Entropic Vectors: Polyhedral Computation & Information Geometry

John MacLaren Walsh
Department of Electrical and Computer Engineering
Drexel University
Philadelphia, PA
jwalsh@ece.drexel.edu

Thanks to NSF CCF-1016588, NSF CCF-1053702, & AFOSR FA9550-12-1-0086.

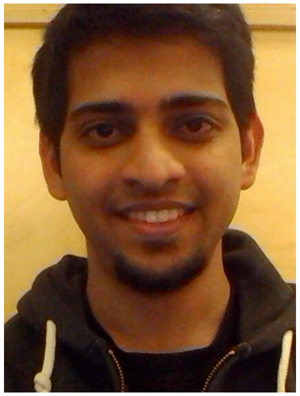
Outline

1. Entropic Vectors and Polyhedral Computation
   (a) One Procedure to Rule Four Problems
      i. Non-Shannon outer bounds for $\bar{\Gamma}^*_N$ (Yeung, Zhang, DFZ, Matúš, ...)
      ii. Vector Matroidal Inner bound for $\bar{\Gamma}^*_N$ (Hassibi)
      iii. Linear Code/Subspace Rank Outer Bound (Hammer, Kinser, DFZ)
      iv. Network Coding/Distributed Storage Regions (Yan, Yeung, Chan, Grant)
   (b) Many Paths Lead to the Same Truth
   (c) The Path Less Laden: Complexity Experiments & Moving Forward
2. Structure of $\bar{\Gamma}^*_4$ & the Case for Non-Polyhedral Tools
3. Characterizing Extremal Entropic Vectors with Information Geometry
   (a) Introduction to Information Geometry
   (b) Decoupling Constraints & Improving Ingleton w/ Binary Inner Bound
   (c) Casting Entropic Vectors as Information Projections
   (d) Information Geometric Properties of Distributions on Shannon Facets
   (e) $\bar{\Gamma}^*_4$ & 4 atom distribution. Information Geometry of Ingleton Violation

Outline

1. **Entropic Vectors and Polyhedral Computation**
   (a) One Procedure to Rule Four Problems
      i. Non-Shannon outer bounds for $\bar{\Gamma}^*_N$ (Yeung, Zhang, DFZ, Matúš, ...)
      ii. Vector Matroidal Inner bound for $\bar{\Gamma}^*_N$ (Hassibi)
      iii. Linear Code/Subspace Rank Outer Bound (Hammer, Kinser, DFZ)
      iv. Network Coding/Distributed Storage Regions (Yan, Yeung, Chan, Grant)
   (b) Many Paths Lead to the Same Truth
   (c) The Path Less Laden: Complexity Experiments & Moving Forward
2. Structure of $\bar{\Gamma}^*_4$ & the Case for Non-Polyhedral Tools
3. Characterizing Extremal Entropic Vectors with Information Geometry
   (a) Introduction to Information Geometry
   (b) Decoupling Constraints & Improving Ingleton w/ Binary Inner Bound
   (c) Casting Entropic Vectors as Information Projections
   (d) Information Geometric Properties of Distributions on Shannon Facets
   (e) $\bar{\Gamma}^*_4$ & 4 atom distribution. Information Geometry of Ingleton Violation

Entropic Vectors and Polyhedral Computation
(Development Team)



Jayant Apte

efficient parallel infinite precision polyhedral computation

Congduan Li

rate regions & rate delay tradeoffs

Daniel Venutolo

computing non-Shannon inequalities

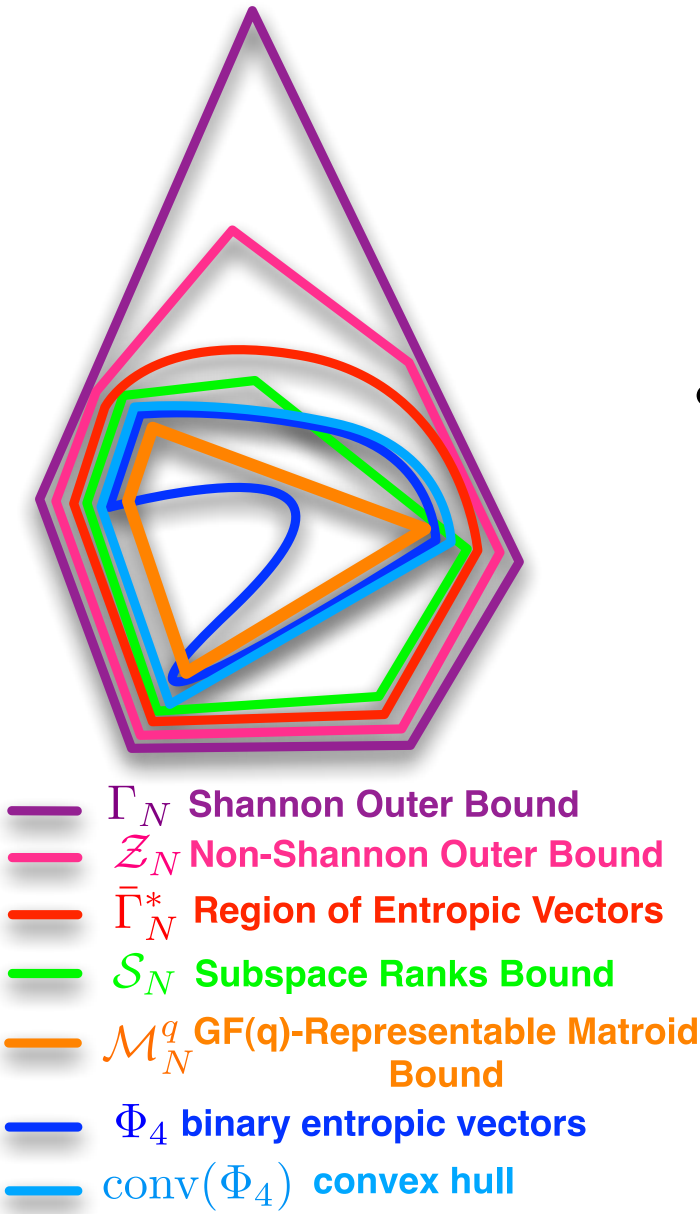
Prof. Steven Weber

Drexel University

Outline

1. Entropic Vectors and Polyhedral Computation
   (a) **One Procedure to Rule Four Problems**
      i. **Non-Shannon outer bounds for $\bar{\Gamma}^*_N$ (Yeung, Zhang, DFZ, Matúš, ...)**
      ii. Vector Matroidal Inner bound for $\bar{\Gamma}^*_N$ (Hassibi)
      iii. Linear Code/Subspace Rank Outer Bound (Hammer, Kinser, DFZ)
      iv. Network Coding/Distributed Storage Regions (Yan, Yeung, Chan, Grant)
   (b) Many Paths Lead to the Same Truth
   (c) The Path Less Laden: Complexity Experiments & Moving Forward
2. Structure of $\bar{\Gamma}^*_4$ & the Case for Non-Polyhedral Tools
3. Characterizing Extremal Entropic Vectors with Information Geometry
   (a) Introduction to Information Geometry
   (b) Decoupling Constraints & Improving Ingleton w/ Binary Inner Bound
   (c) Casting Entropic Vectors as Information Projections
   (d) Information Geometric Properties of Distributions on Shannon Facets
   (e) $\bar{\Gamma}^*_4$ & 4 atom distribution. Information Geometry of Ingleton Violation

Bounding the Region of Entropic Vectors $\bar{\Gamma}_N^*$ from the Outside

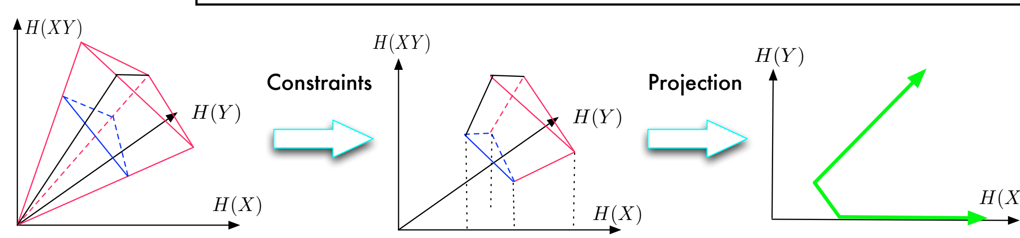


- **Shannon Outer Bound**: Γ_N. entropy is submodular:

$$I(\boldsymbol{X}_{\mathcal{A}}; \boldsymbol{X}_{\mathcal{B}} | \boldsymbol{X}_{\mathcal{C}}) \geq 0 \quad \forall \mathcal{A}, \mathcal{B}, \mathcal{C}$$

  $\Gamma_2 = \bar{\Gamma}_2^*, \Gamma_3 = \bar{\Gamma}_3^*$.
  $\Gamma_N \neq \bar{\Gamma}_N^*, N \geq 4$ $\bar{\Gamma}_N^*$ non-polyhedral convex cone

- **Non-Shannon Outer Bounds**:[1, 2, 3, 4, 5, 6, 7]
  Yeung & Zhang, Dougherty & Freiling & Zeger, Matus
  Start with 4 unconstr. r.v.s
  add rv. obeying distr. match & Markov. cond.
  Intersect Γ_N for $N \geq 5$ w/ Markov & distr. match
  Project back to orig. 4 unconstr. vars.
  obtain new information inequalities this way!
  overall: Shannon $\rightarrow$ linear eq./ineq. $\bigcap$ $\rightarrow$ project



Outline

1. Entropic Vectors and Polyhedral Computation
   (a) One Procedure to Rule Four Problems
      i. Non-Shannon outer bounds for $\bar{\Gamma}_N^*$ (Yeung, Zhang, DFZ, Matúš, ...)
      ii. **Vector Matroidal Inner bound for $\bar{\Gamma}_N^*$ (Hassibi)**
      iii. Linear Code/Subspace Rank Outer Bound (Hammer, Kinser, DFZ)
      iv. Network Coding/Distributed Storage Regions (Yan, Yeung, Chan, Grant)
   (b) Many Paths Lead to the Same Truth
   (c) The Path Less Laden: Complexity Experiments & Moving Forward
2. Structure of $\bar{\Gamma}_4^*$ & the Case for Non-Polyhedral Tools
3. Characterizing Extremal Entropic Vectors with Information Geometry
   (a) Introduction to Information Geometry
   (b) Decoupling Constraints & Improving Ingleton w/ Binary Inner Bound
   (c) Casting Entropic Vectors as Information Projections
   (d) Information Geometric Properties of Distributions on Shannon Facets
   (e) $\bar{\Gamma}_4^*$ & 4 atom distribution. Information Geometry of Ingleton Violation

Bounding $\bar{\Gamma}^*_N$ from the Inside, 1: Representable Matroids

Matroids: $\boldsymbol{r} \in \Gamma_N \cap \mathbb{Z}^{2^N-1}$, $r(\mathcal{A}) \leq |\mathcal{A}|$

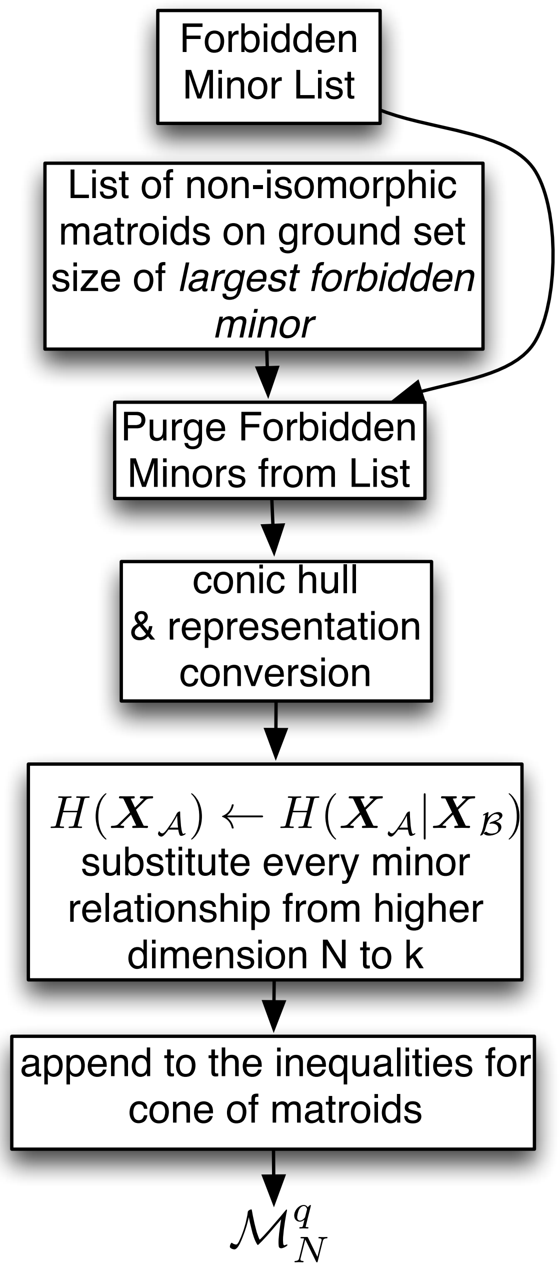
- All non-isomorphic matroids for $N \leq 9$ [8] '08
- enumerating non-iso. matroids is difficult

$GF(q)$-**Representable Matroid**: $\boldsymbol{r} \in \Gamma_N \cap \mathbb{Z}^{2^N-1}$ s.t. $\exists \mathbf{A} \in GF(q)^{M\times N}$ s.t. $r(\mathcal{A}) = \operatorname{rank}(\mathbf{A}_{:,\mathcal{A}})$

- repr. matroid = scaled EV!: $\boldsymbol{u} \sim \mathcal{U}(GF(q)^M)$

$$\boldsymbol{X} = \boldsymbol{u}\mathbf{A} \;\Rightarrow\; h_{\mathcal{A}} = r(\mathcal{A}) \log_2 q$$

- Key: representability $\Leftrightarrow$ no forbidden minors:
  - *complete small list known for* $q \in \{2,3,4\}$ [9, 10, 11, 12, 13] eg.:$GF(2)$ repr. $\Leftrightarrow$ no $U(2,4)$ minor (*Tutte* 1958)
- $\bar{\Gamma}^*_N$ bound: $\mathcal{M}^q_N$ conic hull of $GF(q)$-repr. matroids. (*Hassibi* et. al. 2010 [14]). see right.



Bounding $\bar{\Gamma}^*_N$ from the Inside, 2: Inner Bounds from Subspace Ranks

Subspace Bounds: $\boldsymbol{r} \in \Gamma_N \cap \mathbb{Z}^{2^N-1}$ projections of representable matroids, $N' \geq N$, partition $\{1,\ldots,N'\} = \bigcup_{n=1}^{N} \mathcal{G}_n$, $\mathcal{G}_n \cap \mathcal{G}_k = \emptyset \; n \neq k$

$$r(\mathcal{A}) = \text{rank}([\mathbf{A}_{:,\mathcal{G}_n} | n \in \mathcal{A}]) \qquad (1)$$

subspace ranks = scaled EV!:

$$\boldsymbol{X}_n = \boldsymbol{u}\mathbf{A}_{:,\mathcal{G}_n} \Rightarrow h_{\mathcal{A}} = r(\mathcal{A}) \log_2 q$$

$\mathcal{S}_N$: conic hull of all subspace ranks

$\mathcal{S}_4$: $\Gamma_4 \cap$ Ingleton's [15, 16, 17]

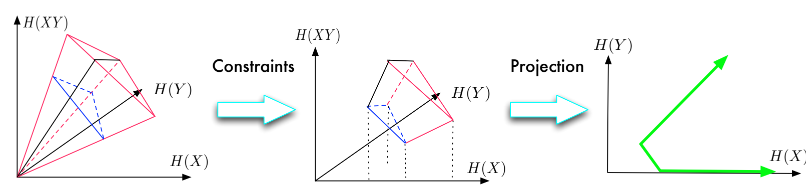
$$I(X_1; X_2) + I(X_3; X_4|X_1) + I(X_3; X_4|X_2) - I(X_3; X_4) \geq 0$$

$\mathcal{S}_5$ recently characterized by *DFZ* + *Kinser* [18, 19]

$\mathcal{S}_N$ unknown for $N \geq 6$, but can inner bounded by projecting $\mathcal{M}^q_N$ (see right)

Build inner bound for $\mathcal{S}_N \subsetneq \bar{\Gamma}^*_N$:

1. Obtain $\mathcal{M}^q_{N'}$ using method from previous slide, i.e. *intersect* Γ_N *with inequalities from forbidden minors & matroids*
2. project (remove all but entropies where each element in $\boldsymbol{X}_n$ appears together)



sound familiar???

Shannon → lin. $\bigcap$ → project

Outline

1. Entropic Vectors and Polyhedral Computation
   (a) One Procedure to Rule Four Problems
      i. Non-Shannon outer bounds for $\bar{\Gamma}_N^*$ (Yeung, Zhang, DFZ, Matúš, ...)
      ii. Vector Matroidal Inner bound for $\bar{\Gamma}_N^*$ (Hassibi)
      iii. **Linear Code/Subspace Rank Outer Bound (Hammer, Kinser, DFZ)**
      iv. Network Coding/Distributed Storage Regions (Yan, Yeung, Chan, Grant)
   (b) Many Paths Lead to the Same Truth
   (c) The Path Less Laden: Complexity Experiments & Moving Forward
2. Structure of $\bar{\Gamma}_4^*$ & the Case for Non-Polyhedral Tools
3. Characterizing Extremal Entropic Vectors with Information Geometry
   (a) Introduction to Information Geometry
   (b) Decoupling Constraints & Improving Ingleton w/ Binary Inner Bound
   (c) Casting Entropic Vectors as Information Projections
   (d) Information Geometric Properties of Distributions on Shannon Facets
   (e) $\bar{\Gamma}_4^*$ & 4 atom distribution. Information Geometry of Ingleton Violation

Bounding $\bar{\Gamma}^*_N$ from the Inside, 3: Outer Bounds for Subspace Ranks $\mathcal{S}_N$

- Hammer et. al.: key [16] r.v.s X, Y made with subspaces: $\exists$ *common information* Z s.t.

$$H(Z|X) = H(Z|Y) = 0 \qquad (2)$$

$$H(Z) = I(X;Y) \qquad (3)$$

- Common information Z obtained by looking at component along intersection of subspaces X Y
- Not all RVs have common information, but rvs from subspaces do

<u>Build outer bound for $\mathcal{S}_N \subsetneq \bar{\Gamma}^*_N$:</u>

1. Shannon
2. Intersect with common information equalities (2)
3. project out the common informations Z

sound familiar?

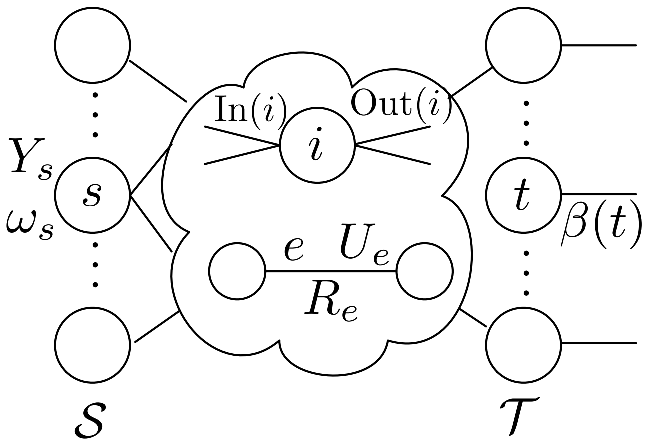
Shannon $\rightarrow$ lin. $\bigcap$ $\rightarrow$ project

Outline

1. Entropic Vectors and Polyhedral Computation
   (a) One Procedure to Rule Four Problems
      i. Non-Shannon outer bounds for $\bar{\Gamma}_N^*$ (Yeung, Zhang, DFZ, Matúš, ...)
      ii. Vector Matroidal Inner bound for $\bar{\Gamma}_N^*$ (Hassibi)
      iii. Linear Code/Subspace Rank Outer Bound (Hammer, Kinser, DFZ)
      iv. **Network Coding/Distributed Storage Regions (Yan, Yeung, Chan, Grant)**
   (b) Many Paths Lead to the Same Truth
   (c) The Path Less Laden: Complexity Experiments & Moving Forward
2. Structure of $\bar{\Gamma}_4^*$ & the Case for Non-Polyhedral Tools
3. Characterizing Extremal Entropic Vectors with Information Geometry
   (a) Introduction to Information Geometry
   (b) Decoupling Constraints & Improving Ingleton w/ Binary Inner Bound
   (c) Casting Entropic Vectors as Information Projections
   (d) Information Geometric Properties of Distributions on Shannon Facets
   (e) $\bar{\Gamma}_4^*$ & 4 atom distribution. Information Geometry of Ingleton Violation

Networking Coding and Distributed Storage Rate Regions

Network Coding



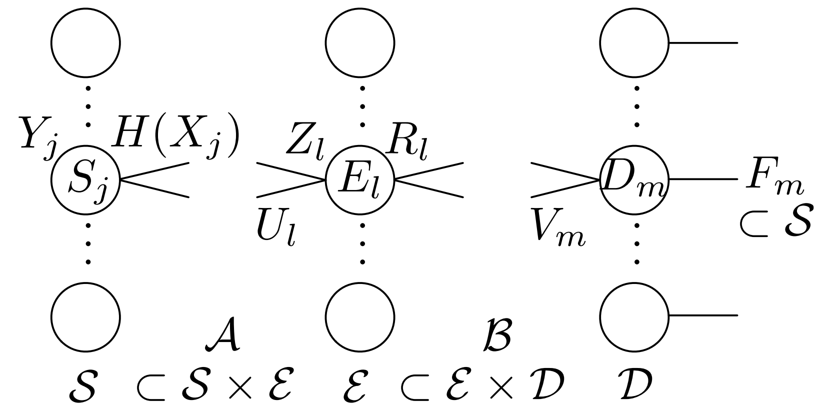
(Roughly) intersect $\bar{\Gamma}^*_N$ with

$$h_{Y_s} \geq \omega_s, \;\; s \in \mathcal{S} \qquad h_{Y_{\mathcal{S}}} = \sum_{s \in \mathcal{S}} h_{Y_s},$$

$$\begin{aligned} h_{U_{\mathrm{Out}(s)}|Y_s} &= 0, \;\; s \in \mathcal{S} \\ h_{U_{\mathrm{Out}(i)}|U_{\mathrm{In}(i)}} &= 0, \;\; i \in \mathcal{V} \setminus (\mathcal{S} \cup \mathcal{T}) \\ h_{U_e} &\leq R_e, \;\; e \in \mathcal{E} \\ h_{Y_{\beta(t)}|U_{\mathrm{In}(t)}} &= 0, \;\; t \in \mathcal{T} \end{aligned}$$

and project onto ω_s, R_e

Distributed Storage (MDCS DSCSC)



(Roughly) intersect $\bar{\Gamma}^*_N$ with

$$\begin{aligned} h_{Y_j, j \in \mathcal{S}} &= \sum_{j \in \mathcal{S}} h_{Y_j} \\ h_{Z_l|(Y_j, j \in U_l)} &= 0, \;\; l \in \mathcal{E} \\ h_{(Y_j, j \in F_m)|(Z_l, l \in V_m)} &= 0, \;\; m \in \mathcal{D} \\ h_{Y_j} &> H(X_j), \;\; j \in \mathcal{S} \\ h_{Z_l} &\leq R_l, \;\; l \in \mathcal{E} \end{aligned}$$

and project onto $\{H(X_i), R_l\}$

Substituting inner/outer bounds for $\bar{\Gamma}^*_N$, we arrive again at

Shannon $\rightarrow$ linear equality/inequality $\bigcap$ $\rightarrow$ project

One Procedure to Rule Four Problems

We've covered:

1. Non-Shannon Outer bounds for $\bar{\Gamma}^*_N$
2. Vector Matroidal Inner Bounds for $\bar{\Gamma}^*_N$
3. Outer Bound for $\mathcal{S}_N$ (Conic hull of subspace ranks)
4. Network Coding/Distributed Storage Rate Regions

(**Our Point**) There were $>$ four jokes... but only one punchline

$$\boxed{\text{Shannon} \rightarrow \text{linear equality/inequality} \bigcap \rightarrow \text{project}}$$

Thus, let's have a look at the general computational structure in this agenda.

Outline

1. Entropic Vectors and Polyhedral Computation
   (a) One Procedure to Rule Four Problems
      i. Non-Shannon outer bounds for $\bar{\Gamma}_N^*$ (Yeung, Zhang, DFZ, Matúš, ...)
      ii. Vector Matroidal Inner bound for $\bar{\Gamma}_N^*$ (Hassibi)
      iii. Linear Code/Subspace Rank Outer Bound (Hammer, Kinser, DFZ)
      iv. Network Coding/Distributed Storage Regions (Yan, Yeung, Chan, Grant)
   (b) **Many Paths Lead to the Same Truth**
   (c) The Path Less Laden: Complexity Experiments & Moving Forward
2. Structure of $\bar{\Gamma}_4^*$ & the Case for Non-Polyhedral Tools
3. Characterizing Extremal Entropic Vectors with Information Geometry
   (a) Introduction to Information Geometry
   (b) Decoupling Constraints & Improving Ingleton w/ Binary Inner Bound
   (c) Casting Entropic Vectors as Information Projections
   (d) Information Geometric Properties of Distributions on Shannon Facets
   (e) $\bar{\Gamma}_4^*$ & 4 atom distribution. Information Geometry of Ingleton Violation

Many Paths Lead to the Same Truth

∃ wide variety of techniques (yielding mathematically equivalent results) for each step:

list of extreme rays

Shannon Outer Bound	list of inequalities		representation conversion		known list of non-isomorphic matroids
Adding linear inequalities/ equalities	append		representation conversion		run intermediate double descriptions step to update ray list with new inequalties
Project onto a subset of variables	Convex Hull Method	Fourier Motzkin	representation conversion	remove elements & convex hull	Extreme Point Method
Representation Conversion	Double Descriptions		Lexicographic Reverse Search	Exploit partial pairs/ lists?	Exploit Symmetries? sympol

Parallelization: message passing VS massively parallel GPU

Outline

1. Entropic Vectors and Polyhedral Computation
   (a) One Procedure to Rule Four Problems
      i. Non-Shannon outer bounds for $\bar{\Gamma}_N^*$ (Yeung, Zhang, DFZ, Matúš, ...)
      ii. Vector Matroidal Inner bound for $\bar{\Gamma}_N^*$ (Hassibi)
      iii. Linear Code/Subspace Rank Outer Bound (Hammer, Kinser, DFZ)
      iv. Network Coding/Distributed Storage Regions (Yan, Yeung, Chan, Grant)
   (b) Many Paths Lead to the Same Truth
   (c) **The Path Less Laden: Complexity Experiments & Moving Forward**
2. Structure of $\bar{\Gamma}_4^*$ & the Case for Non-Polyhedral Tools
3. Characterizing Extremal Entropic Vectors with Information Geometry
   (a) Introduction to Information Geometry
   (b) Decoupling Constraints & Improving Ingleton w/ Binary Inner Bound
   (c) Casting Entropic Vectors as Information Projections
   (d) Information Geometric Properties of Distributions on Shannon Facets
   (e) $\bar{\Gamma}_4^*$ & 4 atom distribution. Information Geometry of Ingleton Violation

The Path Less Laden: Complexity Experiments

- mathematically equal $\neq$ computationally equal

Γ_N^{bin}	$N = 5$	6	7	8
Algorithm 2	1.3 s	11 s	150 s	3800 s
Algorithm 3	1.6 s	45 s	3000 s	36000 s
Algorithm 4	1.4 s	23 s	2000 s	25000 s

- Algorithm 2 uses substitution of conditional entropies into lower dimensional forbidden minor set to get inequality representation.
- Algorithm 3 & 4 work with non-isomorphic matroid list and remove minors (4: minor checking, 3: using 2's inequalities) to get extreme representation.

The Path Less Laden: Complexity Experiments

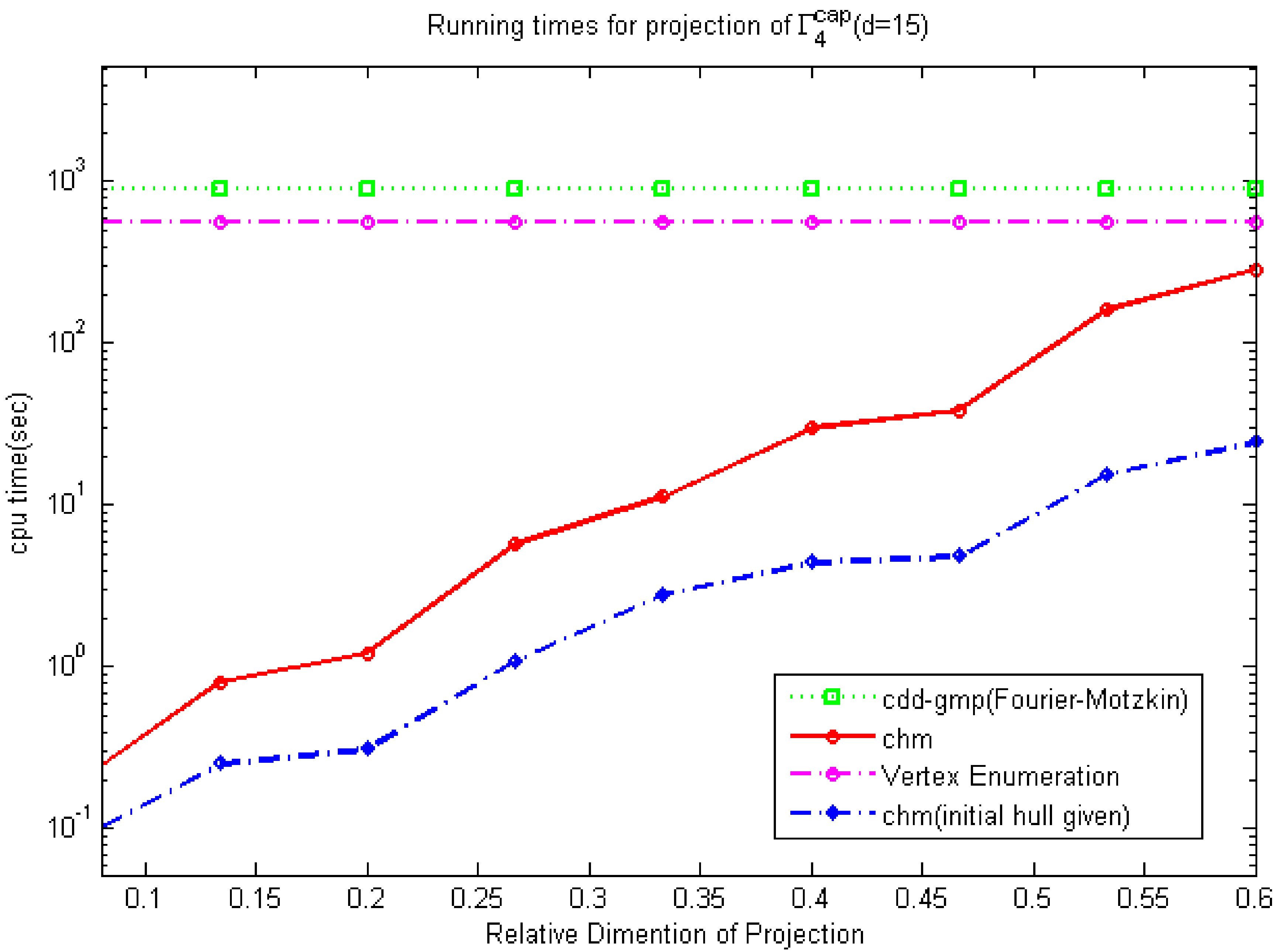
Rate Region	2-level-3-encoder	3-level-3-encoder
Algorithm 6	46 s	3600 s
Algorithm 7	2.9 s	47 s
Algorithm 8	2.7 s	35 s

- Algorithm 6: uses algorithm 2 to get inequalities of inner bound, appends rate region equalities/inequalities, and projects using Fourier Motzkin. (inequality based)

- Algorithm 7 & 8 : adds rate regions ineq.s & eq.s to alg. 3 and 4. inner bound extreme rays via steps of double descriptions, then projects

total time winners are not always concatenation of the winners at each stage

The Path Less Laden: Complexity Experiments



The Path Less Laden: Moving Forward

- There is an absurd amount of symmetry in these problems!
  - Labeling of variables, but also
  - Shannon is only one inequality! conditional mutual info ≥ 0
- Key factor in the computation is exploiting the known symmetry to ease the computation
- 2010 Diploma Thesis from Thomas Rehn Univ. Magdeburg (now at Uni. Rostock) – sympol
  - "representation conversion up to symmetries". D. Bremner & A. Schurmann
- Also, given that there is no unique "best path" and "best representation", yet a rich theory regarding which algorithms are good for which structures, need parallel tools that try the right candidates out then select which ones finish first
  - area of active software development (what our team is developing)
- given proof nature of results, need infinite precision arithmetic

The Path Less Laden: Moving Forward

What we envision:

- Polyhedral Computation Package
- Entropy Vector Bound Package
- Rate Region Package

Would you want to use this? Please tell me what you think about this, as well as any useful features I may have missed, after the talk.

Outline

1. Entropic Vectors and Polyhedral Computation
   (a) One Procedure to Rule Four Problems
      i. Non-Shannon outer bounds for $\bar{\Gamma}_N^*$ (Yeung, Zhang, DFZ, Matúš, ...)
      ii. Vector Matroidal Inner bound for $\bar{\Gamma}_N^*$ (Hassibi)
      iii. Linear Code/Subspace Rank Outer Bound (Hammer, Kinser, DFZ)
      iv. Network Coding/Distributed Storage Regions (Yan, Yeung, Chan, Grant)
   (b) Many Paths Lead to the Same Truth
   (c) The Path Less Laden: Complexity Experiments & Moving Forward
2. **Structure of $\bar{\Gamma}_4^*$ & the Case for Non-Polyhedral Tools**
3. Characterizing Extremal Entropic Vectors with Information Geometry
   (a) Introduction to Information Geometry
   (b) Decoupling Constraints & Improving Ingleton w/ Binary Inner Bound
   (c) Casting Entropic Vectors as Information Projections
   (d) Information Geometric Properties of Distributions on Shannon Facets
   (e) $\bar{\Gamma}_4^*$ & 4 atom distribution. Information Geometry of Ingleton Violation

Structure of $\bar{\Gamma}_4^*$:

- ....Okay.... so, contradiction shows it's not polyhedral. (Matúš ISIT 2007.)
- But, how bad is it?
- In Matúš 1995 Conditional Independence Relations Paper:
  - series showed which Γ_4 faces have entropic point in relative interior, but also
  - Lemma 4: Gap between Shannon & Ingleton: 6 "pyramids" $\mathcal{P}_{ij} := \Gamma_4 \cap \{\text{Ingleton}_{ij} \leq 0\}$. Also, $\bar{\mathcal{P}}_{ij}^* := \bar{\Gamma}_4^* \cap \{\text{Ingleton}_{ij} \leq 0\}$
    - Only one of the 6 Ingletons can be violated at once

$$\Gamma_4 = \mathcal{I} \cup \cup_{ij} \mathcal{P}_{ij} \quad \bar{\Gamma}_4^* = \mathcal{I} \cup \cup_{ij} \bar{\mathcal{P}}_{ij}^* \tag{4}$$

    - Each of the Pyramids $\mathcal{P}_{ij}$: 1 non entropic extreme ray and 15 *binary* entropic extreme rays
- What can we infer about $\bar{\mathcal{P}}_4^* \cap \{\text{Ingleton}_{ij} \leq 0\}$ from these ingredients?

Structure of $\bar{\Gamma}_4^*$: Hey, maybe it's not so bad

h_1	h_2	h_{12}	h_3	h_{13}	h_{23}	h_{123}	h_4	h_{14}	h_{24}	h_{124}	h_{34}	h_{134}	h_{234}	h_{1234}
2	2	3	2	3	3	4	2	3	3	4	4	4	4	4
1	0	1	1	1	1	1	0	1	0	1	1	1	1	1
0	1	1	1	1	2	2	1	1	2	2	2	2	2	2
1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
0	1	1	0	0	1	1	1	1	1	1	1	1	1	1
1	0	1	0	1	0	1	1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	0	0	1	1	1	1	1	1
1	0	1	1	2	1	2	1	2	1	2	2	2	2	2
0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
1	1	2	1	2	2	3	1	2	2	3	2	3	3	3
0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

h_1	h_2	h_{12}	h_3	h_{13}	h_{23}	h_{123}	h_4	h_{14}	h_{24}	h_{124}	h_{34}	h_{134}	h_{234}	h_{1234}
0	0	0	1	0	0	0	1	0	0	0	−1	0	0	0
−1	0	0	0	1	0	0	0	1	0	0	0	−1	0	0
−1	0	1	0	0	0	0	0	1	0	−1	0	0	0	0
−1	0	1	0	1	0	−1	0	0	0	0	0	0	0	0
0	−1	0	0	0	1	0	0	0	1	0	0	0	−1	0
0	−1	1	0	0	0	0	0	0	1	−1	0	0	0	0
0	−1	1	0	0	1	−1	0	0	0	0	0	0	0	0
0	0	0	−1	1	1	−1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	−1	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	−1	1	1	−1	0	0	0	0
0	0	0	0	0	0	0	0	0	0	−1	0	0	0	1
0	0	0	0	0	0	0	0	0	0	0	−1	1	1	−1
0	0	0	0	0	0	0	0	0	0	0	0	−1	0	1
0	0	0	0	0	0	0	0	0	0	0	0	0	−1	1
1	1	−1	0	−1	−1	1	0	−1	−1	1	1	0	0	0

- Drop/Project out, e.g. $H(X_1)$ from $\mathcal{P}_{ij}$: the bad ray falls into the conic hull of the good ones $\Longrightarrow$ $\pi_{\setminus \mathcal{A}} \bar{\mathcal{P}}_4^* = \pi_{\setminus \mathcal{A}} \mathcal{P}_{ij}$
- happens if you drop any one of the 10 entropies $h_{\mathcal{A}} \in \mathrm{Ingleton}_{ij}$
- Dropping this entropy makes $\pi_{\setminus \mathcal{A}} \bar{\mathcal{P}}_{ij}^* = \{\mathbf{h}_{\setminus \mathcal{A}} | \mathbf{A}\mathbf{h}_{\setminus \mathcal{A}} \leq \mathbf{b}\}$ polyhedral

Structure of $\bar{\Gamma}_4^*$:

- Implication: for any $\mathcal{A}$ s.t. $h_{\mathcal{A}} \in \mathrm{Ingleton}_{ij}$, one way to express $\bar{\mathcal{P}}_{i,j}^*$ is

$$\bar{\mathcal{P}}_{i,j}^* = \left\{ \mathbf{h} \in \mathbb{R}^{15} \;\middle|\; \begin{array}{c} \mathbf{A}\mathbf{h}_{\setminus\mathcal{A}} \leq \mathbf{b} \ (= \mathrm{Shannon}) \\ h_{\mathcal{A}} \geq g_{\mathrm{low}}(\mathbf{h}_{\setminus\mathcal{A}}) \\ h_{\mathcal{A}} \leq g_{\mathrm{up}}(\mathbf{h}_{\setminus\mathcal{A}}) \end{array} \right\} \tag{5}$$

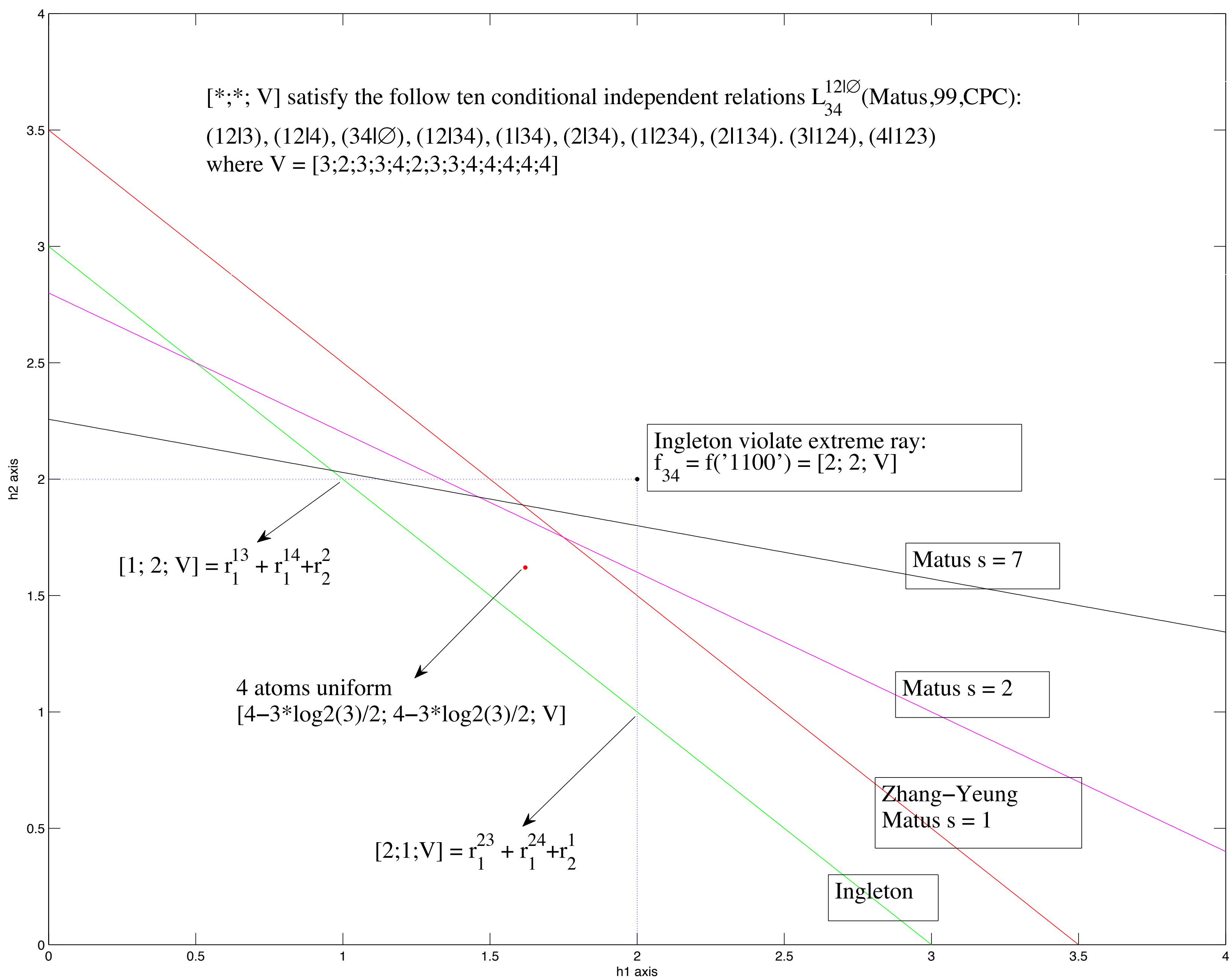
- Sign of $h_{\mathcal{A}} \in \mathrm{Ingleton}_{ij} \implies$ one of $g_{\mathrm{low}}(\mathbf{h}_{\setminus\mathcal{A}})$ or $g_{\mathrm{up}}(\mathbf{h}_{\setminus\mathcal{A}})$ from $\mathrm{Ingleton}_{ij} = 0$
- E.g. $\mathcal{A} = \{1\} \implies$
  $g_{\mathrm{low}}(\mathbf{h}_{\setminus 1}) = -h_2 + h_{12} + h_{23} + h_{13} - h_{123} + h_{14} + h_{24} - h_{124} - h_{34}$

The problem of determining $\bar{\Gamma}_4^*$ is equivalent, e.g., to determining a *single nonlinear function* $g_{\mathrm{up}} : \pi_{\setminus 1}\mathcal{P}_{12} \to \mathbb{R}_+$

$$\boxed{g_{\mathrm{up}}(\mathbf{h}_{\setminus 1}) := \max_{h_1 \;\mid\; \left[h_1, \mathbf{h}_{\setminus 1}^T\right]^T \in \Gamma_4^* \cap \{\mathrm{Ingleton}_{12} \leq 0\}} h_1}$$

(the solution to an optimization problem)

Structure of $\bar{\Gamma}_4^*$: Example



Structure of $\bar{\Gamma}_4^*$: Dropping h_{123}

h_1	h_2	h_{12}	h_3	h_{13}	h_{23}	h_{123}	h_4	h_{14}	h_{24}	h_{124}	h_{34}	h_{134}	h_{234}	h_{1234}
2	2	3	2	3	3	4	2	3	3	4	4	4	4	4
1	0	1	1	1	1	1	0	1	0	1	1	1	1	1
0	1	1	1	1	2	2	1	1	2	2	2	2	2	2
1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
0	1	1	0	0	1	1	1	1	1	1	1	1	1	1
1	0	1	0	1	0	1	1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	0	0	1	1	1	1	1	1
1	0	1	1	2	1	2	1	2	1	2	2	2	2	2
0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
1	1	2	1	2	2	3	1	2	2	3	2	3	3	3
0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

h_1	h_2	h_{12}	h_3	h_{13}	h_{23}	h_{123}	h_4	h_{14}	h_{24}	h_{124}	h_{34}	h_{134}	h_{234}	h_{1234}
0	0	0	1	0	0	0	1	0	0	0	−1	0	0	0
−1	0	0	0	1	0	0	0	1	0	0	0	−1	0	0
−1	0	1	0	0	0	0	0	1	0	−1	0	0	0	0
−1	0	1	0	1	0	−1	0	0	0	0	0	0	0	0
0	−1	0	0	0	1	0	0	0	1	0	0	0	−1	0
0	−1	1	0	0	0	0	0	0	1	−1	0	0	0	0
0	−1	1	0	0	1	−1	0	0	0	0	0	0	0	0
0	0	0	−1	1	1	−1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	−1	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	−1	1	1	−1	0	0	0	0
0	0	0	0	0	0	0	0	0	0	−1	0	0	0	1
0	0	0	0	0	0	0	0	0	0	0	−1	1	1	−1
0	0	0	0	0	0	0	0	0	0	0	0	−1	0	1
0	0	0	0	0	0	0	0	0	0	0	0	0	−1	1
1	1	−1	0	−1	−1	1	0	−1	−1	1	1	0	0	0

$$\bar{\mathcal{P}}_{12}^* = \left\{ \begin{array}{c} \mathbf{A}\mathbf{h}_{\backslash 123} \leq \mathbf{b} \ (= \text{Shannon}) \\ h_{123} \geq g_{\text{low}}(\mathbf{h}_{\backslash 123}) \\ h_{123} \leq g_{\text{up}}(\mathbf{h}_{\backslash 123}) \end{array} \right\}$$

$$g_{\text{low}}(\mathbf{h}_{\backslash 123}) = -h_1 - h_2 + h_{12} + h_{23} + h_{13} + h_{14} + h_{24} - h_{124} - h_{34}$$

The problem of determining $\bar{\Gamma}_4^*$ is equivalent, e.g., to determining a single nonlinear function:

$g_{\text{up}} : \pi_{\backslash 123}\mathcal{P}_{12} \to \mathbb{R}_+$

$$g_{\text{up}}(\mathbf{h}_{\backslash 123}) := \max_{\left[h_{123}, \mathbf{h}_{\backslash 123}^T\right]^T \in \bar{\mathcal{P}}_{12}^*} h_{123}$$

E.g. Shannon says $g_{\text{up}}(\mathbf{h}_{\backslash 123}) \leq \min\{h_{2|1} + h_{13}, h_{2|3} + h_{13}, h_{1|2} + h_{23}, h_{1234}\}$

The lists of non-Shannon inequalities make the list of linear equations in the $\min$ larger.

The Case for Non-Polyhedral Tools:

Since $\bar{\Gamma}_N^*$ is a non-polyhedral convex cone:

- Need a tool that is not limited to (tightening of) polyhedral bounds
- Need a tool to handle non-linear codes!: $\mathcal{S}_N \subsetneq \bar{\Gamma}_N^* \; \forall N \geq 4$
  - Bye bye (linear) representable matroids. more general matroids promising path, but:
    - a pain to enumerate (list gigantic and unknown $N \geq 10$)
    - discrete $\implies$ conic hulls & rep. conv. nec. for REV. also expensive
    - algebraic matroids are far less understood than representable. other tools?

Extreme rays of $\bar{\Gamma}_N^*$ and $\cap$ correspond to efficient codes, hence:

- Want to parameterize the EVs and PMFs on boundary of $\bar{\Gamma}_N^*$
  - (esp. new extreme rays not shared with Shannon)

Information geometry:

- endows differential geometric structure to set of joint PMFs (parameterizations!)
- coord.'s flatness & certain affine sets = familiar properties
  - marginals, independence, conditional independence
- studies divergences (incl. KL) & projections that are easily related to entropy

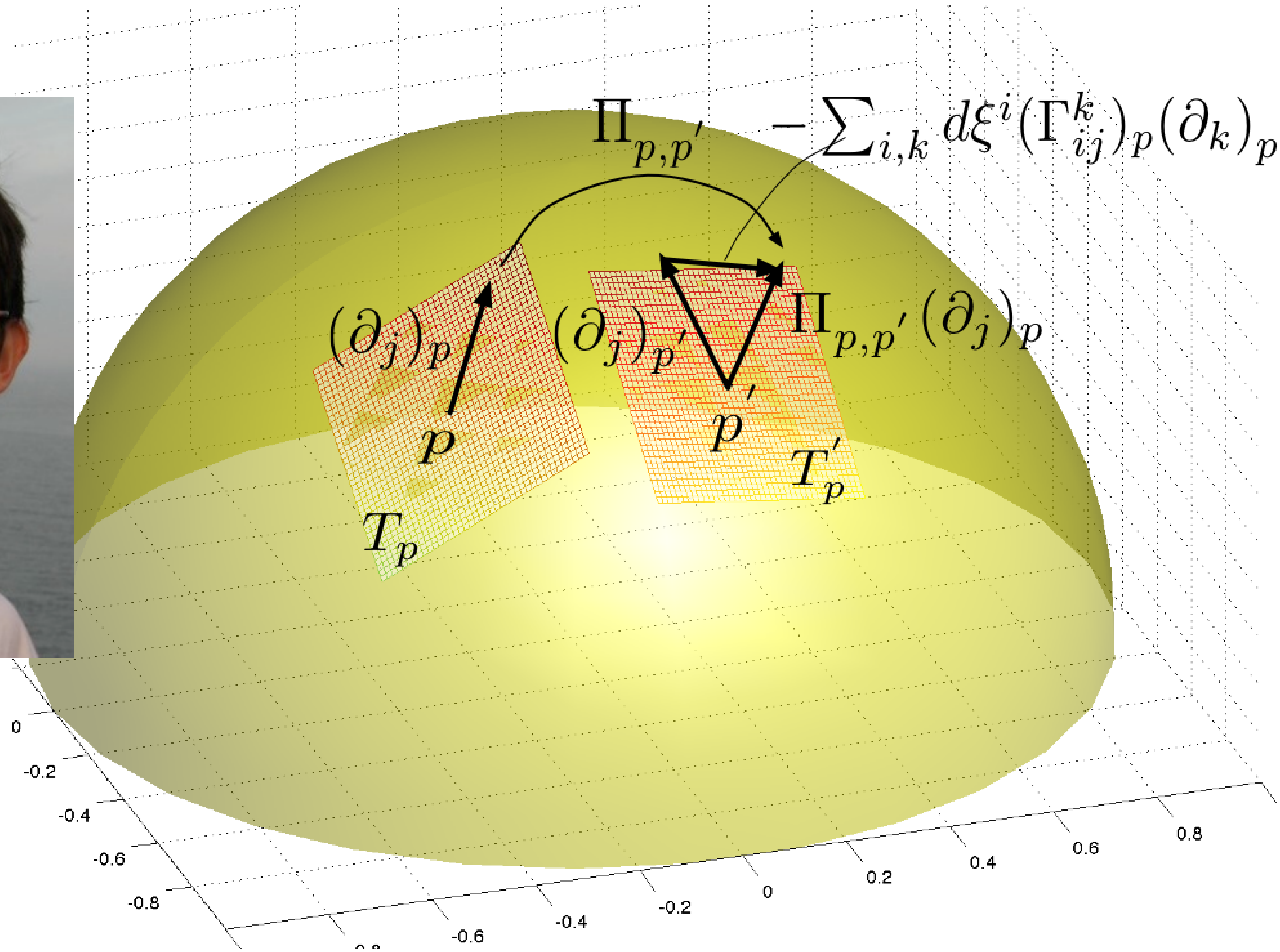
Hence, information geometry seems a potential candidate to deal with these questions

Outline

1. Entropic Vectors and Polyhedral Computation
   (a) One Procedure to Rule Four Problems
      i. Non-Shannon outer bounds for $\bar{\Gamma}_N^*$ (Yeung, Zhang, DFZ, Matúš, ...)
      ii. Vector Matroidal Inner bound for $\bar{\Gamma}_N^*$ (Hassibi)
      iii. Linear Code/Subspace Rank Outer Bound (Hammer, Kinser, DFZ)
      iv. Network Coding/Distributed Storage Regions (Yan, Yeung, Chan, Grant)
   (b) Many Paths Lead to the Same Truth
   (c) The Path Less Laden: Complexity Experiments & Moving Forward
2. Structure of $\bar{\Gamma}_4^*$ & the Case for Non-Polyhedral Tools
3. **Characterizing Extremal Entropic Vectors with Information Geometry**
   (a) Introduction to Information Geometry
   (b) Decoupling Constraints & Improving Ingleton w/ Binary Inner Bound
   (c) Casting Entropic Vectors as Information Projections
   (d) Information Geometric Properties of Distributions on Shannon Facets
   (e) $\bar{\Gamma}_4^*$ & 4 atom distribution. Information Geometry of Ingleton Violation

Characterizing Extremal Entropic Vectors with Information Geometry (Lead Student)

Yunshu Liu



Outline

1. Entropic Vectors and Polyhedral Computation
   (a) One Procedure to Rule Four Problems
      i. Non-Shannon outer bounds for $\bar{\Gamma}_N^*$ (Yeung, Zhang, DFZ, Matúš, ...)
      ii. Vector Matroidal Inner bound for $\bar{\Gamma}_N^*$ (Hassibi)
      iii. Linear Code/Subspace Rank Outer Bound (Hammer, Kinser, DFZ)
      iv. Network Coding/Distributed Storage Regions (Yan, Yeung, Chan, Grant)
   (b) Many Paths Lead to the Same Truth
   (c) The Path Less Laden: Complexity Experiments & Moving Forward
2. Structure of $\bar{\Gamma}_4^*$ & the Case for Non-Polyhedral Tools
3. Characterizing Extremal Entropic Vectors with Information Geometry
   (a) **Introduction to Information Geometry**
   (b) Decoupling Constraints & Improving Ingleton w/ Binary Inner Bound
   (c) Casting Entropic Vectors as Information Projections
   (d) Information Geometric Properties of Distributions on Shannon Facets
   (e) $\bar{\Gamma}_4^*$ & 4 atom distribution. Information Geometry of Ingleton Violation

(Formal) Introduction to Information Geometry [20] – Notation

- Overall idea: treat family of probability distributions as a differentiable manifold: $p(x;\xi)$ is parameterized by ξ
- Endow w/ Riemannian metric (inner product between Tangent vectors) given by Fisher Information Matrix $g_{i,j}(\xi) = \mathbb{E}_\xi[\partial_i \ell_\xi \partial_j \ell_\xi]$ w/ $\ell_\xi = \log p(x;\xi)$, $\partial_i = \frac{\partial}{\partial \xi_i}$.
- Select α-affine connections $\nabla^{(\alpha)}$ such that $\left\langle \nabla^{(\alpha)}_{\partial_i} \partial_j, \partial_k \right\rangle = \Gamma^{(\alpha)}_{ij,k}$

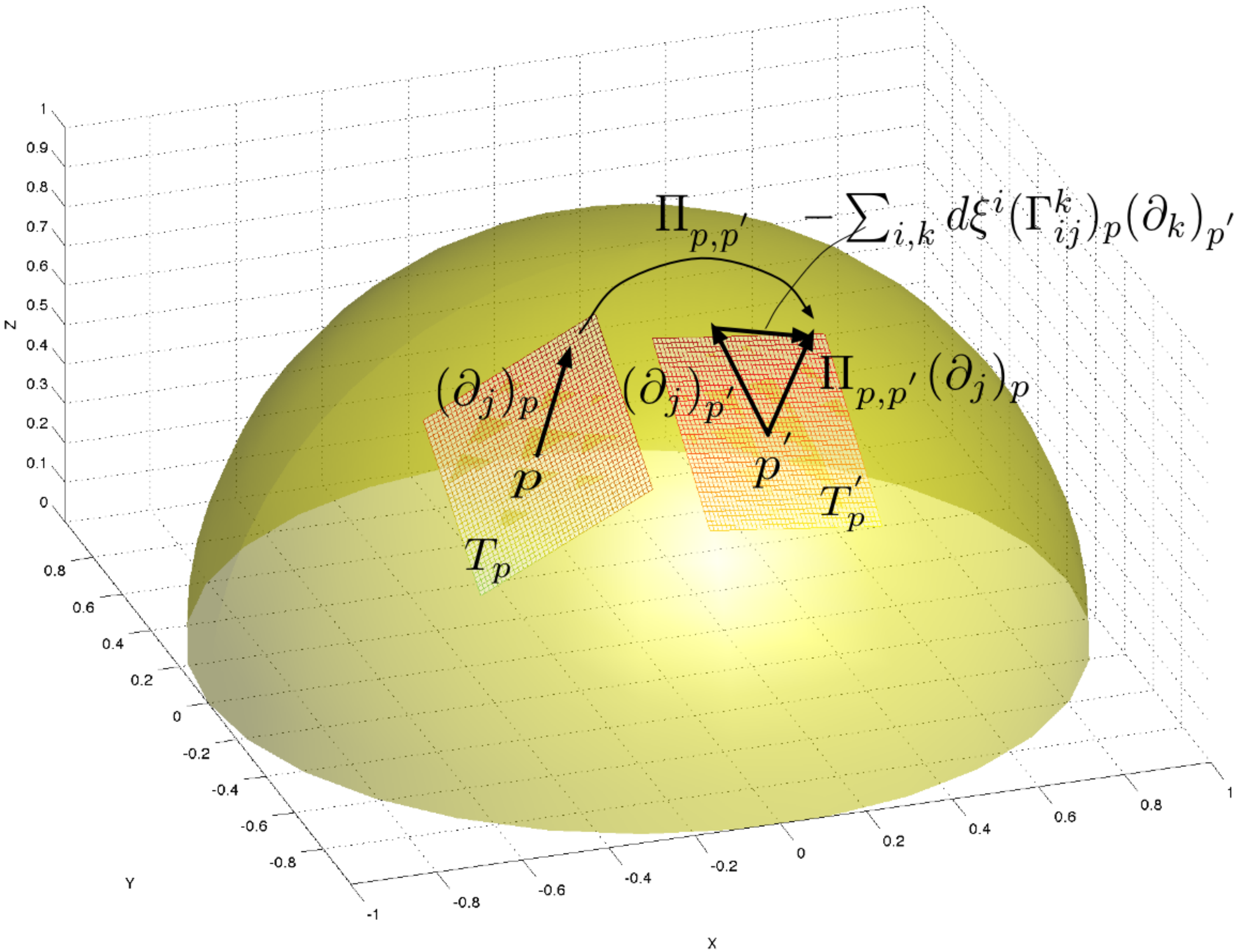
$$\Gamma^{(\alpha)}_{ij,k} = \mathbb{E}\left[\left(\partial_i \partial_j \ell_\xi + \frac{1-\alpha}{2} \partial_i \ell_\xi \partial_j \ell_\xi\right)(\partial_k \ell_\xi)\right] \tag{6}$$

- purpose of affine connection: define parallel translation $\Pi_{p,p'} : T_p \to T_{p'}$ to correspond tangent vectors along curves $\gamma : [a,b] \to \mathcal{P}$

$$\Pi_{\gamma(t),\gamma(t+dt)}(X(t)) = \sum_{ijk} \left\{ X^k(t) - dt \dot{\gamma}^i(t) X^j(t) \left(\Gamma_{ij,k}\right)_{\gamma(t)} \right\} (\partial_k)_{\gamma(t+dt)} \tag{7}$$

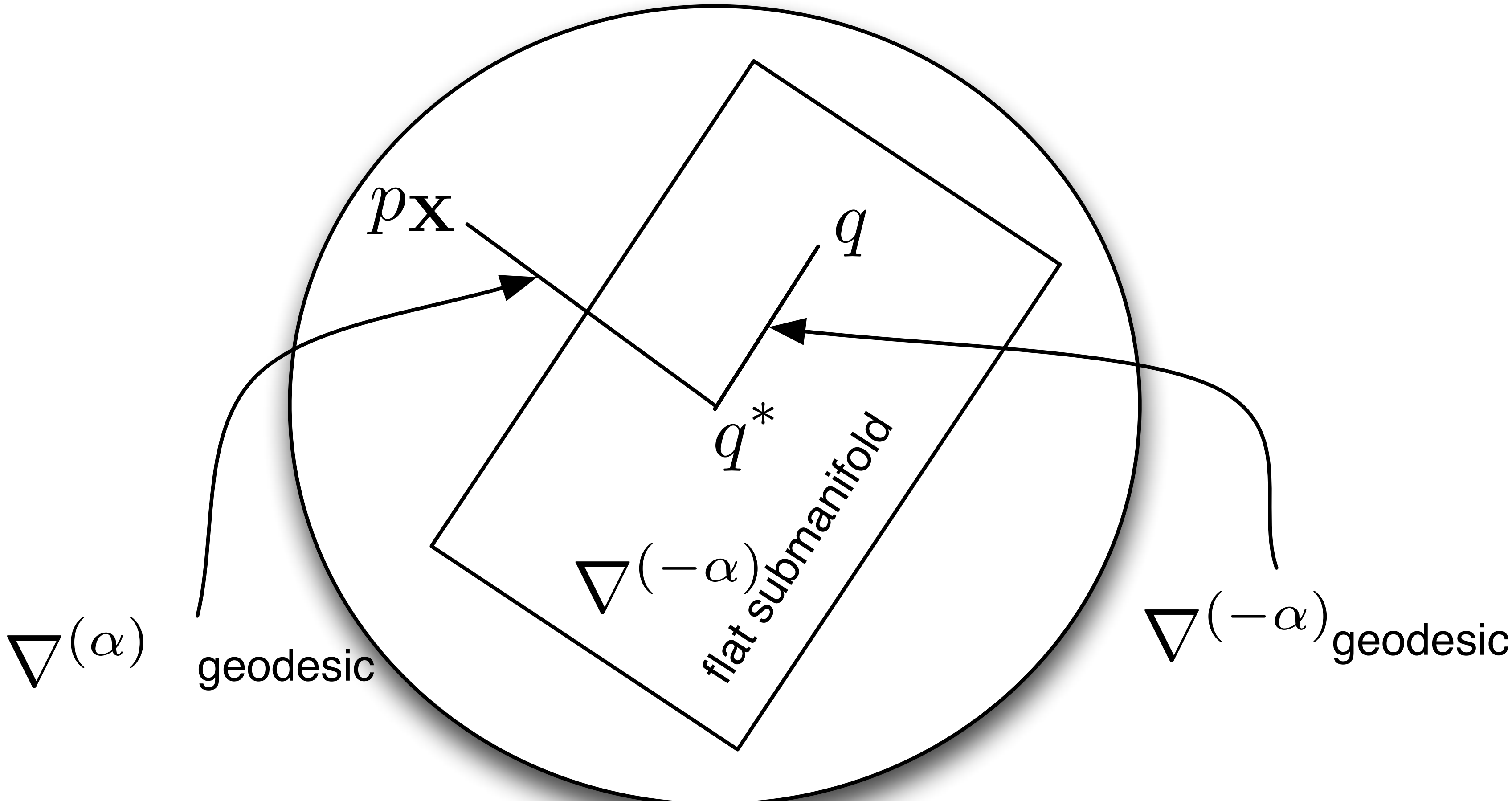
- Curve w/ tangent vector transported by parallel transl. w/ $\nabla^{(\alpha)}$ is $\nabla^{(\alpha)}$ *geodesic*
- If there is a coordinate system in which every parallel translation under $\nabla^{(\alpha)}$ leaves coefficients in Tangent vector unchanged, the manifold is said to be α-flat, and associated coordinate system is an affine coordinate system.
- $\nabla^{(\alpha)}$ has property $\langle X, Y \rangle_p = \langle \Pi^{(\alpha)}_{p,p'}(X), \Pi^{(-\alpha)}_{p,p'}(Y) \rangle_{p'}$

(Formal) Introduction to Information Geometry [20] – Parallel Translation



$$\nabla_{\partial_i}\partial_j = \sum_k \Gamma_{ij,k}\partial_k \quad \Gamma_{ij,k} = 0 \text{ if "flat"} \tag{8}$$

(Formal) Introduction to Information Geometry [20] – Information Projection



$$D^{(\alpha)}(p_{\mathbf{x}}||q) = D^{(\alpha)}(p_{\mathbf{x}}||q^*) + D^{(\alpha)}(q^*||q)$$

(Informal) Introduction to Information Geometry [20] – Examples: Coordinates

m-coordinates:

$$\boldsymbol{\eta} = \left[p_{\mathbf{X}}(v_{i_1,1}, \ldots, v_{i_N,N}) \middle| \; i_k \in \{2, \ldots, |\mathcal{X}_k|\}, k \in \{1, \ldots, N\} \; \right]$$

e-coordinates: $\prod_{n=1}^{N} |\mathcal{X}_n| - 1$ elements take the form

$$\boldsymbol{\theta} = \left[\log\left(\frac{p_{\mathbf{X}}(v_{i_1,1}, \ldots, v_{i_N,N})}{p_{\mathbf{X}}(v_{1,1}, \ldots, v_{1,N})} \right) \middle| \begin{array}{c} i_k \in \{2, \ldots, |\mathcal{X}_k|\}, \\ k \in \{1, \ldots, N\} \end{array} \right]$$

m-autoparallel submanifold (affine subset of m-coords) fix $\mathbf{A}, \mathbf{b}$ all $\boldsymbol{\eta}$ of the form

$$\boldsymbol{\eta} = \mathbf{A}\boldsymbol{p} + \mathbf{b} \tag{9}$$

e-autoparallel submanifold (affine subset of e-coords) fix $\mathbf{A}, \mathbf{b}$ all $\boldsymbol{\theta}$ of the form

$$\boldsymbol{\theta} = \mathbf{A}\boldsymbol{\lambda} + \mathbf{b} \tag{10}$$

properties of affine sets $\implies$ intersections also affine, thus e/m affine closed under intersection.

e-geodesic/m-geodesic: one dimensional affine manifolds

(Informal) Introduction to Information Geometry [20] – Examples: Affine Sets

Examples of e-autoparallel submanifold:

- Set of joint distributions $p_{X,Y}$ s.t. X, Y indep.
- Set of joint distributions $p_{X,Y,Z}$ s.t. X,Y,Z indep. (etc)
- Set of joint distributions s.t. $X \leftrightarrow Y \leftrightarrow Z$

Examples of m-autoparallel submanifold

- Set of joint distributions $p_{X,Y}$ with a particular marginal distribution p_X
- Set of joint distributions $p_{X,Y}$ with a particular marginal distributions p_X, p_Y

(Informal) Introduction to Information Geometry [20] – Examples: Projections

- **e-flat submanifold:** set of all product distributions

$$\mathcal{E}_0 = \left\{ p_{\mathbf{X}} \middle| p_{\mathbf{X}}(x_1, \ldots, x_N) = \prod_{i=1}^{N} p_{X_i}(x_i) \right\} \tag{11}$$

- **m-flat submanifold:** set of joint distributions with given marginals

$$\mathcal{M}_0 = \left\{ p_{\mathbf{X}} \middle| \sum_{\mathbf{x}_{\setminus i}} p_{\mathbf{X}}(\mathbf{x}) = q_i(x_i) \quad \forall i \in \{1, \ldots, N\} \right\} \tag{12}$$

- **Information Projections & Pythagorean Relation:**

$$q^* = \arg\min_{q \in \mathcal{E}_0} D(p_{\mathbf{X}}||q), \quad D(p_{\mathbf{X}}||q) = D(p_{\mathbf{X}}||q^*) + D(q^*||q) \; \forall q \in \mathcal{E}_0 \tag{13}$$

$$q^* = \arg\min_{q \in \mathcal{M}_0} D(q||p_{\mathbf{X}}), \quad D(q||p_{\mathbf{X}}) = D(q^*||p_{\mathbf{X}}) + D(q||q^*) \; \forall q \in \mathcal{M}_0 \tag{14}$$

Information Geometry [20] – What has it been used for?

- re-interpretation of EM algorithm [20]
- acceleration of Blahut Arimoto algorithm [21]
- learning algorithms in Neural Networks [22]
- analysis of Belief propagation & Turbo Decoding [23, 24, 25, 26]

Outline

1. Entropic Vectors and Polyhedral Computation
   (a) One Procedure to Rule Four Problems
      i. Non-Shannon outer bounds for $\bar{\Gamma}^*_N$ (Yeung, Zhang, DFZ, Matúš, ...)
      ii. Vector Matroidal Inner bound for $\bar{\Gamma}^*_N$ (Hassibi)
      iii. Linear Code/Subspace Rank Outer Bound (Hammer, Kinser, DFZ)
      iv. Network Coding/Distributed Storage Regions (Yan, Yeung, Chan, Grant)
   (b) Many Paths Lead to the Same Truth
   (c) The Path Less Laden: Complexity Experiments & Moving Forward
2. Structure of $\bar{\Gamma}^*_4$ & the Case for Non-Polyhedral Tools
3. Characterizing Extremal Entropic Vectors with Information Geometry
   (a) Introduction to Information Geometry
   (b) **Decoupling Constraints & Improving Ingleton w/ Binary Inner Bound**
   (c) Casting Entropic Vectors as Information Projections
   (d) Information Geometric Properties of Distributions on Shannon Facets
   (e) $\bar{\Gamma}^*_4$ & 4 atom distribution. Information Geometry of Ingleton Violation

Decoupling Constraints & Improving Ingleton w/ Binary Inner Bound

$$g_{\text{up}}(\mathbf{h}^o_{\setminus 123}) = \max_{\mathbf{h}\in\bar{\mathcal{P}}^*_{12} \big| \mathbf{h}_{\setminus 123} = \boldsymbol{h}^o_{\setminus 123}} h_{123} \textcolor{red}{\geq} \max_{\alpha_k \geq 0,\ \boldsymbol{\mathcal{X}},\ p^k_{\mathbf{X}} | \sum_k \alpha_k \mathbf{h}_{\setminus 123}(p^k_{\mathbf{X}}) = \mathbf{h}^o_{123}} \sum_k \alpha_k\ h_{123}(p^k_{\mathbf{X}})$$

$$= \max_{\alpha_k \geq 0,\ \boldsymbol{\mathcal{X}},\ \{p^k_{\mathbf{X}_{\mathcal{A}}} | \mathcal{A}\subset[4]\} \big| \sum_k \alpha_k H(p^k_{\mathbf{X}_{\mathcal{A}}}) = h^o_{\mathcal{A}},\ \sum_{\mathcal{A}^c} p^k_{\mathbf{X}} = p^k_{\mathbf{X}_{\mathcal{A}}} \forall \mathcal{A}\subset[4]} \sum_k \alpha_k H(p^k_{\mathbf{X}_{123}}) =$$

$$\max_{\substack{\alpha_k \geq 0, \boldsymbol{\mathcal{X}}, \\ p^k_{\mathbf{X}_{\mathcal{A}}} | \mathcal{A} \neq \{123\}, [4]\}}} \Bigg| \substack{\sum_k \alpha_k H(p^k_{\mathbf{X}_{\mathcal{A}}}) = h^o_{\mathcal{A}} \\ \exists p^k_{\mathbf{X}}, \sum_{\mathcal{A}^c} p^k_{\mathbf{X}} = p^k_{\mathbf{X}_{\mathcal{A}}}} \quad \max_{p^k_{\mathbf{X}_{123}}, p^k_{\mathbf{X}}} \Bigg| \substack{\sum_{\mathcal{A}^c} p^k_{\mathbf{X}} = p^k_{\mathbf{X}_{\mathcal{A}}}, \\ \sum_k \alpha_k H(p^k_{\mathbf{X}}) = h^o_{[4]}} \quad \sum_k \alpha_k H(p^k_{123})$$

(Think red term is actually equality. Matúš?) If we restrict domain to $\pi_{\setminus 123}\Phi_4$, restrict to a single fixed non-zero α $k = 1$ and $\boldsymbol{X} = \{0, 1\}$, then outer optimization has calculable finite # of points in feasible set [27]. Relies on handy m-affine re-parametrization that decouples the marginal constraints

$$q_{\mathcal{A}} = \mathbb{P}[\mathbf{X}_{\mathcal{A}} = \mathbf{1}_{|\mathcal{A}|}] \quad p_{\mathcal{A}}(\boldsymbol{x}_{\mathcal{A}}) = \sum_{\mathcal{C}|\mathcal{A}\subseteq\mathcal{C}\subseteq\mathcal{I}(\boldsymbol{x}_{\mathcal{A}})} (-1)^{|\mathcal{C}|-|\mathcal{I}(\boldsymbol{x}_{\mathcal{A}})|} q_{\mathcal{B}} \qquad (15)$$

which makes $h_{\mathcal{A}} = f(\{q_{\mathcal{B}} | \mathcal{B} \subseteq \mathcal{A}\})$ and turns solving inner optimization into only two parameter optimization problem: q_{123} and q_{1234}, i.e. can calculate this lower bound for g_{up}.

Decoupling Constraints & Improving Ingleton w/ Binary Inner Bound

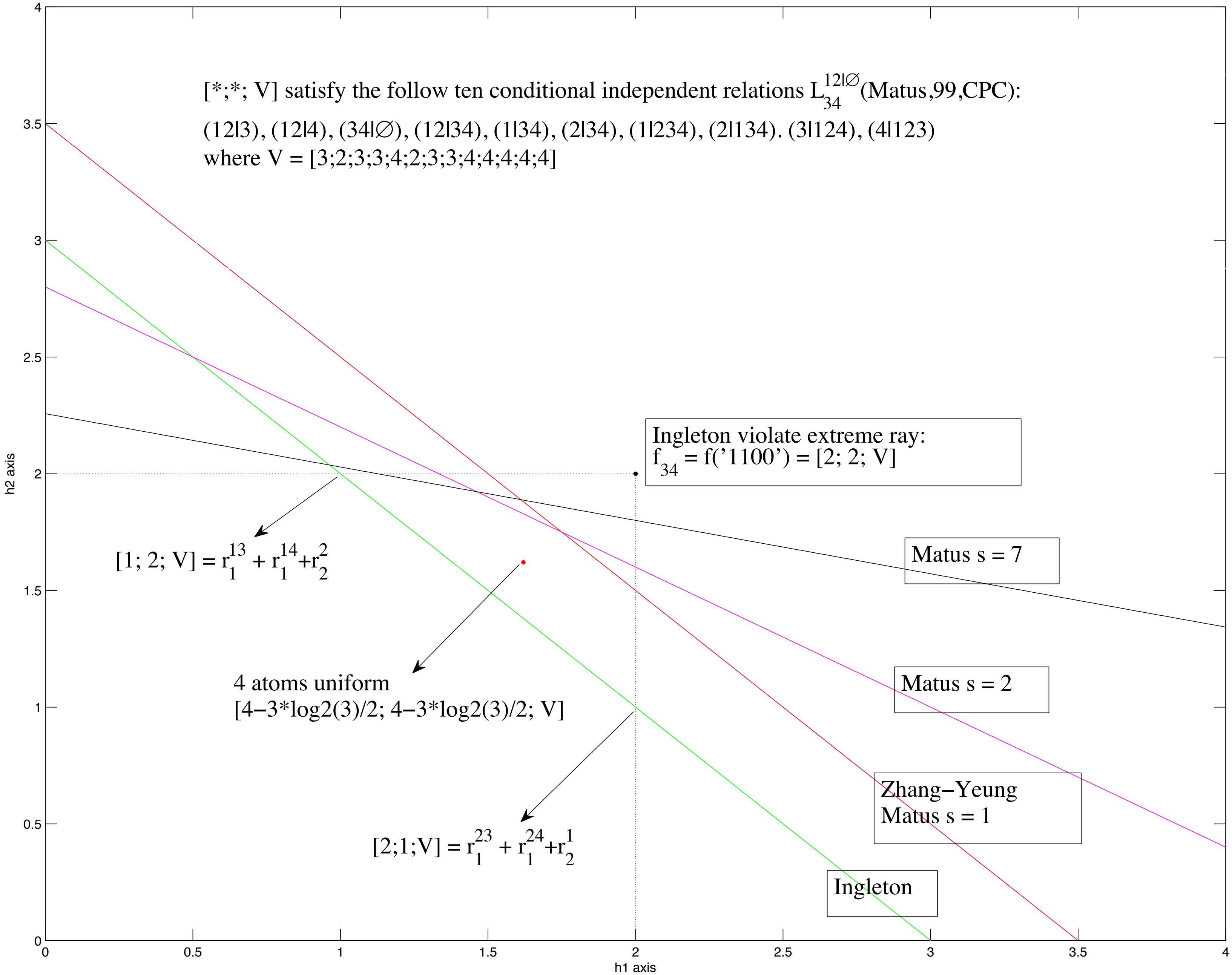
Why this may be a pretty good bound:

- All entropic extreme rays of $\mathcal{P}_{12}$ (Shannon rays on bottom of pyramid) are *binary*.
- Many/most Ingleton violating constructions have made use of non-unif binary r.v.s
- DFZ 4-atom conjecture about maximal Ingleton violation.

Moving forward

- The decoupling trick can be placed in an information geometric framework and generalized beyond binary.
- Inner optimization is almost convex (only one convex equality constraint is the problem). Just a little more transformation?

Structure of $\bar{\Gamma}_4^*$: Example



Outline

1. Entropic Vectors and Polyhedral Computation
   (a) One Procedure to Rule Four Problems
      i. Non-Shannon outer bounds for $\bar{\Gamma}^*_N$ (Yeung, Zhang, DFZ, Matúš, ...)
      ii. Vector Matroidal Inner bound for $\bar{\Gamma}^*_N$ (Hassibi)
      iii. Linear Code/Subspace Rank Outer Bound (Hammer, Kinser, DFZ)
      iv. Network Coding/Distributed Storage Regions (Yan, Yeung, Chan, Grant)
   (b) Many Paths Lead to the Same Truth
   (c) The Path Less Laden: Complexity Experiments & Moving Forward
2. Structure of $\bar{\Gamma}^*_4$ & the Case for Non-Polyhedral Tools
3. Characterizing Extremal Entropic Vectors with Information Geometry
   (a) Introduction to Information Geometry
   (b) Decoupling Constraints & Improving Ingleton w/ Binary Inner Bound
   (c) **Casting Entropic Vectors as Information Projections**
   (d) Information Geometric Properties of Distributions on Shannon Facets
   (e) $\bar{\Gamma}^*_4$ & 4 atom distribution. Information Geometry of Ingleton Violation

Casting Entropic Vectors as Information Projections

Easy to relate Shannon entropy to rel. entropy/ KL Divergence:

$$
\begin{aligned}
D(p_{\mathbf{X}} || \mathcal{U}_{|\boldsymbol{\mathcal{X}}|}) &= \sum_{\mathbf{x} \in \boldsymbol{\mathcal{X}}} p_{\mathbf{X}}(\mathbf{x}) \log_2 \left(\frac{p_{\mathbf{X}}(\mathbf{x})}{1/|\boldsymbol{\mathcal{X}}|} \right) && (16) \\
&= \log_2(|\boldsymbol{\mathcal{X}}|) - H(p_{\mathbf{X}}) = H(\mathcal{U}_{\boldsymbol{\mathcal{X}}}) - H(p_{\mathbf{X}}) && (17)
\end{aligned}
$$

Casting Entropic Vectors as Information Projections

Next consider the family of distributions

$$\mathcal{H}_i := \left\{ p_{\mathbf{X}} \middle| p(\mathbf{X}) = \frac{1}{|\mathcal{X}_i|} q(\mathbf{X}_{\setminus i}), \text{ some } q(\mathbf{X}_{\setminus i}) \right\} \tag{18}$$

Observe:

- $\mathcal{U}_{\boldsymbol{\mathcal{X}}} \in \mathcal{H}_i$
- $\mathcal{H}_i$ is *both* an e-affine and m-affine submanifold.
- Defining $q^*_{\mathcal{H}_i}(p_{\mathbf{X}}) = \arg\min_{q \in \mathcal{H}_i} D(p_{\mathbf{X}} || q)$, have Pythagorean relation:

$$D(p_{\mathbf{X}} || \mathcal{U}_{\boldsymbol{\mathcal{X}}}) = \underbrace{D(p_{\mathbf{X}} || q^*_{\mathcal{H}_i}(p_{\mathbf{X}}))}_{\log_2 |\mathcal{X}_i| - H(X_i | \mathbf{X}_{\setminus i})} + \underbrace{D(q^*_{\mathcal{H}_i}(p_{\mathbf{X}}) || \mathcal{U}_{\mathcal{X}})}_{\log_2 |\boldsymbol{\mathcal{X}}| - \log_2 |\mathcal{X}_i| - H(\mathbf{X}_{\setminus i})} \tag{19}$$

(erm... $H(\mathbf{X}) = H(X_i) + H(\mathbf{X}_{\setminus i} | X_i)$ tyco)

Moving this around, we have

$$H(\mathbf{X}_{\setminus i}) = D(p_{\mathbf{X}} || q^*_{\mathcal{H}_i}(p_{\mathbf{X}})) - D(p_{\mathbf{X}} || \mathcal{U}_{\boldsymbol{\mathcal{X}}}) + \log_2 |\boldsymbol{\mathcal{X}}| - \log_2 |\mathcal{X}_i| \tag{20}$$

Casting Entropic Vectors as Information Projections

Generalizing this idea, consider the family of distributions

$$\bigcap_{i \in \mathcal{A}^c} \mathcal{H}_i = \left\{ p_{\mathbf{X}} = \frac{q(\mathbf{X}_{\mathcal{A}})}{\prod_{i \in \mathcal{A}^c} |\mathcal{X}_i|} \right\} \tag{21}$$

Observe:

- $\mathcal{U}_{\boldsymbol{\mathcal{X}}} \in \bigcap_{i \in \mathcal{A}^c} \mathcal{H}_i$
- $\bigcap_{i \in \mathcal{A}^c} \mathcal{H}_i$ is *both* an e-affine and m-affine submanifold
- Defining $q^*_{\mathcal{A}}(p_{\mathbf{X}}) = \arg\min_{q \in \bigcap_{i \in \mathcal{A}^c} \mathcal{H}_i} D(p_{\mathbf{X}} || q)$, have Pythagorean relation:

$$D(p_{\mathbf{X}} || \mathcal{U}_{\boldsymbol{\mathcal{X}}}) = \underbrace{D(p_{\mathbf{X}} || q^*_{\mathcal{A}}(p_{\mathbf{X}}))}_{\sum_{i \in \mathcal{A}^c} \log_2 |\mathcal{X}_i| - H(\mathbf{X}_{\mathcal{A}^c} | \mathbf{X}_{\mathcal{A}})} + \underbrace{D(q^*_{\mathcal{A}}(p_{\mathbf{X}}) || \mathcal{U}_{\boldsymbol{\mathcal{X}}})}_{\log_2 |\boldsymbol{\mathcal{X}}| - \sum_{i \in \mathcal{A}^c} \log_2 |\mathcal{X}_i| - H(\mathbf{X}_{\mathcal{A}})} \tag{22}$$

(erm... $H(\mathbf{X}) = H(\mathbf{X}_{\mathcal{A}}) + H(\mathbf{X}_{\mathcal{A}^c} | \mathbf{X}_{\mathcal{A}})$ tyco)

From which we observe that

$$H(\mathbf{X}_{\mathcal{A}}) = D(p_{\mathbf{X}} || q^*_{\mathcal{A}}(p_{\mathbf{X}})) - D(p_{\mathbf{X}} || \mathcal{U}_{\boldsymbol{\mathcal{X}}}) - \sum_{i \in \mathcal{A}^c} \log_2 |\mathcal{X}_i| + \log_2 |\boldsymbol{\mathcal{X}}| \tag{23}$$

Casting Entropic Vectors as Information Projections

Defining the set function (then stack into a vector $\mathbf{d}$)

$$d_{\mathcal{A}} := \min_{q \in \bigcap_{i \in \mathcal{A}^c} \mathcal{H}_i} D(p_{\mathbf{X}} || q) = D(p_{\mathbf{X}} || q_{\mathcal{A}}(p_{\mathbf{X}})) \quad \forall \mathcal{A} \subsetneq \{1, \ldots, N\} =: [N] \tag{24}$$

and $d_{\emptyset} = D(p_{\mathbf{X}} || \mathcal{U}_{\boldsymbol{\mathcal{X}}})$. It is evident from the relation we derived

$$H(\mathbf{X}_{\mathcal{A}}) = D(p_{\mathbf{X}} || q^*_{\mathcal{A}}(p_{\mathbf{X}})) - D(p_{\mathbf{X}} || \mathcal{U}_{\boldsymbol{\mathcal{X}}}) - \sum_{i \in \mathcal{A}^c} \log_2 |\mathcal{X}_i| + \log_2 |\boldsymbol{\mathcal{X}}| \tag{25}$$

that

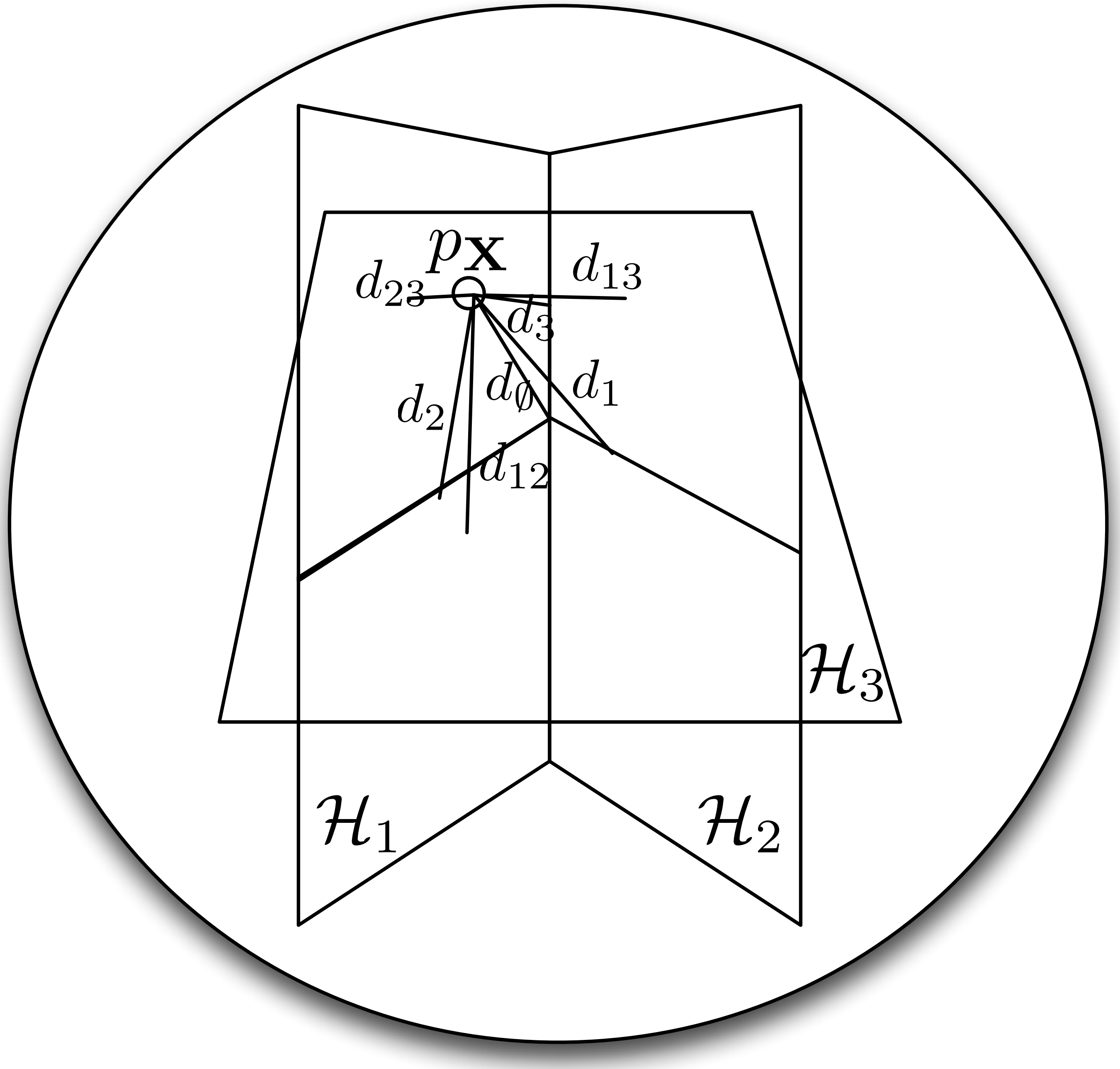
$$h_{\mathcal{A}} = d_{\mathcal{A}} - d_{\emptyset} - \sum_{i \in \mathcal{A}^c} \log_2 |\mathcal{X}_i| + \log_2 |\boldsymbol{\mathcal{X}}| \quad \forall \mathcal{A} \subsetneq [N] \tag{26}$$

and $h_{[N]} = -d_{\emptyset} + \log_2 |\boldsymbol{\mathcal{X}}|$, thus we can express entropic vector in terms of $\mathbf{d}$ via

$$\mathbf{h}(\mathbf{d}) = \mathbf{A}\mathbf{d} + \mathbf{b} \tag{27}$$

Region of entropic vectors is affine transformation of region of simultaneous divergences between submanifolds $\mathcal{H}_i$ and their intersections!

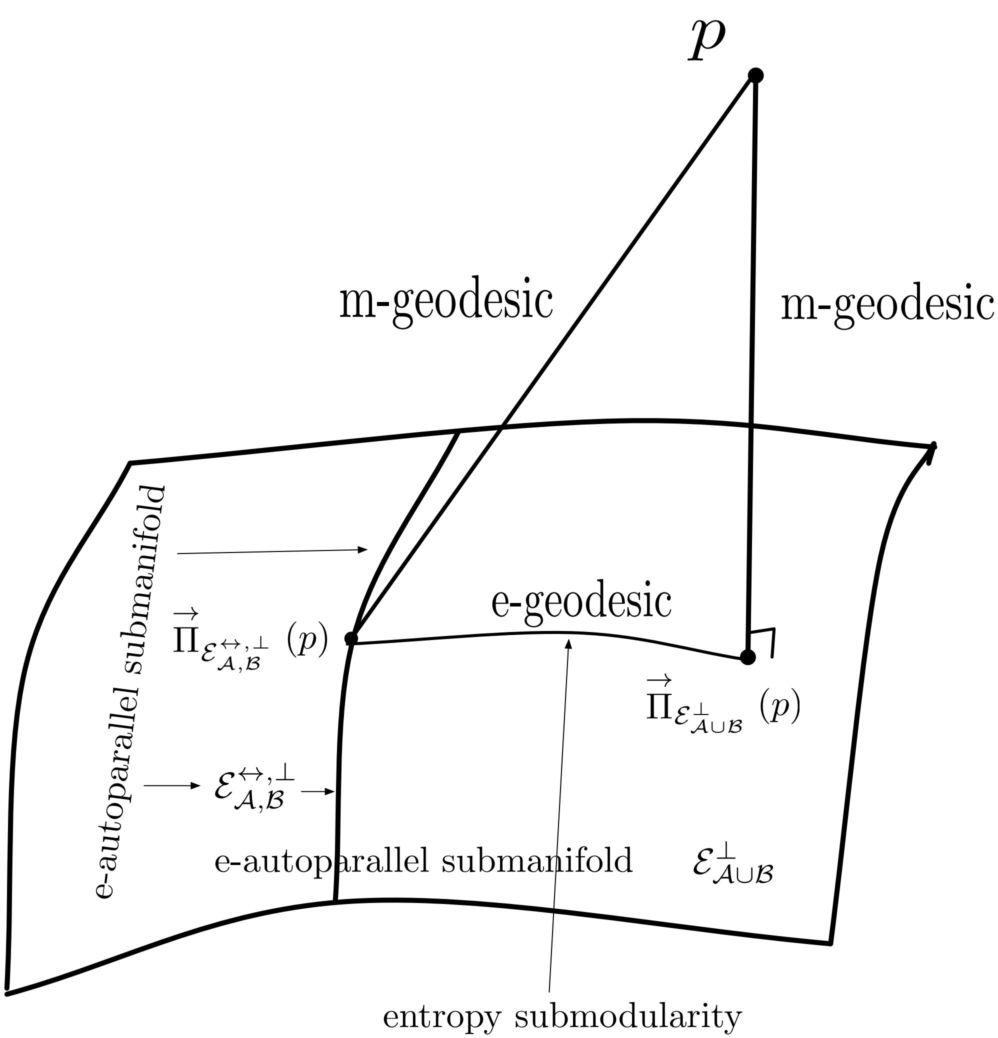
Casting Entropic Vectors as Information Projections



Outline

1. Entropic Vectors and Polyhedral Computation
   (a) One Procedure to Rule Four Problems
      i. Non-Shannon outer bounds for $\bar{\Gamma}_N^*$ (Yeung, Zhang, DFZ, Matúš, ...)
      ii. Vector Matroidal Inner bound for $\bar{\Gamma}_N^*$ (Hassibi)
      iii. Linear Code/Subspace Rank Outer Bound (Hammer, Kinser, DFZ)
      iv. Network Coding/Distributed Storage Regions (Yan, Yeung, Chan, Grant)
   (b) Many Paths Lead to the Same Truth
   (c) The Path Less Laden: Complexity Experiments & Moving Forward
2. Structure of $\bar{\Gamma}_4^*$ & the Case for Non-Polyhedral Tools
3. Characterizing Extremal Entropic Vectors with Information Geometry
   (a) Introduction to Information Geometry
   (b) Decoupling Constraints & Improving Ingleton w/ Binary Inner Bound
   (c) Casting Entropic Vectors as Information Projections
   (d) **Information Geometric Properties of Distributions on Shannon Facets**
   (e) $\bar{\Gamma}_4^*$ & 4 atom distribution. Information Geometry of Ingleton Violation

Information Geometric Properties of Distributions on Shannon Facets



$$h_{\mathcal{A}} + h_{\mathcal{B}} \geq h_{\mathcal{A}\cup\mathcal{B}} + h_{\mathcal{A}\cap\mathcal{B}}$$

$$\mathcal{E}^{\leftrightarrow,\perp}_{\mathcal{A},\mathcal{B}} = \left\{ \boldsymbol{\theta} \middle| p_{\boldsymbol{X}} = p_{\boldsymbol{X}_{\mathcal{A}\setminus\mathcal{B}}|\boldsymbol{X}_{\mathcal{A}\cap\mathcal{B}}} p_{\boldsymbol{X}_{\mathcal{B}}} p_{\boldsymbol{X}_{(\mathcal{A}\cup\mathcal{B})^c}} \right\}$$

$$\mathcal{E}^{\perp}_{\mathcal{A},\mathcal{B}} = \left\{ \boldsymbol{\theta} \middle| p_{\boldsymbol{X}} = p_{\boldsymbol{X}_{\mathcal{A}\cup\mathcal{B}}} p_{\boldsymbol{X}_{(\mathcal{A}\cup\mathcal{B})^c}} \right\}$$

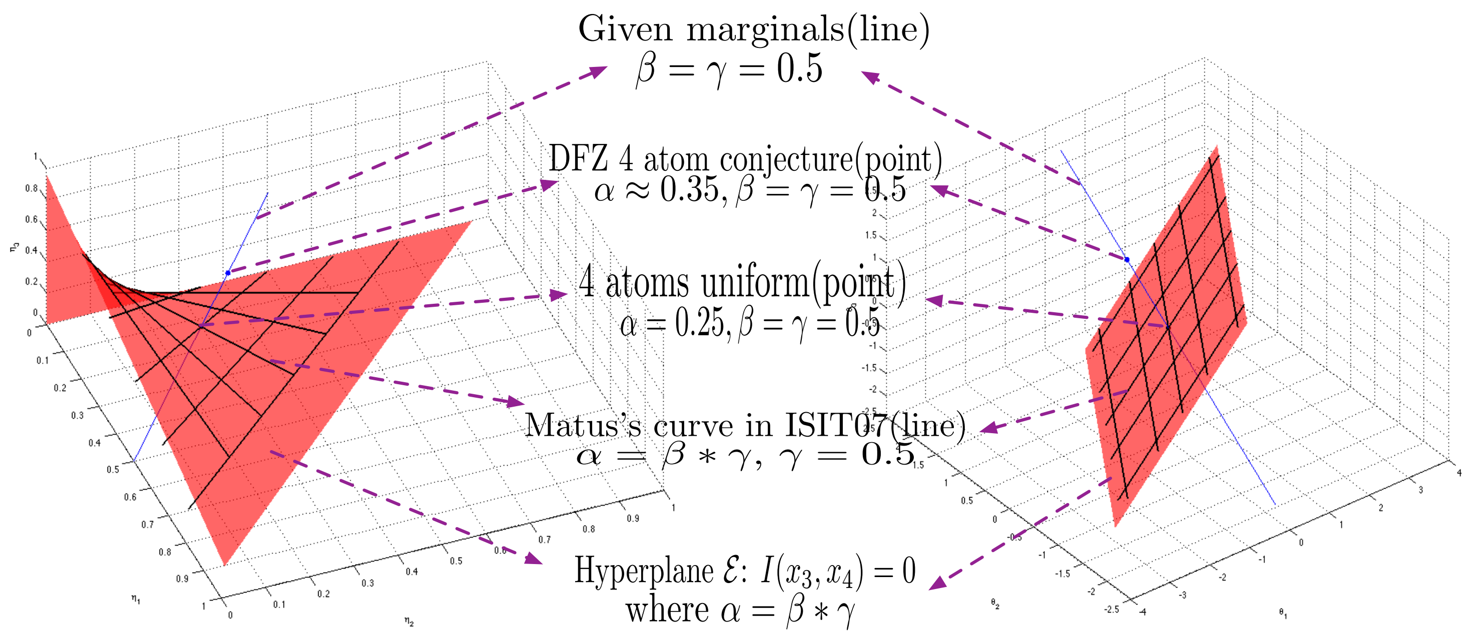
- Shannon outer bound: $I(\boldsymbol{X}_{\mathcal{A}}; \boldsymbol{X}_{\mathcal{B}} | \boldsymbol{X}_{\mathcal{C}}) \geq 0$
- Hence, on the Shannon facet: $I(\boldsymbol{X}_{\mathcal{A}}; \boldsymbol{X}_{\mathcal{B}} | \boldsymbol{X}_{\mathcal{C}}) = 0$
- means $\boldsymbol{X}_{\mathcal{A}} \leftrightarrow \boldsymbol{X}_{\mathcal{C}} \leftrightarrow \boldsymbol{X}_{\mathcal{B}}$
- This is an e-autoparallel submanifold of $p_{\boldsymbol{X}_{\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}}}$!
- $\implies$ those $p_{\boldsymbol{X}_{\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}}}$ on this boundary (affine set) of entropy have a parameterization in which they are also affine (known $\mathbf{A}, \mathbf{b}$)
- Sometimes $\boldsymbol{X} \neq \boldsymbol{X}_{\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}}$, so also need the structure having a particular marginal $p_{\boldsymbol{X}_{\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}}}$ (m-autoparallel)
- mutually dual foliations

Outline

1. Entropic Vectors and Polyhedral Computation
   (a) One Procedure to Rule Four Problems
      i. Non-Shannon outer bounds for $\bar{\Gamma}_N^*$ (Yeung, Zhang, DFZ, Matúš, ...)
      ii. Vector Matroidal Inner bound for $\bar{\Gamma}_N^*$ (Hassibi)
      iii. Linear Code/Subspace Rank Outer Bound (Hammer, Kinser, DFZ)
      iv. Network Coding/Distributed Storage Regions (Yan, Yeung, Chan, Grant)
   (b) Many Paths Lead to the Same Truth
   (c) The Path Less Laden: Complexity Experiments & Moving Forward
2. Structure of $\bar{\Gamma}_4^*$ & the Case for Non-Polyhedral Tools
3. Characterizing Extremal Entropic Vectors with Information Geometry
   (a) Introduction to Information Geometry
   (b) Decoupling Constraints & Improving Ingleton w/ Binary Inner Bound
   (c) Casting Entropic Vectors as Information Projections
   (d) Information Geometric Properties of Distributions on Shannon Facets
   (e) $\bar{\Gamma}_4^*$ **& 4 atom distribution. Information Geometry of Ingleton Violation**

Left: (0000)(0110)(1010)(1111) in m-coordinate

Right: (0000)(0110)(1010)(1111) in e-coordinate



The whole 3D space

$p(0000) = \alpha$

$p(0110) = \beta - \alpha$

$p(1010) = \gamma - \alpha$

$p(1111) = 1 + \alpha - \gamma - \beta$

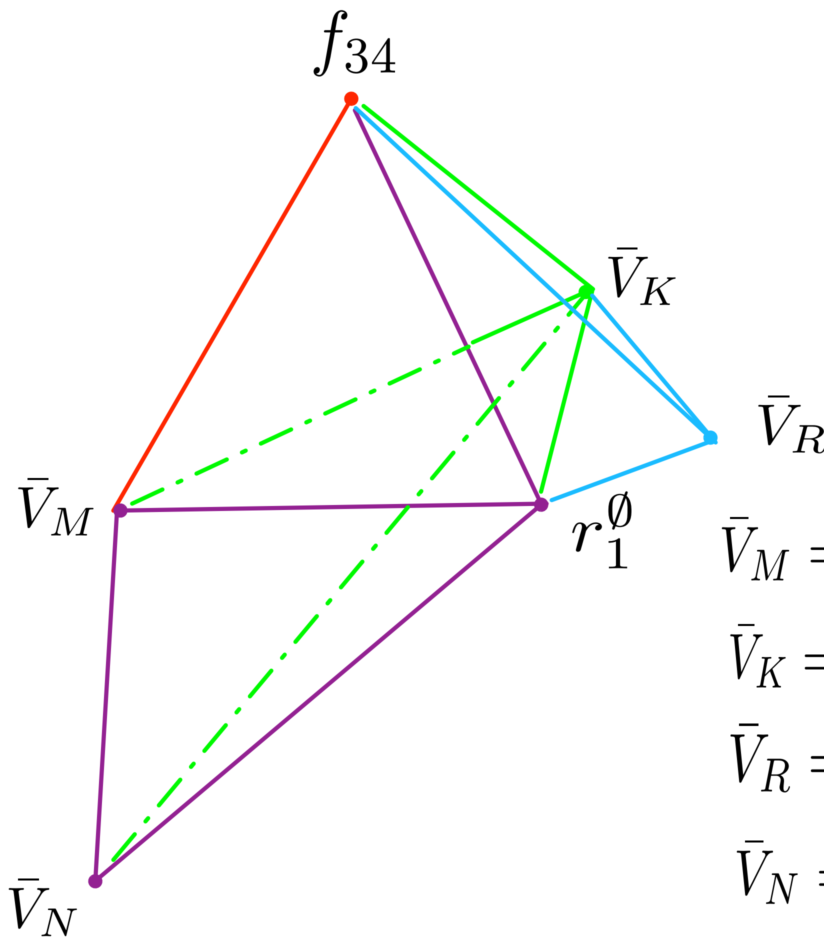
Marginal distribution of x_3 and x_4

$p(x_3 = 0) = \gamma$

$p(x_4 = 0) = \beta$

Structure of $\bar\Gamma_4^*$: Matúš Notation for $\mathcal{P}_{ij}$

$$\begin{aligned}\bar V_P &= (r_1^{\emptyset}, r_1^{3}, r_1^{4}, r_1^{13}, r_1^{14}, r_1^{23}, r_1^{24}, r_1^{123}, r_1^{124}, r_1^{134}, r_1^{234}, r_2^{1}, r_2^{2}, r_3^{\emptyset}, f_{34}) \\ &= (\bar V_M, \bar V_K, \bar V_R, r_1^{\emptyset}, f_{34})\end{aligned}$$

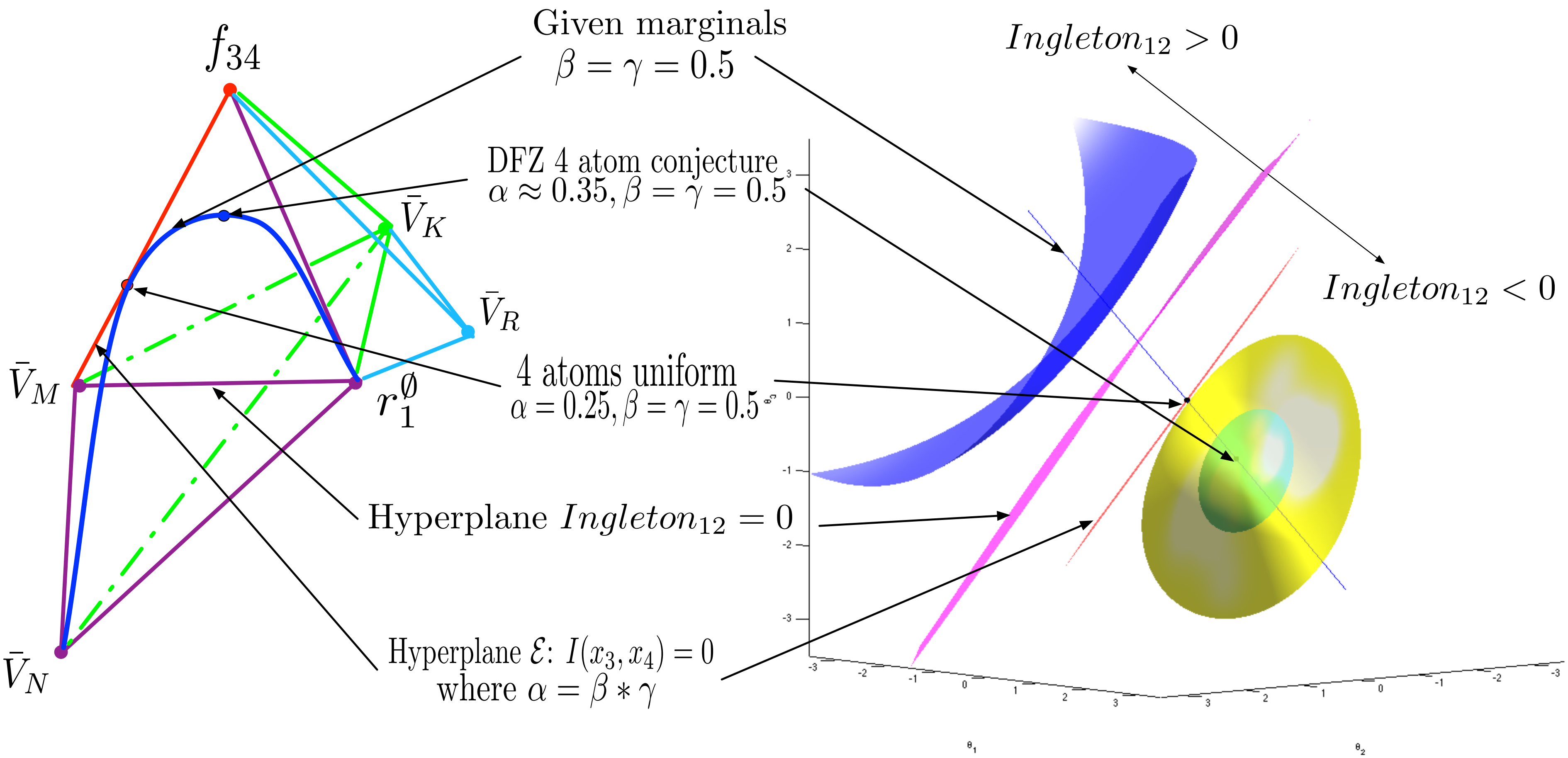


$$\bar V_M = (r_1^{13}, r_1^{14}, r_1^{23}, r_1^{24}, r_2^{1}, r_2^{2})$$

$$\bar V_K = (r_1^{123}, r_1^{124}, r_1^{134}, r_1^{234})$$

$$\bar V_R = (r_1^{3}, r_1^{4}, r_3^{\emptyset})$$

$$\bar V_N = (r_1^{1}, r_1^{2}, r_1^{12})$$



$$\bar{V}_P = (\bar{V}_M, \bar{V}_K, \bar{V}_R, r_1^{\emptyset}, f_{34})$$

where $\bar{V}_M = (r_1^{13}, r_1^{14}, r_1^{23}, r_1^{24}, r_2^{1}, r_2^{2})$ $\quad \bar{V}_K = (r_1^{123}, r_1^{124}, r_1^{134}, r_1^{234})$ $\quad \bar{V}_R = (r_1^{3}, r_1^{4}, r_3^{\emptyset})$ $\quad \bar{V}_N = (r_1^{1}, r_1^{2}, r_1^{12})$

References

[1] Raymond W. Yeung, "A Framework for Linear Information Inequalities," *IEEE Trans. on Information Theory*, vol. 43, no. 6, Nov. 1997.

[2] Zhen Zhang and Raymond W. Yeung, "On Characterization of Entropy Function via Information Inequalities," *IEEE Trans. on Information Theory*, vol. 44, no. 4, Jul. 1998.

[3] ——, "A Non-Shannon-Type Conditional Inequality of Information Quantities," *IEEE Trans. on Information Theory*, vol. 43, no. 6, Nov. 1997.

[4] K. Makarychev, Y. Makarychev, A. Romashchenko, and N. Vereshchagin, "A new class of non-Shannon-type inequalities for entropies," *Communication in Information and Systems*, vol. 2, no. 2, pp. 147–166, December 2002.

[5] Weidong Xu, Jia Wang, Jun Sun, "A projection method for derivation of non-Shannon-type information inequalities," in *IEEE International Symposium on Information Theory (ISIT)*, 2008, pp. 2116 – 2120.

[6] František Matúš, "Infinitely Many Information Inequalities," in *IEEE Int. Symp. Information Theory (ISIT)*, Jun. 2007, pp. 41–44.

[7] Randall Dougherty, Chris Freiling, Kenneth Zeger, "Non-Shannon Information Inequalities in Four Random Variables," Apr. 2011, arXiv:1104.3602v1. [Online]. Available: http://arxiv.org/pdf/1104.3602.pdf

[8] Dillon Mayhew, Gordon F. Royle, "Matroids with nine elements," *Journal of Combinatorial Theory, Series B*, vol. 98, no. 2, pp. 415–431, 2008.

[9] James Oxley, *Matroid Theory, 2nd. Ed.* Oxford University Press, 2011.

[10] W. T. Tutte, "A homotopy theorem for matroids, I, II." *Trans. American Mathematical Society*, vol. 88, pp. 144–174, 1958.

[11] R. E. Bixby, "On Reid's Characterization of the Ternary Matroids," *J. Combin. Theory Ser. B*, no. 26, pp. 174–204, 1979.

[12] P. D. Seymour, "Matroid Representation over GF(3)," *J. Combin. Theory Ser. B*, no. 26, pp. 159–173, 1979.

[13] A. M. H. Gerards, "A short proof of Tutte's characterization of totally unimodular matroids," *Linear Algebra Appl.*, no. 114/115, pp. 207–212, 1989.

[14] Babak Hassibi, Sormeh Shadbakht, Matthew Thill, "On Optimal Design of Network Codes," in *Information Theory and Applications*, UCSD, Feb. 2010, presentation.

[15] A. W. Ingleton, "Representation of Matroids," in *Combinatorial Mathematics and its Applications*, D. J. A. Welsh, Ed. San Diego: Academic Press, 1971, pp. 149–167.

[16] D. Hammer, A. Romashschenko, A. Shen, N. Vereshchagin, "Inequalities for Shannon Entropy and Kolmogorov Complexity," *Journal of Computer and System Sciences*, vol. 60, pp. 442–464, 2000.

[17] F. Matúš and M. Studený, "Conditional Independences among Four Random Variables I," *Combinatorics, Probability and Computing*, no. 4, pp. 269–278, 1995.

[18] Randall Dougherty, Chris Freiling, Kenneth Zeger, "Linear rank inequalities on five or more variables," submitted to SIAM J. Discrete Math. arXiv:0910.0284.

[19] Ryan Kinser, "New Inequalities for Subspace Arrangements," *J. of Comb. Theory Ser. A*, vol. 188, no. 1, pp. 152–161, Jan. 2011.

[20] S. Amari and H. Nagaoka, *Methods of Information Geometry*. American Mathematical Society Translations of Mathematical Monographs, 2004, vol. 191.

[21] G. Matz and P. Duhamel, "Information Geometric Formulation and Interpretation of Accelerated Blahut-Arimoto-Type Algorithms," in *IEEE Information Theory Workshop (ITW-2004)*, Oct. 2004, pp. 66 – 70.

[22] S. Amari, "Information geometry of the EM and em algorithms for neural networks," *Neural Networks*, vol. 8, no. 9, pp. 1379–1408, 1996.

[23] S. Ikeda, T. Tanaka, and S. Amari, "Stochastic reasoning, free energy and information geometry," *Neural Computation*, vol. 16, no. 9, pp. 1779–1810, Sep. 2004.

[24] ——, "Information geometry of turbo and low-density parity-check codes," *IEEE Trans. Inform. Theory*, vol. 50, pp. 1097 – 1114, Jun. 2004.

[25] B. Muquet, P. Duhamel, and M. de Courville, "Geometrical interpretations of iterative 'turbo' decoding," in *Proceedings ISIT*, Jun. 2002.

[26] A. J. Grant, "Information geometry and iterative decoding," in *Proceedings IEEE Communication Theory Workshop*, may 1999.

[27] J. M. Walsh and S. Weber, "A Recursive Construction of the Set of Binary Entropic Vectors and Related Inner Bounds for the Entropy Region," *IEEE Trans. Info. Theory*, vol. 57, no. 10, pp. 6356–6363, Oct. 2011. [Online]. Available: http://dx.doi.org/10.1109/TIT.2011.2165817