Recent progress in multiuser information theory with correlated sources

Venkat Anantharam
University of California, Berkeley

Presenting joint work with

Amin Aminzadeh Gohari
Outline

- Introduction
- “Common Information” of dependent random variables
- Some applications of the “Potential function method”
- Conclusions
Multiuser Information Theory

- One studies the rate region of achievable communication rates, typically assuming independent messages from the sources.

- However
  - The messages may be correlated because of a network state affecting multiple nodes
  - Correlated messages may arise in sensor networks
  - The use of multiple paths by a single source may be modeled as correlated sources
  - Correlation may be deliberately introduced as common randomness
Point-to-point communication: separation is optimal

- Shannon showed that in a point-to-point scenario separation is optimal.
In a network, separation is **not necessarily optimal**

- In a network scenario, separation may fail (Cover, El Gamal, Salehi (1980))
  - The partial information of the two transmitters about each other is destroyed in separation.
  - Correlation allows for partial cooperation between the transmitters.

![Diagram of network encoding and decoding process](image-url)
Correlated sources in the literature

Huge amount of work... various scenarios: channel fading; receiver feedback; side information at the receiver...

<table>
<thead>
<tr>
<th>Broadcast channel</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Multiple-Access Channels</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Interference channels</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Arbitrarily varying channel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahlswede, Cai (1997)</td>
</tr>
</tbody>
</table>
Outline

- Introduction

- “Common Information” of dependent random variables
  - Definition
  - Historical development and known results
  - Our contribution (new upper bound and lower bound)

- Some applications of “Potential function method”

- Conclusions
Common Information of a pair of random variables private from an eavesdropper

- Given random variables \((X, Y, Z)\), how can one quantify the common part of r.v.s \(X\) and \(Y\) that is independent of \(Z\)?

  - Application in a private communication system: secret key generation
Common Information of a pair of random variables private from an eavesdropper

- Given random variables \((X, Y, Z)\), how can one quantify the common part of r.v.s \(X\) and \(Y\) that is independent of \(Z\)?

- Special cases:
  - If \(Z\) is independent of \((X, Y)\) \(\rightarrow I(X; Y)\)
Common Information of a pair of random variables private from an eavesdropper

- Given random variables $(X, Y, Z)$, how can one quantify the common part of r.v.s $X$ and $Y$ that is independent of $Z$?

- Special cases:
  - If $Z$ is independent of $(X, Y)$ $\rightarrow I(X; Y)$
  - If $X = Y = K$ $\rightarrow H(K|Z)$.
Common Information of a pair of random variables private from an eavesdropper

- Given random variables \((X, Y, Z)\), how can one quantify the common part of r.v.s \(X\) and \(Y\) that is independent of \(Z\)?

- Special cases:
  - If \(Z\) is independent of \((X, Y)\) → \(I(X; Y)\)
  - If \(X = Y = K\) → \(H(K|Z)\).

- What about \(I(X; Y|Z)\)?
Common Information of a pair of random variables private from an eavesdropper

- Given random variables \((X, Y, Z)\), how can one quantify the common part of r.v.s \(X\) and \(Y\) that is independent of \(Z\)?

- What about \(I(X; Y | Z)\)?

- But \(I(X; Y | Z)\) can be greater than \(I(X; Y)\)!

\[
X \sim B\left(\frac{1}{2}\right), \quad Y \sim B\left(\frac{1}{2}\right), \quad X \perp Y, \quad Z = X \oplus Y
\]

\[
I(X; Y) = 0 < I(X; Y | Z) = 1
\]
Common Information of a pair of random variables private from an eavesdropper

- Given random variables \((X, Y, Z)\), how can one quantify the common part of r.v.s \(X\) and \(Y\) that is independent of \(Z\)?

- What about \(I(X; Y | Z)\)?

- \(S(X; Y || Z)\) defined later in the talk works.
For two random variables $(X, Y)$: a sense in which $I(X; Y)$ represents the common part of $X$ and $Y$
For two random variables $(X, Y)$: a sense in which $I(X; Y)$ represents the common part of $X$ and $Y$. 

| 1 | 2 | ... | i | ... | $2^nH(X|Y)$ |
For two random variables \((X, Y)\): a sense in which \(I(X; Y)\) represents the common part of \(X\) and \(Y\)
For two random variables \((X, Y)\): a sense in which \(I(X; Y)\) represents the common part of \(X\) and \(Y\).

\[
X^n \quad H(X^n) = H(X^n | Y^n) + I(X^n; Y^n)
\]

Total Common Randomness

Common Randomness due to communication

Extracted Common Rand.

\[
F
\]

Bin Index

\[
I(X^n; Y^n)
\]

Index within the bin
For two random variables \((X, Y)\): a sense in which \(I(X; Y)\) represents the common part of \(X\) and \(Y\).

\[
I(K; F) \approx 0 \\
\frac{1}{n} H(K)
\]
For two random variables \((X, Y)\): a sense in which \(I(X; Y)\) represents the common part of \(X\) and \(Y\)

\[
I(K; F) \cong 0
\]

\[
\frac{1}{n} H(K)
\]
For two random variables \((X, Y)\): a sense in which 
\(I(X; Y)\) represents the common part of \(X\) and \(Y\)

\[
I(K; \overrightarrow{F}) \approx 0 \\
\frac{1}{n} H(K)
\]
The general case requiring secrecy

$$I(K; \overrightarrow{FZ^n}) \cong 0,$$

$$\frac{1}{n} H(K)$$
Definition of $S(X; Y\parallel Z)$

Alice $\rightarrow$ Bob
$F_1(X^{1:n})$
Bob $\rightarrow$ Alice
$F_2(Y^{1:n}, F_1)$
Alice $\rightarrow$ Bob
$F_3(X^{1:n}, F_1, F_2)$

Alice creates
$K_A(X^{1:n}, F)$
Bob creates
$K_B(Y^{1:n}, F)$
Requirements:
\[
P(K_A=K_B=k) > 1 - \varepsilon
\]
\[
\frac{1}{n} I(K; Z^{1:n}, F) \leq \varepsilon
\]
Outline

○ Introduction

○ “Common Information” of dependent random variables
  ○ Definition
  ○ Historical development and known results
  ○ Our contribution (new upper bound and lower bound)

○ Some applications of “Potential function method”

○ Conclusions
Historical development

- Known as the **Source model**, the model was developed in the context of “Information theoretic security”

- Evolutionary ancestors of the model

<table>
<thead>
<tr>
<th>Model</th>
<th>Type of advantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shannon’s one time pad</td>
<td>Common secret key</td>
</tr>
<tr>
<td>Wyner’s Wire-Tap Channel</td>
<td>Eve’s channel degraded</td>
</tr>
<tr>
<td>Csiszar &amp; J. Korner’s Broadcast channel</td>
<td>Directions at which Eve’s channel is worst</td>
</tr>
<tr>
<td>Maurer’s model</td>
<td>Public discussion</td>
</tr>
<tr>
<td>Maurer, Alhswede, Csiszar’s models: Source Model and Channel model</td>
<td>Public discussion and/or Quality of observations</td>
</tr>
</tbody>
</table>
Power of public discussion

\[ E \sim B(\epsilon), D \sim B(\delta), \delta < \epsilon < 0.5 \]

\[ X \rightarrow Y = X \oplus E \]

\[ Z = X \oplus D \]
Power of public discussion

\[ E \sim B(\epsilon), D \sim B(\delta), \delta < \epsilon < 0.5 \]
Power of public discussion

\[ E \sim B(\epsilon), \ D \sim B(\delta), \ \delta < \epsilon < 0.5 \]
Power of public discussion

\[ E \sim B(\epsilon), \quad D \sim B(\delta), \quad \delta < \epsilon < 0.5 \]

\[ V \oplus X \oplus E \text{ sent on the public channel} \]

\[ Z = X \oplus D \]

\[ V \oplus X \oplus E \]
Power of public discussion

\[ E \sim B(\epsilon), \; D \sim B(\delta), \; \delta < \epsilon < 0.5 \]
Power of public discussion

\[ E \sim B(\epsilon), D \sim B(\delta), \delta < \epsilon < 0.5 \]

- Sent on the public channel:
  - \( V \oplus X \oplus E \)
  - \( V \oplus X \oplus E \)
  - \( V \oplus E \)
  - \( V \oplus X \oplus E \)
  - \( V \oplus E \oplus D \)

\[ Z = X \oplus D \]

\[ (V \oplus X \oplus E) \oplus (X \oplus D) = V \oplus E \oplus D \]
<table>
<thead>
<tr>
<th>Authors</th>
<th>Lower bounds on $S(X; Y \parallel Z)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maurer (1993)</td>
<td>$\max{I(X; Y) - I(X; Z), I(Y; X) - I(Y; Z)}$</td>
</tr>
<tr>
<td>Ahlswede and Csiszár (1993)</td>
<td>$\max\left( \sup_{U} I(U; Y</td>
</tr>
</tbody>
</table>

Proof idea for $I(X; Y) - I(X; Z)$:

The $X^n$ space is first partitioned into $2^{n \cdot H(Y | X)}$ bins of size $2^{n \cdot I(X; Y)}$, i.e. the Slepian-Wolf binning strategy; each bin is then further partitioned into $2^{n [I(X; Y) - I(X; Z)]}$ bins of size $2^{n \cdot I(X; Z)}$. 
### Known upper bounds on $S(X; Y\|Z)$

<table>
<thead>
<tr>
<th>Authors</th>
<th>Upper bounds on $S(X; Y|Z)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maurer (1993)</td>
<td>$\min(I(X; Y), I(X; Y</td>
</tr>
<tr>
<td></td>
<td>Idea: classical arguments, e.g.</td>
</tr>
<tr>
<td></td>
<td>$H(K_A) = nI(X;Y</td>
</tr>
<tr>
<td></td>
<td>$H(K_A) = nI(X;Y) + H(K_A</td>
</tr>
<tr>
<td>Maurer and Wolf (1999)</td>
<td>$I(X; Y \downarrow Z) := \inf_{XY-Z-T} I(X; Y</td>
</tr>
<tr>
<td></td>
<td>Idea: decreasing the information of Eve cannot decrease the common private information</td>
</tr>
<tr>
<td>Renner and Wolf (2003)</td>
<td>$\inf_U (H(U) + I(X; Y \downarrow ZU))$</td>
</tr>
<tr>
<td></td>
<td>Idea: providing Eve with a random variable $U$ cannot decrease the common private information</td>
</tr>
<tr>
<td></td>
<td>by more than $H(U)$ bits.</td>
</tr>
</tbody>
</table>
Outline

- Introduction
- “Common Information” of dependent random variables
  - Definition
    - Historical development and known results
    - Our contribution (new upper bound and lower bound)
  - Some applications of “Potential function method”
- Conclusions
The Goal

- Given $\psi(X; Y \| Z)$, we would like to show that

$$\psi(X; Y \| Z) \geq S(X; Y \| Z)$$
The Goal

- Given $\psi(X; Y\|Z)$, we would like to show that

$$\psi(X; Y\|Z) \geq S(X; Y\|Z)$$

- Find properties that $S(X; Y\|Z)$ has
The Goal

- Given $\psi(X; Y \parallel Z)$, we would like to show that

\[ \psi(X; Y \parallel Z) \geq S(X; Y \parallel Z) \]

- Find properties that $S(X; Y \parallel Z)$ has

- Consider the set of all functions that have those properties
The Goal

- Given $\psi(X; Y \| Z)$, we would like to show that

$$\psi(X; Y \| Z) \geq S(X; Y \| Z)$$

- Find properties that $S(X; Y \| Z)$ has

- Consider the set of all functions that have those properties

- Prove that each of them is an upper bound
Some properties of $S(X; Y \parallel Z)$

1) $n \cdot S(X; Y \parallel Z) \geq S(X^n; Y^n \parallel Z^n), \ \forall n, p(x, y, z)$
Some properties of $S(X; Y\|Z)$

1) $n \cdot S(X; Y\|Z) \geq S(X^n; Y^n\|Z^n), \quad \forall n, p(x, y, z)$
Some properties of $S(X; Y \parallel Z)$

1) $n \cdot S(X; Y \parallel Z) \geq S(X^n; Y^n \parallel Z^n)$, $\forall n, p(x, y, z)$

2) $\forall F : H(F|X) = 0$ or $H(F|Y) = 0$,

$\rightarrow S(X; Y \parallel Z) \geq S(XF; YF \parallelZF)$
Some properties of $S(X; Y \parallel Z)$

1) $n \cdot S(X; Y \parallel Z) \geq S(X^n; Y^n \parallel Z^n), \ \forall n, p(x, y, z)$
2) $\forall F : H(F|X) = 0$ or $H(F|Y) = 0,$
   $\rightarrow S(X; Y \parallel Z) \geq S(XF; YF \parallel ZF')$
Some properties of $S(X; Y \| Z)$

1) $n \cdot S(X; Y \| Z) \geq S(X^n; Y^n \| Z^n), \; \forall n, p(x, y, z)$

2) $\forall F: H(F|X) = 0$ or $H(F|Y) = 0,$

   $\rightarrow S(X; Y \| Z) \geq S(XF; YF \| ZF')$

3) $\forall X', Y': H(X'|X) = 0, H(Y'|Y) = 0,$

   $\rightarrow S(X; Y \| Z) \geq S(X'; Y'|Z)$
Some properties of $S(X; Y\|Z)$

1) $n \cdot S(X; Y\|Z) \geq S(X^n; Y^n\|Z^n)$, $\forall n, p(x, y, z)$

2) $\forall F : H(F|X) = 0$ or $H(F|Y) = 0$, 
   $\rightarrow S(X; Y\|Z) \geq S(XF; YF\|ZF)$

3) $\forall X', Y' : H(X'|X) = 0, H(Y'|Y) = 0$, 
   $\rightarrow S(X; Y\|Z) \geq S(X'; Y'|Z)$

4) $S(X; Y\|Z) \geq H(X|Z) - H(X|Y) = I(X; Y) - I(X; Z)$
$S'(\text{Alice’s information}; \text{Bob’s information} \parallel \text{Eve’s information})$ is a non-increasing potential function

Take an arbitrary $p(x, y, z)$ and an arbitrary strategy of length $n$

\[ n \cdot S(X; Y \parallel Z) \geq S(X^n; Y^n \parallel Z^n) \]

Property used here: 1) $n \cdot S(X; Y \parallel Z) \geq S(X^n; Y^n \parallel Z^n)$
$S(\text{Alice’s information; Bob’s information} \parallel \text{Eve’s information})$ is a non-increasing potential function

Take an arbitrary $p(x, y, z)$ and an arbitrary strategy of length $n$

\[ n \cdot S(X; Y \parallel Z) \geq S(X^n; Y^n \parallel Z^n) \geq S(X^nF_1; Y^nF_1 \parallel Z^nF_1) \]

Property used here: 2) $\forall F : H(F|X) = 0$ or $H(F|Y) = 0$,
\[ \rightarrow S(X; Y \parallel Z) \geq S(XF; YF \parallelZF) \]
$S$(Alice’s information; Bob’s information|| Eve’s information) is a non-increasing potential function

Take an arbitrary $p(x, y, z)$ and an arbitrary strategy of length $n$

\[ n \cdot S(X; Y\| Z) \geq S(X^n; Y^n\| Z^n) \]
\[ \geq S(X^nF_1; Y^nF_1\| Z^nF_1) \]
\[ \geq S(X^nF_1F_2; Y^nF_1F_2\| Z^nF_1F_2) \]

Property used here: 2) $\forall F : H(F|X) = 0$ or $H(F|Y) = 0$, \[ \rightarrow S(X; Y\| Z) \geq S(XF; YF\|ZF) \]
\( S(\text{Alice’s information}; \text{Bob’s information} \parallel \text{Eve’s information}) \) is a non-increasing potential function

Take an arbitrary \( p(x, y, z) \) and an arbitrary strategy of length \( n \)

\[
\begin{align*}
n \cdot S(X; Y \parallel Z) &\geq S(X^n; Y^n \parallel Z^n) \\
&\geq S(X^n F_1; Y^n F_1 \parallel Z^n F_1) \\
&\geq S(X^n F_1 F_2; Y^n F_1 F_2 \parallel Z^n F_1 F_2) \geq \ldots \\
&\geq S(X^n \overrightarrow{F}; Y^n \overrightarrow{F} \parallel Z^n \overrightarrow{F})
\end{align*}
\]

Property used here: 2) \( \forall F : H(F|X) = 0 \) or \( H(F|Y) = 0 \),

\[
\rightarrow S(X; Y \parallel Z) \geq S(X F; Y F \parallel Z F')
\]
$S(\text{Alice’s information; Bob’s information} \parallel \text{Eve’s information})$ is a non-increasing potential function

Take an arbitrary $p(x, y, z)$ and an arbitrary strategy of length $n$

$n \cdot S(X; Y \parallel Z) \geq S(X^n; Y^n \parallel Z^n)$

$\geq S(X^n F_1; Y^n F_1 \parallel Z^n F_1)$

$\geq S(X^n F_1 F_2; Y^n F_1 F_2 \parallel Z^n F_1 F_2) \geq ...$

$\geq S(X^n \overrightarrow{F}; Y^n \overrightarrow{F} \parallel Z^n \overrightarrow{F})$

$\geq S(K_A; K_B \parallel Z^n \overrightarrow{F})$

Property used here: 3) $\forall X', Y': H(X'|X) = 0, H(Y'|Y) = 0,$

$\rightarrow S(X; Y \parallel Z) \geq S(X'; Y' \parallel Z)$
$S(\text{Alice’s information}; \text{Bob’s information}\parallel \text{Eve’s information})$ is a non-increasing potential function

Take an arbitrary $p(x, y, z)$ and an arbitrary strategy of length $n$

$n \cdot S(X; Y \parallel Z) \geq S(X^n; Y^n \parallel Z^n)$

$\geq S(X^nF_1; Y^nF_1 \parallel Z^nF_1)$

$\geq S(X^nF_1F_2; Y^nF_1F_2 \parallel Z^nF_1F_2) \geq \ldots$

$\geq S(X^n\overrightarrow{F}; Y^n\overrightarrow{F} \parallel Z^n\overrightarrow{F})$

$\geq S(K_A; K_B \parallel Z^n\overrightarrow{F})$

$\geq H(K_A | Z^n\overrightarrow{F}) - H(K_A | K_B Z^n\overrightarrow{F})$

Property used here: 4) $S(X; Y \parallel Z) \geq H(X | Z) - H(X | Y)$
$S(Alice's \text{ information}; Bob's \text{ information} || Eve's \text{ information})$ is a non-increasing potential function

Take an arbitrary $p(x, y, z)$ and an arbitrary strategy of length $n$

\[
n \cdot S(X; Y \| Z) \geq S(X^n; Y^n \| Z^n) \geq S(X^n F_1; Y^n F_1 \| Z^n F_1) \geq S(X^n F_1 F_2; Y^n F_1 F_2 \| Z^n F_1 F_2) \geq \ldots \\
\geq S(X^n \overrightarrow{F}; Y^n \overrightarrow{F} \| Z^n \overrightarrow{F}) \geq S(K_A; K_B \| Z^n \overrightarrow{F}) \geq H(K_A | Z^n \overrightarrow{F}) - H(K_A | K_B Z^n \overrightarrow{F}) \approx H(K_A)
\]
Properties required of the functions of interest

1) \( n \cdot \psi(X; Y\|Z) \geq \psi(X^n; Y^n\|Z^n), \ \forall n, p(x, y, z) \)

2) \( \forall F : H(F|X) = 0 \text{ or } H(F|Y) = 0, \)
   \( \rightarrow \psi(X; Y\|Z) \geq \psi(XF; YF\|ZF) \)

3) \( \forall X', Y' : H(X'|X) = 0, H(Y'|Y) = 0, \)
   \( \rightarrow \psi(X; Y\|Z) \geq \psi(X'; Y'|Z) \)

4) \( \psi(X; Y\|Z) \geq H(X|Z) - H(X|Y) \)
Proving that any function that satisfies the properties is an upper bound

Take an arbitrary \( p(x, y, z) \) and an arbitrary strategy of length \( n \)

Can write the same chain of inequalities:

\[
\begin{align*}
n \cdot \psi(X; Y \parallel Z) & \geq \psi(X^n; Y^n \parallel Z^n) \\
& \geq \psi(X^n F_1; Y^n F_1 \parallel Z^n F_1) \\
& \geq \psi(X^n F_1 F_2; Y^n F_1 F_2 \parallel Z^n F_1 F_2) \geq \ldots \\
& \geq \psi(X^n \overrightarrow{F}; Y^n \overrightarrow{F} \parallel Z^n \overrightarrow{F}) \\
& \geq \psi(K_A; K_B \parallel Z^n \overrightarrow{F}) \\
& \geq H(K_A|Z^n \overrightarrow{F}) - H(K_A|K_B Z^n \overrightarrow{F}) \equiv H(K_A)
\end{align*}
\]

Conclusion: \( \forall p(x, y, z), n: n \cdot \psi(X; Y \parallel Z) \geq H(K_A) \)
Example: $I(X; Y|Z)$ is an upper bound

1) $n \cdot I(X; Y|Z) \geq I(X^n; Y^n|Z^n), \ \forall n, p(x, y, z)$  

2) $\forall F : H(F|X) = 0$ or $H(F|Y) = 0,$

   $\rightarrow I(X; Y|Z) \geq I(XF; YF|ZF)$  

   since if $H(F|X) = 0$:

   $I(X; Y|Z) = I(XF; Y|Z) = I(F; Y|Z) + I(XF; YF|ZF)$

3) $\forall X', Y' : H(X'|X) = 0, H(Y'|Y) = 0,$

   $\rightarrow I(X; Y|Z) \geq I(X'; Y'|Z)$  

4) $I(X; Y|Z) \geq H(X|Z) - H(X|Y)$
Strategy for finding a new upper bound

○ Take an existing outer bound that verifies the properties

○ Perturb the expression of the outer bound

○ Check whether the properties are still satisfied:
Strategy for finding a new upper bound

- Take an existing outer bound that verifies the properties
- Perturb the expression of the outer bound
- Check whether the properties are still satisfied:
  - Yes!
    - Hopefully it is strictly better than the existing bound
  - No.
    - See which property is violated and why?
    - Trial and error: Try to change the perturbation in a way that it works
Our new upper bound

For any increasing convex function $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, $S(X; Y\|Z)$ is bounded from above by

$$\inf_J f^{-1}\{f(S(X; Y\|J)) + S_{f-\text{one-way}}(XY; J^{(s)}\|Z)\}$$

where

$$S_{f-\text{one-way}}(A; B^{(s)}\|C) = \sup_{U-V-A-BC}[f(H(U|ZV)) - f(H(U|YV))]$$

leads to an upper bound when $S(X; Y\|J)$ is bounded from above by $I(X; Y|J)$.
Comparison with Renner and Wolf upper bound (I)

\[ \inf_J \left( I(XY; J|Z) + I(X; Y|J) \right) \] is a computable expression that is greater than or equal to our new upper bound.

\[ \bullet \inf_J \left( I(XY; J|Z) + I(X; Y|J) \right) \leq \inf_U (H(U) + I(X; Y \downarrow ZU)) \]

\[ \inf_U (H(U) + I(X; Y \downarrow ZU)) = \inf_U \left( H(U) + \min_J - ZU - XY I(X; Y|J) \right) \]

Assume minimum occurs at \( J_U \):

\[ \inf_U (H(U) + I(X; Y \downarrow ZU)) = \inf_U \left( H(U) + I(X; Y|J_U) \right) \geq \inf_U \left( I(XY; ZU|Z) + I(X; Y|J_U) \right) \geq \inf_U \left( I(XY; J_U|Z) + I(X; Y|J_U) \right) \geq \inf_J \left( I(XY; J|Z) + I(X; Y|J) \right) \]
Comparison with Renner and Wolf upper bound (II)

• There is an example for which the inequality is strict

$$\inf_J \left( I(XY; J|Z) + I(X; Y|J) \right) < \inf_U \left( H(U) + I(X; Y \downarrow ZU) \right)$$

Idea: perturbing the Renner-Wolf example:

<table>
<thead>
<tr>
<th></th>
<th>X</th>
<th>Y</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1/8</td>
<td>1/8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1/8</td>
<td>1/8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1/4</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1/4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$$Z = \left\{ \begin{array}{ll} X + Y \pmod{2} & \text{if } X, Y \in \{0, 1\} \\ X \pmod{2} & \text{if } X \in \{2, 3\} \end{array} \right. \quad U = \left\lfloor \frac{X}{2} \right\rfloor$$
Comparison with Renner and Wolf upper bound (III)

We perturb the mentioned example. Since the RW bound does not behave as smoothly as the new bound behaves, the new bound outperforms the RW bound.

We find a binary random variable $V$ of small entropy satisfying $V - U - XYZ$ such that the new bound is strictly better than the double intrinsic information bound for the triple $\tilde{X} = X$, $\tilde{Y} = Y$, $\tilde{Z} = (Z, V)$. Proof idea:

- Assuming that the RW bound and ours are the same at $(\tilde{X}, \tilde{Y}, \tilde{Z})$, we prove that: For any sequence of $U_i$'s such that $H(U_i) + I(X; Y \downarrow ZU_i) \rightarrow \inf_U [H(U) + I(X; Y \downarrow ZU)]$ as $i \rightarrow \infty$, we must have $H(U_i) \rightarrow 0$.

- Since the intrinsic information is a continuous function, $H(U_i) + I(X; Y \downarrow ZU_i)$ must converge to $I(X; Y \downarrow Z)$ which is equal to $\frac{3}{2}$. Hence the RW bound would be around $\frac{3}{2}$ at $(\tilde{X}, \tilde{Y}, \tilde{Z})$.

- Since our bound is continuous, and is around 1 at $(X, Y, Z)$, it has to be close to 1 at $(\tilde{X}, \tilde{Y}, \tilde{Z})$. Contradiction!
Outline

- Introduction
- “Common Information” of dependent random variables
  - Definition
  - Historical development and known results
  - Our contribution (new upper bound and lower bound)
- Some applications of “Potential function method”
- Conclusions
New Lower Bound

Given

\[ U_1 - X - YZ; t_1 := I(U_1; Y) - I(U_1; Z) \]

\[ U_2 - YU_1 - XZU_1; t_2 := I(U_2; Y|U_1) - I(U_2; Z|U_1) \]

\[ U_3 - XU_1U_2 - YZU_1U_2; t_3 := I(U_3; Y|U_1U_2) - I(U_3; Z|U_1U_2) \]

\[ \ldots \]

\[ S(X; Y||Z) \geq \sum_{i=p}^{q} t_i \]
Comparison with Ahlswede and Csiszárs lower bound

- A generalization of Ahlswede and Csiszárs lower bound (new feature: interactive communication): $U_1 = V, U_2 = 0, U_3 = U, p = q = 3$
Comparison with Ahlswede and Csiszárs lower bound

• A generalization of Ahlswede and Csiszárs lower bound (new feature: interactive communication): \( U_1 = V, U_2 = 0, U_3 = U, p = q = 3 \)

• There is an example for which the new bound is strictly better: it is tight for the example provided in Ahlswede and Csisz’ar to show that their bound is not tight: Choice of \( X = (X_1, X_2), Y = (Y_1, Y_2), Z = (Z_1, Z_2) \)

\[
X_1 - Y_1 - Z_1, \quad Y_2 - X_2 - Z_2, \quad I(X_1Y_1Z_1; X_2Y_2Z_2) = 0
\]
Proof idea

For simplicity assume $p = 1$, $q = 2$:

\[
U^n_1 - X^n - Y^n Z^n \\
U^n_2 - Y^n U^n_1 - X^n Z^n
\]

- $U^n_1$ is created by $X^n$ and transmitted to $Y^n$ using Slepian-Wolf bin index $F_1$
- $U^n_2$ is created by $Y^n U^n_1$ transmitted to $X^n$ using Slepian-Wolf bin index $F_2$
- Generated key:

\[
H(U^n_1 U^n_2 | F_1 F_2 Z^n) = H(U^n_1 | F_1 F_2 Z^n) + H(U^n_2 | U^n_1 F_1 F_2 Z^n) = \\
H(U^n_1 | F_1 F_2 Z^n) + H(U^n_2 | U^n_1 F_2 Z^n)
\]

If $t_2 > 0$, $F_2 \perp (Z^n U^n_1) \Rightarrow H(U^n_1 | F_1 F_2 Z^n) = H(U^n_1 | F_1 Z^n) = [t_1] + \\
H(U^n_2 | U^n_1 F_2 Z^n) = t_2 \implies \frac{1}{n} H(U^n_1 U^n_2 | F_1 F_2 Z^n) \geq t_1 + t_2$

If $t_2 \leq 0$, $H(U^n_1 | F_1 F_2 Z^n) \geq H(U^n_1 | F_1 Z^n) - n|t_2| \implies \frac{1}{n} H(U^n_1 U^n_2 | F_1 F_2 Z^n) \geq t_1 + t_2$
Outline

- Introduction
- “Common Information” of dependent random variables
- Some applications of “Potential function method”
  - Multiterminal channel coding
    - General multiterminal network with correlated sources (a generalized cut-set bound)
      - Broadcast channel with correlated sources
      - Strong Interference channel with correlated sources
  - Source coding
  - Channel Model
- Conclusions
Review: Traditional Cut-Set Bound

Given $q(y^{(1)}, y^{(2)}, \ldots, y^{(m)} | x^{(1)}, x^{(2)}, \ldots, x^{(m)})$, and rates $R^{(i,j)} : i^{th} \rightarrow j^{th}$

$$\exists p(x^{(1)}, x^{(2)}, \ldots, x^{(m)}) : \forall K \subset \{1, 2, \ldots, m\}$$
$$\sum_{i \in K, j \in K^c} R^{(i,j)} \leq I \left( X^{(i)} : i \in K ; Y^{(j)} : j \in K^c | X^{(j)} : j \in K^c \right)$$

where

$X^{(1)}, \ldots, X^{(m)}, Y^{(1)}, \ldots, Y^{(m)} \sim p(x^{(1)}, \ldots, x^{(m)}) q(y^{(1)}, \ldots, y^{(m)} | x^{(1)}, \ldots, x^{(m)})$
Intuitive description of the new setup: Channel specifications

\[ q(y^{(1)}, y^{(2)}, \ldots, y^{(m)} | x^{(1)}, x^{(2)}, \ldots, x^{(m)}) \]

\( \Psi \) : A permissible set of input distributions, e.g. coupled magnitude constraint across inputs; MAC or IC with independent messages
Intuitive description of the new setup:
The communication task

\[ p(w^{(1)}, w^{(2)}, \ldots, w^{(m)}) \]

\[ M^{(i)} = f^{(i)}(W^{(1)}, W^{(2)}, W^{(3)}, \ldots, W^{(m)}) \]

\[ \Delta^{(i)} : M^{(i)} \times M^{(i)} \to [0, \infty) : \Delta^{(i)}(m^{(i)}, m^{(i)}) = 0 \]

\[ D^{(i)} : \text{average distortion constraint} \]
Admissibility of a source marginal distribution

- Given \( q(y^{(1)}, y^{(2)}, ..., y^{(m)}|x^{(1)}, x^{(2)}, ..., x^{(m)}), \Psi, \)

  \[ f^{(i)}: \mathcal{W}^{(1)} \times \mathcal{W}^{(2)} \times ... \times \mathcal{W}^{(m)} \rightarrow \mathcal{M}^{(i)}, \]

  \( \Delta^{(i)} \) and \( D^{(i)} \) for \( i = 1, 2, ..., m, \)

- source marginal distribution \( p(w^{(1)}, w^{(2)}, ..., w^{(m)}) \) is called admissible if

  - \( \forall \epsilon : \exists n : \)

  - The \( i^{th} \) party observes \( n \) i.i.d. copies of \( W^{(i)} \), i.e. \( W_{1:n}^{(i)}, \)

  - \( n \)-use of the multiterminal network is allowed,

  - Following the communication, the \( i^{th} \) party creates \( \hat{M}_{1:n}^{(i)} \) such that

  \[
  \frac{1}{n} \sum_{j=1}^{n} \mathbb{E} \left[ \Delta^{(i)}_{n} \left( \hat{M}_{j}^{(i)}, M_{j}^{(i)} \right) \right] \leq D^{(i)} + \epsilon, \]

  where \( M_{k}^{(i)} = f^{(i)}(W_{k}^{(1)}, W_{k}^{(2)}, ..., W_{k}^{(m)}). \)
A Generalized Cut-Set Bound

- if source marginal distribution \( p(w^{(1)}, w^{(2)}, ..., w^{(m)}) \) is admissible then
  - \( \exists \) joint distribution \( p(x^{(1)}, x^{(2)}, ..., x^{(m)}, z) \) such that \( p(x^{(1)}, x^{(2)}, ..., x^{(m)}|z) \in \Psi \) for any value \( z \)
  - \( \exists \) conditional distribution \( p(\hat{m}^{(1)}, \hat{m}^{(2)}, ..., \hat{m}^{(m)}|w^{(1)}, w^{(2)}, ..., w^{(m)}) \)
    such that \( \Delta^{(i)}(M^{(i)}, \hat{M}^{(i)}) \leq D^{(i)} \) where \( M^{(i)} = f^{(i)}(W^{(1)}, ..., W^{(m)}) \)
  - such that for any arbitrary \( K \subset \{1, 2, 3, ..., m\} \) the following inequality holds:
    \[
    I(W^{(i)} : i \in K; \hat{M}^{(j)} : j \in K^c|W^{(j)} : j \in K^c) \leq I(X^{(i)} : i \in K; Y^{(j)} : j \in K^c|X^{(j)} : j \in K^c, Z),
    \]
  where \( Y^{(1)}, Y^{(2)}, ..., Y^{(m)}, X^{(1)}, X^{(2)}, ..., X^{(m)} \) and \( Z \) are jointly distributed according to
  \[
  q(y^{(1)}, y^{(2)}, ..., y^{(m)}|x^{(1)}, x^{(2)}, ..., x^{(m)}) \cdot p(x^{(1)}, x^{(2)}, ..., x^{(m)}, z).\]
Example

\[ W(i) = (E^{(i,1)}, E^{(i,2)}, \ldots, E^{(i,m)}) , \]
\[ E^{(i,j)} \text{ mutually independent} \]
\[ M(i) = (E^{(1,i)}, E^{(2,i)}, \ldots, E^{(m,i)}) \]
\[ D(i) = 0, \Delta^{(i)}(m(i), m'(i)) = 1[m(i) \neq m'(i)] , \]
\[ \Psi = \text{ set of all joint distributions} \]

Then
\[ I\left(W(i) : i \in K ; \widehat{M}(j) : j \in K^c | W(j) : j \in K^c \right) \leq \]
\[ I\left(X(i) : i \in K ; Y(j) : j \in K^c | X(j) : j \in K^c, Z \right) , \]

becomes
\[ \sum_{i \in K, j \in K^c} H(E^{(i,j)}) \leq I\left(X(i) : i \in K ; Y(j) : j \in K^c | X(j) : j \in K^c, Z \right) \leq \]
\[ I\left(X(i) : i \in K ; Y(j) : j \in K^c | X(j) : j \in K^c \right) , \]

since \( Y(1), Y(2), \ldots, Y(m) - X(1), X(2), \ldots, X(m) - Z \)
Intuitive description of the result

Traditional cut-set bound ($m = 3$): $\exists p(x^{(1)}, x^{(2)}, x^{(3)})$ such that:

$$
\begin{bmatrix}
R^{(1,2)} + R^{(1,3)} \\
R^{(2,1)} + R^{(2,3)} \\
R^{(3,1)} + R^{(3,2)} \\
R^{(1,3)} + R^{(2,3)} \\
R^{(1,2)} + R^{(3,2)} \\
R^{(2,1)} + R^{(3,1)}
\end{bmatrix}
\leq
\begin{bmatrix}
I(X^{(1)}; Y^{(2)}Y^{(3)}|X^{(2)}X^{(3)}) \\
I(X^{(2)}; Y^{(1)}Y^{(3)}|X^{(1)}X^{(3)}) \\
I(X^{(3)}; Y^{(1)}Y^{(2)}|X^{(1)}X^{(2)}) \\
I(X^{(1)}X^{(2)}; Y^{(3)}|X^{(3)}) \\
I(X^{(1)}X^{(3)}; Y^{(2)}|X^{(2)}) \\
I(X^{(2)}X^{(3)}; Y^{(1)}|X^{(1)})
\end{bmatrix}
$$
Intuitive description of the result

Traditional cut-set bound \((m = 3)\): \(\exists p(x^{(1)}, x^{(2)}, x^{(3)})\) such that:

\[
\begin{bmatrix}
R^{(1,2)} + R^{(1,3)} \\
R^{(2,1)} + R^{(2,3)} \\
R^{(3,1)} + R^{(3,2)} \\
R^{(1,3)} + R^{(2,3)} \\
R^{(1,2)} + R^{(3,2)} \\
R^{(2,1)} + R^{(3,1)}
\end{bmatrix}
\in \Pi \left( \begin{bmatrix}
I(X^{(1)}; Y^{(2)} Y^{(3)} | X^{(2)} X^{(3)}) \\
I(X^{(2)}; Y^{(1)} Y^{(3)} | X^{(1)} X^{(3)}) \\
I(X^{(3)}; Y^{(1)} Y^{(2)} | X^{(1)} X^{(2)}) \\
I(X^{(1)} X^{(2)}; Y^{(3)} | X^{(3)}) \\
I(X^{(1)} X^{(3)}; Y^{(2)} | X^{(2)}) \\
I(X^{(2)} X^{(3)}; Y^{(1)} | X^{(1)})
\end{bmatrix} \right)
\]

where \(\Pi(A) := \{ \vec{v} \in \mathbb{R}^6_+ : \vec{v} \leq \vec{w} \text{ for some } \vec{w} \in A \} \)
Intuitive description of the result

Traditional cut-set bound ($m = 3$):

$$
\begin{bmatrix}
R^{(1,2)} + R^{(1,3)} \\
R^{(2,1)} + R^{(2,3)} \\
R^{(3,1)} + R^{(3,2)} \\
R^{(1,3)} + R^{(2,3)} \\
R^{(1,2)} + R^{(3,2)} \\
R^{(2,1)} + R^{(3,1)}
\end{bmatrix}
\in \bigcup_{\Psi} \Pi\left(\left\{ \begin{bmatrix}
I(X^{(1)}; Y^{(2)}Y^{(3)} | X^{(2)}X^{(3)}) \\
I(X^{(2)}; Y^{(1)}Y^{(3)} | X^{(1)}X^{(3)}) \\
I(X^{(3)}; Y^{(1)}Y^{(2)} | X^{(1)}X^{(2)}) \\
I(X^{(1)}X^{(2)}; Y^{(3)} | X^{(3)}) \\
I(X^{(1)}X^{(3)}; Y^{(2)} | X^{(2)}) \\
I(X^{(2)}X^{(3)}; Y^{(1)} | X^{(1)})
\end{bmatrix} \right\} \right)
$$

where $\Pi(A) := \{ \vec{v} \in \mathbb{R}_+^6 : \vec{v} \leq \vec{w} \text{ for some } \vec{w} \in A \}$

Here $\Psi =$ set of all possible input distributions
Intuitive description of the result

Traditional cut-set bound \((m = 3)\):

\[
\begin{bmatrix}
R^{(1,2)} + R^{(1,3)} \\
R^{(2,1)} + R^{(2,3)} \\
R^{(3,1)} + R^{(3,2)} \\
R^{(1,3)} + R^{(2,3)} \\
R^{(1,2)} + R^{(3,2)} \\
R^{(2,1)} + R^{(3,1)}
\end{bmatrix}
\in \bigcup_{\Psi} \prod_{\left\{ \begin{bmatrix}
I(X^{(1)}; Y^{(2)}Y^{(3)}|X^{(2)}X^{(3)}) \\
I(X^{(2)}; Y^{(1)}Y^{(3)}|X^{(1)}X^{(3)}) \\
I(X^{(3)}; Y^{(1)}Y^{(2)}|X^{(1)}X^{(2)}) \\
I(X^{(1)}X^{(2)}; Y^{(3)}|X^{(3)}) \\
I(X^{(1)}X^{(3)}; Y^{(2)}|X^{(2)}) \\
I(X^{(2)}X^{(3)}; Y^{(1)}|X^{(1)})
\end{bmatrix} \right\}} \Phi \left( q(y^{(1)},y^{(2)},y^{(3)}|x^{(1)},x^{(2)},x^{(3)}), \Psi \right)
\]

Here \(\Psi = \) set of all possible input distributions
Intuitive description of the result

Traditional cut-set bound \((m = 3)\):

\[
\Pi \left( \begin{array}{cccc}
R^{(1,2)} + R^{(1,3)} \\
R^{(2,1)} + R^{(2,3)} \\
R^{(3,1)} + R^{(3,2)} \\
R^{(1,3)} + R^{(2,3)} \\
R^{(1,2)} + R^{(3,2)} \\
R^{(2,1)} + R^{(3,1)} \\
\end{array} \right) = \Phi \left( \text{virtual channel, \{uniform distribution\}} \right)
\]
Intuitive description of the result

Traditional cut-set bound ($m = 3$):

$$
\Phi\left(\text{virtual channel} , \{\text{uniform distribution}\}\right) \subset \Phi\left(\text{physical channel} , \text{set of all input distributions}\right)
$$
Intuitive description of the result

Traditional cut-set bound ($m = 3$):

$$\Phi\left(\text{virtual channel }, \{\text{uniform distribution}\}\right) \subset \Phi\left(\text{physical channel channel }, \text{set of all input distributions}\right)$$

Generalized cut-set bound ($m = 3$):

There must exist a virtual channel corresponding to the communication task

$$p(\hat{m}(1), \hat{m}(2), \hat{m}(3)|w(1), w(2), w(3))$$

such that

$$\triangle^{(i)}(M^{(i)}, \hat{M}^{(i)}) \leq D^{(i)}; \quad M^{(i)} = f^{(i)}(W^{(1)}, W^{(2)}, W^{(3)})$$

where

$$\Phi\left(p(\hat{m}(1), \hat{m}(2), \hat{m}(3)|w(1), w(2), w(3)), \{p(w(1), w(2), w(3))\}\right) \subset \text{convex hull } \left\{ \Phi\left(q(y(1), y(2), y(3)|x(1), x(2), x(3)), \psi\right) \right\}$$
Example: Nair and El Gamal’s outer bound on the general broadcast channel

\[ \Phi\left( \text{virtual channel}, \{ \text{uniform distribution} \} \right) \subset \Phi\left( \text{physical channel}, \text{set of all input distributions} \right) \]

where \( \Phi(q(y^{(1)}, y^{(2)}, y^{(3)}|x^{(1)}, x^{(2)}, x^{(3)}), \Psi) \) is equal to

\[ \Pi\left( \bigcup_{UX^{(1)} - Y^{(2)}Y^{(3)}; q(x^{(1)}) \in \Psi} \left[ \begin{array}{c} I(U; Y^{(2)}) \\ I(V; Y^{(3)}) \\ I(U; Y^{(2)}) + I(V; Y^{(3)}|U) \\ I(V; Y^{(3)}) + I(U; Y^{(2)}|V) \end{array} \right] \right) \]
Structure of the proof (assuming m=3)

Generalized cut-set bound ($m = 3$):
There must exist a virtual channel corresponding to the communication task
\[ p(\hat{m}^{(1)}, \hat{m}^{(2)}, \hat{m}^{(3)} | w^{(1)}, w^{(2)}, w^{(3)}) \]
where
\[ \Phi(p(\hat{m}^{(1)}, \hat{m}^{(2)}, \hat{m}^{(3)} | w^{(1)}, w^{(2)}, w^{(3)}), \{p(w^{(1)}, w^{(2)}, w^{(3)})\}) \subset \text{convex hull} \left\{ \Phi(q(y^{(1)}, y^{(2)}, y^{(3)} | x^{(1)}, x^{(2)}, x^{(3)}), \Psi) \right\} \] (1)

- \( \Phi \): the set of all 3-input/3-output discrete memoryless networks, and a subset of input distributions \( \rightarrow \) subsets of \( \mathbb{R}_+^6 \)

- The function \( \Phi \) has some properties, e.g. it involves mutual information terms that are continuous in its first argument

- **Question**: what are properties of \( \Phi \) that imply equation 1

  - Our perspective: \( \Phi \) as a function from all discrete memoryless networks, and all permissible input distributions; rather than \( \Phi \) evaluated at an arbitrary but fixed network and permissible input distribution
Structure of the proof (assuming m=3)

Generalized cut-set bound \((m = 3)\):

There must exist a virtual channel corresponding to the communication task

\[ p(\hat{m}^{(1)}, \hat{m}^{(2)}, \hat{m}^{(3)} | w^{(1)}, w^{(2)}, w^{(3)}) \]

where

\[ \Phi \left( p(\hat{m}^{(1)}, \hat{m}^{(2)}, \hat{m}^{(3)} | w^{(1)}, w^{(2)}, w^{(3)}), \{ p(w^{(1)}, w^{(2)}, w^{(3)}) \} \right) \subset \]

\[ \text{convex hull} \left\{ \Phi \left( q(y^{(1)}, y^{(2)}, y^{(3)} | x^{(1)}, x^{(2)}, x^{(3)}), \Psi \right) \right\} \] (2)

- \( \Phi \): the set of all 3-input/3-output discrete memoryless networks, and a subset of input distributions \( \rightarrow \) subsets of \( \mathbb{R}^6_+ \)

- **Identify properties** on such functions that imply equation 2

- **Verify** that the given expression belongs to the class of functions that satisfy those properties
The first three properties (applicable to specific multi-terminal problems)

1) If

\[ p(y^{(1)}, y^{(2)}, y^{(3)} | x^{(1)}, x^{(2)}, x^{(3)}) = \prod_{i=1}^{3} 1[y^{(i)} = x^{(i)}]. \]

Then \( \forall q(x^{(1)}, x^{(2)}, x^{(3)}) \),

\[ \phi\left( p(y^{(1)}, y^{(2)}, y^{(3)} | x^{(1)}, x^{(2)}, x^{(3)}), \{q(x^{(1)}, x^{(2)}, x^{(3)})\} \right) = \{(0, 0, 0, 0, 0, 0)\} \]
The first three properties (applicable to specific multi-terminal problems)

2) $\forall q(x^{(1)}, x^{(2)}, x^{(3)}), \forall \Psi \text{ s.t.} q(x'^{(1)}, x'^{(2)}, x'^{(3)}) \in \Psi$:

$$
\phi \left( p(y^{(1)}y'^{(1)}, y^{(2)}y'^{(2)}, y^{(3)}y'^{(3)}|x^{(1)}, x^{(2)}, x^{(3)}), \{q(x^{(1)}, x^{(2)}, x^{(3)})\} \right) \subseteq \\
\phi \left( p(y^{(1)}, y^{(2)}, y^{(3)}|x^{(1)}, x^{(2)}, x^{(3)}), \{q(x^{(1)}, x^{(2)}, x^{(3)})\} \right) \\
\oplus \phi \left( p(y'^{(1)}, y'^{(2)}, y'^{(3)}|x'^{(1)}, x'^{(2)}, x'^{(3)}), \Psi \right).
$$
The first three properties (applicable to specific multi-terminal problems)

3) If

\[
p(z^{(1)}, z^{(2)}, z^{(3)}, y^{(1)}, y^{(2)}, y^{(3)}|x^{(1)}, x^{(2)}, x^{(3)}) =
\]

\[
p(y^{(1)}, y^{(2)}, y^{(3)}|x^{(1)}, x^{(2)}, x^{(3)}) \prod_{i=1}^{3} p(z^{(i)}|y^{(i)}).
\]

Then \(\forall q(x^{(1)}, x^{(2)}, x^{(3)}),\)

\[
\phi\left(p(z^{(1)}, z^{(2)}, z^{(3)}|x^{(1)}, x^{(2)}, x^{(3)}), \{q(x^{(1)}, x^{(2)}, x^{(3)})\}\right) \subseteq
\]

\[
\phi\left(p(y^{(1)}, y^{(2)}, y^{(3)}|x^{(1)}, x^{(2)}, x^{(3)}), \{q(x^{(1)}, x^{(2)}, x^{(3)})\}\right).
\]
Implication of the first three properties

Take some $\epsilon$ and a code of length $n$ for admissible source $p(w^{(1)}, w^{(2)}, w^{(3)})$.

- Initial observations: $(W^{(1)}_{1:n}, W^{(2)}_{1:n}, W^{(3)}_{1:n})$

- $j^{th}$ stage: $X^{(1)}_j, X^{(2)}_j, X^{(3)}_j$ created from $Y^{(1)}_{1:j-1} W^{(1)}, Y^{(2)}_{1:j-1} W^{(2)}, Y^{(3)}_{1:j-1} W^{(3)}$

- $\hat{M}^{(1)}_{1:n}, \hat{M}^{(2)}_{1:n}, \hat{M}^{(3)}_{1:n}$ created from $Y^{(1)}_{1:n} W^{(1)}, Y^{(2)}_{1:n} W^{(2)}, Y^{(3)}_{1:n} W^{(3)}$
Quantifying the evolution of the state

\[ \Phi\left( p(w_1^{(1)}, w_1^{(2)}, w_1^{(3)} | w_1^{(1)}, w_1^{(2)}, w_1^{(3)}), \{p(w_1^{(1)}, w_1^{(2)}, w_1^{(3)})\} \right) \]
Quantifying the evolution of the state

Initial observation:
\[ W_{1:n}^{(1)} \]
Gained knowledge:
\[ W_{1:n}^{(1)} Y_{1}^{(1)} \]

Channel

Initial observation:
\[ W_{1:n}^{(2)} \]
Gained knowledge:
\[ W_{1:n}^{(2)} Y_{1}^{(2)} \]

Initial observation:
\[ W_{1:n}^{(3)} \]
Gained knowledge:
\[ W_{1:n}^{(3)} Y_{1}^{(3)} \]

\[ \Phi \left( p(y_{1}^{(1)}, w_{1:n}^{(1)}, y_{1}^{(2)}, w_{1:n}^{(2)}, y_{1}^{(3)}, w_{1:n}^{(3)} | w_{1:n}^{(1)}, \ldots w_{1:n}^{(3)}), \{ p(w_{1:n}^{(1)}, w_{1:n}^{(2)}, w_{1:n}^{(3)}) \} \right) \]
Quantifying the evolution of the state

\[ \Phi(p(y_{1:n}^{(1)}w_{1:n}^{(1)}, y_{1:n}^{(2)}w_{1:n}^{(2)}, y_{1:n}^{(3)}w_{1:n}^{(3)}|w_{1:n}^{(1)}, \ldots w_{1:n}^{(3)}), \{p(w_{1:n}^{(1)}, w_{1:n}^{(2)}, w_{1:n}^{(3)})\}) \]
Quantifying the evolution of the state

\[ \Phi\left(p(y_{1:n}^{(1)}w_{1:n}^{(1)}, y_{1:n}^{(2)}w_{1:n}^{(2)}, y_{1:n}^{(3)}w_{1:n}^{(3)}|w_{1:n}^{(1)}, \ldots w_{1:n}^{(3)}), \{p(w_{1:n}, w_{1:n}, w_{1:n})\}\right) \]
Quantifying the evolution of the state

\[ \Phi(p(\hat{m}_1^{(1)}, \hat{m}_1^{(2)}, \hat{m}_1^{(3)}|w_1^{(1)}, w_1^{(2)}, w_1^{(3)}), \{p(w_1^{(1)}, w_1^{(2)}, w_1^{(3)})\}) \]
Implication of the three properties

\[ \Phi \left( p(w_{1:n}^{(1)}, w_{1:n}^{(2)}, w_{1:n}^{(3)} | w_{1:n}^{(1)}, w_{1:n}^{(2)}, w_{1:n}^{(3)}), \{p(w_{1:n}^{(1)}, w_{1:n}^{(2)}, w_{1:n}^{(3)}) \} \right) = \{(0, 0, 0, 0, 0, 0)\} \]

\[ \Phi \left( p(y_{1:j}^{(1)} w_{1:n}^{(1)}, y_{1:j}^{(2)} w_{1:n}^{(2)}, y_{1:j}^{(3)} w_{1:n}^{(3)} | w_{1:n}^{(1)}, w_{1:n}^{(2)}, w_{1:n}^{(3)}), \{p(w_{1:n}^{(1)}, \ldots w_{1:n}^{(3)}) \} \right) \subset \Phi \left( p(y_{1:j-1}^{(1)} w_{1:n}^{(1)}, y_{1:j-1}^{(2)} w_{1:n}^{(2)}, y_{1:j-1}^{(3)} w_{1:n}^{(3)} | w_{1:n}^{(1)}, w_{1:n}^{(2)}, w_{1:n}^{(3)}), \{p(w_{1:n}^{(1)}, \ldots w_{1:n}^{(3)}) \} \right) \oplus \Phi \left( p(y_j^{(1)}, y_j^{(2)}, y_j^{(3)} | x_j^{(1)}, x_j^{(2)}, x_j^{(3)}), \Psi \right) \]

\[ \Phi \left( p(\hat{m}_{1:n}^{(1)}, \hat{m}_{1:n}^{(2)}, \hat{m}_{1:n}^{(3)} | w_{1:n}^{(1)}, w_{1:n}^{(2)}, w_{1:n}^{(3)}), \{p(w_{1:n}^{(1)}, w_{1:n}^{(2)}, w_{1:n}^{(3)}) \} \right) \subset \Phi \left( p(y_{1:n}^{(1)} w_{1:n}^{(1)}, y_{1:n}^{(2)} w_{1:n}^{(2)}, y_{1:n}^{(3)} w_{1:n}^{(3)} | w_{1:n}^{(1)}, w_{1:n}^{(2)}, w_{1:n}^{(3)}), \{p(w_{1:n}^{(1)}, w_{1:n}^{(2)}, w_{1:n}^{(3)}) \} \right) \]
Implication of the three properties

\[
\Phi(p(\hat{m}_{1:n}, \hat{m}_{1:n}, \hat{m}_{1:n}|w_{1:n}, w_{1:n}, w_{1:n}), \{p(w_{1:n}, w_{1:n}, w_{1:n})\}) \\
\subset \Phi(p(y_1w_{1:n}, y_{1:n}w_{1:n}, y_{1:n}w_{1:n}|w_{1:n}, w_{1:n}, w_{1:n}), \{p(w_{1:n}, w_{1:n}, w_{1:n})\}) \\
\subset \Phi(p(y_1, y_{1}, y_1|x_1, x_1, x_1), \psi) \oplus \\
\Phi(p(y_2, y_{2}, y_2|x_2, x_2, x_2), \psi) \oplus \\
\Phi(p(y_n, y_{n}, y_n|x_n, x_n, x_n), \psi) \oplus \\
\Phi(p(w_1, w_{1}, w_1|w_{1:n}, w_{1:n}, w_{1:n}), \{p(w_{1:n}, w_{1:n}, w_{1:n})\}) \\
\subset n \times \text{convex hull}\{\Phi(q(y_1, y_2, y_3|x_1, x_2, x_3), \psi)\}
\]

Hence:

\[
\Phi(p(\hat{m}_{1:n}, \hat{m}_{1:n}, \hat{m}_{1:n}|w_{1:n}, w_{1:n}, w_{1:n}), \{p(w_{1:n}, w_{1:n}, w_{1:n})\}) \\
\subset n \times \text{convex hull}\{\Phi(q(y_1, y_2, y_3|x_1, x_2, x_3), \psi)\}
\]
Outline

- Introduction
- “Common Information” of dependent random variables
- Some applications of “Potential function method”
  - Multiterminal channel coding
    - General multiterminal network with correlated sources (a generalized cut-set bound)
  - Broadcast channel with correlated sources
  - Strong Interference channel with correlated sources
  - Source coding
- Channel Model
- Conclusions
Definition

- General broadcast channel:

\[ M_0, M_1, M_2 \rightarrow q(y, z|x) \rightarrow \widehat{M}_1 \widehat{M}_0 \]

\[ \widehat{M}_2 \widehat{M}_0 \]
Known Outer Bounds

Two receiver broadcast channel

- Outer bounds for independent source messages
  - Körner-Marton (1979)
  - Nair and El Gamal (2007)
  - Liang, Kramer and Shamai (2008)
  - Nair (2008)
- Correlated sources
  - No interesting outer bound before us
  - Kramer, Liang and Shamai in ITA 2009 independently came up with a non-single letter outer bound
  - Paolo Minero and Young-Han Kim (ISIT 2009), an alternative characterization for the coding theorem by Han and Costa
\[
\bigcup_{W_0, W_1, W_2, U, V - X - YZ} \begin{cases} 
R_0 \geq 0, R_1 \geq 0, R_2 \geq 0; \\
R_0 \leq \min\{I(W_0; Y|U), I(W_0; Z|V)\}; \\
R_1 \leq I(W_1; Y|U); \quad R_2 \leq I(W_2; Z|V); \\
R_0 + R_1 \leq \min(I(W_0W_1; Y|U), I(W_1; Y|W_0UV) + I(W_0U; Z|V)); \\
R_0 + R_2 \leq \min(I(W_0W_2; Z|V), I(W_2; Z|W_0UV) + I(W_0V; Y|U)); \\
R_0 + R_1 + R_2 \leq \min(I(W_1; Y|W_0W_2UV) + I(W_0W_2U; Z|V), \\
I(W_2; Z|W_0W_1UV) + I(W_0W_1V; Y|U), \\
I(W_0UV; Y) + I(W_1; Y|W_0W_2UV) + I(W_2; Z|W_0UV), \\
I(W_0UV; Z) + I(W_2; Z|W_0W_1UV) + I(W_1; Y|W_0UV)). 
\end{cases}
\]

- \(W_0, W_1\) and \(W_2\) mutually independent and uniform;
- \(H(X|W_0, W_1, W_2, U, V) = 0.\)
Simplifying the outer bound

\[ \bigcup_{W_0, W_1, W_2, U, V - X - Y Z} \begin{cases} 
R_0 \geq 0, R_1 \geq 0, R_2 \geq 0; \\
R_0 \leq \min\{I(W_0; Y|U), I(W_0; Z|V)\}; \\
R_1 \leq I(W_1; Y|U); \quad R_2 \leq I(W_2; Z|V); \\
R_0 + R_1 \leq \min(I(W_0W_1; Y|U), I(W_1; Y|W_0UV) + I(W_0U; Z|V)); \\
R_0 + R_2 \leq \min(I(W_0W_2; Z|V), I(W_2; Z|W_0UV) + I(W_0V; Y|U)); \\
R_0 + R_1 + R_2 \leq \\
\min(I(W_1; Y|W_0W_2UV) + I(W_0W_2U; Z|V), \\
I(W_2; Z|W_0W_1UV) + I(W_0W_1V; Y|U), \\
I(W_0UV; Y) + I(W_1; Y|W_0W_2UV) + I(W_2; Z|W_0UV), \\
I(W_0UV; Z) + I(W_2; Z|W_0W_1UV) + I(W_1; Y|W_0UV)). \end{cases} \]

- \( W_0, W_1 \) and \( W_2 \) mutually independent and uniform;

- \( H(X|W_0, W_1, W_2, U, V) = 0 \). Add \( H(W_0|W_1U) = H(W_0|W_2V) = 0 \).
Relaxing $W_0, W_1, W_2$ being mutually independent and uniform

- Take arbitrary $(U, V, W_0, W_1, W_2, X, Y, Z)$
- Assume $\mathcal{W}_i = \{1, 2, 3, \ldots, |\mathcal{W}_i|\}$ for $i = 1, 2, 3$.
- $(A_0, A_1, A_2)$ mutually independent, and independent of all previously defined random variables. $A_0$ defined on $\mathcal{W}_0$, $A_1$ defined on $\mathcal{W}_1$, $A_2$ defined on $\mathcal{W}_2$
- Let $\Delta_i = 1 + (A_i + W_i \mod |\mathcal{W}_i|)$
- Set

$$\left(\tilde{U}, \tilde{V}, \tilde{W}_0, \tilde{W}_1, \tilde{W}_2, \tilde{X}, \tilde{Y}, \tilde{Z}\right) := (U \Delta_0 \Delta_1 \Delta_2, V \Delta_0 \Delta_1 \Delta_2, A_0, A_1, A_2, X, Y, Z)$$
Adding the constraints \( H(W_0|W_1U) = H(W_0|W_2V) = 0 \)

- Take arbitrary \((U, V, W_0, W_1, W_2, X, Y, Z)\)
- Assume \(\mathcal{W}_0 = \{1, 2, 3, \ldots, |\mathcal{W}_0|\}\).
- Will define \(A_0, A_1, A_2\) and \(A_3\) uniform on \(\mathcal{W}_0\)
- Let \(A_2\) and \(A_3\) and \((W_0, W_1, W_2, U, V, X, Y, Z)\) be mutually independent.
- \(A_0\) and \(A_1\) are then defined as follows:
  \[ A_i = 1 + (W_0 + A_{i+2} \mod |\mathcal{W}_0|) \quad i = 0, 1. \]
- Let
  \[(\tilde{U}, \tilde{V}, \tilde{W}_0, \tilde{W}_1, \tilde{W}_2, \tilde{X}, \tilde{Y}, \tilde{Z}) := (UA_0, VA_1, W_0, W_1A_2, W_2A_3, X, Y, Z)\]
An equivalent version of the outer bound

\[ \bigcup_{W_0, W_1, W_2, U, V - X - Y Z} \begin{cases} R_0 \geq 0, R_1 \geq 0, R_2 \geq 0; \\
R_0 \leq \min\{I(W_0; Y|U), I(W_0; Z|V)\}; \\
R_0 + R_1 \leq I(W_1; Y|U); \quad R_0 + R_2 \leq I(W_2; Z|V); \\
R_0 + R_1 \leq I(W_1; Y|W_0 UV) + I(W_0 U; Z|V); \\
R_0 + R_2 \leq I(W_2; Z|W_0 UV) + I(W_0 V; Y|U); \\
R_0 + R_1 + R_2 \leq \min(I(W_1; Y|W_0 W_2 UV) + I(W_2 U; Z|V), \\
I(W_2; Z|W_0 W_1 UV) + I(W_1 V; Y|U), \\
I(W_0 UV; Y) + I(W_1; Y|W_0 W_2 UV) + I(W_2; Z|W_0 UV), \\
I(W_0 UV; Z) + I(W_2; Z|W_0 W_1 UV) + I(W_1; Y|W_0 UV)). \end{cases} \]

- \( H(W_0|W_1 U) = H(W_0|W_2 V) = 0; \)

- \( H(X|W_0, W_1, W_2, U, V) = 0. \)
Definition

- Broadcast channel with arbitrarily correlated sources
  - Observe \( n \) i.i.d. repetitions of \((L_1, L_2)\). Allowed to use the broadcast channel \( n \) times
  - \((L_1, L_2)\) is admissible if the first receiver can successfully decode \( L_1 \) and the second decoder can successfully decode \( L_2 \)
Recall \( \bigcup_{W_0, W_1, W_2, U, V \rightarrow X \rightarrow Y \rightarrow Z} \) 
\[
\begin{align*}
R_0 & \geq 0, R_1 \geq 0, R_2 \geq 0; \\
R_0 & \leq \min\{I(W_0; Y|U), I(W_0; Z|V)\}; \\
R_0 + R_1 & \leq I(W_1; Y|U); \quad R_0 + R_2 \leq I(W_2; Z|V); \\
R_0 + R_1 & \leq I(W_1; Y|W_0UV) + I(W_0U; Z|V); \\
R_0 + R_2 & \leq I(W_2; Z|W_0UV) + I(W_0V; Y|U); \\
R_0 + R_1 + R_2 & \leq \min(I(W_1; Y|W_0W_2UV) + I(W_2U; Z|V), \\
& \quad I(W_2; Z|W_0W_1UV) + I(W_1V; Y|U), \\
& \quad I(W_0UV; Y) + I(W_1; Y|W_0W_2UV) + I(W_2; Z|W_0UV), \\
& \quad I(W_0UV; Z) + I(W_2; Z|W_0W_1UV) + I(W_1; Y|W_0UV)).
\end{align*}
\]

- \( H(W_0|W_1U) = H(W_0|W_2V) = 0; \)

- \( H(X|W_0, W_1, W_2, U, V) = 0. \)
Broadcast channel with arbitrarily correlated sources

$L_0$ be the common part of $L_1$, $L_2$ in the Gács and Körner sense. $H(L_0|L_1) = H(L_0|L_2) = 0$

\[ \exists W_0, W_1, W_2, U, V - X - Y Z \left\{ \begin{array}{l}
H(L_0) \leq \min\{I(W_0; Y|U), I(W_0; Z|V)\}; \\
H(L_1) \leq I(W_1; Y|U); \quad H(L_2) \leq I(W_2; Z|V);
\end{array} \right. \]

\[ \begin{array}{l}
H(L_1) \leq I(W_1; Y|W_0UV) + I(W_0U; Z|V);
H(L_2) \leq I(W_2; Z|W_0UV) + I(W_0V; Y|U);
H(L_1L_2) \leq \min(I(W_1; Y|W_0W_2UV) + I(W_2U; Z|V),
I(W_2; Z|W_0W_1UV) + I(W_1V; Y|U),
I(W_0V; Y|U) + I(W_1; Y|W_0W_2UV) + I(W_2; Z|W_0UV),
I(W_0V; Z|V) + I(W_2; Z|W_0W_1UV) + I(W_1; Y|W_0UV)).
\end{array} \]

- $H(W_0|W_1U) = H(W_0|W_2V) = 0$;
- $H(X|W_0, W_1, W_2, U, V) = 0$. 
Proof sketch (I)

\[ \Phi(q(y, z|x), \Psi) = \]

\[ \bigcup_{p(x) \in \Psi} \bigcup W_0 W_1 W_2 U V - X - Y Z \text{ s.t.} \]
\[ H(W_0|W_1 U) = H(W_0|W_2 V) = 0; \]
\[ H(X|W_0, W_1, W_2, U, V) = 0 \]

\[ \Pi \left( \{ (I(W_0; Y|U), I(W_0; Z|V), I(W_1; Y|U), I(W_2; Z|V), \right. \]
\[ I(W_1; Y|W_0 U V) + I(W_0 U; Z|V), I(W_2; Z|W_0 U V) + I(W_0 V; Y|U), \]
\[ I(W_1; Y|W_0 W_2 U V) + I(W_0 W_2 U; Z|V), \]
\[ I(W_2; Z|W_0 W_1 U V) + I(W_0 W_1 V; Y|U), \]
\[ I(W_0 V; Y|U) + I(W_1; Y|W_0 W_2 U V) + I(W_2; Z|W_0 U V), \]
\[ I(W_0 U; Z|V) + I(W_2; Z|W_0 W_1 U V) + I(W_1; Y|W_0 U V)) \} \) \]
Proof sketch (II)

The permissible set of input distribution $\Psi = \text{the set of all distributions.}$

The same three conditions are imposed. Using the same chain of equations:

$$\Phi(p(y_1 \ldots y_{i+1}, z_1 \ldots z_{i+1} | l_1^n l_2^n), \{p(l_1^n, l_2^n)\})$$

$$\subseteq \Phi(p(y_1 \ldots y_i, z_1 \ldots z_i | l_1^n l_2^n), \{p(l_1^n, l_2^n)\}) \oplus \Phi(q(., |.), \Psi)$$

Can conclude

$$\Phi(p(l_1^n, l_2^n | l_1^n l_2^n), \{p(l_1^n, l_2^n)\})$$

$$\subseteq \underbrace{\Phi(q(., |.), \Psi) \oplus \Phi(q(., |.), \Psi) \oplus \Phi(q(., |.), \Psi)}_{n \text{ times}}$$

$$\subseteq n \times \Phi(q(., |.), \Psi).$$  \hspace{1cm} (3)
Proof sketch (III)

\[ n \times \left( H(L_0), H(L_0), H(L_1), H(L_2), H(L_1), H(L_2), H(L_1 L_2), H(L_1 L_2), H(L_1 L_2), H(L_1 L_2) \right) \in \Phi(p(l_1^n, l_2^n, l_1^n l_2^n), \{p(l_1^n, l_2^n)\}). \]  

(3), (4) \Rightarrow

\[ \left( H(L_0), H(L_0), H(L_1), H(L_2), H(L_1), H(L_2), H(L_1 L_2), H(L_1 L_2), H(L_1 L_2), H(L_1 L_2) \right) \in \Phi(q(., .|.,.), \Psi). \]
Outline

- Introduction
- “Common Information” of dependent random variables
- Some applications of “Potential function method”
  - Multiterminal channel coding
    - General multiterminal network with correlated sources (a generalized cut-set bound)
    - Broadcast channel with correlated sources
    - Strong Interference channel with correlated sources
  - Source coding
  - Channel Model
- Conclusions
Known results for memoryless discrete Strong Interference Channel (IC)

∀p(x_1)p(x_2) : I(X_1; Y_1|X_2) ≤ I(X_1; Y_2|X_2)
I(X_2; Y_2|X_1) ≤ I(X_2; Y_1|X_1)

- Strong IC with ind. messages: solved by Costa and El Gamal in 1987.
- Includes MAC as a special case
Known results for memoryless discrete Strong Interference Channel (IC)

The capacity region with common information (Maric, Yates and Kramer 2005): union over \( p(u, x_1, x_2, y_1, y_2) = p(u)p(x_1|u)p(x_2|u)q(y_1, y_2|x_1, x_2) \), of

\[
\begin{align*}
R_0, R_1, R_2 & \geq 0; \\
R_1 & \leq I(X_1; Y_1|X_2, U); \\
R_2 & \leq I(X_2; Y_2|X_1, U); \\
R_1 + R_2 & \leq \min(I(X_1X_2; Y_1|U), I(X_1X_2; Y_2|U)); \\
R_0 + R_1 + R_2 & \leq \min(I(X_1X_2; Y_1), I(X_1X_2; Y_2)).
\end{align*}
\]
Definition

- Interference channel with arbitrarily correlated sources
  - The first party observes $n$ i.i.d. copies of $L_1$, the second party $n$ i.i.d. copies of $L_1$
  - Allowed to use the interference channel $n$ times
  - $(L_1, L_2)$ is admissible if the first receiver can successfully decode $L_1$ and the second decoder can successfully decode $L_2$
Memoryless discrete Strong Interference Channel with correlated sources

Strong interference channel with correlated source: open problem.

A generalization of Maric, Yates and Kramer (2005):

Given admissible pair \((L_1, L_2)\) and any \(W\) where \(L_1 - W - L_2\) holds, there must exist \(p(u, x_1, x_2, y_1, y_2) = p(u)p(x_1|u)p(x_2|u)q(y_1, y_2|x_1, x_2)\) where

\[
\begin{align*}
H(L_1|W) &\leq I(X_1; Y_1|X_2, U); \\
H(L_2|W) &\leq I(X_2; Y_2|X_1, U); \\
H(L_1) + H(L_2|W) &\leq \min(I(X_1X_2; Y_1|U), I(X_1X_2; Y_2|U)); \\
H(L_1) + H(L_2) &\leq I(X_1X_2; Y_1); \\
H(L_2) + H(L_1|W) &\leq I(X_1X_2; Y_2).
\end{align*}
\]
Proof sketch (I)

\[
\Phi(q(y_1, y_2|x_1, x_2), \Psi) = \\
\bigcup_{p(x_1, x_2) \in \Psi} \bigcup_{U \text{ s.t.}} U \quad \text{s.t.} \\
X_1 - U - X_2, \\
U - X_1X_2 - Y_1Y_2 \\
\Pi\left(\left\{ (I(X_1; Y_1|X_2, U), I(X_2; Y_2|X_1, U), \\
I(X_1X_2; Y_1|U), I(X_1X_2; Y_2|U), \\
I(X_1X_2; Y_1), I(X_1X_2; Y_2) \right\} \right)
\]
Proof sketch (II)

The permissible set of input distributions is $\Psi = \text{the set of all distributions}$.

The same three conditions are imposed. Using the same chain of equations:

$$\Phi(p(y_1 \ldots y_{i+1}, z_1 \ldots z_{i+1} | l_1^n, l_2^n), \{p(l_1^n, l_2^n)\})$$

$$\subseteq \Phi(p(y_1 \ldots y_i, z_1 \ldots z_i | l_1^n, l_2^n), \{p(l_1^n, l_2^n)\}) \oplus \Phi(q(\ldots|\ldots), \Psi)$$

Can conclude

$$\Phi(p(y_1 \ldots y_n, z_1 \ldots z_n | l_1^n, l_2^n), \{p(l_1^n, l_2^n)\})$$

$$\subseteq \underbrace{\Phi(q(\ldots|\ldots), \Psi) \oplus \Phi(q(\ldots|\ldots), \Psi) \oplus \Phi(q(\ldots|\ldots), \Psi)}_{\text{n times}}$$

$$\subseteq n \times \Phi(q(\ldots|\ldots), \Psi).$$

(5)
Proof sketch (III)

\[ n \times \left( H(L_1|W), H(L_2|W), H(L_1|W) + H(L_2|W), H(L_1|W) + H(L_2|W), \right. \\
\left. H(L_1) + H(L_2|W), H(L_2) + H(L_1|W) \right) \in \\
\Phi \left( p(y_1...y_n, z_1...z_n | l_1^n, l_2^n), \{ p(l_1^n, l_2^n) \} \right). \quad (6) \]

\[(5), (6) \Rightarrow \]

\[ \left( H(L_1|W), H(L_2|W), H(L_1|W) + H(L_2|W), H(L_1|W) + H(L_2|W), \right. \\
\left. H(L_1) + H(L_2|W), H(L_2) + H(L_1|W) \right) \in \\
\Phi (q(,.|,.), \Psi). \]
Outline

- Introduction
- “Common Information” of dependent random variables
- Some applications of “Potential function method”
  - Multiterminal channel coding
  - Source coding
  - Channel Model
- Conclusions
Problem definition

- $q(x, y, z)$ is given; The $Z$ party wants to reconstruct the i.i.d. copies of $X$ with high probability.
- The answer is known:

$$\bigcup_{U-Y-XZ} \{(R_1, R_2) : R_1 \geq H(X|ZU), R_2 \geq I(U; Y|Z)\}.$$
Application of the potential function method

Would like to show the following region is an outer bound

\[ \mathcal{R}(q(x, y, z)) = \bigcup_{U-Y-XZ} \{(R_1, R_2) : R_1 \geq H(X|ZU), R_2 \geq I(U; Y|Z)\}. \]

- \( \mathcal{R}(\cdot) \): the set of all joint distributions on all triples of discrete sets \( \rightarrow \) subsets of \( \mathbb{R}^2_+ \)

- **Identify properties** on such functions that imply \( \mathcal{R}(\cdot) \) being an outer bound

- **Verify** that the given expression belongs to the class of functions that satisfy those properties
The properties

(1) \( \bigcup_{U \rightarrow Y \rightarrow XZ} \{(R_1, R_2) : R_1 \geq H(X|ZU), R_2 \geq I(U; Y|Z)\} \)

is a subset of \( \mathcal{R}(q(x, y, z)) \).

(2) For any arbitrary channel \( q(x\tilde{x}, y\tilde{y}, z\tilde{z}) \) that factorizes as \( q(\tilde{x}, \tilde{y}, \tilde{z}) \cdot q(x, y, z) \), we have

\[
\mathcal{R}(q(x\tilde{x}, y\tilde{y}, z\tilde{z})) \subset \mathcal{R}(q(\tilde{x}, \tilde{y}, \tilde{z})) \oplus \mathcal{R}(q(x, y, z)).
\]

(3) \( \mathcal{R}(q(x, y, z)) \) is convex for all \( q(x, y, z) \).
Outline

- Introduction
- “Common Information” of dependent random variables
- Some applications of “Potential function method”
  - Multiterminal channel coding
  - Source coding
  - Channel Model
- Conclusions
Channel Model: Definition of $C_{CH}(p(yz|x))$

Alice puts $X_1$, Bob and Eve receive $Y_1$, $Z_1$; a round of public discussion; Alice puts $X_2$, Bob and Eve receive $Y_2$, $Z_2$; a round of public discussion... Alice creates $K_A(X^n, \overrightarrow{F})$, Bob creates $K_B(Y^n, \overrightarrow{F})$.

$$P(K_A = K_B = K) > 1 - \epsilon, \frac{1}{n}I(K; Z^n \overrightarrow{F}) < \epsilon$$
## Known results on $C_{CH}(p(yz|x))$

| Authors       | Lower bounds on $C_{CH}(p(yz|x))$                                                                                                                                                                                                 |
|---------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Maurer (1993) | $\sup_{p(x)}(\max\{I(X; Y) - I(X; Z), I(Y; X) - I(Y; Z)\})$  
Idea: $C_{CH}(p(yz|x)) \geq \sup_{p(x)} S(X; Y \parallel Z)$  
Remark: Can use the lower bound of Ahlswede and Csisz’ar to get a better lower bound |
| Authors       | Upper bounds on $C_{CH}(p(yz|x))$                                                                                                                                                                                                 |
| Maurer (1993) | $\min(\sup_{p(x)} I(X; Y|Z), \sup_{p(x)} I(X; Y))$  


### Known results on $C_{CH}(p(yz|x))$

| Authors       | Lower bounds on $C_{CH}(p(yz|x))$                                                                 |
|---------------|-------------------------------------------------------------------------------------------------|
| Maurer (1993) | $\sup_{p(x)}(\max\{I(X; Y) - I(X; Z), I(Y; X) - I(Y; Z)\})$                                  |
|               | Idea: $C_{CH}(p(yz|x)) \geq \sup_{p(x)} S(X; Y\|Z)$                                           |
|               | Remark: Can use the lower bound of Ahlswede and Csisz’ar to get a better lower bound            |

| Authors       | Upper bounds on $C_{CH}(p(yz|x))$                                                                |
|---------------|-------------------------------------------------------------------------------------------------|
| Maurer (1993) | $\min(\sup_{p(x)} I(X; Y|Z), \sup_{p(x)} I(X; Y))$                                               |

We prove new lower and upper bounds
Application of the potential function method to Channel Model

New upper bound (can be shown to be strictly better than the best known upper bound):

\[ \sup_{p(x)} \inf_{J} \left[ I(X; Y|J) + I(XY; J|Z) \right] \]

Idea: Find properties that imply an expression is an upper bound

Verify that the given expression satisfies these properties.

Would like to prove that

\[ \Psi(\psi(q(x)|y|z)) = \sup_{q(x)} \psi(q(x)q(y, z|x)) \]

is an outer bound.
Sufficient conditions for a function to be an upper bound for the Channel Model

1) Whenever $H(X'|X) = 0$ and $XYZ - X - X' - X'Y'Z'$ and $p(y', z'|x') = q(y', z'|x')$ are true, we have:

\[ \psi(XX'; YY'||ZZ') \geq \psi(X; Y||Z) + \Psi(q(xy|z)) \]

2) $\forall F : H(F|X) = 0$ or $H(F|Y) = 0$,

\[ \rightarrow \psi(X; Y||Z) \geq \psi(XF; YF||ZF) \]

3) $\forall X', Y' : H(X'|X) = 0, H(Y'|Y) = 0$,

\[ \rightarrow \psi(X; Y||Z) \geq \psi(X'; Y'||Z) \]

4) $\psi(X; Y||Z) \geq H(X|Z) - H(X|Y) = I(X; Y) - I(X; Z)$
Outline

- Introduction
- “Common Information” of dependent random variables
- Some applications of “Potential function method”
  - Channel Model
  - Multiterminal channel coding
  - Source coding
- Conclusions
Conclusions

- Derived a new upper bound on private common information and discussed a technique for proving outer bounds.

- Demonstrated the applicability of the technique to:
  - Channel Model
  - Multiterminal channel coding (generalized cut-set bound, broadcast channel, strong interference channel)
  - Source coding
Backup slide: A few definitions

Backup slides
(1): $K, L$ sets of points in $\mathbb{R}^c_+$:

$$K \oplus L := \{v_1 + v_2 : v_1 \in K, v_2 \in L\}.$$ 

(2): $\overrightarrow{v_1}, \overrightarrow{v_2} \in \mathbb{R}^c_+$:

$$\overrightarrow{v_1} \geq \overrightarrow{v_2} \iff$$ 

each coordinate of $\overrightarrow{v_1}$ is greater than or equal to the corresponding coordinate of $\overrightarrow{v_2}$.

(3): $K, L$ sets of points in $\mathbb{R}^c_+$:

$$K \leq L \iff \forall \overrightarrow{k} \in K, \exists \overrightarrow{l} \in L : \overrightarrow{k} \leq \overrightarrow{l}.$$