Fixed Block-Length coding for Joint Source-Channel Communication

Arun Padakandla

August 27, 2019
Setting: Transmitting Correlated sources over 2-user MAC and IC

\[ S_1 \rightarrow W_{S_1S_2} \rightarrow S_2 \]

\[ X_1, X_2 \rightarrow 2\text{-MAC} \]

\[ Y \rightarrow S_1S_2 \]
Setting: Transmitting Correlated sources over 2-user MAC and IC
Problem Setup: Transmitting Correlated sources over 2-user MAC and IC

Shannon-theoretic study

Optimal coding technique

Necessary and sufficient conditions (in terms of $W_{S_1S_2}, W_{Y|X_1X_2}$, $W_{Y_1|Y_2, X_1X_2}$)

Plain Vanilla Lossless source-channel coding over 2-user MAC or IC

1. A new coding technique.
2. Characterize performance, derive new sufficient conditions.

What is the New idea?
Problem Setup: Transmitting Correlated sources over 2-user MAC and IC

Shannon-theoretic study

?? Optimal coding technique ??

?? Necessary and sufficient conditions (in terms of $W_{S_1S_2}, W_{Y|X_1X_2}$ (MAC) $W_{Y_1Y_2|X_1X_2}$ (IC)) ??

Plain Vanilla Lossless source-channel coding over 2-user MAC or IC.
Problem Setup: Transmitting Correlated sources over 2–user MAC and IC

Shannon-theoretic study

?? Optimal coding technique ??

?? Necessary and sufficient conditions (in terms of $W_{S_1S_2}$, $W_{Y|X_1X_2}$ (MAC) $W_{Y_1Y_2|X_1X_2}$ (IC)) ??

Plain Vanilla Lossless source-channel coding over 2-user MAC or IC.

1. A new coding technique.
2. Characterize performance, derive new sufficient conditions.

?? What is the New idea ??
What are the tasks involved?

1. Source Compression/Coding (Lossless)

2. Channel Coding (Over MAC/IC)

3. Co-ordinate channel inputs
What are the tasks involved?

1. Source Compression/Coding (Lossless)

2. Channel Coding (Over MAC/IC)

3. Co-ordinate channel inputs  
   
   Separation sub-optimal

Slepian-Wolf bin indices independent. *Disables* co-ordination.
Co-ordination can be critical to communication: Toy Example

Noiseless if $X_1 = X_2$ and Erasure if $X_1 \neq X_2$

Source correlation has to be extracted for co-ordination.
Three Tasks - What is the block-length (B-L) of the codes?

1. Source Compression/Coding (Lossless)

2. Channel Coding (Over MAC/IC)

3. Co-ordinate channel inputs
Three Tasks - What is the block-length (B-L) of the codes?

1. Source Compression/Coding (Lossless)

2. Channel Coding (Over MAC/IC)

3. Co-ordinate channel inputs
Three Tasks - What is the block-length (B-L) of the codes?

1. Source Compression/Coding (Lossless)

2. Channel Coding (Over MAC/IC)

3. Co-ordinate channel inputs
Three Tasks - What is the block-length (B-L) of the codes?

1. Source Compression/Coding (Lossless)

2. Channel Coding (Over MAC/IC)

3. Co-ordinate channel inputs
Three Tasks - What is the block-length (B-L) of the codes?

1. Source Compression/Coding (Lossless)

   Large block-length is good

2. Channel Coding (Over MAC/IC)

   Large block-length is good

3. Co-ordinate channel inputs

   ?? block-length ??
Co-ordinating Channel Inputs is by Extracting Correlation from Distributed Sources.

?? optimal block-length for extracting correlation from distributed sources ??
Witsenhausen [1975] Problem Setup

$X_1 X_2 \ldots X_n \ldots \xrightarrow{\text{IID } P_{XY}} Y_1 Y_2 \ldots Y_n \ldots$

Best $f_n, g_n$ for 1-bit agreement.

$P(n) =$ Best Agreement probability for $n$ symbol processing.

$P(1) > P(2) > P(3) > \cdots$
Witsenhausen [1975] Problem Setup

IID $P_{XY}$

Agreement on a non-trivial RV is a measure of correlation.
Witsenhausen [1975] Problem Setup

\[ X_1 X_2 \ldots X_n \ldots \rightarrow f_n(X_1 X_2 \ldots X_n) = A \in \{0,1\} \]

\[ Y_1 Y_2 \ldots Y_n \ldots \rightarrow g_n(Y_1 Y_2 \ldots Y_n) = B \in \{0,1\} \]

IID \( \mathbb{P}_{XY} \)

Best \( f_n, g_n \) for 1-bit agreement.

\( P(n) = \text{Best Agreement probability for } n \text{ symbol processing.} \)

\( P(1) > P(2) > P(3) > \cdots \)
Witsenhausen [1975] Problem Setup

$X_1 X_2 \ldots X_n \ldots \xrightarrow{} \text{Alice} \xrightarrow{f_n(X_1 X_2 \ldots X_n)} A \in \{0,1\}$

$IID \quad P_{XY}$

$Y_1 Y_2 \ldots Y_n \ldots \xrightarrow{} \text{Bob} \xrightarrow{g_n(Y_1 Y_2 \ldots Y_n)} B \in \{0,1\}$

Prob. of Agreement $P(A = B)$

Is a measure of correlation.

Binary RVs $A$ and $B$ non-trivial and agree with high prob.
Witsenhausen [1975] Problem Setup

\[ X_1 X_2 \ldots X_n \ldots \xrightarrow{\text{IID } \mathbb{P}_{XY}} Y_1 Y_2 \ldots Y_n \ldots \]

Best \( f_n, g_n =: P(n) \) for 1-bit agreement.

\[ P(n) = \text{Best Agreement probability for } n \text{ symbol processing.} \]

\[ P(1) > P(2) > P(3) > \cdots \]
Witsenhausen [1975] Problem Setup

\[ X_1 X_2 \ldots X_n \ldots \xrightarrow{\text{IID}} Y_1 Y_2 \ldots Y_n \ldots \]

\( \mathbb{P}_{XY} \)

\[ f_n \colon \mathbb{X}^n \to \{0,1\} \]

\[ g_n \colon \mathbb{Y}^n \to \{0,1\} \]

\( \mathbb{P}(f_n(X^n) = g_n(Y^n)) \)
Witsenhausen [1975] Problem Setup

\[ X_1 X_2 \ldots X_n \ldots \xrightarrow{\text{IID } P_{XY}} Y_1 Y_2 \ldots Y_n \ldots \]

\[
\sup_{f_n, g_n} P(f_n(X^n) = g_n(Y^n)) = P_{XY} \]

Best \( f_n, g_n \) for 1-bit agreement.
Witsenhausen [1975] Problem Setup

\[ X_1X_2\ldots X_n\ldots \rightarrow \{0,1\} \]

\[ Y_1Y_2\ldots Y_n\ldots \rightarrow \{0,1\} \]

IID \( P_{XY} \)

\[ \sup_{f_n, g_n} \mathbb{P}(f_n(X^n) = g_n(Y^n)) =: \mathcal{P}(n) \]

Best \( f_n, g_n \) for 1-bit agreement.

\( \mathcal{P}(n) = \) Best Agreement probability for \( n \) symbol processing.
Witsenhausen [1975] Problem Setup

\[ X_1X_2\ldots X_n\ldots \rightarrow \text{Alice} \]

IID \( P_{XY} \)

\[ Y_1Y_2\ldots Y_n\ldots \rightarrow \text{Bob} \]

\[ \sup_{f_n, g_n} \mathbb{P}(f_n(X^n) = g_n(Y^n)) =: \mathcal{P}(n) \]

\( \mathcal{P}(n) = \text{Best Agreement probability for } n \text{ symbol processing.} \)

\[ \mathcal{P}(1) > \mathcal{P}(2) > \mathcal{P}(3) > \cdots \]
Extracting Correlation favors Short Block-Lengths!!!
Indeed, ⋅⋅⋅ a binary symmetric source ⋅⋅⋅

<table>
<thead>
<tr>
<th>$p_{S_1S_2}$</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.495</td>
<td>0.005</td>
</tr>
<tr>
<td>1</td>
<td>0.005</td>
<td>0.495</td>
</tr>
</tbody>
</table>

IID Source $S_1, S_2$ with $P(S_1 = S_2) = 0.99$
Indeed, a binary symmetric source

\[
\begin{array}{c|cc}
  & 0 & 1 \\
\hline
0 & 0.495 & 0.005 \\
1 & 0.005 & 0.495 \\
\end{array}
\]

IID Source \( S_1, S_2 \) with \( P(S_1 = S_2) = 0.99 \)
Indeed, ··· a binary symmetric source ···

\[ p_{S_1S_2} = \begin{array}{cc}
0 & 1 \\
0 & 0.495 & 0.005 \\
1 & 0.005 & 0.495
\end{array} \]

IID Source \( S_1, S_2 \) with \( P(S_1 = S_2) = 0.99 \)
Indeed, a binary symmetric source is such that

\[
p_{S_1S_2} = \begin{pmatrix}
0 & 1 \\
0 & 0.495 & 0.005 \\
1 & 0.005 & 0.495
\end{pmatrix}
\]

where \(S_1, S_2\) are IID sources with \(P(S_1 = S_2) = 0.99\).
Indeed, a binary symmetric source …

\begin{align*}
p_{S_1S_2} & \begin{pmatrix} 0 & 1 \\ 0 & 0.495 \\ 0.005 & 0.495 \end{pmatrix} \\
S_1 & 010100101001010101010100101100 \ldots \\
0 & \text{As } n \to \infty \\
1 & \text{One among the exponentially large conditional typical sequences} \\
S_2 & \\
\text{IID Source } S_1, S_2 \text{ with } P(S_1 = S_2) = 0.99
Indeed, ... a binary symmetric source ...  

\begin{array}{c|c|c}
| & 0 & 1 \\
\hline
0 & 0.495 & 0.005 \\
\hline
1 & 0.005 & 0.495 \\
\end{array}

\text{S}_1 \quad 010100101001010101001011101100 \ldots 

As \ n \to \infty 

One among the \textcolor{red}{exponentially} large 
conditional typical sequences 

\text{S}_2

IID Source \ S_1, S_2 \text{ with } P(S_1 = S_2) = 0.99

A shorter block agrees with higher prob.

\[ P(S_1^n \neq S_2^n) = 1 - P(S_1^n = S_2^n) = 1 - (P(S_1 = S_2))^n = 1 - 0.99^n \xrightarrow{n \to \infty} 1 \]
Communicating Distributed Correlated Sources over Networks
Communicating Distributed Correlated Sources over Networks

\[01010010100100101010101101100011100101101001010\]

\[010100101001001010\]

\[S_1 \rightarrow \]

\[\downarrow \]

\[W_{S_1S_2} \rightarrow \]

\[\downarrow \]

\[S_2 \rightarrow \]

\[X_1 \leftarrow \]

\[2\text{-MAC} \]

\[\rightarrow Y \]

\[W_{Y|X_1X_2} \]

\[X_2 \leftarrow \]

\[\rightarrow S_1S_2 \]

\[010100101001010010010101001011011000111001011010010100\]

\[01010010100100101000101101100011100101101001010\]
Communicating Distributed Correlated Sources over Networks
Communicating Distributed Correlated Sources over Networks

01010010100100101010101101100011100101101001010

S₁ →

\[ W_{S₁S₂} \]

S₂ →

01010010100100101000101101100011100101101001010

S₁S₂

X₁ → 2-MAC → Y → X₂

W_{Y|X₁X₂}
Communicating Distributed Correlated Sources over Networks

$S_1 \xrightarrow{W_{S_1S_2}} S_2$

Source–Chnl Mapping

$010100101001001010$

$010100101001001010$

$0111011$

$0111011$

$11011000111$

$0111011$

$11011000111$

$0111011$

$11011000111$

$0111011$

$2$-MAC

$X_1 \xrightarrow{W_1} X_2$
Communicating Distributed Correlated Sources over Networks

\[ S_1 \xrightarrow{W_{S_1S_2}} S_2 \]

\[ X_1 \quad 2-MAC \quad X_2 \]

\[ 010100101001001010 \]

\[ 01011011000111 \]

\[ 00101101001 \]

\[ 00101101 \]

\[ 01010010100100101010101101100011100101101001010 \]

\[ 01010010100100101000101101100011100101101001010 \]

\[ 10101101 \]

\[ 10101101 \]

\[ 00101101001 \]
Shorter Blocks

Efficient transfer of correlation.

Longer Block

Efficient Chnl Coding, Source Compression.

A tension in choice of Block-Length (B-L)

Tension in choice of Block-Length (B-L)

Optimal B-L neither too big nor too small

*Fixed Block-Length (B-L)* coding

A New approach in information theory.
Challenges in the Design and Information-theoretic Analysis


2. Information - theoretic performance analysis requires block-length (B-L) of overall coding scheme to $\to \infty$.

3. Fixing the B-L results in multi-letter distribution. How do you get a single-letter characterization?

?? Design ?? Analysis ?? Performance Characterization ??
Two-layer coding - one fixed B-L, one $\infty$ – B-L

B-L $ml$. 

0101001010010010101010110110001110010110……..1001010

1 Source 1

$ml$
Two-layer coding - one fixed B-L, one $\infty$ – B-L

B-L $ml$.

Source 1

$0101001010010010101010110110001110010110 \ldots \ldots 1001010$

$m$

0101001
0100100
1010101
0110110
0011100
1011010

$l$

$m$

0101001
Two-layer coding - one fixed B-L, one $\infty$ – B-L

B-L $ml$. 

Source 1

Layer 1 code
Maps row-by-row
Fixed B-L $l$
Two-layer coding - one fixed B-L, one $\infty$ – B-L

B-L $ml$. 

Layer 2 code
Block mapping
B-L $ml$

1010111
1110111
0001011
1111111
1101011
1101000
0101001

0101001010010010101010110110001110010110……..1001010

Source 1

ml
Two-layer coding - one fixed B-L, one $\infty$ – B-L

B-L $ml$.

Layer 1 code
Maps row-by row

Fixed B-L $l$

Layer 2 code
Block mapping

B-L $ml$

Multiplexing
Two layers into
Channel inputs

Source 1

Layer 1 code
Maps row-by row

Fixed B-L $l$

Multiplexing
Two layers into
Channel inputs

1

1

1

1

1

Source 1
Two-layer coding - one fixed B-L, one $\infty$— B-L

1. Fixed B-L Layer 1 transfers correlation efficiently

2. $\infty$— B-L Layer 2 performs source compression, channel coding.
Challenges in Information-theoretic Performance Analysis

1. Multiplexing multiple codewords of Layer 1 with single codeword of Layer 2.
   ▶ Shirani and Pradhan’s technique of interleaving.

2. Distribution induced by the coding scheme not the same as the chosen test channel
   ▶ Constant composition codes and bounding the divergence between the respective distributions.
Fixed Block-Length Coding

A New Framework in Information Theory


Builds on recent findings by Shirani Pradhan [2014] on Distributed Source Coding.
Communicating Distributed Correlated Sources over Networks

New Coding Structure for

that strictly outperform all previous known.

Derived new sufficient conditions.

Thank You
A new coding theorem in a simplified form

Theorem

\((S, W_S)\) is transmissible over IC \((X, Y, W_{Y|X})\) if there exists (i) a finite set \(\mathcal{K}\), maps \(f_j : S_j \rightarrow \mathcal{K}\), with \(K_j = f_j(S_j)\) for \(j \in [2]\), (ii) \(l \in \mathbb{N}, \delta > 0\), (iii) finite set \(\mathcal{U}, \mathcal{V}_1, \mathcal{V}_2\) and pmf \(p_U p_{V_1} p_{V_2} p_{X_1|U} p_{X_2|U} p_{V_2}\) defined on \(\mathcal{U} \times \mathcal{V} \times \mathcal{X}\), where \(p_U\) is a type of sequences in \(\mathcal{U}^l\), (iv) \(A, B \geq 0, \rho \in (0, A)\) such that \(\phi \in [0, 0.5]\),

\[
A + B \geq (1 + \delta) H(K_1), \quad \text{and for } j \in [2],
\]

\[
B + H(S_j|K_1) + \mathcal{L}^S_i(\phi, |S_j|) < I(V_j; Y_j) - \mathcal{L}^C_j(\phi, |\mathcal{V}|),
\]

where,
A new coding theorem in a simplified form

Theorem

\((S, W_S)\) is transmissible over IC \((X, Y, W_{Y|X})\) if there exists (i) a finite set \(K\), maps \(f_j : S_j \rightarrow K\), with \(K_j = f_j(S_j)\) for \(j \in [2]\), (ii) \(l \in \mathbb{N}, \delta > 0\), (iii) finite set \(\mathcal{U}, \mathcal{V}_1, \mathcal{V}_2\) and pmf \(p_{U|V_1}p_{V_2|X_1|U}p_{X_2|U|V_2}\) defined on \(\mathcal{U} \times \mathcal{V} \times \mathcal{X}\), where \(p_U\) is a type of sequences in \(\mathcal{U}^l\), (iv) \(A, B \geq 0, \rho \in (0, A)\) such that \(\phi \in [0, 0.5)\),

\[
A + B \geq (1 + \delta)H(K_1), \text{ and for } j \in [2],
\]

\[
B + H(S_j|K_1) + \mathcal{L}_i^S(\phi, |S_j|) < I(V_j; Y_j) - \mathcal{L}_j^C(\phi, |\mathcal{V}|),
\]

where, \(\phi := g_{\rho, l} + \xi^{[l]}(K) + \tau_{l, \delta}(K_1)\), \(\xi^{[l]}(K) := P(K_1^l \neq K_2^l)\)

\[
g_{\rho, l} := \sum_{j=1}^{2} \exp\{-l(E_r(A + \rho, p_U, p_{Y_j|U}) - \rho)\}, \text{ (Channel coding error)}
\]

\[
\tau_{l, \delta}(K) = 2|\mathcal{K}| \exp\{-2\delta^2 p_K^2(a^*)l\} \text{ (Prob of non-typicality)}
\]

\[
\mathcal{L}_j^C(\phi, |\mathcal{U}|) = h_b(\phi) + \phi \log |\mathcal{U}| + |\mathcal{V}| |\mathcal{U}| \phi \log \frac{1}{\phi},
\]

\[
\mathcal{L}_i^S(\phi, |\mathcal{K}|) := \frac{1}{l} h_b(\phi) + \phi \log |\mathcal{K}|.
\]
Problem Setup: Transmitting Correlated sources over 2-user MAC and IC

Shannon-theoretic study
- Optimal coding technique
- Necessary and sufficient conditions (in terms of $W_{S_1S_2}, W_{Y|X_1X_2}(MAC)$, $W_{Y_1Y_2|X_1X_2}(IC)$)

Plain Vanilla Lossless source-channel coding over 2-user MAC or IC.
1. A new coding technique.
2. Characterize performance, derive new sufficient conditions.
Problem Setup: Transmitting Correlated sources over 2–user MAC and IC

Shannon-theoretic study

?? Optimal coding technique ??

?? Necessary and sufficient conditions (in terms of $W_{S_1S_2}$, $W_{Y|X_1X_2}$ (MAC) $W_{Y_1Y_2|X_1X_2}$ (IC)) ??

Plain Vanilla Lossless source-channel coding over 2-user MAC or IC.
Problem Setup: Transmitting Correlated sources over 2-user MAC and IC

Shannon-theoretic study

?? Optimal coding technique ??

?? Necessary and sufficient conditions (in terms of $W_{S_1S_2}, W_{Y|X_1X_2}$ (MAC) $W_{Y_1Y_2|X_1X_2}$ (IC)) ??

Plain Vanilla Lossless source-channel coding over 2-user MAC or IC.

1. A new coding technique.

2. Characterize performance, derive new sufficient conditions.
Central element of both problems

? Optimally transferring source correlation onto co-ordinated channel inputs ?
Prior work: Cover, El Gamal and Salehi (CES) [Nov. 1980] technique. Dueck’s example

\[ H(S_1|S_2) < I(X_1; Y|X_2, S_2, U), \quad H(S_2|S_1) < I(X_2; Y|X_1, S_1, U) \]
\[ H(S_1, S_2|K) < I(X_1X_2; Y|U), \quad H(S_1, S_2) < I(X_1X_2; Y) \]

for a valid pmf \( W_{S_1S_2} p_U p_{X_1|US_1} p_{X_2|US_2} W_{Y|X_1X_2} \).
Prior work: Liu and Chen [Dec. 2011]

Incorporated

1. CES technique
2. Random source Partitioning [Han Costa 1987]
3. Message Splitting Via super position coding [Han-Kobayashi 1981]

into a single coding (LC) technique and derived sufficient (LC) conditions.
CES and LC techniques: A common thread/constraint

\[ X_1 - S_1 - S_2 - X_2 \]

\[ X_1 = g_1(S_1, W_1), \quad X_2 = g_2(S_2, W_2). \quad W_1 \perp W_2. \]
CES and LC conditions are sub-optimal


IC: LC conditions are sub-optimal. (Adapt Dueck’s argument).

This talk:

Understand why CES and LC are sub-optimal via Dueck’s example and develop a new coding technique.
Begin with a very simple MAC example

$S_1, S_2 \equiv \text{Sources. Source alphabet} = \{0, 1, \cdots, J\}$.

Channel inputs $\equiv X_1, X_2 \in \{0, 1, \cdots, L - 1\}$

Channel output $\equiv Y \in \{0, 1, \cdots, L - 1, *\}$.
Begin with a very simple example
Source alphabet $= \{0, 1, \cdots, J\}$. Source PMF

\[
W_{S_1S_2} = \begin{pmatrix}
0 & 1 & J-1 & J \\
0 & 1-\varepsilon & 0 & \cdots & 0 & 0 \\
1 & 0 & \frac{\varepsilon}{J} & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
J-1 & 0 & 0 & \frac{\varepsilon}{J} & 0 \\
J & 0 & 0 & 0 & \frac{\varepsilon}{J}
\end{pmatrix}
\]

$S_1 = S_2$ w.p. 1.
Begin with a very simple example
Source alphabet = \{0, 1, \cdots, J\}. Source PMF

\begin{align*}
W_{S_1, S_2} & \begin{array}{cccc}
0 & 1 & J-1 & J \\
0 & 1-\epsilon & 0 & \cdots & 0 & 0 \\
1 & 0 & \frac{\epsilon}{J} & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
J-1 & 0 & 0 & \frac{\epsilon}{J} & 0 \\
J & 0 & 0 & 0 & \frac{\epsilon}{J}
\end{array}
\end{align*}

\[ S_1 = S_2 \text{ w.p. } 1. \quad H(S_1) = h_b(\epsilon) + \epsilon \log J \]
Begin with a very simple example

Source alphabet = \{0, 1, \cdots, J\}. Source PMF

\[
\begin{array}{cccc|c}
W_{S_1, S_2} & 0 & 1 & \cdots & J-1 & J \\
\hline
0 & 1 - \epsilon & 0 & \cdots & 0 & 0 \\
1 & 0 & \frac{\epsilon}{J} & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
J-1 & 0 & 0 & \frac{\epsilon}{J} & 0 \\
J & 0 & 0 & 0 & \frac{\epsilon}{J} \\
\end{array}
\]

\[
H(S_1) = h_b(\epsilon) + \epsilon \log J
\]

\[
S_1 = S_2 \text{ w.p. 1.}
\]

MAC Channel

\[
H(S_1) = h_b(\epsilon) + \epsilon \log J
\]
Begin with a very simple example

Source alphabet = \{0, 1, \cdots, J\}. Source PMF

\[ W_{S_1, S_2} = \begin{pmatrix} 0 & 1 & \cdots & J-1 & J \\ 0 & 1-\epsilon & 0 & \cdots & 0 & 0 \\ 1 & 0 & \frac{\epsilon}{J} & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ J-1 & 0 & 0 & \frac{\epsilon}{J} & 0 \\ J & 0 & 0 & 0 & \frac{\epsilon}{J} \end{pmatrix} \]

\[ S_1 = S_2 \text{ w.p. } 1. \quad H(S_1) = h_b(\epsilon) + \epsilon \log J \]
Begin with a very simple example
Source alphabet = \{0, 1, \cdots, J\}. Source PMF

\[ W_{S_1, S_2} \]

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>\cdots</th>
<th>J-1</th>
<th>J</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1-\epsilon &amp; 0 &amp; \cdots &amp; 0 &amp; 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0   &amp; \epsilon/J &amp; 0   &amp; \cdots &amp; 0 &amp; 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J-1</td>
<td>0   &amp; 0   &amp; \epsilon/J &amp; 0   &amp; \cdots &amp; 0 &amp; 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>0   &amp; 0   &amp; 0   &amp; \epsilon/J &amp; \cdots &amp; 0 &amp; 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ H(S_1) = h_b(\epsilon) + \epsilon \log J \]

MAC Channel

\[ X_1 \rightarrow Y \]
\[ X_2 \]

\[ X_1, X_2 \in \{0, 1, \ldots, L-1\} \]
\[ Y = X_1 = X_2 \text{ if } X_1 = X_2 \]
\[ Y = * \text{ if } X_1 \neq X_2 \]
\[ \log L = H(S_1) \]

Source is transmissible as long as \( H(S_1) = H(S_2) \leq \log L \).
Begin with a very simple example

Source alphabet = \{0, 1, \cdots, J\}. Source PMF

\[
W_{S_1, S_2} = \begin{bmatrix}
0 & 1-\epsilon & 0 & \cdots & 0 & 0 \\
1 & 0 & \frac{\epsilon}{J} & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
J-1 & 0 & 0 & \frac{\epsilon}{J} & 0 \\
J & 0 & 0 & 0 & \frac{\epsilon}{J}
\end{bmatrix}
\]

\[
S_1 = S_2 \text{ w.p 1. } H(S_1) = h_b(\epsilon) + \epsilon \log J
\]

Above source, is transmissible as long as

\[
H(S_1) = H(S_2) \leq \sup_{p_{X_1X_2}} I(X_1X_2; Y)
\] (1)
Begin with a very simple example
Source alphabet = \{0, 1, \cdots, J\}. Source PMF

\[
W_{S_1, S_2} = \begin{array}{c|cccc}
0 & 1 & J-1 & J \\
\hline
0 & 1-\epsilon & 0 & \cdots & 0 & 0 \\
1 & 0 & \frac{\epsilon}{J} & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
J-1 & 0 & 0 & \frac{\epsilon}{J} & 0 \\
J & 0 & 0 & 0 & \frac{\epsilon}{J} \\
\end{array}
\]

\[S_1 = S_2 \text{ w.p. } 1. \ H(S_1) = h_b(\epsilon) + \epsilon \log J\]

MAC Channel

Above source, is transmissible as long as

\[H(S_1) = H(S_2) \leq \sup_{p_{X_1 X_2}} I(X_1 X_2; Y) \quad (1)\]

ACHIEVABLE (CO-ORDINATED) INPUT PMF IS UNCONSTRAINED BY SOURCE PMF.

Can Squeeze \(\epsilon\), blow up \(J\), hold \(H(S_1)\). Source is transmissible if (??) holds.
Gács Körner part facilitates co-ordination and efficient communication

Gács Körner part provides considerable flexibility to co-ordinate and communicate efficiently.

\[ H(S_1) = \log L \]

Need \( X_1 = X_2 \) uniform and \( X_1 = X_2 \).

For small \( \epsilon \), source highly non-uniform.
What if we tweak the source just a little?

\[ W_{S_1,S_2} \]

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>( J-1 )</th>
<th>( J )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1-( \varepsilon )</td>
<td>0</td>
<td>( \ldots )</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>( \frac{\varepsilon}{J} )</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
<td>( \vdots )</td>
</tr>
<tr>
<td>( J-1 )</td>
<td>0</td>
<td>0</td>
<td>( \frac{\varepsilon}{J} )</td>
<td>0</td>
</tr>
<tr>
<td>( J )</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>( \frac{\varepsilon}{J} )</td>
</tr>
</tbody>
</table>
What if we tweak the source just a little?

\[
\begin{array}{cccc}
W_{S_1,S_2} & 0 & 1 & J-1 & J \\
0 & 1-\varepsilon & \theta & \theta & \theta \\
1 & 0 & \frac{\varepsilon}{J-\theta} & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
J-1 & 0 & 0 & \frac{\varepsilon}{J-\theta} & 0 \\
J & 0 & 0 & 0 & \frac{\varepsilon}{J-\theta}
\end{array}
\]
What if we tweak the source just a little?

<table>
<thead>
<tr>
<th>$W_{S_1, S_2}$</th>
<th>0</th>
<th>1</th>
<th>$J-1$</th>
<th>$J$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$1-\epsilon$</td>
<td>$\theta$</td>
<td>...</td>
<td>$\theta$</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>$\frac{\epsilon}{J-\theta}$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$J-1$</td>
<td>0</td>
<td>0</td>
<td>$\frac{\epsilon}{J-\theta}$</td>
<td>0</td>
</tr>
<tr>
<td>$J$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\frac{\epsilon}{J-\theta}$</td>
</tr>
</tbody>
</table>

MAC Channel remains same.

Noiseless if $X_1 = X_2$ and Erasure if $X_1 \neq X_2$

Need $X_1 = X_2$ uniform and $X_1 = X_2$. 
What if we tweak the source just a little?

MAC Channel remains same.

<table>
<thead>
<tr>
<th>$W_{S_1,S_2}$</th>
<th>0</th>
<th>1</th>
<th>$J-1$</th>
<th>$J$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$1-\varepsilon$</td>
<td>$\theta$</td>
<td>$\theta$</td>
<td>$\theta$</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>$\frac{\varepsilon}{J-\theta}$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$J-1$</td>
<td>0</td>
<td>0</td>
<td>$\frac{\varepsilon}{J-\theta}$</td>
<td>0</td>
</tr>
<tr>
<td>$J$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\frac{\varepsilon}{J-\theta}$</td>
</tr>
</tbody>
</table>

Source stripped off its Gács-Körner part.

Noiseless if $X_1 = X_2$ and Erasure if $X_1 \neq X_2$

Need $X_1 = X_2$ uniform and $X_1 = X_2$. 
What if we tweak the source just a little?

MAC Channel remains same.

\[
W_{S_1, S_2} | 0 & 1 & J-1 & J \\
0 & 1-\varepsilon & \theta & \theta & \theta \\
1 & 0 & \frac{\varepsilon}{J-\theta} & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
J-1 & 0 & 0 & \frac{\varepsilon}{J-\theta} & 0 \\
J & 0 & 0 & 0 & \frac{\varepsilon}{J-\theta} \\
\]

Source stripped off its Gács-Körner part.

Single-Letter (S-L) coding is constrained to

\[
X_1 - S_1 - S_2 - X_2.
\]

Noiseless if \(X_1 = X_2\) and

Erasure if \(X_1 \neq X_2\)

Need \(X_1 = X_2\) uniform and \(X_1 = X_2\).
What if we tweak the source just a little?

MAC Channel remains same.

\[
\begin{array}{cccc|c}
W_{S_1,S_2} & 0 & 1 & J-1 & J \\
\hline
0 & 1-\varepsilon & \theta & \cdots & \theta & \theta \\
1 & 0 & \frac{\varepsilon}{J-\theta} & 0 & 0 & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \\
J-1 & 0 & 0 & \frac{\varepsilon}{J-\theta} & 0 & \\
J & 0 & 0 & 0 & \frac{\varepsilon}{J-\theta} & \\
\end{array}
\]

\[X_1 = g_1(S_1, W_1), \ X_2 = g_2(S_2, W_2). \ W_1 \perp W_2.\]

\[W_1 \text{ and/or } W_2 \text{ non-trivial RVs reduces } P(X_1 = X_2).\]

\[X_1 = X_2 \text{ uniform and } X_1 = X_2.\]
What if we tweak the source just a little?

MAC Channel remains same.

Noiseless if \( X_1 = X_2 \) and Erasure if \( X_1 \neq X_2 \)

Need \( X_1 = X_2 \) uniform and \( X_1 = X_2 \).

### Table:

<table>
<thead>
<tr>
<th>( W_{S_1,S_2} )</th>
<th>0</th>
<th>1</th>
<th>( J-1 )</th>
<th>( J )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1-( \varepsilon )</td>
<td>( \theta )</td>
<td>( \theta )</td>
<td>( \theta )</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>( \frac{\varepsilon}{J-\theta} )</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( J-1 )</td>
<td>0</td>
<td>0</td>
<td>( \frac{\varepsilon}{J-\theta} )</td>
<td>0</td>
</tr>
<tr>
<td>( J )</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>( \frac{\varepsilon}{J-\theta} )</td>
</tr>
</tbody>
</table>

### Diagram:

\( X_1 = g_1(S_1, \phi), X_2 = g_2(S_2, \phi) \).

?Trivializing \( W_1, W_2 \) and requiring \( X_1 = X_2 \) uniform?
What if we tweak the source just a little?

<table>
<thead>
<tr>
<th>$W_{S_1,S_2}$</th>
<th>0</th>
<th>1</th>
<th>$J-1$</th>
<th>$J$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$1-\varepsilon$</td>
<td>$\theta$</td>
<td>$\cdots$</td>
<td>$\theta$</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>$\frac{\varepsilon}{J-\theta}$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$J-1$</td>
<td>0</td>
<td>0</td>
<td>$\frac{\varepsilon}{J-\theta}$</td>
<td>0</td>
</tr>
<tr>
<td>$J$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\frac{\varepsilon}{J-\theta}$</td>
</tr>
</tbody>
</table>

MAC Channel remains same.

$X_1 = g_1(S_1, \phi), \ X_2 = g_2(S_2, \phi)$.

?Trivializing $W_1, W_2$ and requiring $X_1 = X_2$ uniform?

Pool all less likely symbols together.
What if we tweak the source just a little?

\[
\begin{array}{c|c|c|c|c}
W_{S_1S_2} & 0 & 1 & J-1 & J \\
\hline
0 & 1-\varepsilon & \theta & \cdots & \theta \\
\hline
1 & 0 & \frac{\varepsilon}{J-\theta} & 0 & 0 \\
\hline
\vdots & \vdots & \ddots & \vdots & \vdots \\
\hline
J-1 & 0 & 0 & \frac{\varepsilon}{J-\theta} & 0 \\
\hline
J & 0 & 0 & 0 & \frac{\varepsilon}{J-\theta} \\
\end{array}
\]

\[X_1 = g_1(S_1, \phi), \quad X_2 = g_2(S_2, \phi).\]

?Trivializing \(W_1, W_2\) and requiring \(X_1 = X_2\) uniform?

Pool all less likely symbols together.

Can get only \((1 - \varepsilon, \varepsilon)\)
What if we tweak the source just a little?

<table>
<thead>
<tr>
<th>$W_{S_1,S_2}$</th>
<th>0</th>
<th>1</th>
<th>$J-1$</th>
<th>$J$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1-$\epsilon$</td>
<td>$\theta$</td>
<td>$\theta$</td>
<td>$\theta$</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>$\frac{\epsilon}{J-\theta}$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$J-1$</td>
<td>0</td>
<td>0</td>
<td>$\frac{\epsilon}{J-\theta}$</td>
<td>0</td>
</tr>
<tr>
<td>$J$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\frac{\epsilon}{J-\theta}$</td>
</tr>
</tbody>
</table>

Carefully choose values for $\epsilon, \theta, J, L$, Dueck [1981] proves


2. Proposes mapping fixed length blocks of source to channel codewords. Proves $S_1, S_2$ is transmissible.
Mimic conditional coding via fixed block-length coding

When \( S_1 = S_2 \).

<table>
<thead>
<tr>
<th>( W_{S_1,S_2} )</th>
<th>0</th>
<th>1</th>
<th>( J-1 )</th>
<th>( J )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1-( \varepsilon )</td>
<td>0</td>
<td>\ldots</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>( \varepsilon )</td>
<td>( J-1 )</td>
<td>0</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots \ddots \vdots</td>
<td>\vdots</td>
<td>\vdots</td>
<td></td>
</tr>
<tr>
<td>( J-1 )</td>
<td>0</td>
<td>0</td>
<td>( \varepsilon )</td>
<td>( J-1 )</td>
</tr>
<tr>
<td>( J )</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>( \varepsilon )</td>
</tr>
</tbody>
</table>

Map large (typical) source sequences to codewords chosen uniformly at random.

We can ensure \( X_1 = X_2 \) uniform and \( X_1 = X_2 \).
Mimic conditional coding via fixed block-length coding

\[ P(S_1 \neq S_2) = J\theta > 0. \]

<table>
<thead>
<tr>
<th>( W_{S_1, S_2} )</th>
<th>0</th>
<th>1</th>
<th>( J-1 )</th>
<th>( J )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1-( \varepsilon )</td>
<td>( \theta )</td>
<td>( \theta )</td>
<td>( \theta )</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>( \frac{\varepsilon}{J-\theta} )</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
<td>\vdots</td>
</tr>
<tr>
<td>( J-1 )</td>
<td>0</td>
<td>0</td>
<td>( \frac{\varepsilon}{J-\theta} )</td>
<td>0</td>
</tr>
<tr>
<td>( J )</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>( \frac{\varepsilon}{J-\theta} )</td>
</tr>
</tbody>
</table>

With increasing block-length \( n \), \( P(X_1^n \neq X_2^n) \to 1. \)

Conditional coding with \textbf{FIXED BLOCK-LENGTH} \( l \).

\[ P(S_1^l \neq S_2^l) = 1 - (1 - J\theta)^l \leq lJ\theta. \]
Fixed block-length coding forces a concatenated coding scheme

Source PMF

<table>
<thead>
<tr>
<th>W_{S_1,S_2}</th>
<th>0</th>
<th>1</th>
<th>J-1</th>
<th>J</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1-\epsilon</td>
<td>\theta</td>
<td>\theta</td>
<td>\theta</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>\epsilon/J-\theta</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>J-1</td>
<td>0</td>
<td>\epsilon/J-\theta</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>J</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>\epsilon/J-\theta</td>
</tr>
</tbody>
</table>

MAC Channel

- X_1, X_2 \in \{0,1,\ldots, L-1\}
- Y = X_1=X_2 if X_1=X_2
- Y = * if X_1 \neq X_2
- \log L = H(S_1)

Block-length of coding is determined by \epsilon, \theta, J, L, not be the desired prob. of error.

This fixed block-length code results in non-vanishing probability of error.

Necessitates an outer code of arbitrarily large block-length to

1. Correct for errors in the fixed block-length code.
2. Communicate rest of information.

This results in concatenated coding structure.
Part 2 and 3

Concatenated coding scheme for a general problem instance whose info-theoretic performance can be characterized?
What challenges does the finite block-length code throw up?

Revisit CES scheme with common part $K$

Two phase coding, in the presence of GK common part $K$.

Phase 1: Common part $K^n$ is decoded perfectly (owing to infinite block-length)
What challenges does the finite block-length code throw up?

Revisit CES scheme with common part $K$

Two phase coding, in the presence of GK common part $K$.

Phase 1: Common part $K^n$ is decoded perfectly (owing to infinite block-length)

Phase 2: Rest of Information: $S^n_1, S^n_2$ conditioned on common part $K^n$ is encoded in $X^n_1, X^n_2$ and decoded by Rx.

$K^n$ can be viewed as side-information available at all terminals.
What challenges does the finite block-length code throw up?

Suppose $S_1, S_2$ do NOT have common part, but highly correlated.

$m \times l$ matrix encoding. $l$ remains fixed, $m \to \infty$.

Two phase coding.

Phase 1: Rows of $K_1^{m \times l}$ ($K_2^{m \times l}$) mapped to $U_1^{m \times l}$ ($U_2^{m \times l}$) via finite block-length code. Decoded erroneously as $V^{m \times l}$.
What challenges does the finite block-length code throw up?

Suppose $S_1, S_2$ do NOT have common part, but highly correlated.

$m \times l$ matrix encoding. $l$ remains fixed, $m \to \infty$.

Two phase coding.

Phase 1: Rows of $K_1^m \times l, K_2^m \times l$ mapped to $U_1^m \times l, U_2^m \times l$ via finite block-length code. Decoded erroneously as $V^m \times l$.

Phase 2: Rest of Information: $S_1^m \times l, S_2^m \times l$ conditioned on common part $V^m \times l$ is encoded in $X_1^m \times l, X_2^m \times l$ and decoded by Rx.
Multi-letter distribution induced by the fixed block-length code

Each row of matrices $U_1^{m \times l}$, $U_2^{m \times l}$, $V^{m \times l}$, $G^{m \times l}$ have a multi-letter pmf. How do you perform superposition coding?

Leverage the technique of interleaving suggested in Shirani and Pradhan [ISIT 2014].

Note that the inner code of finite block-length is fixed upfront.

As a consequence, the rows of each of the $m \times l$ matrices are iid. However, each row has a product distribution.
What is Interleaving?
Interleaving [Shirani and Pradhan 2014]


tabular content:

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$E(1,1)$</td>
<td>$E(1,2)$</td>
<td>$E(1,l)$</td>
</tr>
<tr>
<td>$E(t,1)$</td>
<td>$E(t,2)$</td>
<td>$*$</td>
</tr>
<tr>
<td>$E(m,1)$</td>
<td>$E(m,2)$</td>
<td></td>
</tr>
</tbody>
</table>

$E(1,1:l) \ldots E(m,1:l) \sim \Pi_{E}^{l}$
Interleaving [Shirani and Pradhan 2014]

\[ \prod_{E}^{l} \]

<table>
<thead>
<tr>
<th>E(1,1)</th>
<th>E(1,2)</th>
<th>E(1,l)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E(t,1)</td>
<td>E(t,2)</td>
<td>E(t,l)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E(m,1)</td>
<td>E(m,2)</td>
<td>E(m,l)</td>
</tr>
</tbody>
</table>

Let \( \lambda_1, \lambda_2, ..., \lambda_m \) be unif. And randomly chosen permutations in \{1, 2, ..., l\}.
The same argument holds for each sub-vectors $E(1,\lambda_1(i)) \cdots E(m,\lambda_m(i)) : i = 1, \cdots, l$.

We therefore have $l$ iid sub-vectors $E(1,\lambda_1(i)) \cdots E(m,\lambda_m(i)) : i = 1, \cdots, l$. 

Let $\lambda_1, \lambda_2, \ldots, \lambda_m$ be unif. And randomly chosen permutations in $\{1, 2, \ldots, l\}$

$E(1,\lambda_1(1)) E(2,\lambda_2(1)) \cdots E(m,\lambda_m(1)) \sim \prod_{t=1}^m \frac{1}{l} \sum_{i=1}^l p_{E_i}$
Theorem

$(S, W_S)$ is transmissible over $IC (X, Y, W_Y|X)$ if there exists (i) a finite set $K$, maps $f_j : S_j \to K$, with $K_j = f_j (S_j)$ for $j \in [2]$ with

$P(K_1 = K_2) = \phi + \xi[l] = 0$, (i) finite set $U, V_1, V_2$ and pmf $p_{uv_1p_{v_2}p_x_1|uv_1p_{x_2|uv_2}$ defined on $U \times V \times X$, (iii) $A, B \geq 0$,

$$A + B \geq H(K_1) = H(K_2),$$

$$A \leq \min\{I(U; Y_1), I(U; Y_2)\} \text{ and for } j \in [2],$$

$$B + H(S_j|K_1) < I(V_j; Y_j).$$
A new coding theorem in a simplified form

Theorem

$$(S, W_S)$$ is transmissible over IC $$(X, Y, W_Y|X)$$ if there exists (i) a finite set $$K$$, maps $$f_j : S_j \rightarrow K$$, with $$K_j = f_j(S_j)$$ for $$j \in [2]$$, (ii) $$l \in \mathbb{N}, \delta > 0$$, (iii) finite set $$U, V_1, V_2$$ and pmf $$p_U p_{V_1} p_{V_2} p_{X_1|UV_1} p_{X_2|UV_2}$$ defined on $$U \times V \times X$$, where $$p_U$$ is a type of sequences in $$U^l$$, (iv) $$A, B \geq 0$$, $$\rho \in (0, A)$$ such that $$\phi \in [0, 0.5)$$,

$$A + B \geq (1 + \delta)H(K_1), \text{ and for } j \in [2],$$

$$B + H(S_j|K_1) + \mathcal{L}_i^S(\phi, |S_j|) < I(V_j; Y_j) - \mathcal{L}_j^C(\phi, |V|),$$

where,
A new coding theorem in a simplified form

Theorem

(\mathcal{S}, \mathbb{W}_S) is transmissible over IC (\mathcal{X}, \mathcal{Y}, \mathbb{W}_{Y|X}) if there exists (i) a finite set \mathcal{K}, maps \( f_j : S_j \to \mathcal{K} \), with \( K_j = f_j(S_j) \) for \( j \in [2] \), (ii) \( l \in \mathbb{N}, \delta > 0 \), (iii) finite set \( \mathcal{U}, \mathcal{V}_1, \mathcal{V}_2 \) and pmf \( p_U p_{V_1} p_{V_2} p_{X_1|U} p_{X_2|U} p_{V_2} \) defined on \( \mathcal{U} \times \mathcal{V}_1 \times \mathcal{X} \), where \( p_U \) is a type of sequences in \( \mathcal{U}^l \), (iv) \( A, B \geq 0, \rho \in (0, A) \) such that \( \phi \in [0, 0.5) \),

\[
A + B \geq (1 + \delta)H(K_1), \quad \text{and for } j \in [2],
\]

\[
B + H(S_j|K_1) + \mathcal{L}_S^l(\phi, |S_j|) < I(V_j; Y_j) - \mathcal{L}_j^C(\phi, |\mathcal{V}|),
\]

where, \( \phi := g_{\rho, l} + \xi^{[l]}(K) + \tau_{\rho, \delta}(K_1) \), \( \xi^{[l]}(K) := P(K_1 \neq K_2) \)

\[
g_{\rho, l} := \sum_{j=1}^{2} \exp\{-l(E_\tau(A + \rho, p_U, p_{Y_j}|U) - \rho)\}, \quad (\text{Channel coding error})
\]

\[
\tau_{\rho, \delta}(K) = 2|\mathcal{K}| \exp\{-2\delta^2 p_K^2(a^*)l\} \quad (\text{Prob of non-typicality})
\]

\[
\mathcal{L}_j^C(\phi, |\mathcal{U}|) = h_b(\phi) + \phi \log |\mathcal{U}| + |\mathcal{Y}| |\mathcal{U}| \phi \log \frac{1}{\phi},
\]

\[
\mathcal{L}_S^l(\phi, |\mathcal{K}|) := \frac{1}{l} h_b(\phi) + \phi \log |\mathcal{K}|.
\]
Backup
Finite block-length inner code concatenated with an large block-length outer code

Let $U_l^j$ be $C_U$ codeword chosen by encoder $j$.

$$P(U_1^l \neq U_2^l) \leq P(K_1^l \neq K_2^l) + P(K_1^l \notin T_{\delta}(K_1))$$
$$\leq 1 - (1 - \xi)^l + 2|\mathcal{K}|- \exp\{-2p_{K_1}(a^*)l\} =: \beta$$

Suppose $g$ is the prob. of error of a PTP channel decoder operating over a PTP channel $p_{Y|U}$. 
Inner block code of finite block length

<table>
<thead>
<tr>
<th>$S_1(1,1)$</th>
<th>$S_1(1,2)$</th>
<th>$S_1(1,l)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_1(t,1)$</td>
<td>$S_1(t,2)$</td>
<td>$* \ldots *$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$S_1(t,l)$</td>
</tr>
<tr>
<td>$S_1(m,1)$</td>
<td>$S_1(m,2)$</td>
<td>$S_1(m,l)$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$K_1(1,1)$</th>
<th>$K_1(1,2)$</th>
<th>$K_1(1,l)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_1(t,1)$</td>
<td>$K_1(t,2)$</td>
<td>$* \ldots *$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$K_1(t,l)$</td>
</tr>
<tr>
<td>$K_1(m,1)$</td>
<td>$K_1(m,2)$</td>
<td>$K_1(m,l)$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$G(1,1)$</th>
<th>$G(1,2)$</th>
<th>$G(1,l)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$G(t,1)$</td>
<td>$G(t,2)$</td>
<td>$* \ldots *$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$G(t,l)$</td>
</tr>
<tr>
<td>$G(m,1)$</td>
<td>$G(m,2)$</td>
<td>$G(m,l)$</td>
</tr>
</tbody>
</table>

Encoder 1

Encoder 2

Decoder

Assuming encoders agree, decoder decodes into $C_U$-codebook with a PTP decoder on $p_{Y|U}$ chnl.
Inner block code of finite block length

$V$ is decoded version of $U$. $G(t, 1 : l) = e_U^{-1}(V(t, 1 : l))$ is decoded version of $K$. 

<table>
<thead>
<tr>
<th>$S_1(1,1)$</th>
<th>$S_1(1,2)$</th>
<th>$S_1(1,l)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_1(t,1)$</td>
<td>$S_1(t,2)$</td>
<td>$S_1(t,l)$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_1(m,1)$</td>
<td>$S_1(m,2)$</td>
<td>$S_1(m,l)$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$K_1(1,1)$</th>
<th>$K_1(1,2)$</th>
<th>$K_1(1,l)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_1(t,1)$</td>
<td>$K_1(t,2)$</td>
<td>$K_1(t,l)$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_1(m,1)$</td>
<td>$K_1(m,2)$</td>
<td>$K_1(m,l)$</td>
</tr>
</tbody>
</table>

Encoder 1

Encoder 2

Decoder

$V(t, 1 : l)$ is input to decoder. $V(t, 1 : l) = e_U^{-1}(V(t, 1 : l))$ is decoded version of $K$.
Encoders and decoder agree on a fraction $1 - (g + \beta)$ of the rows of $U$-matrices.
Outer code: The approach

Encoder 1 has
\[ S_1(1 : l) \cdots S_1((m - 1)l + 1 : ml) \]
\[ K_1(1 : l) \cdots K_1((m - 1)l + 1 : ml) \]
\[ U_1(1 : l) \cdots U_1((m - 1)l + 1 : ml) \]

Encoder 2 has
\[ S_2(1 : l) \cdots S_2((m - 1)l + 1 : ml) \]
\[ K_2(1 : l) \cdots K_2((m - 1)l + 1 : ml) \]
\[ U_2(1 : l) \cdots U_2((m - 1)l + 1 : ml) \]

Decoder has
\[ G(1 : l) \cdots G((m - 1)l + 1 : ml) \]
\[ V(1 : l) \cdots V((m - 1)l + 1 : ml) \]

If these had a single-letter form i.e., distributed as \( \prod_{t=1}^{ml} p_S S_1 K_1 U_1 S_2 K_2 U_2 V G \), then one could apply standard info-theoretic techniques.
Challenges

Encoder 1 has
\[ S_1(1 : l) \cdots S_1((m - 1)l + 1 : ml) \]
\[ K_1(1 : l) \cdots K_1((m - 1)l + 1 : ml) \]
\[ U_1(1 : l) \cdots U_1((m - 1)l + 1 : ml) \]

Encoder 2 has
\[ S_2(1 : l) \cdots S_2((m - 1)l + 1 : ml) \]
\[ K_2(1 : l) \cdots K_2((m - 1)l + 1 : ml) \]
\[ U_2(1 : l) \cdots U_2((m - 1)l + 1 : ml) \]

Decoder has
\[ G(1 : l) \cdots G((m - 1)l + 1 : ml) \]
\[ V(1 : l) \cdots V((m - 1)l + 1 : ml) \]

- Not guaranteed to have a iid form.
- Even if we can extract sub-vectors that have product iid form, we do not know the underlying pmf. Hence, cannot express rates in terms of these pmfs.
The rows are independent and identically distributed.

The rows are independent.
The rows are independent and identically distributed.

The rows are independent. And identically distributed.
The rows are independent and identically distributed.

Encoder 1

Encoder 2

Decoder

The rows are independent and identically distributed.
The rows are independent and identically distributed.

\[
\begin{array}{ccc}
S_1(t,1) & S_1(t,2) & \cdots & S_1(t,l) \\
K_1(t,1) & K_1(t,2) & \cdots & K_1(t,l) \\
U_1(t,1) & U_1(t,2) & \cdots & U_1(t,l) \\
S_2(t,1) & S_2(t,2) & \cdots & S_2(t,l) \\
K_2(t,1) & K_2(t,2) & \cdots & K_2(t,l) \\
U_2(t,1) & U_2(t,2) & \cdots & U_2(t,l) \\
V(t,1) & V(t,2) & \cdots & V(t,l) \\
G(t,1) & G(t,2) & \cdots & G(t,l)
\end{array}
\]

\[= \mathbf{E}(t,1:l) \sim \mathbf{p}_{\mathbf{E}}^l\]

\[\mathbf{E}(1,1:l) \ldots \mathbf{E}(m,1:l) \sim \prod \mathbf{p}_{\mathbf{E}}^l\]

The rows are independent and identically distributed.
Randomly chosen co-ordinates are iid!!!

For each row $t$, choose column $\pi_t(1)$, where $\pi_1, \cdots, \pi_m$ are random permutations.
Randomly chosen co-ordinates are iid!!!
Randomly chosen co-ordinates are iid!!!

Let $\prod_1, \prod_2, ..., \prod_m$ be unif. And randomly chosen permutations in $\{1, 2, ..., l\}$

$E(1, \prod_1(1)) \ E(2, \prod_2(1)) \ ... \ E(m, \prod_m(1)) \sim \prod_{t=1 \ldots m} (1/l) \sum_{i=1 \ldots l} p_{E_i}$

The rows are independent and identically distributed.

We therefore have $l$ iid sub-vectors $E(1, \pi_1(i)) \cdots E(m, \pi_m(i)) : i = 1, \cdots, l$. 

| $E(1,1)$ | $E(1,2)$ |   | $E(1,l)$ |
|-----------------|-----------------|   |-----------------|
| $E(t,1)$ | $E(t,2)$ |   | $E(t,l)$ |
|-----------------|-----------------|   |-----------------|
| $E(m,1)$ | $E(m,2)$ |   | $E(m,l)$ |
Randomly chosen co-ordinates are iid!!!

Co-ordinates with same color coded together within a block in the outer code. $l$ such blocks.
Randomly chosen co-ordinates are iid!!!

Co-ordinates with same color coded together within a block in the outer code. $l$ such blocks.
Randomly chosen co-ordinates are iid!!!
Randomly chosen co-ordinates are iid!!!

Co-ordinates with same color coded together within a block in the outer code. $l$ such blocks.
The induced single-letter distributions are unknown!!!
The induced single-letter distributions are unknown!!!

\[
\begin{align*}
S_1(1, \pi_1(1)) & S_1(2, \pi_2(1)) \cdots S_1(m, \pi_m(1)) \\
K_1(1, \pi_1(1)) & K_1(2, \pi_2(1)) \cdots K_1(m, \pi_m(1)) \\
U_1(1, \pi_1(1)) & U_1(2, \pi_2(1)) \cdots U_1(m, \pi_m(1)) \\
S_2(1, \pi_1(1)) & S_2(2, \pi_2(1)) \cdots S_2(m, \pi_m(1)) \\
K_2(1, \pi_1(1)) & K_2(2, \pi_2(1)) \cdots K_2(m, \pi_m(1)) \\
U_2(1, \pi_1(1)) & U_2(2, \pi_2(1)) \cdots U_2(m, \pi_m(1)) \\
V(1, \pi_1(1)) & V(2, \pi_2(1)) \cdots V(m, \pi_m(1)) \\
G(1, \pi_1(1)) & G(2, \pi_2(1)) \cdots G(m, \pi_m(1))
\end{align*}
\]

\[
\approx \prod_{t=1}^{m} p_{\hat{S}_1 \hat{K}_1 \hat{U}_1 \hat{S}_2 \hat{K}_2 \hat{U}_2 \hat{V} \hat{G}}
\]

The sufficient conditions will look like

\[
H(\hat{S}_1 | \hat{S}_2, G) < I(\hat{X}_1; Y \hat{X}_2 | \hat{V} \hat{S}_2), \quad H(\hat{S}_2 | \hat{S}_1, G) < I(\hat{X}_2; Y \hat{X}_1 | \hat{V} \hat{S}_1) \cdots
\]

But, we do not know this pmf. This pmf depends on the pmf of the empirical distribution of the inner layer code and the errors in that layer.
Upperbound the difference using relationship between chosen and actual pmf

Let \( p_U p_{X_1|U S_1} p_{X_2|U S_2} \) be chosen pmf.

\( p_{\hat{U}_j} = p_U \) since symbols of \( C_U \) chosen iid \( p_U \).

\[ P(\hat{U}_1 \neq \hat{U}_2) \leq \beta \]

\( p_{\hat{X}_j|\hat{U}_j \hat{S}_j} = p_{X_j|U S_j} \) : By choice of coding technique. One can derive an upper bound using these relations.
Upperbound the difference using relationship between chosen and actual pmf

Let $p_U p_{X_1|U S_1} p_{X_2|U S_2}$ be chosen pmf.

$p_{\hat{U}_j} = p_U$ since symbols of $C_U$ chosen iid $p_U$.

$P(\hat{U}_1 \neq \hat{U}_2) \leq \beta$

$p_{X_j|\hat{U}_j \hat{S}_j} = p_{X_j|U S_j}$: By choice of coding technique. One can derive an upper bound using these relations.