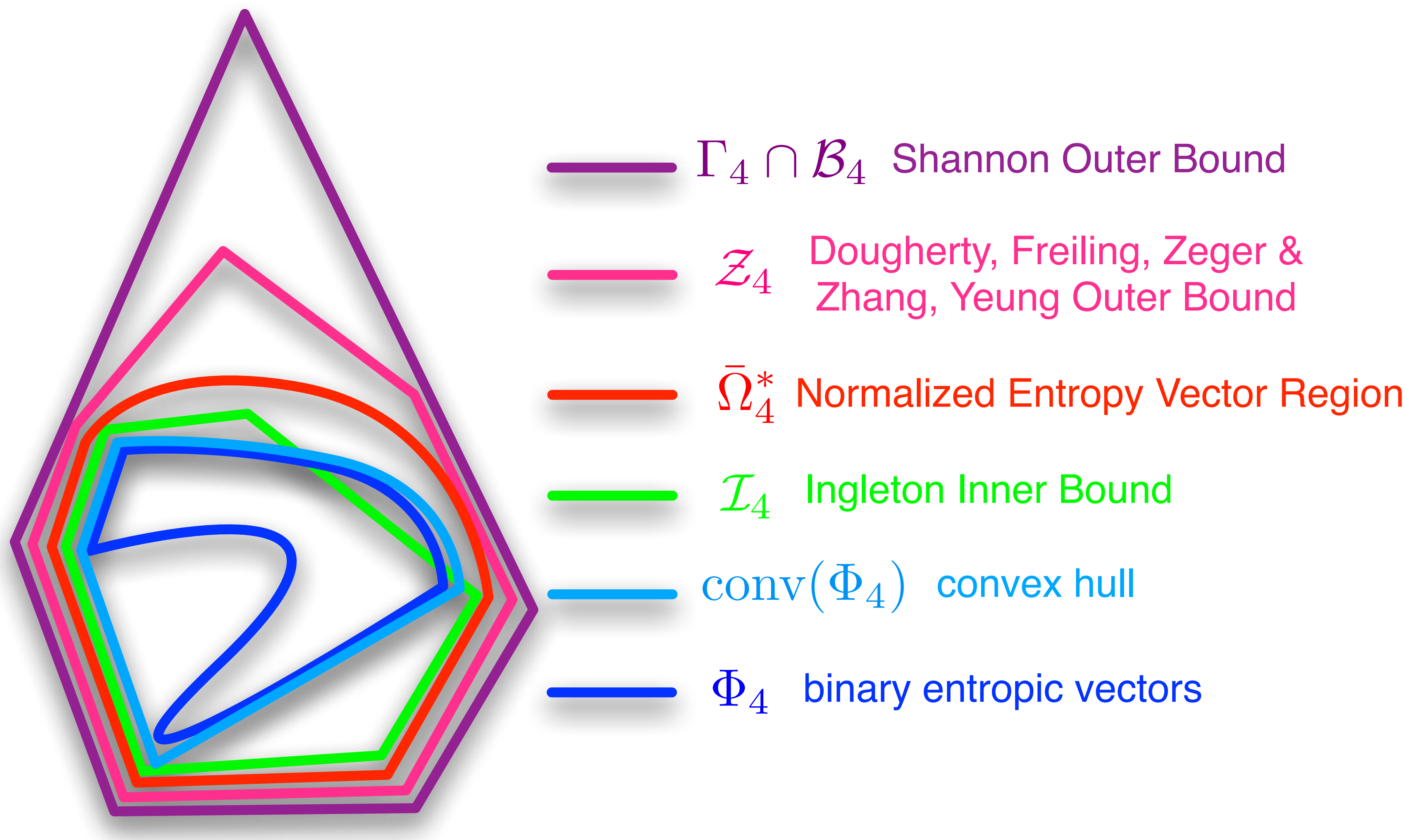
Tunable Inner Bounds for the Region of Entropy Vectors

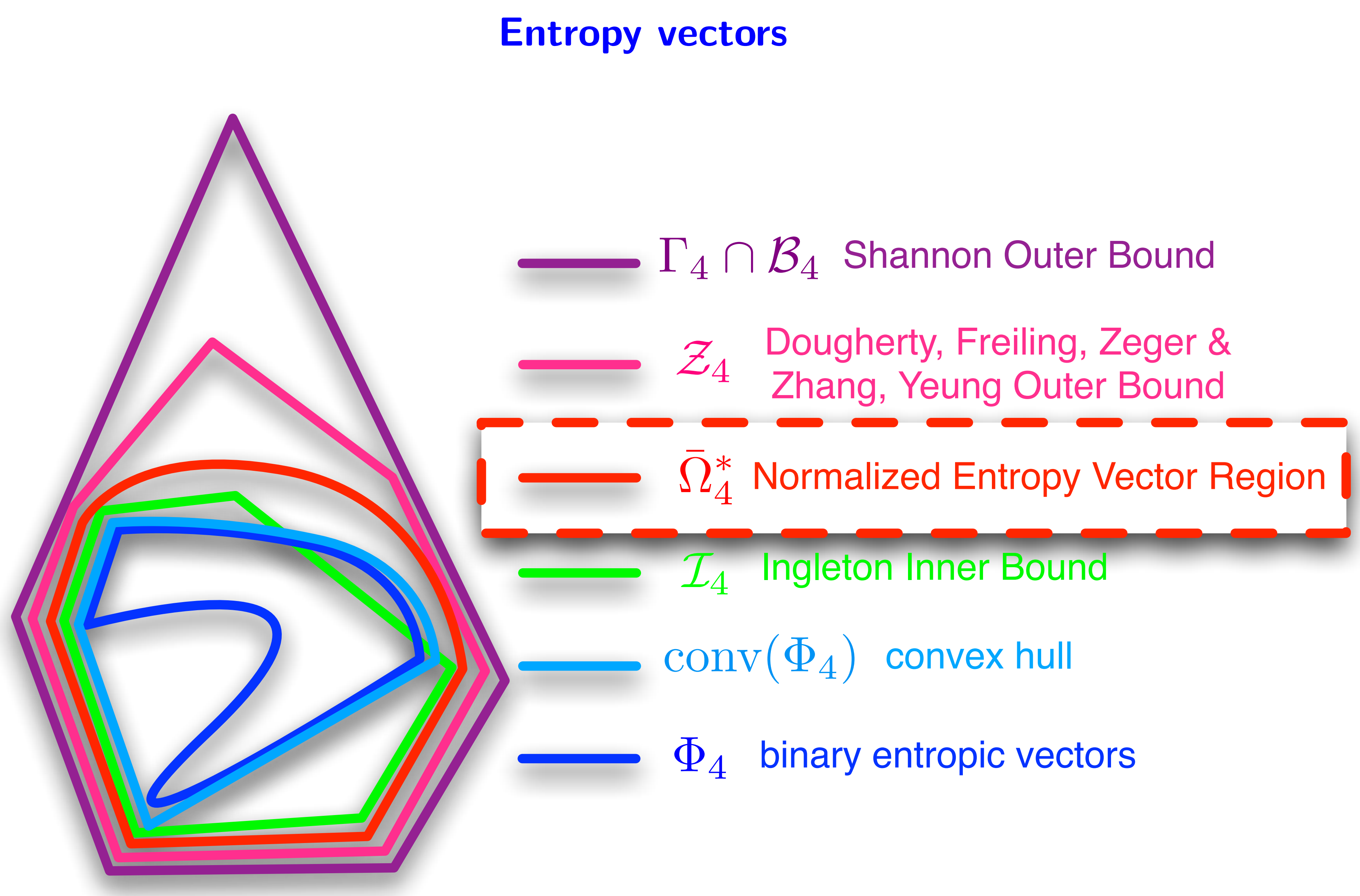
John M. Walsh, & Steven P. Weber
Department of Electrical and Computer Engineering
Drexel University
Philadelphia, PA
{jwalsh,sweber}@ece.drexel.edu

Outline

- What are entropy vectors, $\bar{\Gamma}_N^*, \bar{\Omega}_N^*$?
- Why are they important?
- Introducing the Set Φ_N of Binary Entropy Vectors
- Algorithm for Checking Membership in Φ_N
- Inner Bound for $\bar{\Omega}_N^*$ tuned to *any* Polyhedral Outer Bound
- Reinterpreting the Shannon, Ingleton, & DFZ,ZY bounds in this light.
- Extensions and Future Work

Outline


$\Gamma_4 \cap \mathcal{B}_4$ Shannon Outer Bound
$\mathcal{Z}_4$ Dougherty, Freiling, Zeger & Zhang, Yeung Outer Bound
$\bar{\Omega}_4^*$ Normalized Entropy Vector Region
$\mathcal{I}_4$ Ingleton Inner Bound
$\mathrm{conv}(\Phi_4)$ convex hull
Φ_4 binary entropic vectors


Entropy vectors
$\Gamma_4 \cap \mathcal{B}_4$ Shannon Outer Bound
$\mathcal{Z}_4$ Dougherty, Freiling, Zeger & Zhang, Yeung Outer Bound
$\bar{\Omega}_4^*$ Normalized Entropy Vector Region
$\mathcal{I}_4$ Ingleton Inner Bound
$\mathrm{conv}(\Phi_4)$ convex hull
Φ_4 binary entropic vectors

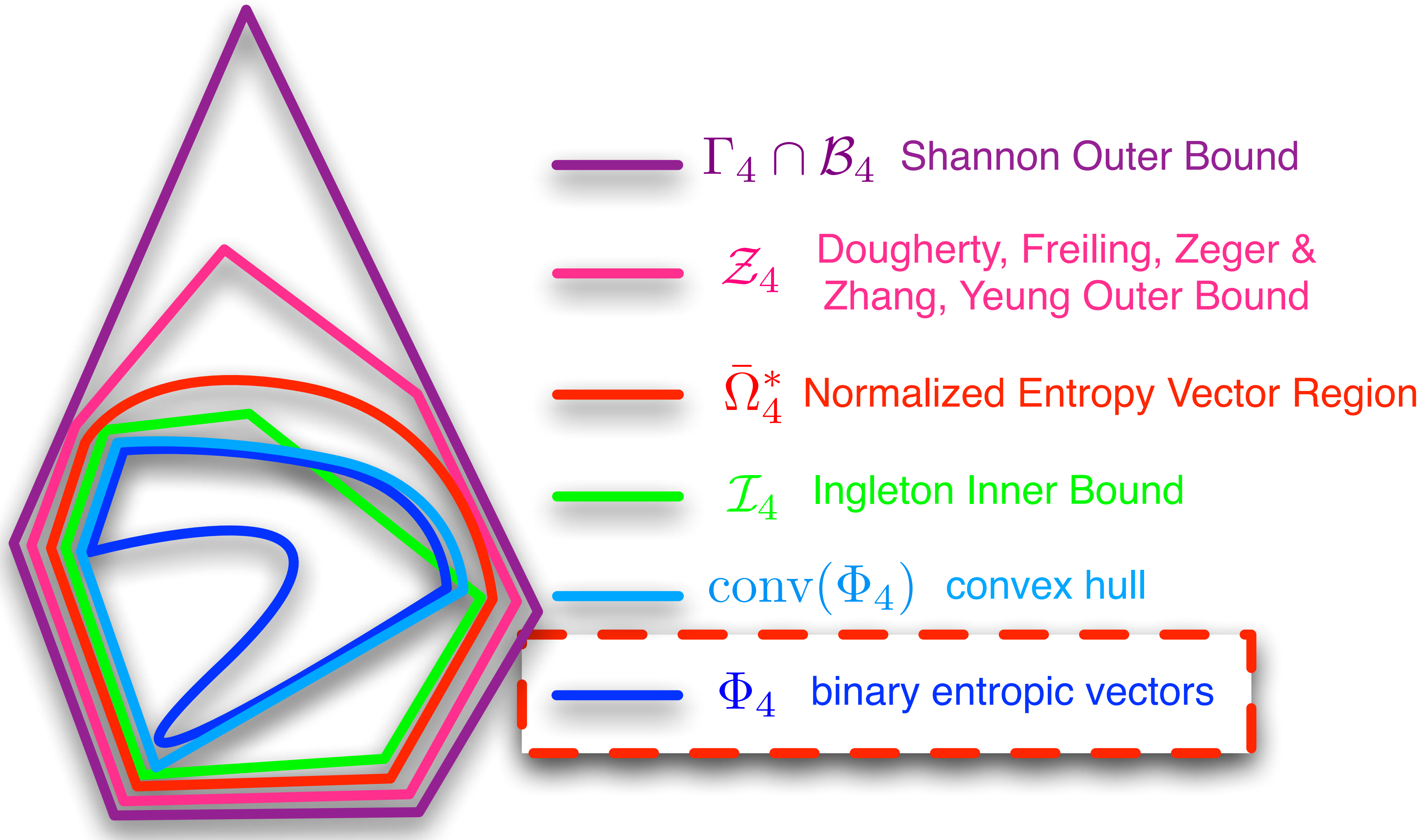
Entropy vectors

1. Let $\mathbf{X} = (X_1, \ldots, X_N)$ be N discrete random variables with finite support.
2. Let $h(\mathbf{X}_\mathcal{A})$ be the entropy of the subset of rvs $\mathbf{X}_\mathcal{A} = (X_i, i \in \mathcal{A})$ for some non-empty subset $\mathcal{A} \subseteq \{1, \ldots, N\} \equiv [N]$.
3. Let $\mathbf{h} = (h(\mathbf{X}_\mathcal{A}),\ \mathcal{A} \subseteq [N])$ be the vector of entropies of each non-empty subset $\mathcal{A} \subseteq [N]$. Note $\mathbf{h}$ has $2^N - 1$ entries.
   - Example: for $N = 3$, $\mathbf{h} = (h_1, h_2, h_3, h_{12}, h_{13}, h_{23}, h_{123})$.
4. A vector $\mathbf{h} \in \mathbb{R}^{2^N - 1}$ is called entropic if its elements are the entropies for some joint distribution $\mathbf{p}$ on the N rvs $\mathbf{X}$.
5. The entropy vector region (EVR) $\bar{\Gamma}^*_N$ is the closure of the set of all entropic vectors. [1]
6. Normalize by the number of bits for the support m: $\tilde{\mathbf{h}} = \mathbf{h}/\log_2 m$, and define $\bar{\Omega}^*_N$ as the set of normalized entropy vectors (Hassibi and Shadbackt 2007). [2, 3]

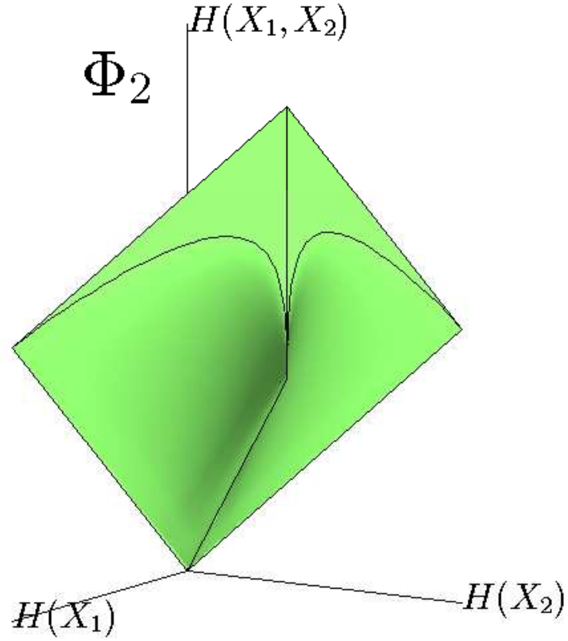
Why are Entropy Vectors and the Entropy Vector Region Important?

- 1:1 correspondence between every nontrivial linear information inequality and tangent hyperplanes to $\bar{\Gamma}^*_N$ [1]
  - Zhang & Yeung [4, 5] showed that there were non-Shannon linear information inequalities for $N \geq 4$. [6]
  - Matùš recently definitively proved [7] there are an infinite number of information inequalities for $N \geq 4$. ($\bar{\Gamma}^*_4$ is not polyhedral)
- multi-user rate regions can be expressed as a function (linear projection) of $\bar{\Gamma}^*_N$ or $\bar{\Omega}^*_N$.
  - Chan & Grant showed that for every non-Shannon inequality face of $\bar{\Gamma}^*_N$ there is a network whose capacity region depends on this non-Shannon inequality [8]
  - Network coding capacity region [1, 9] is a linear projection of $\bar{\Gamma}^*_N$ intersected with a vector subspace
- if willing to limit to distributions obeying constraints $\mathcal{C}$, any rate region in information theory is a linear projection of some $\bar{\Gamma}^*_N(\mathcal{C})$ $\bar{\Omega}^*_N(\mathcal{C})$.
- Hence, most fundamental questions in information theory come back to these sets.

Introducing the set of Binary Entropy Vectors Φ_N



Introducing the set of Binary Entropy Vectors Φ_N



- Very few inner bounds for the EVR exist. [8]
- Even fewer *computable* inner bounds.
- Can not even determine membership in $\bar{\Gamma}^*_N$ or $\bar{\Omega}^*_N$ easily.
- Need finite terminating algorithms to attack these problems.
- Know that all discrete random variables can be represented as bits, and binary is fundamental in information theory.
- Consider restriction of $\bar{\Gamma}^*_N$ to EVs assoc. w/ N *binary* r.v.s!

Two Representations of Joint Distributions on Binary RVs

Lem. 1: The joint distribution $\mathbf{p}$ on N binary rvs

$$\mathbf{p} = (p_{[N]}(\mathbf{x}) = \mathbb{P}(\mathbf{X} = \mathbf{x}),\ \mathbf{x} \in \{0,1\}^N),$$

and the collection of probabilities of each subset of rvs each taking the value zero

$$\mathbf{q} = (p_{\mathcal{A}}(\mathbf{0}) = \mathbb{P}(\mathbf{X}_{\mathcal{A}} = \mathbf{0}),\ \mathcal{A} \subseteq [N])$$

are equivalent in that there is a bijection between them.

Specifying the zero probabilities for all *strict* subsets leaves one degree of freedom, the joint probability of zero.

Example: $N = 2$ rvs. There exists an invertible matrix such that $\mathbf{q} = \mathbf{M}\mathbf{p}$:

$$\begin{bmatrix} p_{(1)}(0) \\ p_{(2)}(0) \\ p_{(1,2)}(00) \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & & \\ 1 & & 1 & \\ 1 & & & \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} p_{(1,2)}(00) \\ p_{(1,2)}(01) \\ p_{(1,2)}(10) \\ p_{(1,2)}(11) \end{bmatrix}.$$

A finite terminating algorithm to determine membership in the set of binary entropy vectors

1. The algorithm considers a sequence of subsets of increasing cardinality and finds the marginal distributions on these subsets consistent with the given entropy vector candidate.

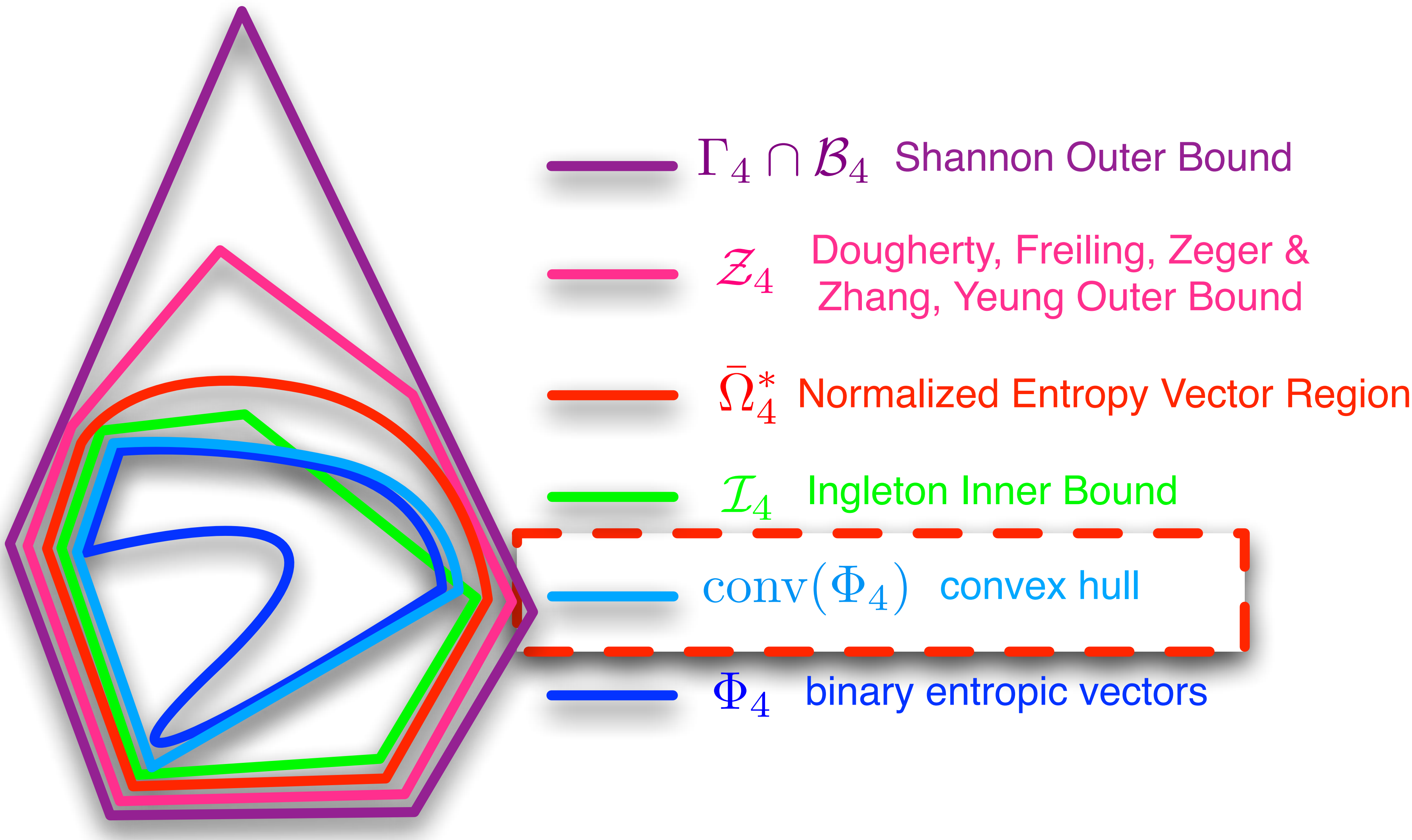
2. The set of all possible binary distributions on a given set $\mathcal{B} \subset \mathcal{A}$ of size $|\mathcal{B}| = k - 1$ consistent with the given entropies is stored in the set $\mathcal{Q}_{\mathcal{B}}$.

3. Find candidate joint distributions $\mathbf{p}_{\mathcal{A}}$ on $\mathbf{X}_{\mathcal{A}}$ consistent with their marginals $\mathbf{p}_{\mathcal{B}}$ for each $\mathcal{B} \subseteq \mathcal{A}$.

4. For each candidate joint disbn $\mathbf{p}_{\mathcal{A}}$ set the remaining free variable using the specified entropy $h_{\mathcal{A}}$ (either 0,1,2 values).

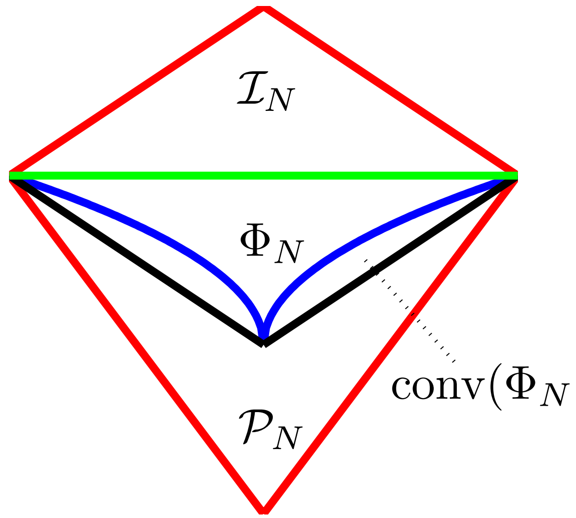
5. Return all joint disbns $\mathbf{p}$ on $\mathbf{X}$ consistent with the given entropy vector $\mathbf{h}$ (if any).

Bounding $\mathrm{conv}(\Phi_N)$



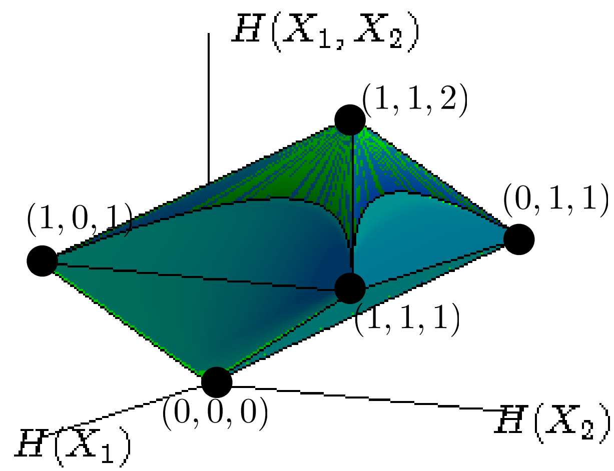
An algorithm to generate an inner bound for a given outer bound

1. Enumerate the extreme points of the polytope $\mathcal{P}_N$, by using a double description algorithm to convert the linear inequality representation into the generating vertices representation.
2. For each of these vertices, determine if they lie in Φ_N, using the membership algorithm. Keep only those vertices lying in Φ_N.
3. Take the convex hull of these vertices to get the polytope $\mathcal{I}_N$. $\mathcal{I}_N$ can be expressed in normal linear inequality form by using the double description method again.



An easy proof of Yeung's result that $\mathcal{P}_2 = \bar{\Omega}_2^*$ (and similarly $\mathcal{P}_3 = \bar{\Omega}_3^*$)

1. Let Φ_N be the collection of all entropy vectors for N *binary* rvs.
2. The Shannon-type inequalities Γ_N plus cardinality constraints ($\mathcal{B}_N = \{\mathbf{h} : h_i \leq 1, i = 1, \ldots, N\}$), denoted $\mathcal{P}_N = \Gamma_N \cap \mathcal{B}_N$, form a polytope in entropy space that outer bounds $\mathrm{conv}(\Phi_N)$.
3. For $N = 2$, $\mathcal{P}_2$ is generated by the vertices below, each of which is in Φ_2 (using the algo.), as can be seen with the following constructions:
   - $(0, 0, 0)$: $(X_1, X_2) = (0, 0)$ with probability one.
   - $(0, 1, 1)$: X_1 deterministic, X_2 uniform.
   - $(1, 0, 1)$: X_1 uniform, X_2 deterministic.
   - $(1, 1, 1)$: $X_1 = X_2$ with probability one, X_1 uniform.
   - $(1, 1, 2)$: X_1, X_2 independent and uniform.



   Therefore $\mathcal{P}_2 = \mathrm{conv}(\Phi_2)$.
4. Further, since $\mathrm{conv}(\Phi_N) \subseteq \bar{\Omega}_N^* \subseteq \mathcal{P}_N$, we have $\mathcal{P}_2 = \bar{\Omega}_2^*$.
5. Similar result and proof for $N = 3$ (both results originally due to Yeung).

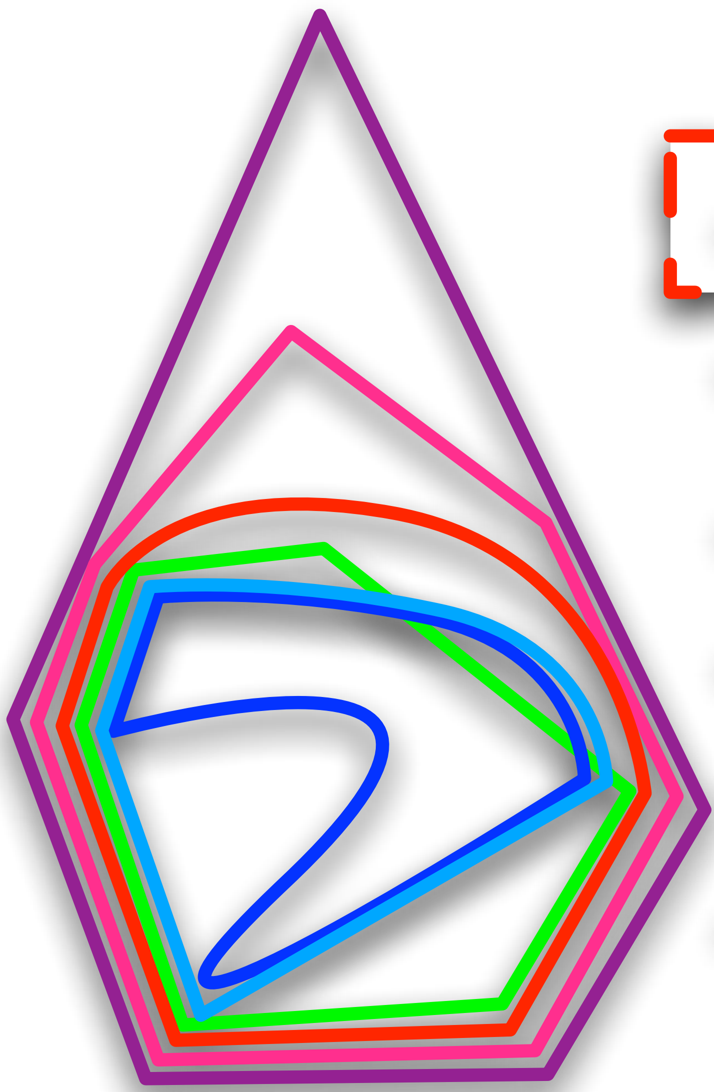
Key properties of the inner bound algorithm

Thm. 1: Given any polytopic outer bound $\mathcal{O}_N$ to $\bar{\Omega}_N^*$, the inner bound algorithm provides after a finite number of computations a polytopic inner bound $\mathcal{I}_N(\mathcal{O}_N)$ to $\mathrm{conv}(\Phi_N)$ and hence $\bar{\Omega}_N^*$ and $\bar{\Gamma}_N^*$. Every exposed face of $\mathcal{O}_N$ which is also an exposed face of $\mathrm{conv}(\Phi_N)$ will also be an exposed face of $\mathcal{I}_N(\mathcal{O}_N)$. Such an exposed face will also necessarily be an exposed face of $\bar{\Omega}_N^*$.

The inner bound algorithm applies to *any* outer bound, i.e., not just $\mathcal{P}_N$. In particular, augmenting the Shannon outer bound with the recently discovered non-Shannon-type inequalities yields a better outer bound.

The quality of the inner bound improves with the quality of the outer bound. That is, given a sequence of increasingly tight outer bounds, our algorithm generates a corresponding sequence of increasingly tight inner bounds, in the sense that any face of the outer bound that is tight on an exposed face of Φ_N will generate a face of the inner bound that is also tight.

Reinterpreting the Shannon Outer Bound $\Gamma_4 \cap \mathcal{B}_4$



$\Gamma_4 \cap \mathcal{B}_4$ Shannon Outer Bound

$\mathcal{Z}_4$ Dougherty, Freiling, Zeger & Zhang, Yeung Outer Bound

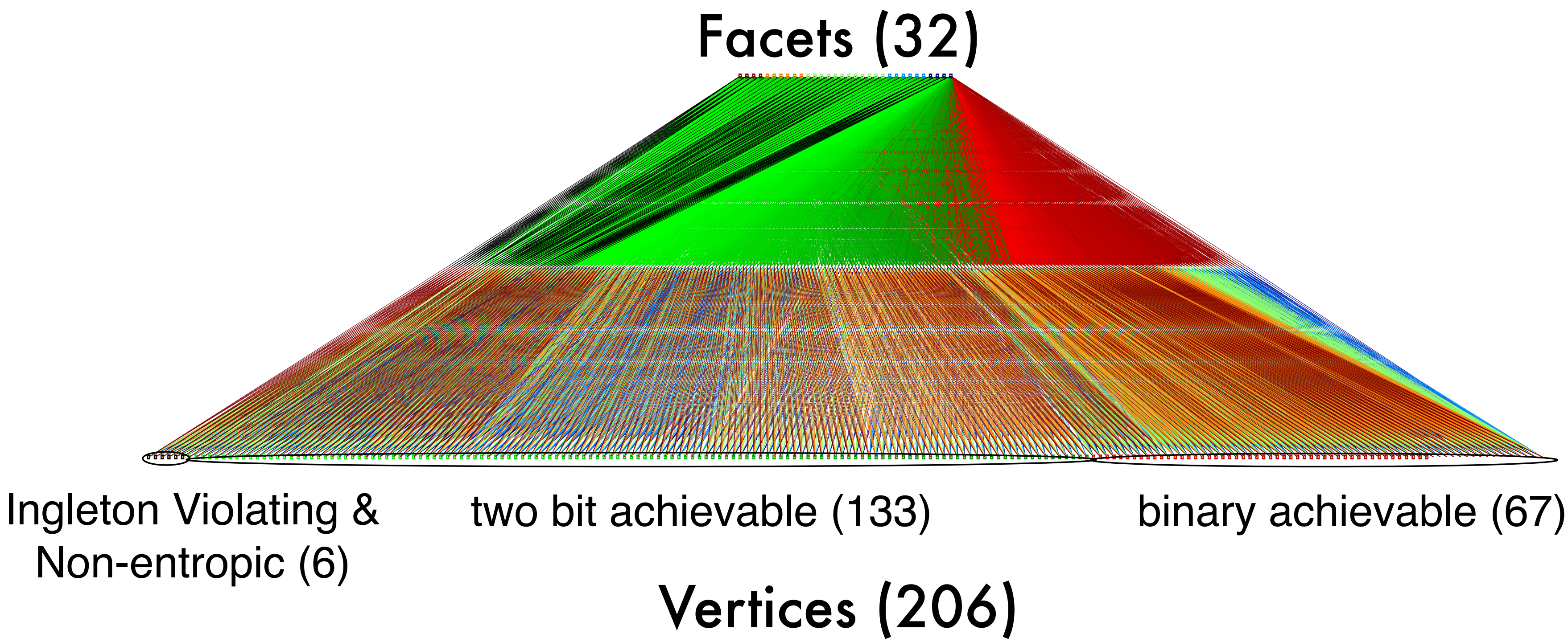
$\bar{\Omega}_4^*$ Normalized Entropy Vector Region

$\mathcal{I}_4$ Ingleton Inner Bound

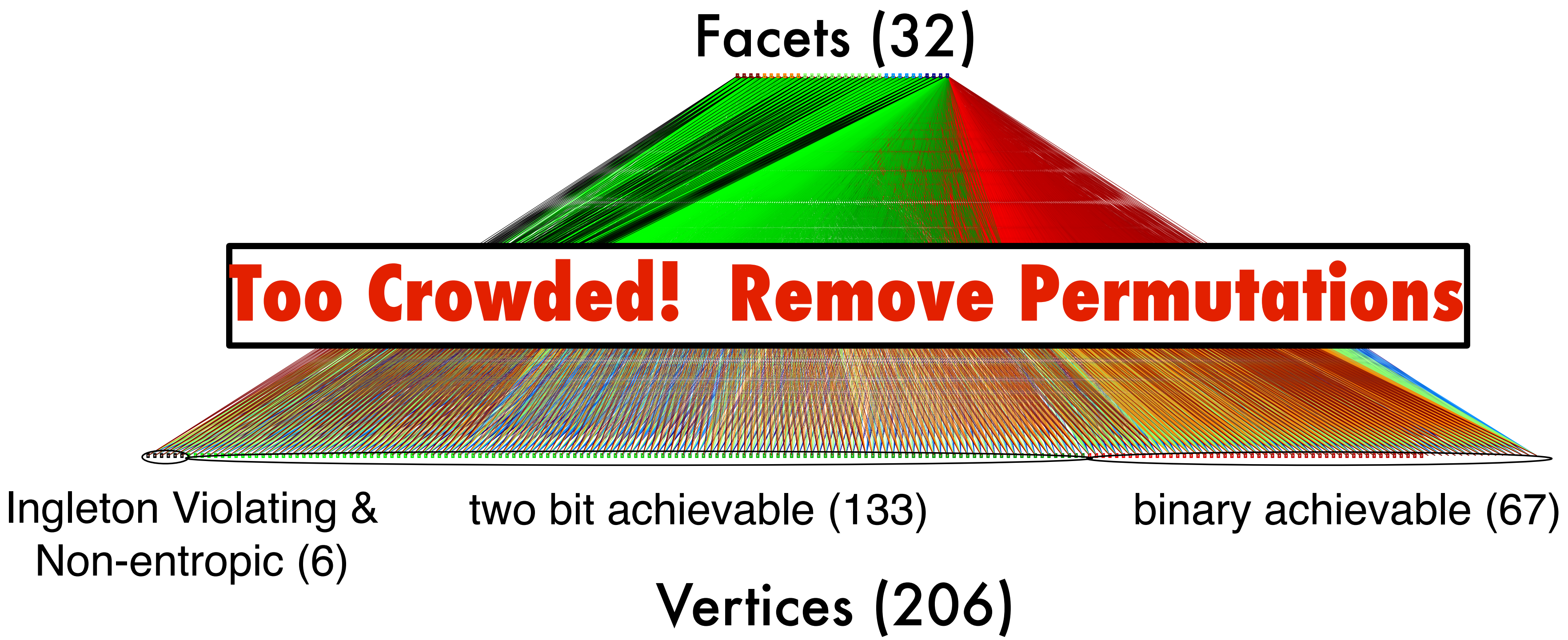
$\mathrm{conv}(\Phi_4)$ convex hull

Φ_4 binary entropic vectors

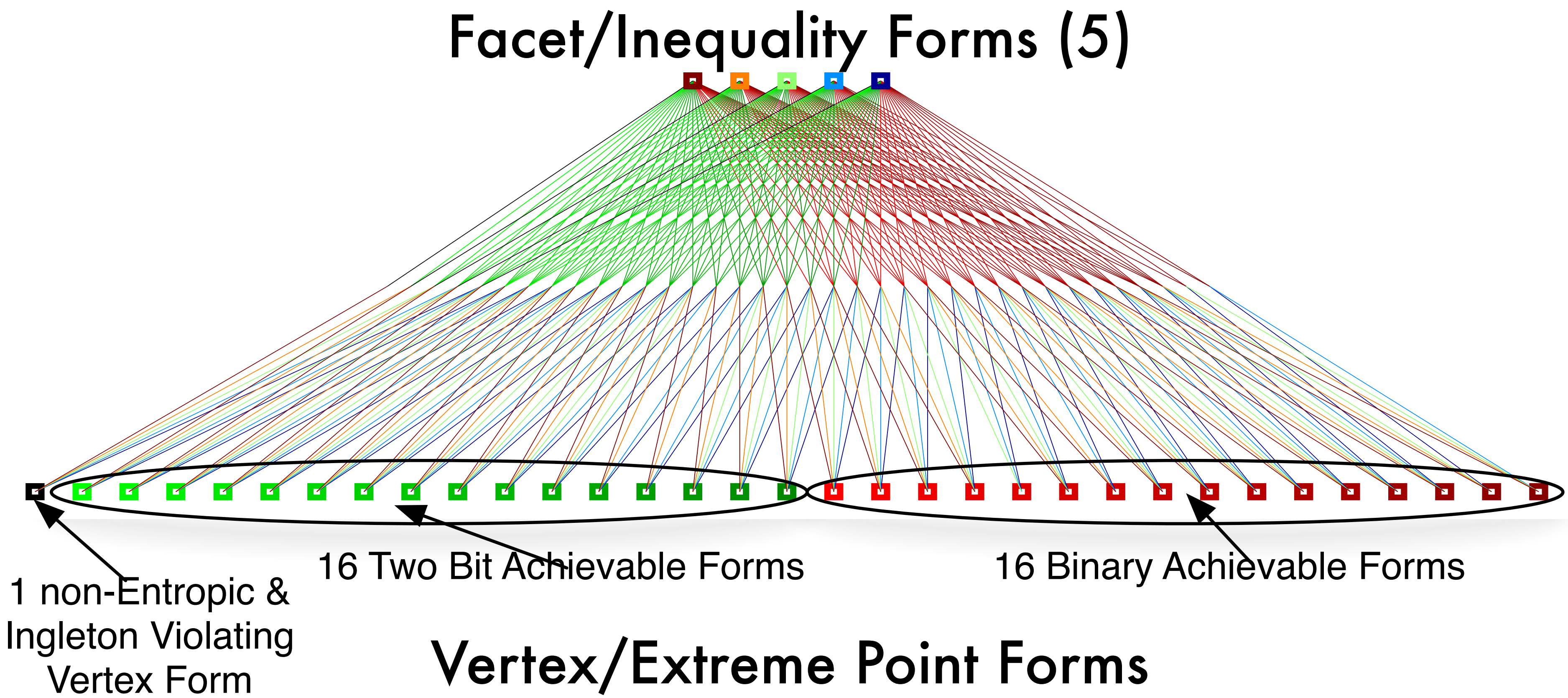
Reinterpreting the Shannon Outer Bound $\Gamma_4 \cap \mathcal{B}_4$



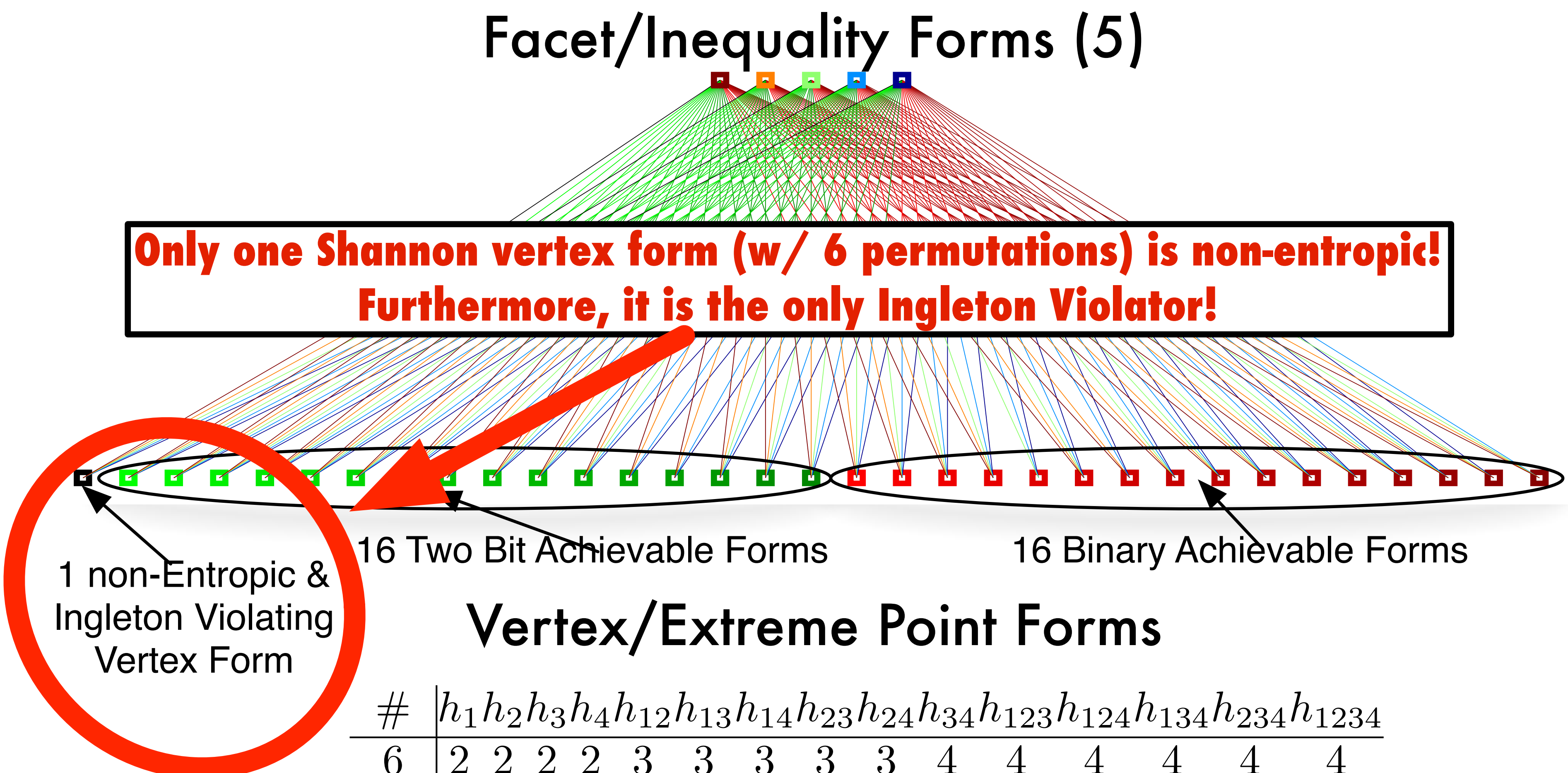
Reinterpreting the Shannon Outer Bound $\Gamma_4 \cap \mathcal{B}_4$



Reinterpreting the Shannon Outer Bound $\Gamma_4 \cap \mathcal{B}_4$

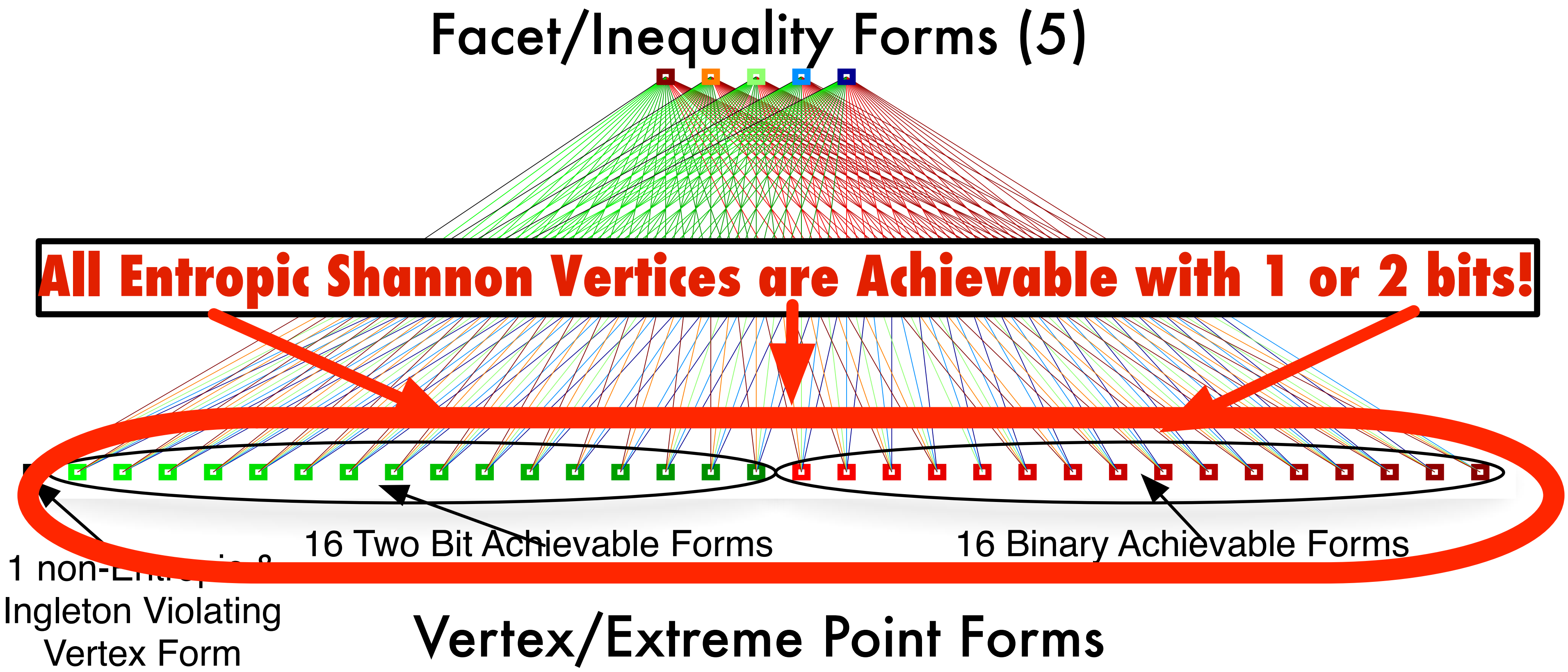


Reinterpreting the Shannon Outer Bound $\Gamma_4 \cap \mathcal{B}_4$: Consequences - 1

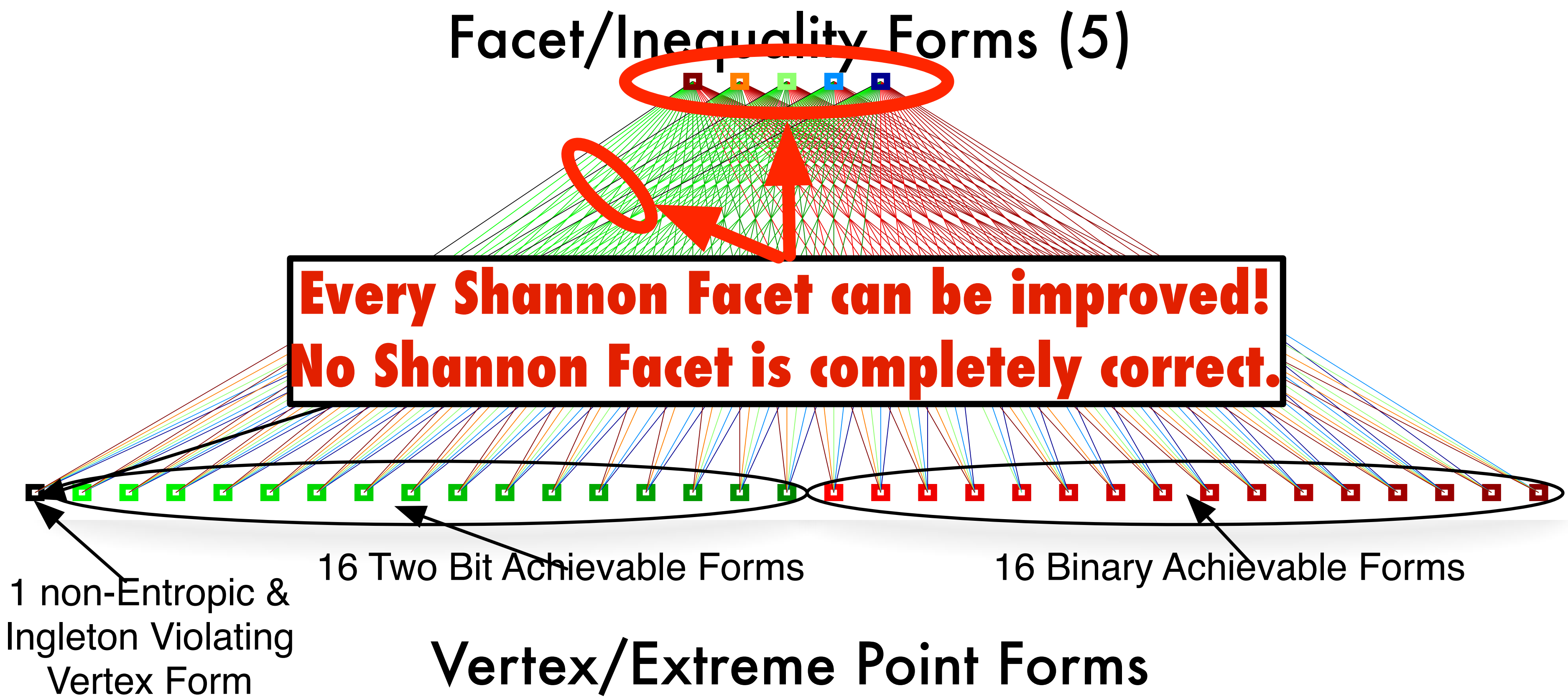


#	h_1	h_2	h_3	h_4	h_{12}	h_{13}	h_{14}	h_{23}	h_{24}	h_{34}	h_{123}	h_{124}	h_{134}	h_{234}	h_{1234}
6	2	2	2	2	3	3	3	3	3	4	4	4	4	4	4

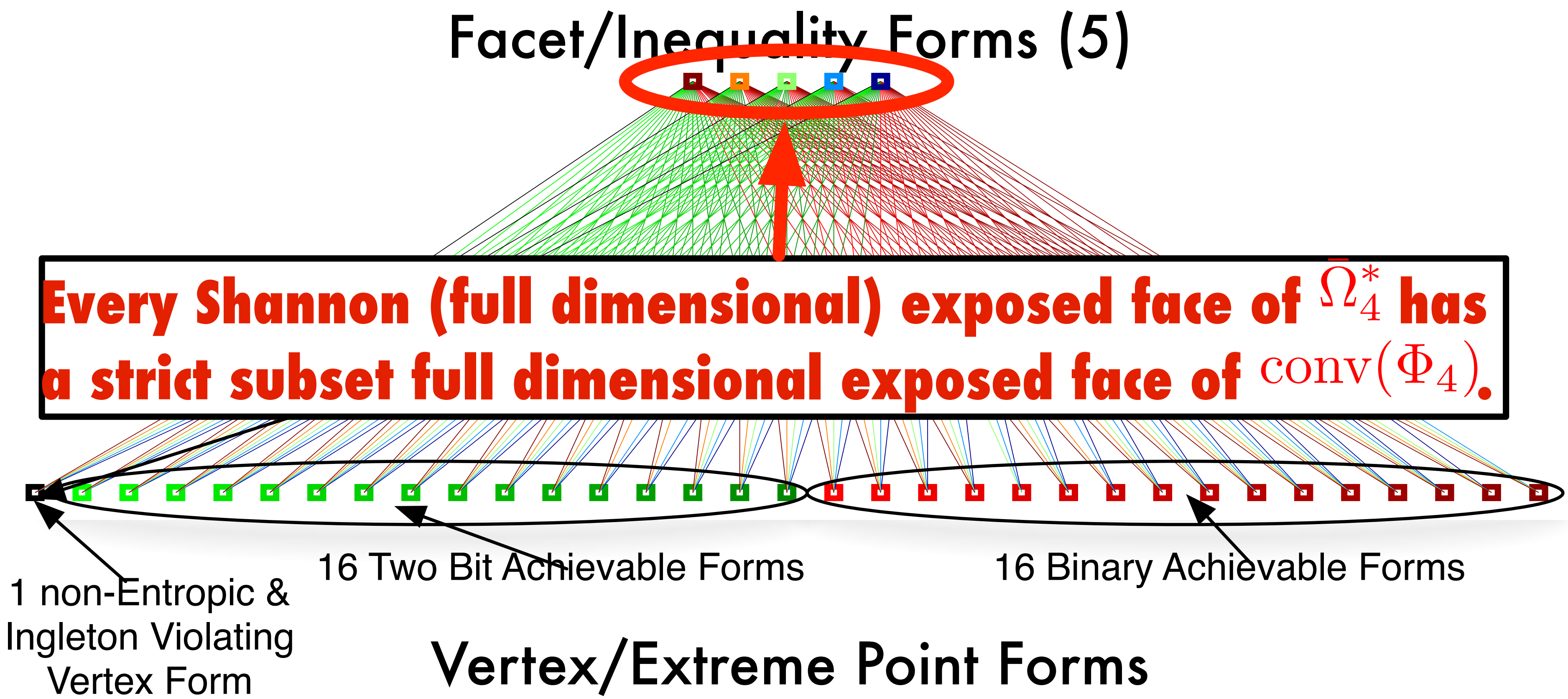
Reinterpreting the Shannon Outer Bound $\Gamma_4 \cap \mathcal{B}_4$: Consequences - 2



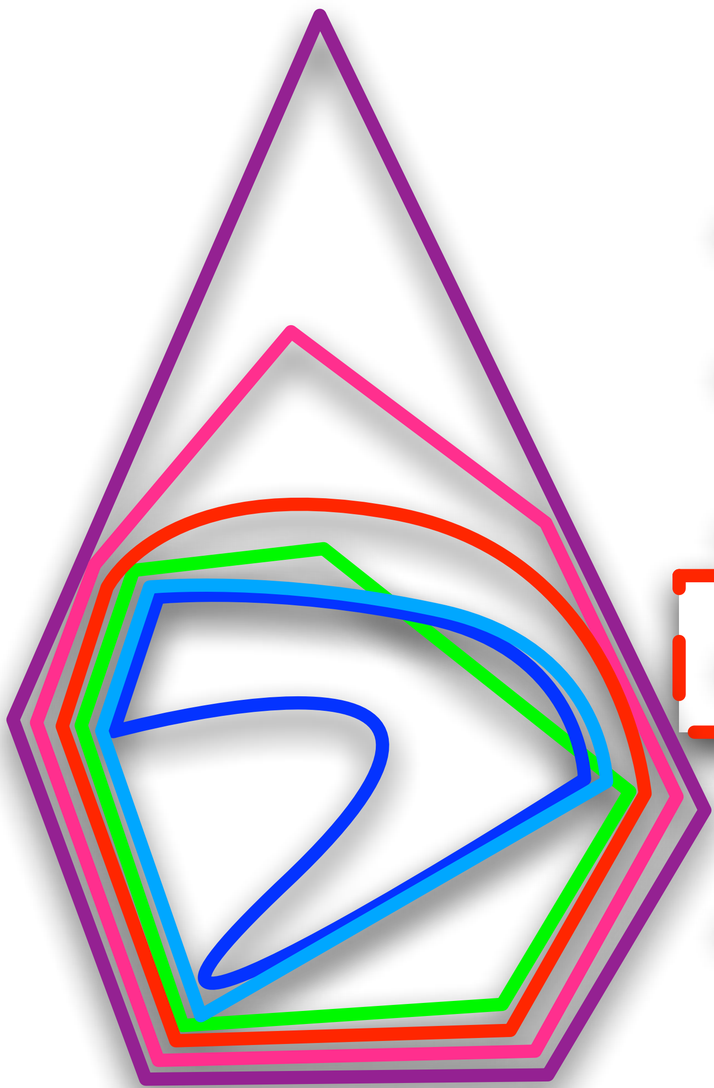
Reinterpreting the Shannon Outer Bound $\Gamma_4 \cap \mathcal{B}_4$: Consequences - 3



Reinterpreting the Shannon Outer Bound $\Gamma_4 \cap \mathcal{B}_4$: Consequences - 4



Reinterpreting the Ingleton Inner Bound $\Gamma_4 \cap \mathcal{B}_4 \cap \mathcal{T}_4$



$\Gamma_4 \cap \mathcal{B}_4$ Shannon Outer Bound

$\mathcal{Z}_4$ Dougherty, Freiling, Zeger & Zhang, Yeung Outer Bound

$\bar{\Omega}_4^*$ Normalized Entropy Vector Region

$\mathcal{I}_4$ Ingleton Inner Bound

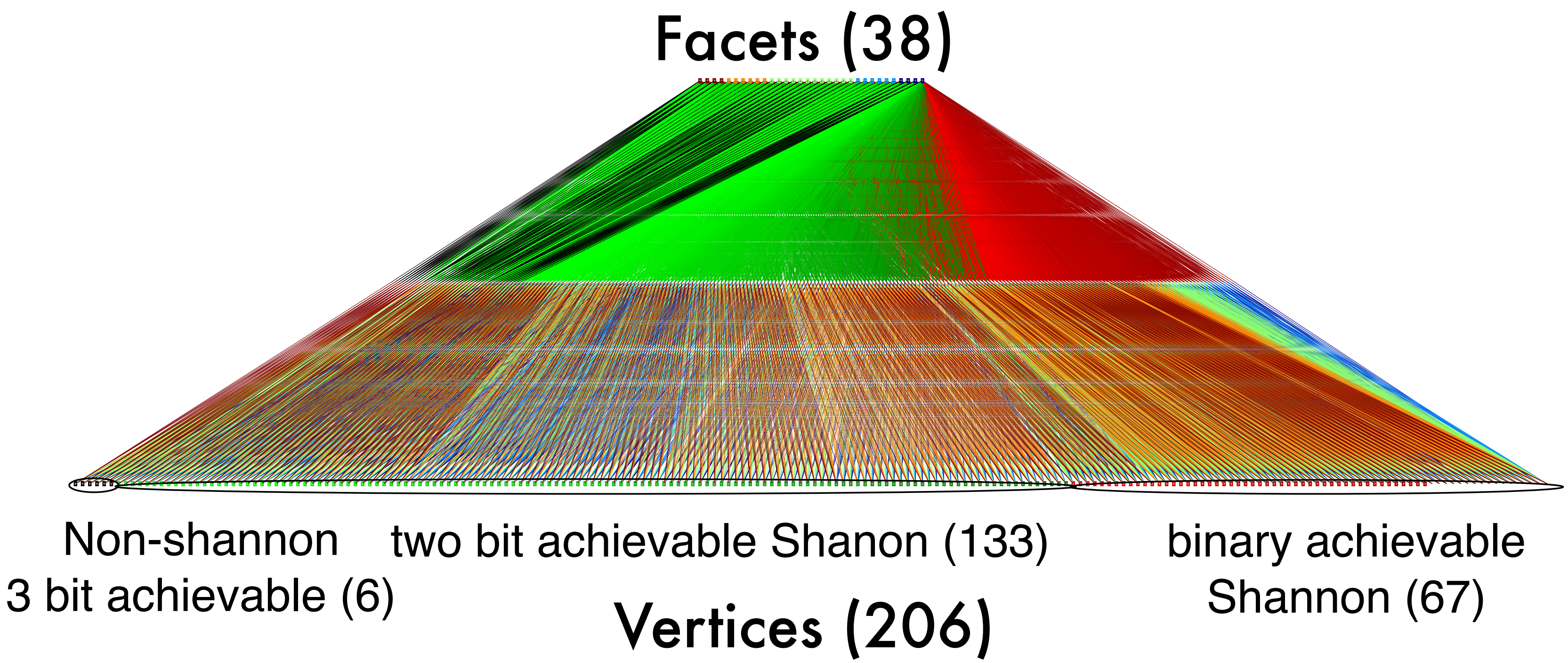
$\text{conv}(\Phi_4)$ convex hull

Φ_4 binary entropic vectors

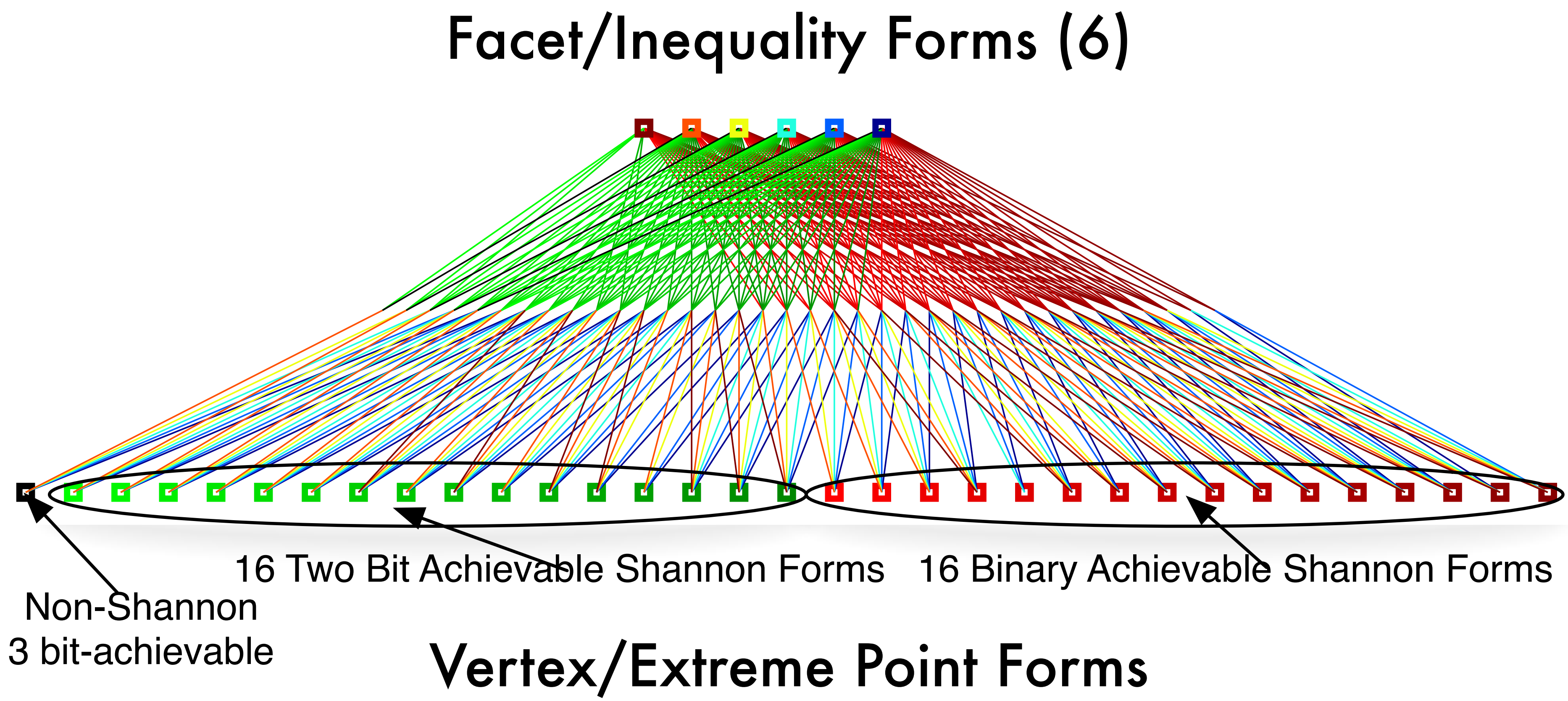
Reinterpreting the Ingleton Inner Bound $\Gamma_4 \cap \mathcal{B}_4 \cap \mathcal{T}_4$

38 inequalities of 6 forms: 32 Shannon outer bound inequalities of 5 forms and 6 Ingleton inequalities of the form

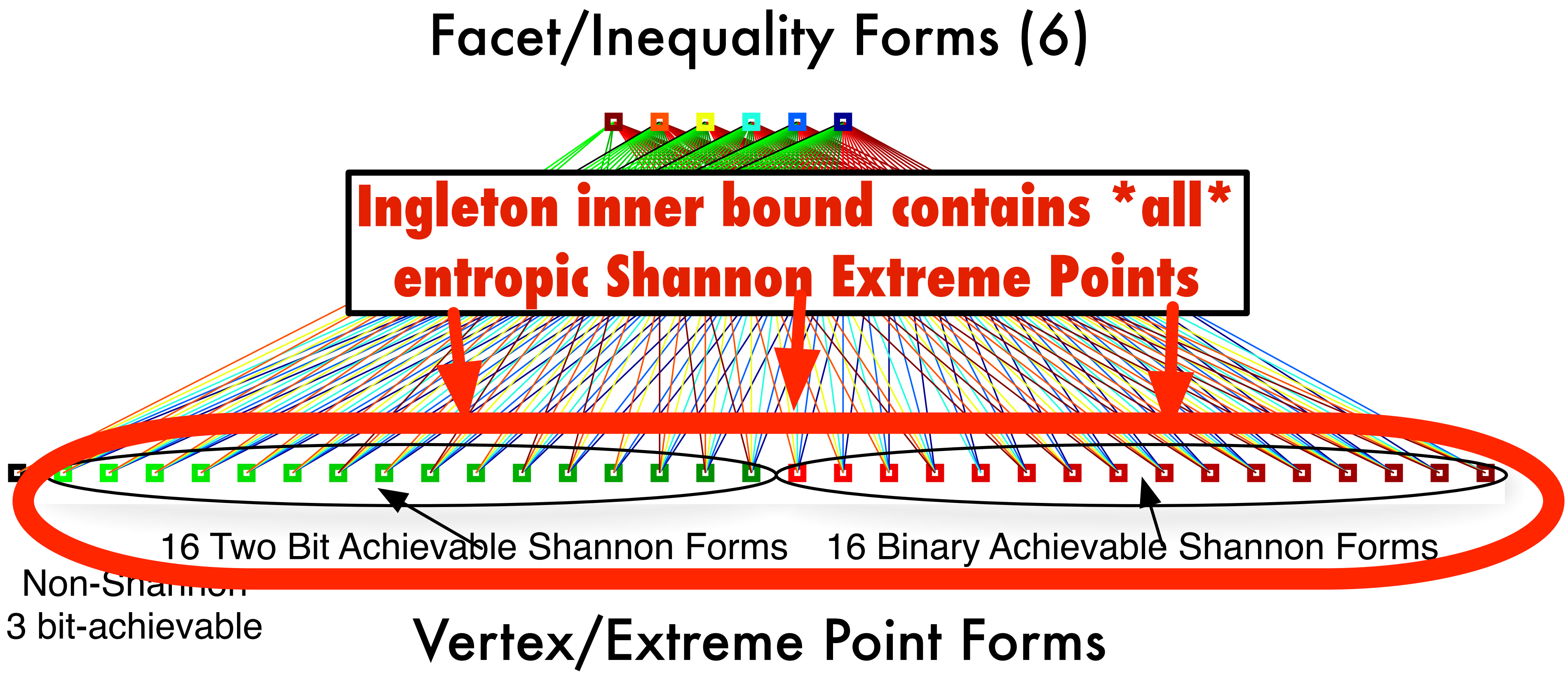
$$h_k + h_l + h_{ij} + h_{ikl} + H_{jkl} - H_{ik} - H_{il} - H_{jk} - H_{jl} - H_{kl} \leq 0 \tag{1}$$



Reinterpreting the Ingleton Inner Bound $\Gamma_4 \cap \mathcal{B}_4 \cap \mathcal{T}_4$

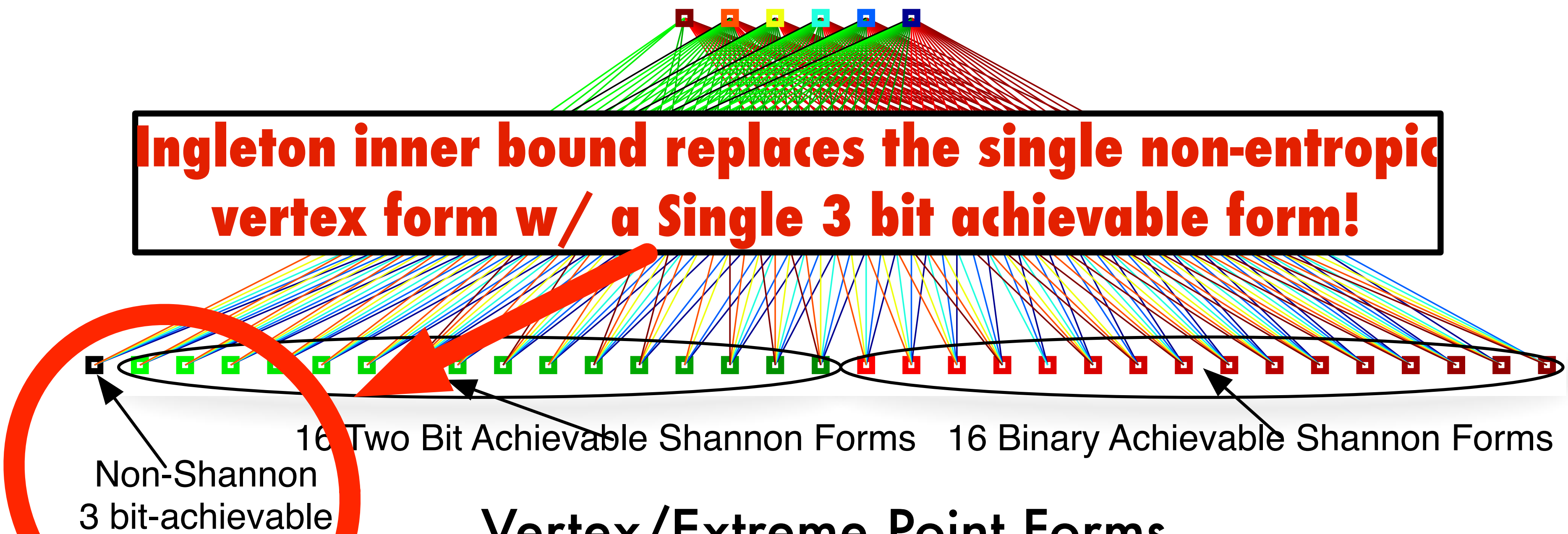


Reinterpreting the Ingleton Inner Bound: Consequences - 1



Reinterpreting the Ingleton Inner Bound: Consequences - 2

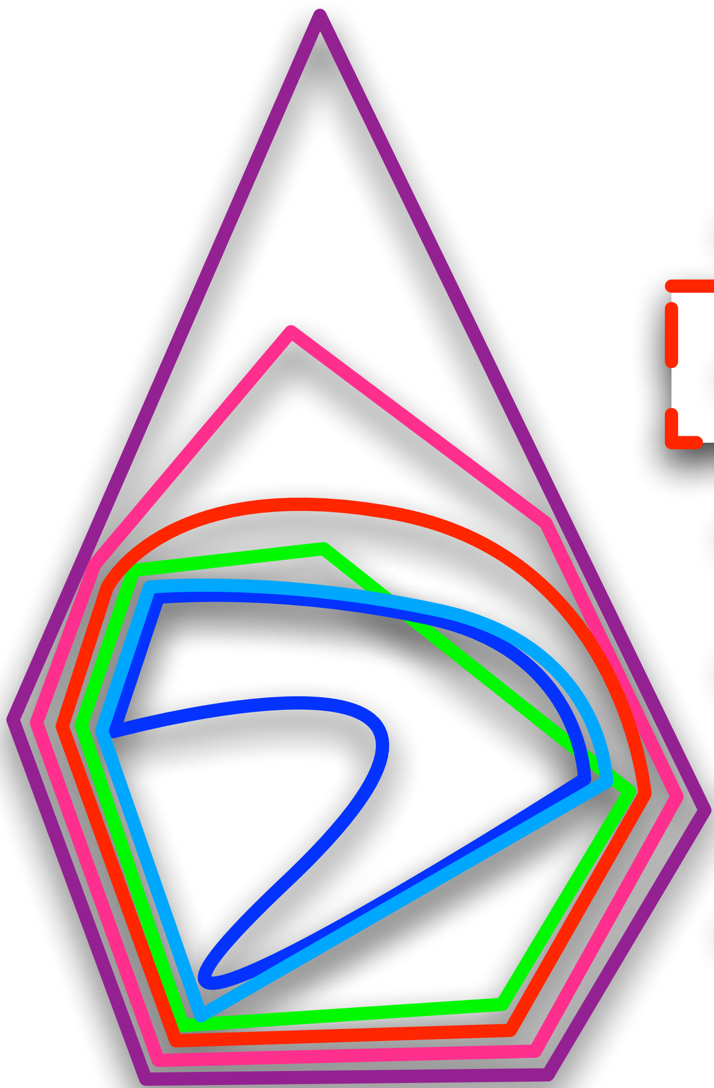
Facet/Inequality Forms (6)



Vertex/Extreme Point Forms

#	h_1	h_2	h_3	h_4	h_{12}	h_{13}	h_{14}	h_{23}	h_{24}	h_{34}	h_{123}	h_{124}	h_{134}	h_{234}	h_{1234}	X_1	X_2	X_3	X_4
6	3	3	3	3	5	4	5	5	6	5	6	6	6	6	6	$b_2 \oplus b_3 \oplus b_5$ $b_4 \oplus b_5$ $b_3 \oplus b_5$	b_1 b_2 b_3	$b_1 \oplus b_4 \oplus b_5 \oplus b_6$ $b_4 \oplus b_5$ b_2	b_4 b_5 b_6

Reinterpreting the DFZ & YZ Non-Shannon Outer Bound



$\Gamma_4 \cap \mathcal{B}_4$ Shannon Outer Bound

$\mathcal{Z}_4$ Dougherty, Freiling, Zeger & Zhang, Yeung Outer Bound

$\bar{\Omega}_4^*$ Normalized Entropy Vector Region

$\mathcal{I}_4$ Ingleton Inner Bound

$\mathrm{conv}(\Phi_4)$ convex hull

Φ_4 binary entropic vectors

Reinterpreting the DFZ & YZ Non-Shannon Outer Bound

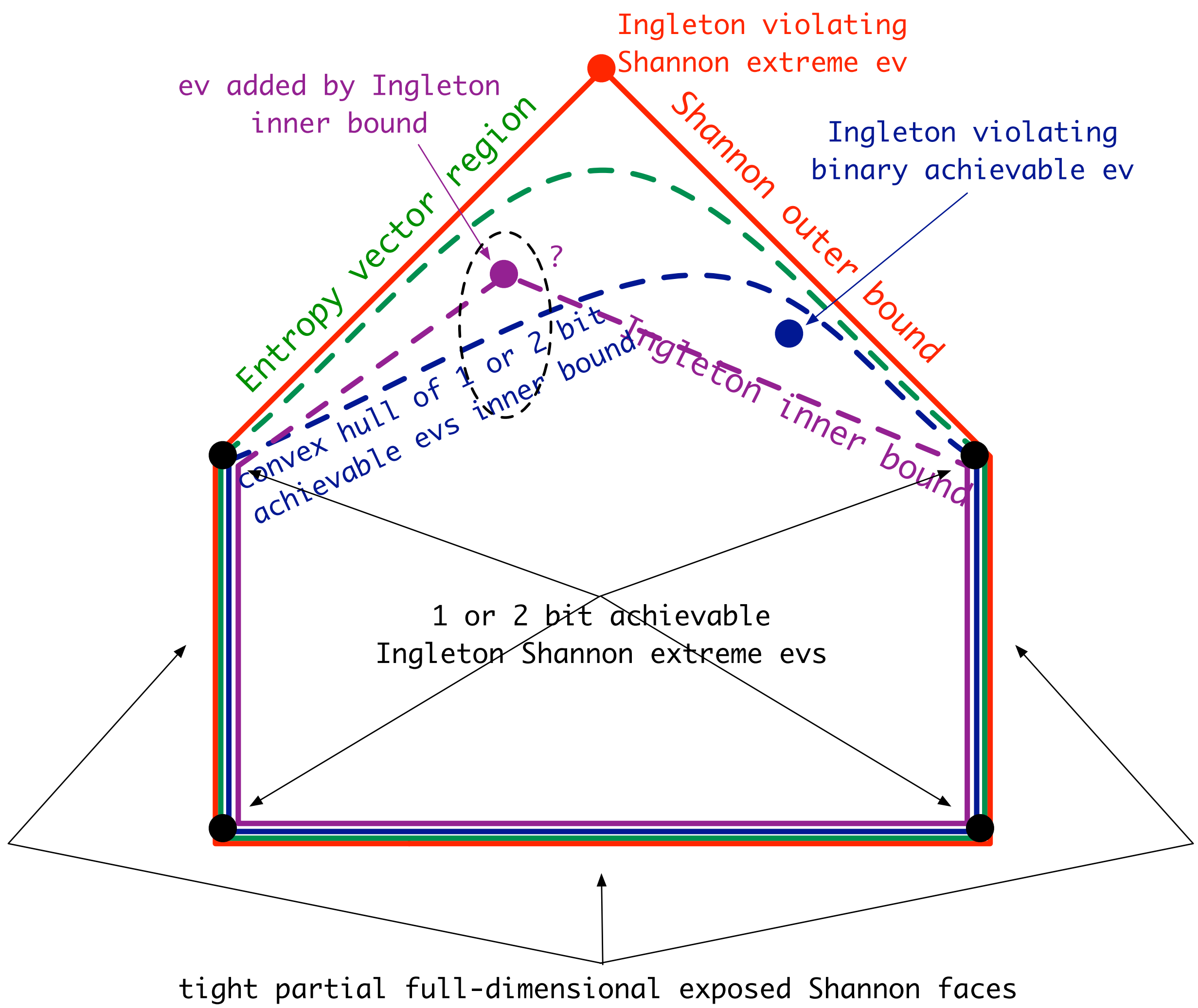
- 2744 vertices of 152 forms:
  - 200 binary & 2-bit achievable Shannon vertices of 32 forms
  - 2544 vertices of 120 new other forms which are
    - all not binary entropic
    - all Ingleton violators
- Important Open Question: which, if any, of these 120 new forms are entropic?
  - current research: looking for a k-bit quasi-uniform construction for these DFZ vertices. Have already been able to prove that no such construction $\exists$ for 1 vertex form.

How far apart are these bounds?

$$\max_{\mathbf{x}\in\mathcal{O}} \min_{\mathbf{y}\in\mathcal{I}} \|\mathbf{x}-\mathbf{y}\|_2 = \max_{\mathbf{v}_i\in\mathcal{V}(\mathcal{O})} \min_{\mathbf{y}\in\mathcal{I}} \|\mathbf{x}-\mathbf{y}\|_2 \tag{2}$$

	Binary Ach. Shan.	2-Bit Ach. Shan.	Ingleton Inner Bound
Shan. Outer	0.3982	$\frac{1}{2\sqrt{10}} \approx .1581$	$\frac{1}{2\sqrt{10}} \approx .1581$
DFZ & YZ + Shan.	0.3982	$\frac{1}{3\sqrt{10}} \approx .1054$	$\frac{1}{3\sqrt{10}} \approx .1054$

Note that the extra achievable vertex form the Ingleton inner bound adds is not where the biggest gap between the Shannon / DFZ&YZ outer bound and the convex hull of the 2-bit achievable Shannon vertices is.


Ingleton violating
Shannon extreme ev
ev added by Ingleton
inner bound
Entropy vector region
Shannon outer bound
Ingleton violating
binary achievable ev
?
convex hull of 1 or 2 bit
achievable evs inner bound
Ingleton inner bound
1 or 2 bit achievable
Ingleton Shannon extreme evs
tight partial full-dimensional exposed Shannon faces

Extensions and future work

1. We further applied this general technique to the Shannon outer bound augmented with the non-Shannon type inequality due to Zhang and Yeung (1998), and established that none of the non-Shannon exposed faces are tight on a full face of $\mathrm{conv}(\Phi_4)$.

2. We have also extended the inner bound algorithm to the set of entropy vectors under one or more distribution constraints, meaning we are restricted to a subset of the full entropy vector space.

3. An important open question is whether or not $\mathrm{conv}(\Phi_N)$ is a polytope. If it is, then it admits a finite characterization in both a listing of its generating vertices and the inequalities characterizing its exposed faces, and it is of interest to obtain these listings. This is the subject of our current work.

References

[1] Raymond W. Yeung, *Information Theory and Network Coding*. Springer, 2008.

[2] B. Hassibi and S. Shadbakht, "Normalized Entropy Vectors, Network Information Theory and Convex Optimization," in *IEEE Information Theory Workshop on Information Theory for Wireless Networks*, Jul. 2007, pp. 1 – 5.

[3] ——, "On a Construction of Entropic Vectors Using Lattice-Generated Distributions," in *IEEE International Symposium on Information Theory (ISIT)*, Jun. 2007, pp. 501 – 505.

[4] Zhen Zhang and Raymond W. Yeung, "A Non-Shannon-Type Conditional Inequality of Information Quantities," *IEEE Transactions on Information Theory*, vol. 43, no. 6, Nov. 1997.

[5] ——, "On Characterization of Entropy Function via Information Inequalities," *IEEE Transactions on Information Theory*, vol. 44, no. 4, Jul. 1998.

[6] R. Dougherty, C. Freiling, and K. Zeger, "Networks, Matroids, and Non-Shannon Information Inequalities," *IEEE Transactions on Information Theory*, vol. 53, no. 6, pp. 1949–1969, Jun. 2007.

[7] František Matúš, "Infinitely Many Information Inequalities," in *IEEE International Symposium on Information Theory (ISIT)*, Jun. 2007, pp. 41–44.

[8] T. Chan and A. Grant, "Entropy Vectors and Network Codes," in *IEEE International Symposium on Information Theory*, Jun. 2007.

[9] Xijin Yan, Raymond W. Yeung, and Zhen Zhang, "The Capacity Region for Multi-source Multi-sink Network Coding," in *IEEE International Symposium on Information Theory (ISIT)*, 2007.

Thanks to NSF CCF-0728496 and AFOSR FA9550-09-C-0014.