]: For any $\lambda \in [0, 1]$ the following function factorizes

$$\mathfrak{C}[-\lambda H(Y_1) - (1 - \lambda)H(Y_2) + T(X)]$$

If true, then M-IB would be sum-rate optimal

The union of rate pairs (R_1, R_2) satisfying

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

over all $(U,V,W) \to X \to (Y_1, Y_2)$ is achievable.

$T(X) := \max_{p(uv|x)} I(U;Y_1) + I(V;Y_2) - I(U;V)$

Conjecture [Gohari-Nair-Anantharam'12]: For any $\lambda \in [0,1]$ the following function factorizes

$$\mathfrak{C}[-\lambda H(Y_1) - (1-\lambda)H(Y_2) + T(X)]$$

If true, then M-IB would be sum-rate optimal

Marton's inner bound - broadcast channel

The union of rate pairs (R_1, R_2) satisfying

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

over all $(U,V,W) \to X \to (Y_1, Y_2)$ is achievable.

$T(X) := \max_{p(uv|x)} I(U;Y_1) + I(V;Y_2) - I(U;V)$

Conjecture [Gohari-Nair-Anantharam'12]: For any $\lambda \in [0,1]$ the following function factorizes

$$\mathfrak{C}[-\lambda H(Y_1) - (1-\lambda) H(Y_2) + T(X)]$$

If true, then M-IB would be sum-rate optimal

An information inequality

Recall $T(X) := \max_{p(uv|x)} I(U;Y_1) + I(V;Y_2) - I(U;V)$

A 5-variable inequality [Geng-Jog-Nair-Wang '11]

For any $(U,V) \to X \to (Y_1,Y_2)$ and $|X| = 2$, the following holds:

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

In other words when $|X| = 2$

$$T(X) = \max\{I(X;Y_1), I(X;Y_2)\}$$

Remarks

- The inequality is false when $|X| = 3$
- Not quite in the framework of Shannon/non-Shannon type inequalities
  - The cardinality constraint (natural under a channel coding setting) destroys the convex cone property

- Conjectured for a particular binary input BC [Nair-Wang '08]
  - Motivation: to exhibit gap between inner and outer bounds for this channel
- Used perturbation analysis and obtained cardinality bounds for M-IB [Gohari-Anantharam '12]
  - They numerically verified the plausibility of this conjecture
- The conjecture was established for the channel [Jog-Nair '09] extending the perturbation techniques
- The proof was generalized for all binary input broadcast channels [Geng-Nair-Wang '10]

What about beyond sum-rate?

Is it true that when $(U,V) \to X \to (Y_1,Y_2)$, $|X| = 2$, and $\alpha > 1$

$$\alpha I(U;Y) + I(V;Z) - I(U;V) \le \max\{\alpha I(X;Y), I(X;Z)\}$$

False (counterexample to this inequality due to Geng)

REMARKS ABOUT THIS INEQUALITY

- Conjectured for a particular binary input BC [Nair-Wang '08]
  - Motivation: to exhibit gap between inner and outer bounds for this channel
- Used perturbation analysis and obtained cardinality bounds for M-IB [Gohari-Anantharam '12]
  - They numerically verified the plausibility of this conjecture
- The conjecture was established for the channel [Jog-Nair '09] extending the perturbation techniques
- The proof was generalized for all binary input broadcast channels [Geng-Nair-Wang '10]

What about beyond sum-rate?

Is it true that when $(U, V) \to X \to (Y_1, Y_2)$, $|X| = 2$, and $\alpha > 1$

$$\alpha I(U;Y) + I(V;Z) - I(U;V) \le \max\{\alpha I(X;Y), I(X;Z)\}$$

False (counterexample to this inequality due to Geng)

Remarks about this inequality

- Conjectured for a particular binary input BC [Nair-Wang '08]
  - Motivation: to exhibit gap between inner and outer bounds for this channel
- Used perturbation analysis and obtained cardinality bounds for M-IB [Gohari-Anantharam '12]
  - They numerically verified the plausibility of this conjecture
- The conjecture was established for the channel [Jog-Nair '09] extending the perturbation techniques
- The proof was generalized for all binary input broadcast channels [Geng-Nair-Wang '10]

What about beyond sum-rate?

Is it true that when $(U,V) \to X \to (Y_1, Y_2)$, $|X| = 2$, and $\alpha > 1$

$$\alpha I(U;Y) + I(V;Z) - I(U;V) \leq \max\{\alpha I(X;Y), I(X;Z)\}$$

False (counterexample to this inequality due to Geng)

$$T_\alpha(X) := \max_{p(u,v|x)} \alpha I(U;Y) + I(V;Z) - I(U;V)$$

Conjecture [Gohari-Nair-Anantharam '12]: For $\lambda \in [0,1], \alpha \geq 1$, $|X| = 2$ and $(U,V) \to X \to (Y_1, Y_2)$

$$\begin{aligned} &\mathfrak{C}[(\alpha - \lambda)H(Y) - \lambda H(Z) + T_\alpha(X)] \\ &\quad \leq \mathfrak{C}[(\alpha - \lambda)H(Y) - \lambda H(Z) + \max\{\alpha I(X;Y), I(X;Z)\}] \end{aligned}$$

Remarks

- Solved affirmatively [Gohari-Nair-Anantharam '13]
- Indeed suffices to consider $|U| + |V| \leq |X| + 1$ to compute

$$\mathfrak{C}[(\alpha - \lambda)H(Y) - \lambda H(Z) + T_\alpha(X)]$$

BEYOND SUM-RATE

$$T_\alpha(X) := \max_{p(u,v|x)} \alpha I(U;Y) + I(V;Z) - I(U;V)$$

Conjecture [Gohari-Nair-Anantharam '12]: For $\lambda \in [0,1], \alpha \geq 1$, $|X| = 2$ and $(U,V) \to X \to (Y_1, Y_2)$

$$\begin{aligned} &\mathfrak{C}[(\alpha - \lambda)H(Y) - \lambda H(Z) + T_\alpha(X)] \\ &\quad \leq \mathfrak{C}[(\alpha - \lambda)H(Y) - \lambda H(Z) + \max\{\alpha I(X;Y), I(X;Z)\}] \end{aligned}$$

Remarks

- Solved affirmatively [Gohari-Nair-Anantharam '13]
- Indeed suffices to consider $|U| + |V| \leq |X| + 1$ to compute

$$\mathfrak{C}[(\alpha - \lambda)H(Y) - \lambda H(Z) + T_\alpha(X)]$$

ANOTHER FACTORIZATION CONJECTURE

$$T_\alpha(X) := \max_{p(u,v|x)} \alpha I(U;Y) + I(V;Z) - I(U;V)$$

For $\lambda \in [0,1], \alpha \geq 1$, define

$$S_{\alpha,\lambda}(X) := \mathfrak{C}[(\alpha-\lambda)H(Y) - \lambda H(Z) + T_\alpha(X)]$$

Conjecture [Gohari-Nair-Anantharam '12]: The functional $S_{\alpha,\lambda}(X)$ factorizes over product broadcast channels

Remarks

- If true, would imply that M-IB is the capacity region for a DM-BC (very huge deal)
- Know it is true when $\lambda \in \{0,1\}$
- Numerically verified for product of binary channels
- Feeling: You need new extremal inequalities and progress on factorization techniques to make progress

Another factorization conjecture

$$T_\alpha(X) := \max_{p(u,v|x)} \alpha I(U;Y) + I(V;Z) - I(U;V)$$

For $\lambda \in [0,1], \alpha \geq 1$, define

$$S_{\alpha,\lambda}(X) := \mathfrak{C}[(\alpha - \lambda)H(Y) - \lambda H(Z) + T_\alpha(X)]$$

Conjecture [Gohari-Nair-Anantharam '12]: The functional $S_{\alpha,\lambda}(X)$ factorizes over product broadcast channels

Remarks

- If true, would imply that M-IB is the capacity region for a DM-BC (very huge deal)
- Know it is true when $\lambda \in \{0,1\}$
- Numerically verified for product of binary channels
- Feeling: You need new extremal inequalities and progress on factorization techniques to make progress

Entropy Power Inequality (EPI)

Perhaps one of the most famous information inequalities

EPI: If X, Y are independent and have densities

$$2^{2h(X+Y)} \geq 2^{2h(X)} + 2^{2h(Y)}$$

Conditional EPI: If $X \to U \to Y$ is Markov, and conditioned on U (finite valued) they have densities, then

$$2^{2h(X+Y|U)} \geq 2^{2h(X|U)} + 2^{2h(Y|U)}$$

This is an extremal inequality

- Its utility has been to evaluate extremal auxiliaries
- Several variations of this inequality known and used
- Key to several converses in Gaussian noise settings (until recently)
- Proof uses perturbation ideas [Stam '58]

Entropy Power Inequality (EPI)

Perhaps one of the most famous information inequalities

EPI: If X, Y are independent and have densities

$$2^{2h(X+Y)} \geq 2^{2h(X)} + 2^{2h(Y)}$$

Conditional EPI: If $X \to U \to Y$ is Markov, and conditioned on U (finite valued) they have densities, then

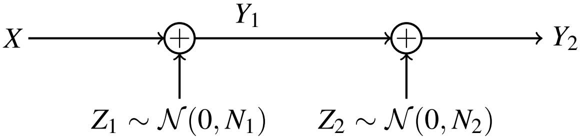
$$2^{2h(X+Y|U)} \geq 2^{2h(X|U)} + 2^{2h(Y|U)}$$

This is an extremal inequality

- Its utility has been to evaluate extremal auxiliaries
- Several variations of this inequality known and used
- Key to several converses in Gaussian noise settings (until recently)
- Proof uses perturbation ideas [Stam '58]

An illustration of EPI's use

Consider the Gaussian degraded broadcast channel



Power constraint: $E(X^2) \leq P$.

The capacity region is the set of rate pairs (R_1, R_2) such that they satisfy

$$R_1 \leq I(X_1; Y_1|U)$$
$$R_2 \leq I(U; Y_2)$$

for some (U, X) with $E(X^2) \leq P$

How would you compute this region? (real-valued variables)

BERGMAN'S PROOF (1973)

Let

$$h(Y_1|U) = \frac{1}{2}\log(2\pi e(N_1 + aP))$$

for some $a \in [0, 1]$

By EPI

$$2^{2h(Y_2|U)} = 2^{2h(Y_1+Z_2|U)} \geq 2^{2h(Y_1|U)} + 2^{2h(Z_2|U)} = 2\pi e(N_1 + aP + N_2)$$

Thus if

$$R_1 \leq I(X_1; Y_1|U) = H(Y_1|U) - h(Y_1|X) = \frac{1}{2}\log\left(1 + \frac{aP}{N_1}\right)$$

then

$$R_2 \leq I(U; Y_2) \leq \frac{1}{2}\log\left(1 + \frac{aP}{N_1 + N_2}\right)$$

Equality: $X = U + V$, $U \perp V$, $V \sim \mathcal{N}(0, aP)$, $U \sim \mathcal{N}(0, (1-a)P)$

BEYOND THE SCOPE OF THIS TALK

Factorization inequalities used to replace EPI in converses [Geng-Nair '12]

- Recover known converse proofs without using EPI (and in a much simpler way)
- In addition solved an *important* open problem of 2-receiver vector Gaussian broadcast channel with private and common messages

Basic idea: Use factorization inequalities to deduce that if X is a maximizer to a relevant optimization problem, X_1, X_2 i.i.d. $\sim X$ then

- Both $\frac{X_1+X_2}{\sqrt{2}}$ and $\frac{X_1-X_2}{\sqrt{2}}$ are also maximizers
- Further $\frac{X_1+X_2}{\sqrt{2}}$ and $\frac{X_1-X_2}{\sqrt{2}}$ are independent

Complete the argument (optimality of Gaussian) either using

- Central limit theorem
- Gaussian characterization: If X, Y are independent and $X+Y, X-Y$ are independent then X, Y are i.i.d. Gaussians [Berstein '40, Skijtovic '54]

Concluding remarks

- New representations using concave envelopes
  - Greatly simplifies some existing proofs
- Optimality using factorization inequalities
  - As a new tool to study optimality of inner bounds
- Extremal inequalities to compute extremal auxiliaries
  - Cardinality constrained inequalities are richer and messier, but essential
- Mentioned some open problems and conjectures
- Few more exciting ideas left out from this talk

WANTED: A new (†) proof of the factorization of

$$T_\lambda(X) := \mathfrak{C}[I(X;Y) - \lambda I(X;Z)], \quad \lambda \geq 1$$

CONCLUDING REMARKS

- New representations using concave envelopes
  - Greatly simplifies some existing proofs
- Optimality using factorization inequalities
  - As a new tool to study optimality of inner bounds
- Extremal inequalities to compute extremal auxiliaries
  - Cardinality constrained inequalities are richer and messier, but essential
- Mentioned some open problems and conjectures
- Few more exciting ideas left out from this talk

WANTED: A new (†) proof of the factorization of

$$T_\lambda(X) := \mathfrak{C}[I(X;Y) - \lambda I(X;Z)], \quad \lambda \geq 1$$

Thank You