A Computational Approach for Determining Rate Regions and Codes using Entropic Vector Bounds

Congduan Li, John MacLaren Walsh, Steven Weber

Drexel University, Dept. of ECE
Philadelphia, PA 19104
congduan.li@drexel.edu

Allerton 2012, UIUC
Oct. 4, 2012

Support by NSF (CCF-1016588) gratefully acknowledged
Motivation

- For which classes of networks are simple codes optimal? (simple codes achieve the entire capacity (coding rate) region)
- Simplest codes: Binary linear codes
- Simple networks: Multilevel Diversity Coding System (MDCS)
Outline

1. Coding rate region for MDCS
   1.1 MDCS configurations and coding rate region

2. Region of entropic vector and its bounds
   2.1 Entropic vector region
   2.2 Shannon outer bound
   2.3 Binary matroid inner bound

3. Calculation of coding rate region and experiments on MDCS
   3.1 General coding rate region expression
   3.2 Calculation flowchart
   3.3 Experiments on 2-level-3-encoder MDCS
Outline

1. Coding rate region for MDCS
   1.1 MDCS configurations

2. Region of entropic vector and its bounds
   2.1 Entropic vector region
   2.2 Shannon outer bound
   2.3 Binary matroid inner bound

3. Calculation of coding rate region and experiments on MDCS
   3.1 General coding rate region expression
   3.2 Calculation flowchart
   3.3 Experiments on 2-level-3-encoder MDCS
Multilevel Diversity Coding System (MDCS): \( K \) independence sources, encoder has access to all sources, decoders are classified using levels

- \( m \)-level decoder can reconstruct \( Y_1, Y_2, \ldots, Y_m \).

\[
\mathbf{R} = \{ R_e, e \in \mathcal{E} \} \\
\mathcal{R} = \{ \mathbf{R} : \mathbf{R} \text{ is admissible} \}
\]
MDCS Example

- 2-level ($K=2$) 3-encoder: 31 non-isomorphic cases in total
- One example:

<table>
<thead>
<tr>
<th>level 1</th>
<th>level 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(1),(2,3)$</td>
<td>$(1,2), (1,3)$</td>
</tr>
</tbody>
</table>

$\mathcal{R} = \{R_1, R_2, R_3\}$

$\mathcal{R} = \{R : R \text{ is admissible}\}$
Outline

1. Coding rate region for MDCS
   1.1 MDCS configurations

2. Region of entropic vector and its bounds
   2.1 Entropic vector region
   2.2 Shannon outer bound
   2.3 Binary matroid inner bound

3. Calculation of coding rate region and experiments on MDCS
   3.1 General coding rate region expression
   3.2 Calculation flowchart
   3.3 Experiments on 2-level-3-encoder MDCS
Entropic Vector

- A collection \( \mathbf{X} \) of \( N \) discrete variables \( X_1, \ldots, X_N \), (joint) entropies of all non-empty subsets \( A \) of \( \mathbf{X} \) form a \( 2^N - 1 \)-dimensional vector \( h = (H(A), \forall A \subseteq \mathbf{X} \setminus \emptyset) \).

- A vector \( h \) is said to be entropic if there exists a joint distribution associated with \( h \).

- Region of entropic vector: \( \Gamma^*_N = \{ h : h \text{ is entropic} \} \).

- \( N = 2 \) example

- Not all points in Euclidean space are entropic.
Outline

1. Coding rate region for MDCS
   1.1 MDCS configurations

2. Region of entropic vector and its bounds
   2.1 Entropic vector region
   2.2 Shannon outer bound
   2.3 Binary matroid inner bound

3. Calculation of coding rate region and experiments on MDCS
   3.1 General coding rate region expression
   3.2 Calculation flowchart
   3.3 Experiments on 2-level-3-encoder MDCS
Entropic Vector Region
For $N \geq 4$

- No computable characterization, but know it is a convex cone with curved faces (?)
- Bounds to approximate it
- Shannon outer bound $\Gamma_N$ (?)
- All Shannon type entropy inequalities can be converted to half-space constraints $I(X_A; X_B|X_C) \geq 0, \forall A, B, C \subseteq X$.

- These basic inequalities form Shannon Outer Bound on entropic vector region
Outline

1. Coding rate region for MDCS
   1.1 MDCS configurations and coding rate region

2. Region of entropic vector and its bounds
   2.1 Entropic vector region
   2.2 Shannon outer bound
   2.3 Binary matroid inner bound

3. Calculation of coding rate region and experiments on MDCS
   3.1 General coding rate region expression
   3.2 Calculation flowchart
   3.3 Experiments on 2-level-3-encoder MDCS
Suppose $A$ is a $M \times N$ ($M \leq N$) binary matrix associated with a binary representable matroid. Collect column indices as a set $\{1, \ldots, N\}$.

Ranks of vectors associated with all subsets of the index set form a $2^N - 1$ dimensional rank vector $r = (r_1, r_2, r_{12}, \ldots r_{12, \ldots, N})$.

This vector is entropic if we define binary independent uniform random variables $U_1, \ldots, U_M$ and then define $(X_1, \ldots, X_N) = (U_1, \ldots, U_M)A$. $X_1, \ldots, X_N$ has entropy vector $h$ same as rank vector $r$.

Varying entries, enumerate all possible such rank vectors.

The conic hull of these vectors gives binary matroid inner bound of entropic vector region for $N$ variables.
Binary Matroid Inner Bound of Entropic Vector Region
Shannon Outer and Binary Inner Bound

- Binary Matroid Inner Bound for $N$ variables $\Gamma_{\text{Bin}}^N$.
- For large $N$, huge computation work to get $H$-representation.
Binary Matroid Inner Bound of Entropic Vector Region

Binary Matroid Inner Bound for Large $N$

- (Tutte 1950's) A matroid is binary representable iff it has no $U_{2,4}$ as a minor:

\[
\begin{bmatrix}
1 & 0 & 1 & ? \\
0 & 1 & 1 & ? \\
\end{bmatrix}
\]

- (?) A rank $r_N \in \mathbb{R}^{2^N-1}_+$ lies in $\Gamma_{\text{Bin}}^{N}$ iff
  - $r_N \in \Gamma_N$
  - for every size 4 subset of indices of $\{1, \ldots, N\}$, say $\{i, j, k, l\}$, for every $A \subseteq \{1, \ldots, N\}\{i, j, k, l\}$, $r_4(B) := r_N(A \cup B) - r_N(A)$, $B \subseteq \{i, j, k, l\}$ is not the same as rank of $U_{2,4}$. 

Outline

1. Coding rate region for MDCS
   1.1 MDCS configurations
   1.2 Rate region for MDCS

2. Region of entropic vector and its bounds
   2.1 Entropic vector region
   2.2 Shannon outer bound
   2.3 Binary matroid inner bound

3. Calculation of coding rate region and experiments on MDCS
   3.1 General coding rate region expression
   3.2 Calculation flowchart
   3.3 Experiments on 2-level-3-encoder MDCS
Coding Rate Region for MDCS

Suppose we know the entropic vector rate region $\Gamma^*_N$. General network coding rate region is expressible using $\Gamma^*_N (\cdot)$. For MDCS, coding rate region is:

Constraints: (Hyperplanes)

$$
\mathcal{L}_1 = \{h \in \mathcal{H}_N : h_{Y_s} \geq \omega_s\} \\
\mathcal{L}_2 = \{h \in \mathcal{H}_N : h_{YS} = \sum_{s \in S} h_{Y_s}\} \\
\mathcal{L}_3 = \{h \in \mathcal{H}_N : h_{X_{Out(k)}}|_{Y_s = 0}\} \\
\mathcal{L}_4 = \{h \in \mathcal{H}_N : h_{Y_{\beta(t)}}|_{U_{In(t)} = 0}\}
$$

$$
\mathcal{R} = \text{Ex}(\text{proj}_{U_e}(\text{con}(\Gamma^*_n \cap \mathcal{L}_{123}) \cap \mathcal{L}_4))
$$
Outline

1. Coding rate region for MDCS
   1.1 MDCS configurations and coding rate region

2. Region of entropic vector and its bounds
   2.1 Entropic vector region
   2.2 Shannon outer bound
   2.3 Binary matroid inner bound

3. Calculation of coding rate region and experiments on MDCS
   3.1 General coding rate region expression
   3.2 Calculation flowchart
   3.3 Experiments on 2-level-3-encoder MDCS
Computational Steps

Replace $\Gamma_N^*$ by $\Gamma_N^{Bin}$

Insert Constraints

Replace $\Gamma_N^*$ by $\Gamma_N$

Add Rate variables

$R_e \geq H(U_e)$

$R_{in}$

Project on rate and source

$R_{out}$

$R_{in} = R = R_{out}$

$R_{in} \subseteq R \subseteq R_{out}$

Yes

Binary Codes Suffice

No

Binary Codes NOT Suffice
Computational Steps Demo

Constraints

No Gap

Gap

Projection
Outline

1. Coding rate region for MDCS
   1.1 MDCS configurations and coding rate region

2. Region of entropic vector and its bounds
   2.1 Entropic vector region
   2.2 Shannon outer bound
   2.3 Binary matroid inner bound

3. Calculation of coding rate region and experiments on MDCS
   3.1 General coding rate region expression
   3.2 Calculation flowchart
   3.3 Experiments on 2-level-3-encoder MDCS
Experiments

- We have developed software to compute inner and outer bounds of rate regions automatically for all 31 cases.
- 2-level ($K=2$) 3-encoder: 6 out of 31 non-isomorphic cases where binary inner bound do not match with Shannon outer bound in general.
Example: Two Bounds Not Match in General

Configuration

<table>
<thead>
<tr>
<th>level 1</th>
<th>level 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>{(1)}</td>
<td>{(1, 2), (1, 3), (2, 3)}</td>
</tr>
</tbody>
</table>

\[
\begin{align*}
  \{X, Y\} & \quad U_1 \quad R_1 & \quad \{X\} \\
  E_1 & \quad & D_1 \\
  E_2 & \quad \quad U_2 \quad R_2 & \quad \{X, Y\} \\
  E_3 & \quad \quad U_3 \quad R_3 & \quad \{X, Y\} \\
  & \quad \quad \quad \quad \quad \quad & \quad \{X, Y\}
\end{align*}
\]
Example: Two Bounds Not Match in General

Outer bound and inner bound

\( \mathcal{R}_{\text{out}} : \)

\[
\begin{align*}
R_1 & \geq H(X) \\
R_2 + R_3 & \geq H(X) + H(Y) \\
R_1 + R_3 & \geq H(X) + H(Y) \\
R_1 + R_2 & \geq H(X) + H(Y) \\
R_1 + R_2 + R_3 & \geq H(X) + 2 \cdot H(Y)
\end{align*}
\]

\( \mathcal{R}_{\text{in}} : \)

\[
\begin{align*}
R_1 & \geq H(X) \\
R_2 + R_3 & \geq H(X) + H(Y) \\
R_1 + R_3 & \geq H(X) + H(Y) \\
R_1 + R_2 & \geq H(X) + H(Y) \\
R_1 + R_2 + R_3 & \geq H(X) + 2 \cdot H(Y)
\end{align*}
\]
Example: Two Bounds Not Match in General

Binary scalar linear codes: when outer bound and inner bound match

\[ R_1 \geq H(X) \]
\[ R_2 + R_3 \geq H(X) + H(Y) \]
\[ R_1 + R_3 \geq H(X) + H(Y) \]
\[ R_1 + R_2 \geq H(X) + H(Y) \]
\[ (R_1 + R_2 + R_3 \geq H(X) + 2 \cdot H(Y)) \]

Suppose \( H(X) = H(Y) = 1 \)

\[
\begin{array}{ccc}
U_1 & U_2 & U_3 \\
1 & 0 & 1 \\
0 & 1 & 1 \\
\end{array}
\]

\( \mathcal{R}_{\text{out}} = \mathcal{R}_{\text{in}} \)
Example: Two Bounds Not Match in General
Non-admissible extreme points of $\mathcal{R}_{\text{out}}$: when outer and inner bound not match

Suppose $H(X) = 1; H(Y) = 2$
Source Vector: $[X, Y^1, Y^2]$

$\frac{3}{2}, \frac{3}{2}, \frac{3}{2}$

NO Binary Linear Scalar Solution!

$R_1 \geq H(X)$
$R_2 + R_3 \geq H(X) + H(Y)$
$R_1 + R_3 \geq H(X) + H(Y)$
$R_1 + R_2 \geq H(X) + H(Y)$
Example: Two Bounds Not Match in General

Binary scalar linear codes: when outer bound and inner bound not match

Suppose $H(X) = 1; H(Y) = 2$

Source Vector: $[X, Y^1, Y^2]$

$R_1 \geq H(X)$

$R_2 + R_3 \geq H(X) + H(Y)$

$R_1 + R_3 \geq H(X) + H(Y)$

$R_1 + R_2 \geq H(X) + H(Y)$

$R_1 + R_2 + R_3 \geq H(X) + 2 \cdot H(Y)$
When $R_{\text{out}} = R_{\text{in}}$ and when not in general?

For 2-level-3-encoder MDCS

- A necessary condition (for a gap): decoders having access to encoders $(1,2),(1,3),(2,3)$ appear as either level 1 or level 2 decoders.
- A sufficient condition (for a gap): decoders having access to encoders $(1,2),(1,3),(2,3)$ appear as all level 1 or all level 2 decoders.
- The gap (exception): one of the two cases where superposition coding is not optimal.
Conclusion

1. Computational approach to determine coding rate region, together with field size and linear codes to achieve it for MDCS, using entropic vector region and its bounds
2. General software extendable to network coding problem in multiple multicast networks
3. Ongoing work: examine more networks, and try to characterize the networks for which binary codes suffice.
References
Thank you!

Thank you! Any questions?