On the Sub-optimality of Single-letter Coding in Multi-terminal Communications

Farhad Shirani, S. Sandeep Pradhan

Advisor: Sandeep Pradhan
University of Michigan

ITA 2017, San Diego
Outline

1. Effective Length
   - Binary Block Encoders

2. Distributed Source Coding
   - Problem Statement
   - Binary One-help-one

3. Boolean Functions
   - Maximum Correlation

4. Single-letter Coding

On the Sub-optimality of Single-letter Coding in Multi-terminal Communications
Binary Block Encoder: A function $e : \{0, 1\}^n \rightarrow \{0, 1\}^k$. 

Effective length

Distributed Source Coding

Boolean Functions

Single-letter Coding

Binary Block Encoders
Effective Length

Binary Block Encoders

- Binary Block Encoder: A function $e : \{0, 1\}^n \to \{0, 1\}^k$.
- The BBE is a vector of Boolean functions.
Binary Block Encoder: A function $e : \{0, 1\}^n \rightarrow \{0, 1\}^k$.

- The BBE is a vector of Boolean functions.
- The blocklength is $n$. 

Effective length
The number of input bits needed to estimate the output of the Boolean function with high precision.

How to determine 'effective length' of the BBE?
- e.g. A concatenation of the BBE doesn't change the effective length.

$f = (e_1, e_2) : \{0, 1\}^2 \rightarrow \{0, 1\}^2$. 

Farhad Shirani, S. Sandeep Pradhan
On the Sub-optimality of Single-letter Coding in Multi-terminal Communications
Binary Block Encoder: A function $e : \{0, 1\}^n \rightarrow \{0, 1\}^k$.

The BBE is a vector of Boolean functions.

The blocklength is $n$.

**Effective length**

The number of input bits needed to estimate the output of the Boolean function with high precision.
- Binary Block Encoder: A function $e : \{0, 1\}^n \rightarrow \{0, 1\}^k$.
- The BBE is a vector of Boolean functions.
- The blocklength is $n$.

**Effective length**

The number of input bits needed to estimate the output of the Boolean function with high precision.

- How to determine 'effective length' of the BBE?
- Binary Block Encoder: A function $e : \{0, 1\}^n \rightarrow \{0, 1\}^k$.
- The BBE is a vector of Boolean functions.
- The blocklength is $n$.

**Effective length**

The number of input bits needed to estimate the output of the Boolean function with high precision.

- How to determine ‘effective length’ of the BBE?
- e.g. A concatenation of the BBE doesn’t change the effective length.

$$f = (e, e) : \{0, 1\}^{2n} \rightarrow \{0, 1\}^{2k}.$$
PtP Optimality

- PtP communication: optimality achieved using large blocklengths.
PtP Optimality

- PtP communication: optimality achieved using large blocklengths.
- Approaching optimality requires large effective-lengths.
PtP Optimality

- PtP communication: optimality achieved using large blocklengths.
- Approaching optimality requires large effective-lengths.
- Law of large numbers $\rightarrow$ hardening of the distribution.
PtP Optimality

- PtP communication: optimality achieved using large blocklengths.
- Approaching optimality requires large effective-lengths.
- Law of large numbers $\rightarrow$ hardening of the distribution.
- Source coding: exploits source redundancy.
PtP Optimality

- PtP communication: optimality achieved using large blocklengths.
- Approaching optimality requires large effective-lengths.
- Law of large numbers $\rightarrow$ hardening of the distribution.
- Source coding: exploits source redundancy.
- Channel coding: exploit typicality of the noise.
PtP Optimality

- PtP communication: optimality achieved using large blocklengths.
- Approaching optimality requires large effective-lengths.
- Law of large numbers $\rightarrow$ hardening of the distribution.
- Source coding: exploits source redundancy.
- Channel coding: exploit typicality of the noise.
Multi-terminal Communications

Almost all multi-terminal schemes use random, large blocklength codes.
Almost all multi-terminal schemes use random, large blocklength codes. Performance is super-additive with blocklength. 

\[ C_{2n} \supseteq C_n, \quad RD_{2n} \supseteq RD_n \]
Multi-terminal Communications

- Almost all multi-terminal schemes use random, large blocklength codes.
- Performance is super-additive with blocklength.

\[ C_{2n} \supseteq C_n, \quad RD_{2n} \supseteq RD_n \]

- These random codes have large effective lengths.
Multi-terminal Communications

- Almost all multi-terminal schemes use random, large blocklength codes.
- Performance is super-additive with blocklength.

\[ C_{2n} \supseteq C_n, \quad RD_{2n} \supseteq RD_n \]

- These random codes have large effective lengths.
- Large Effective length → Loss of correlation.
Multi-terminal Communications

- Almost all multi-terminal schemes use random, large blocklength codes.
- Performance is super-additive with blocklength.

\[ C_{2n} \supseteq C_n, \quad RD_{2n} \supseteq RD_n \]

- These random codes have large effective lengths.
- Large Effective length $\rightarrow$ Loss of correlation.
- In networks, correlation is necessary for cooperation.
Multi-terminal Communications

- Almost all multi-terminal schemes use random, large blocklength codes.
- Performance is super-additive with blocklength.

\[ C_{2n} \supseteq C_n, \quad RD_{2n} \supseteq RD_n \]

- These random codes have large effective lengths.
- Large Effective length $\rightarrow$ Loss of correlation.
- In networks, correlation is necessary for cooperation.
- Performance is **not** super-additive with effective-length.
Multi-terminal Communications

- Almost all multi-terminal schemes use random, large blocklength codes.
- Performance is super-additive with blocklength.

\[ C_{2n} \supseteq C_n, \quad RD_{2n} \supseteq RD_n \]

- These random codes have large effective lengths.
- Large Effective length → Loss of correlation.
- In networks, correlation is necessary for cooperation.
- Performance is not super-additive with effective-length.
Outline

1. Effective Length
   - Binary Block Encoders

2. Distributed Source Coding
   - Problem Statement
   - Binary One-help-one

3. Boolean Functions
   - Maximum Correlation

4. Single-letter Coding

F. Shirani Chaharsooghi, S. Pradhan, *Finite-Length Gains in Distributed Source Coding*, ISIT 2014

F. Shirani Chaharsooghi, A. Ghasemian Sahebi, S. Pradhan, *Distributed Source Coding in Absence of Common Components*, ISIT 2013
Cooperation in Data Compression

Problem Statement

Y₁ and Y₂ are two DMS's. The encoders do not communicate. The decoder is to reconstruct each source based on distortion criteria.

Question: How can the encoders cooperate to exploit correlation?
Cooperation in Data Compression

- $Y_1$ and $Y_2$ are two DMS's.
Cooperation in Data Compression

- $Y_1$ and $Y_2$ are two DMS’s.
- The encoders do not communicate.
Cooperation in Data Compression

- $Y_1$ and $Y_2$ are two DMS’s.
- The encoders do not communicate.
- The decoder is to reconstruct each source based on distortion criteria.
Cooperation in Data Compression

- $Y_1$ and $Y_2$ are two DMS’s.
- The encoders do not communicate.
- The decoder is to reconstruct each source based on distortion criteria.

**Question:** How can the encoders cooperate to exploit correlation?
Cooperation in Data Compression

- $Y_1$ and $Y_2$ are two DMS’s.
- The encoders do not communicate.
- The decoder is to reconstruct each source based on distortion criteria.
- **Question:** How can the encoders cooperate to exploit correlation?
One method of cooperation: users refine each other’s quantization noises.
One method of cooperation: users refine each other’s quantization noises.

One encoder quantizes its source, the other “guesses” the noise.
One method of cooperation: users refine each other’s quantization noises.

One encoder quantizes its source, the other “guesses” the noise.

The estimation is refined and sent to the decoder.
- One method of cooperation: users refine each other’s quantization noises.
- One encoder quantizes its source, the other “guesses” the noise.
- The estimation is refined and sent to the decoder.
- Example: $Y_1 = X$, $Y_2 = X + E$
  
  $$P(X = 1) = 0.5, \quad P(E = 1) = \epsilon.$$
Problem Statement

- What if we quantize $X$ and $X + E$ using the same large effective-length quantizer?
Problem Statement

- What if we quantize $X$ and $X + E$ using the same large effective-length quantizer?

- Two quantization noises become independent.
Problem Statement

- What if we quantize $X$ and $X + E$ using the same large effective-length quantizer?

- Two quantization noises become independent.
- Most vectors are on the boundaries.
Binary One-help-one Problem

- Let \( X \sim \text{Be}(\frac{1}{2}), E \sim \text{Be}(\epsilon), Z \sim \text{Be}(p) \).
- Define \( Y_1 = X + E \) and \( Y_2 = (X, Z) \).

The decoder wants to reconstruct \( X + Z \) with distortion \( D \).

The first encoder is acting as a helper to the second encoder.

If \( \epsilon = 0 \) then the encoders have the same quantization noise.

If \( \epsilon \neq 0 \) the large blocklength quantizers lose correlation.
Binary One-help-one Problem

- Let $X \sim Be\left(\frac{1}{2}\right)$, $E \sim Be(\epsilon)$, $Z \sim Be(p)$.
- Define $Y_1 = X + E$ and $Y_2 = (X, Z)$.

Encoder Decoder

$X + E$ $X + Z$

$X, Z$

Encoder Decoder

The decoder wants to reconstruct $X + Z$ with distortion $D$. 

Farhad Shirani, S. Sandeep Pradhan
On the Sub-optimality of Single-letter Coding in Multi-terminal Communications
Binary One-help-one Problem

- Let $X \sim Be\left(\frac{1}{2}\right)$, $E \sim Be(\epsilon)$, $Z \sim Be(p)$.
- Define $Y_1 = X + E$ and $Y_2 = (X, Z)$.

The decoder wants to reconstruct $X + Z$ with distortion $D$.
- The first encoder is acting as a helper to the second encoder.
Let $X \sim Be(\frac{1}{2})$, $E \sim Be(\epsilon)$, $Z \sim Be(p)$. Define $Y_1 = X + E$ and $Y_2 = (X, Z)$.

The decoder wants to reconstruct $X + Z$ with distortion $D$. The first encoder is acting as a helper to the second encoder. If $\epsilon = 0$ then the encoders have the same quantization noise.
Let $X \sim B(e(\frac{1}{2}))$, $E \sim B(e)$, $Z \sim B(p)$.

Define $Y_1 = X + E$ and $Y_2 = (X, Z)$.

The decoder wants to reconstruct $X + Z$ with distortion $D$.

The first encoder is acting as a helper to the second encoder.

If $\epsilon = 0$ then the encoders have the same quantization noise.

If $\epsilon \neq 0$ the large blocklength quantizers loose correlation.
We propose a new scheme:

\[
X + E \rightarrow C_f^{(n)} \rightarrow V
\]

\[
Z \rightarrow S \rightarrow \hat{V}
\]

\[
X \rightarrow C_f^{(n)} \rightarrow \hat{V}
\]

\[
X + Z \rightarrow \hat{X} + Z
\]

\[
C(n) \rightarrow f \rightarrow S \hat{V} \rightarrow C(m) \rightarrow \hat{Q} \rightarrow \pi^{-1} \rightarrow Q
\]

\[
\text{Encoding}
\]

\[
\text{Decoding}
\]
Outline

1. Effective Length
   - Binary Block Encoders

2. Distributed Source Coding
   - Problem Statement
   - Binary One-help-one

3. Boolean Functions
   - Maximum Correlation

4. Single-letter Coding

F. Shirani Chaharsooghi, S. Pradhan, *On the Correlation between Functions of Sequences of Random Variables*, ISIT 2017
The agents receive two binary strings.
The agents receive two binary strings.

The sources are DMS’s.
The agents receive two binary strings.
The sources are DMS’s.
First agents makes a Boolean decision.
Problem Statement

The agents receive two binary strings.

The sources are DMS’s.

First agents makes a Boolean decision.

Second agent guesses the decision.
The agents receive two binary strings.
The sources are DMS’s.
First agents makes a Boolean decision.
Second agent guesses the decision.
Effective Length

Distributed Source Coding

Boolean Functions

Single-letter Coding

Maximal Correlation

\[ X_1, X_2, \ldots, X_n \]
\[ e(X^n) \in \{0, 1\} \]

\[ Y_1, Y_2, \ldots, Y_n \]
\[ f(Y^n) \in \{0, 1\} \]

- Best strategy: both output the first element.

\[ e(X^n) = X_1, \quad f(Y^n) = Y_1. \]
Maximum Correlation

Best strategy: both output the first element.

\[ e(X^n) = X_1, \quad f(Y^n) = Y_1. \]

Processing does not increase correlation.
Best strategy: both output the first element.

\[ e(X^n) = X_1, \quad f(Y^n) = Y_1. \]

Processing does not increase correlation.

In information theory efficient communication requires large effective-lengths.
Best strategy: both output the first element.

\[ e(X^n) = X_1, \quad f(Y^n) = Y_1. \]

- Processing does not increase correlation.
- In information theory efficient communication requires large effective-lengths.
Additive Boolean Functions

For $e(X^n) = X_{i_1} \oplus_2 X_{i_2} \oplus_2 \cdots \oplus_2 X_{i_s}$, the effective length is $s$. 
Additive Boolean Functions

For $e(X^n) = X_{i_1} \oplus_2 X_{i_2} \oplus_2 \cdots X_{i_s}$, the effective length is $s$.

- Non-additive functions: define the dependency spectrum.
Additive Boolean Functions

For $e(X^n) = X_{i1} \oplus X_{i2} \oplus \cdots X_{is}$, the effective length is $s$.

1. Non-additive functions: define the dependency spectrum.
   - Map the Boolean function to real functions.

$$\Phi : e \to \tilde{e}.$$
Additive Boolean Functions

For $e(X^n) = X_{i_1} \oplus_2 X_{i_2} \oplus_2 \cdots X_{i_s}$, the effective length is $s$.

- Non-additive functions: define the dependency spectrum.
  1. Map the Boolean function to real functions.
     \[ \Phi : e \to \tilde{e}. \]
  2. Decompose the real function into additive components.
     \[ \tilde{e} = \sum_{i \in \{0,1\}^n} \tilde{e}_i. \]
Additive Boolean Functions

For \( e(X^n) = X_{i_1} \oplus_2 X_{i_2} \oplus_2 \cdots X_{i_s} \), the effective length is \( s \).

1. Non-additive functions: define the dependency spectrum.
   - Map the Boolean function to real functions.
     \[ \Phi : e \to \tilde{e}. \]

2. Decompose the real function into additive components.
   \[ \tilde{e} = \sum_{i \in \{0,1\}^n} \tilde{e}_i. \]

3. Dependency spectrum is the variance of these components.
   \[ (Var(\tilde{e}_i))_{i \in \{0,1\}^n}. \]
Additive Boolean Functions

For \( e(X^n) = X_{i_1} \oplus_2 X_{i_2} \oplus_2 \cdots X_{i_s} \), the effective length is \( s \).

- Non-additive functions: define the dependency spectrum.
  1. Map the Boolean function to real functions.
     \[ \Phi : e \rightarrow \tilde{e}. \]
  2. Decompose the real function into additive components.
     \[ \tilde{e} = \sum_{i \in \{0, 1\}^n} \tilde{e}_i. \]

  3. Dependency spectrum is the variance of these components.
     \[ (Var(\tilde{e}_i))_{i \in \{0, 1\}^n}. \]
We find a decomposition of $\tilde{e}$ into independent functions with fixed length.
We find a decomposition of $\tilde{e}$ into independent functions with fixed length.

Let $n=3$:

$$\tilde{e}(X_1, X_2, X_3) = \tilde{e}_{000} + \tilde{e}_{001} + \tilde{e}_{010} + \tilde{e}_{011} + \tilde{e}_{100} + \tilde{e}_{101} + \tilde{e}_{110} + \tilde{e}_{111}.$$
We find a decomposition of \( \tilde{e} \) into independent functions with fixed length.

Let \( n=3 \):

\[
\tilde{e}(X_1, X_2, X_3) = \tilde{e}_{000} + \tilde{e}_{001} + \tilde{e}_{010} + \tilde{e}_{011} + \tilde{e}_{100} + \tilde{e}_{101} + \tilde{e}_{110} + \tilde{e}_{111}.
\]

\( \tilde{e}_{111} \) corresponds to effective length 3.
We find a decomposition of $\tilde{e}$ into independent functions with fixed length.

Let $n=3$:

$$\tilde{e}(X_1, X_2, X_3) = \tilde{e}_{000} + \tilde{e}_{001} + \tilde{e}_{010} + \tilde{e}_{011} + \tilde{e}_{100} + \tilde{e}_{101} + \tilde{e}_{110} + \tilde{e}_{111}.$$  

- $\tilde{e}_{111}$ corresponds to effective length 3.
- $\tilde{e}_{011}, \tilde{e}_{101}, \tilde{e}_{110}$ correspond to effective length 2.
We find a decomposition of $\tilde{e}$ into independent functions with fixed length.

Let $n=3$:

$$\tilde{e}(X_1, X_2, X_3) = \tilde{e}_{000} + \tilde{e}_{001} + \tilde{e}_{010} + \tilde{e}_{011} + \tilde{e}_{100} + \tilde{e}_{101} + \tilde{e}_{110} + \tilde{e}_{111}.$$

- $\tilde{e}_{111}$ corresponds to effective length 3.
- $\tilde{e}_{011}, \tilde{e}_{101}, \tilde{e}_{110}$ correspond to effective length 2.
Example:
- Let $n = 2$, and $X_1, X_2$ independent BSS's.
Example:

Let \( n = 2 \), and \( X_1, X_2 \) independent BSS's.

Let \( e(X_1, X_2) = X_1 \land X_2 \), then:

\[
\tilde{e}(X_1, X_2) = \begin{cases} 
-\frac{1}{4} & (X_1, X_2) \neq (1, 1), \\
\frac{3}{4} & (X_1, X_2) = (1, 1).
\end{cases}
\]
Example:
- Let $n = 2$, and $X_1, X_2$ independent BSS's.
- Let $e(X_1, X_2) = X_1 \land X_2$, then:

$$
\tilde{e}(X_1, X_2) = \begin{cases} 
-\frac{1}{4} & (X_1, X_2) \neq (1, 1), \\
\frac{3}{4} & (X_1, X_2) = (1, 1).
\end{cases}
$$

- Lagrange interpolation gives $\tilde{e} = X_1 X_2 - \frac{1}{4}$. 
Example:

- Let $n = 2$, and $X_1, X_2$ independent BSS’s.
- Let $e(X_1, X_2) = X_1 \land X_2$, then:

$$
\tilde{e}(X_1, X_2) = \begin{cases} 
-\frac{1}{4} & (X_1, X_2) \neq (1, 1), \\
\frac{3}{4} & (X_1, X_2) = (1, 1).
\end{cases}
$$

- Lagrange interpolation gives $\tilde{e} = X_1 X_2 - \frac{1}{4}$.
- The decomposition is given by:

$$
\tilde{e}_{1,1} = (X_1 - \frac{1}{2})(X_2 - \frac{1}{2}), \quad \tilde{e}_{1,0} = \frac{1}{2}(X_1 - \frac{1}{2}),
$$
$$
\tilde{e}_{0,1} = \frac{1}{2}(X_2 - \frac{1}{2}), \quad \tilde{e}_{0,0} = 0.
$$
Example:

- Let $n = 2$, and $X_1, X_2$ independent BSS’s.
- Let $e(X_1, X_2) = X_1 \land X_2$, then:

$$
\tilde{e}(X_1, X_2) = \begin{cases} 
-\frac{1}{4} & (X_1, X_2) \neq (1, 1), \\
\frac{3}{4} & (X_1, X_2) = (1, 1).
\end{cases}
$$

- Lagrange interpolation gives $\tilde{e} = X_1X_2 - \frac{1}{4}$.
- The decomposition is given by:

$$
\tilde{e}_{1,1} = (X_1 - \frac{1}{2})(X_2 - \frac{1}{2}), \quad \tilde{e}_{1,0} = \frac{1}{2}(X_1 - \frac{1}{2}),
$$

$$
\tilde{e}_{0,1} = \frac{1}{2}(X_2 - \frac{1}{2}), \quad \tilde{e}_{0,0} = 0.
$$

- The variances of these functions are given below:

$$
Var(\tilde{e}) = \frac{3}{16}, \quad Var(\tilde{e}_{0,1}) = Var(\tilde{e}_{1,0}) = Var(\tilde{e}_{1,1}) = \frac{1}{16}.
$$
Example:

- Let \( n = 2 \), and \( X_1, X_2 \) independent BSS’s.
- Let \( e(X_1, X_2) = X_1 \land X_2 \), then:

\[
\tilde{e}(X_1, X_2) = \begin{cases} 
\frac{-1}{4} & (X_1, X_2) \neq (1, 1), \\
\frac{3}{4} & (X_1, X_2) = (1, 1).
\end{cases}
\]

Lagrange interpolation gives \( \tilde{e} = X_1 X_2 - \frac{1}{4} \).

The decomposition is given by:

\[
\tilde{e}_{1,1} = (X_1 - \frac{1}{2})(X_2 - \frac{1}{2}), \quad \tilde{e}_{1,0} = \frac{1}{2}(X_1 - \frac{1}{2}),
\]

\[
\tilde{e}_{0,1} = \frac{1}{2}(X_2 - \frac{1}{2}), \quad \tilde{e}_{0,0} = 0.
\]

The variances of these functions are given below:

\[
Var(\tilde{e}) = \frac{3}{16}, \quad Var(\tilde{e}_{0,1}) = Var(\tilde{e}_{1,0}) = Var(\tilde{e}_{1,1}) = \frac{1}{16}.
\]
Let $P_i$ be the variance of $\tilde{e}_i$. 
Let $P_i$ be the variance of $\tilde{e}_i$.

We lower bound the maximal correlation:

**Theorem**

Let $\epsilon \triangleq P(X \neq Y)$, the following bound holds:

$$2\sqrt{\sum_i P_i} \sqrt{\sum_i Q_i} - 2 \sum_i C_i P_i^{1/2} Q_i^{1/2} \leq P(e(X^n) \neq f(Y^n)),$$

where $C_i \triangleq (1 - 2\epsilon)^{w_H(i)}$. 

Farhad Shirani, S. Sandeep Pradhan
Advisor: Sandeep Pradhan
University of Michigan

On the Sub-optimality of Single-letter Coding in Multi-terminal Communications
Let $P_i$ be the variance of $\tilde{e}_i$.

We lower bound the maximal correlation:

**Theorem**

Let $\epsilon \triangleq P(X \neq Y)$, the following bound holds:

$$2 \sqrt{\sum_i P_i} \sqrt{\sum_i Q_i} - 2 \sum_i C_i P_i^{\frac{1}{2}} Q_i^{\frac{1}{2}} \leq P(e(X^n) \neq f(Y^n)),$$

where $C_i \triangleq (1 - 2\epsilon)^{w_H(i)}$.

This shows that correlation falls with effective length:

![Plot of $C_i$ as a function of $N_i$, $\epsilon = 10^{-3}$](image_url)
Special Cases

- Assume the entropy constraint is $H(e) = H(f) = 1$. 
Special Cases

- Assume the entropy constraint is $H(e) = H(f) = 1$.
- The outputs should be symmetric:

$$\frac{1}{2} - 2 \sum_i C_i P_i^{1/2} Q_i^{1/2} \leq P(e(X^n) \neq f(Y^n))$$
Special Cases

- Assume the entropy constraint is $H(e) = H(f) = 1$.
- The outputs should be symmetric:

$$\frac{1}{2} - 2 \sum_i C_i P_i^{1/2} Q_i^{1/2} \leq P(e(X^n) \neq f(Y^n)).$$

- **Case 1:** All of the variance is on large block components

  $\rightarrow \frac{1}{2} \leq P(e(X^n) \neq f(Y^n)).$
Special Cases

- Assume the entropy constraint is $H(e) = H(f) = 1$.
- The outputs should be symmetric:

$$\frac{1}{2} - 2 \sum_{i} C_i P_i^{\frac{1}{2}} Q_i^{\frac{1}{2}} \leq P(e(X^n) \neq f(Y^n)).$$

- **Case 1:** All of the variance is on large block components
  \[ \rightarrow \frac{1}{2} \leq P(e(X^n) \neq f(Y^n)). \]
- **Case 2:** All of the variance on single-letter component
  \[ \rightarrow \epsilon \leq P(e(X^n) \neq f(Y^n)) = P(X_1 \neq Y_1). \]
Outline

1. Effective Length
   - Binary Block Encoders

2. Distributed Source Coding
   - Problem Statement
   - Binary One-help-one

3. Boolean Functions
   - Maximum Correlation

4. Single-letter Coding

F. Shirani Chaharsooghi, S. Pradhan, *On the Sub-optimality of Single-letter Coding for Multi-terminal Communications*, ISIT 2017
Hypothesis: SLC schemes are sub-optimal for multi-terminal communications.
**Hypothesis:** SLC schemes are sub-optimal for multi-terminal communications.

**Coding scheme:** A probability distribution $P_E(e)P_{D|E}(d|e)$. 

An SLC satisfies three properties:

1. Codewords are chosen pairwise independent.
2. As $n \to \infty$, the output distribution approaches a product distribution.
3. The coding scheme is not sensitive to permutations: $∀\pi ∈ S_n : P(E) = P(E_\pi)$, where $E_\pi(X_n) = \pi^{-1}(E(\pi(X_n)))$. 

The third assumption is true since in typicality encoding we have: $y_n ∈ A_n \leftrightarrow \pi(y_n) ∈ A_n \epsilon(X/x_n)$. 

Farhad Shirani, S. Sandeep Pradhan

On the Sub-optimality of Single-letter Coding in Multi-terminal Communications
- **Hypothesis**: SLC schemes are sub-optimal for multi-terminal communications.
- **Coding scheme**: A probability distribution \( P_E(e)P_D|_E(d|e) \).
- An SLC satisfies three properties:
Hypothesis: SLC schemes are sub-optimal for multi-terminal communications.

Coding scheme: A probability distribution $P_E(e)P_{D|E}(d|e)$.

An SLC satisfies three properties:

1. Codewords are chosen pairwise independent.
Hypothesis: SLC schemes are sub-optimal for multi-terminal communications.

Coding scheme: A probability distribution $P_E(e)P_{D|E}(d|e)$.

An SLC satisfies three properties:

1. Codewords are chosen pairwise independent.
2. As $n \to \infty$, the output distribution approaches a product distribution.
Hypothesis: SLC schemes are sub-optimal for multi-terminal communications.

Coding scheme: A probability distribution $P_E(e)P_{D|E}(d|e)$.

An SLC satisfies three properties:

1. Codewords are chosen pairwise independent.
2. As $n \to \infty$, the output distribution approaches a product distribution.
3. The coding scheme is not sensitive to permutations:
   \[ \forall \pi \in S_n : P(E) = P(E_{\pi}), \text{ where } E_{\pi}(X^n) = \pi^{-1}(E(\pi(X^n))). \]
Hypothesis: SLC schemes are sub-optimal for multi-terminal communications.

Coding scheme: A probability distribution $P_E(e)P_D|_E(d|e)$.

An SLC satisfies three properties:

1. Codewords are chosen pairwise independent.
2. As $n \to \infty$, the output distribution approaches a product distribution.
3. The coding scheme is not sensitive to permutations:
   \[ \forall \pi \in S_n : P(E) = P(E_{\pi}), \text{ where } E_{\pi}(X^n) = \pi^{-1}(E(\pi(X^n))). \]

The third assumption is true since in typicality encoding we have:

\[ y^n \in A^n_\epsilon(X|x^n) \leftrightarrow \pi(y^n) \in A^n_\epsilon(X|\pi(x^n)). \]
We prove that single-letter schemes have two components in the effective length:

1. A single-letter component.
2. An $n$-letter component, $n \to \infty$.
We prove that single-letter schemes have two components in the effective length:

1. A single-letter component.
2. An $n$-letter component, $n \to \infty$.

Theorem

For any $k \in [1, n], m \in \mathbb{N}, \gamma > 0$, $P_S(\sum_{i: N_i \leq m, i \neq i_k} P_{k,i} \geq \gamma) \to 0$, as $n \to \infty$. Where, $i_k$ is the $k$th standard basis element.

The only non-zero finite-letter element is the single-letter element in the decomposition.
We prove that single-letter schemes have two components in the effective length:

1. A single-letter component.
2. An $n$-letter component, $n \to \infty$.

**Theorem**

For any $k \in [1, n], m \in \mathbb{N}, \gamma > 0$, $P_S(\sum_{i:N_i \leq m, i \neq i_k} P_{k,i} \geq \gamma) \to 0$, as $n \to \infty$. Where, $i_k$ is the $k$th standard basis element.

- The only non-zero finite-letter element is the single-letter element in the decomposition.
- This gives an upper-bound on correlation for SLC’s.
We prove that single-letter schemes have two components in the effective length:

1. A single-letter component.
2. An \( n \)-letter component, \( n \to \infty \).

**Theorem**

For any \( k \in [1, n], m \in \mathbb{N}, \gamma > 0 \),

\[
P_S\left( \sum_{i \in \mathbb{N}, i \neq i_k} \mathbb{P} \left( P_{k, i} \geq \gamma \right) \right) \to 0, \quad \text{as} \quad n \to \infty.
\]

Where, \( i_k \) is the \( k \)th standard basis element.

The only non-zero finite-letter element is the single-letter element in the decomposition.

This gives an upper-bound on correlation for SLC’s.

We use this bound to prove sub-optimality in various multi-terminal settings.
Thanks!


Farhad Shirani, S. Sandeep Pradhan

On the Sub-optimality of Single-letter Coding in Multi-terminal Communications