Bounding the Entropic Region via Information Geometry

John MacLaren Walsh & Yunshu Liu
Department of Electrical and Computer Engineering
Drexel University
Philadelphia, PA
jwalsh@ece.drexel.edu & yunshu6@gmail.com

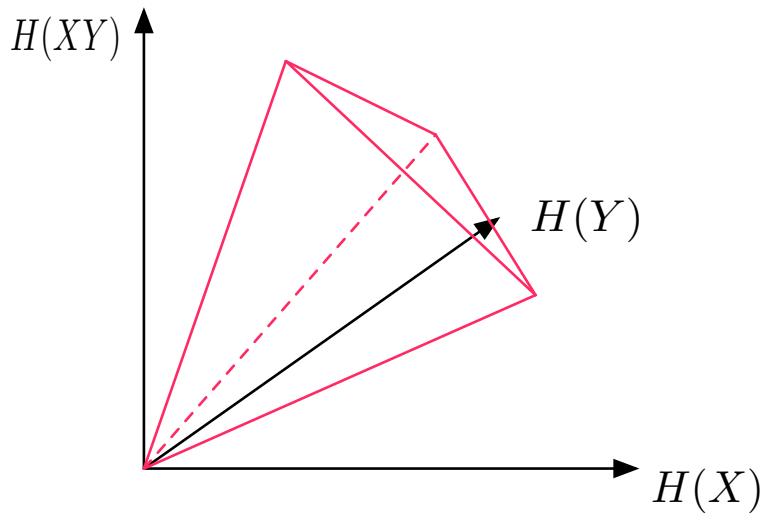
Thanks to NSF CCF-1016588, NSF CCF-1053702, & AFOSR FA9550-12-1-0086.

Outline

1. Entropic Vectors Review: What are they, and why are they important?
2. Entropy Vector Region: What is known/unknown?
3. Structure of the unknown part of $\bar{\Gamma}_4^*$
4. How can one parameterize distributions giving extremal entropic vectors via information geometry?
5. Which distributions give entropy vectors in the unknown part of $\bar{\Gamma}_4^*$?

Region of Entropic Vectors $\bar{\Gamma}_N^*$ – What is it?

1. $\boldsymbol{X} = (X_1, \ldots, X_N)$ N discrete RVs
2. every subset $\boldsymbol{X}_{\mathcal{A}} = (X_i, i \in \mathcal{A})$ $\mathcal{A} \subseteq \{1, \ldots, N\} \equiv [N]$ has joint entropy $h(\boldsymbol{X}_{\mathcal{A}})$.
3. $\boldsymbol{h} = (h(\boldsymbol{X}_{\mathcal{A}}) \,|\mathcal{A} \subseteq [N]) \in \mathbb{R}^{2^N-1}$ *entropic vector*
   - Example: for $N = 3$, $\boldsymbol{h} = (h_1, h_2, h_3, h_{12}, h_{13}, h_{23}, h_{123})$.
4. a $\boldsymbol{h}_o \in \mathbb{R}^{2^N-1}$ is entropic if $\exists$ joint PMF $p_{\boldsymbol{X}}$ s.t. $\boldsymbol{h}(p_{\boldsymbol{X}}) = \boldsymbol{h}_o$.
5. Region of entropic vectors $= \Gamma_N^*$
6. Closure $\bar{\Gamma}_N^*$ is a convex cone [1].



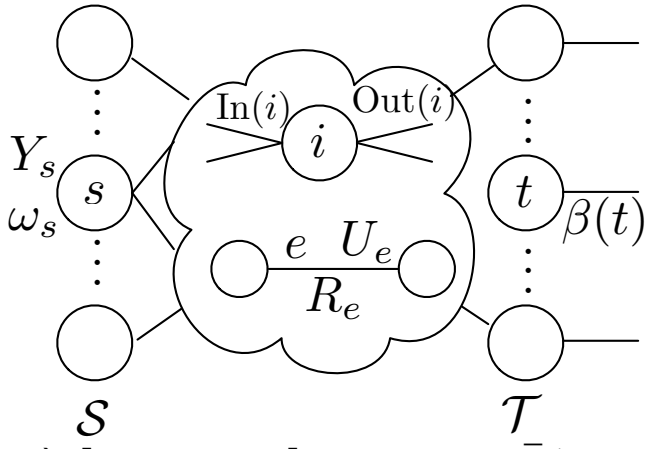
REV for $N = 2$:

$$H(X) \leq H(XY)$$
$$H(Y) \leq H(XY)$$
$$H(XY) \leq H(X) + H(Y)$$

$\bar{\Gamma}_N^*$ *is an* **unknown** *non-polyhedral convex cone for* $N \geq 4$.

Why are Entropic Vectors Important?

Network Coding



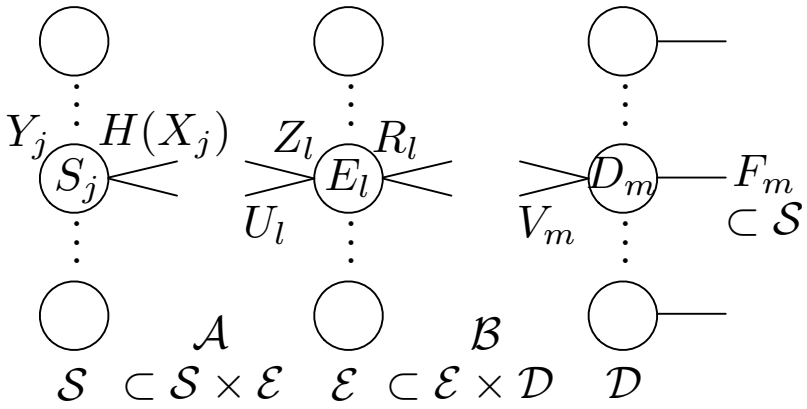
(Roughly) [1, 2, 3, 4] intersect $\bar{\Gamma}_N^*$ with

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

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

and project onto ω_s, R_e

Distributed Storage (MDCS DSCSC)

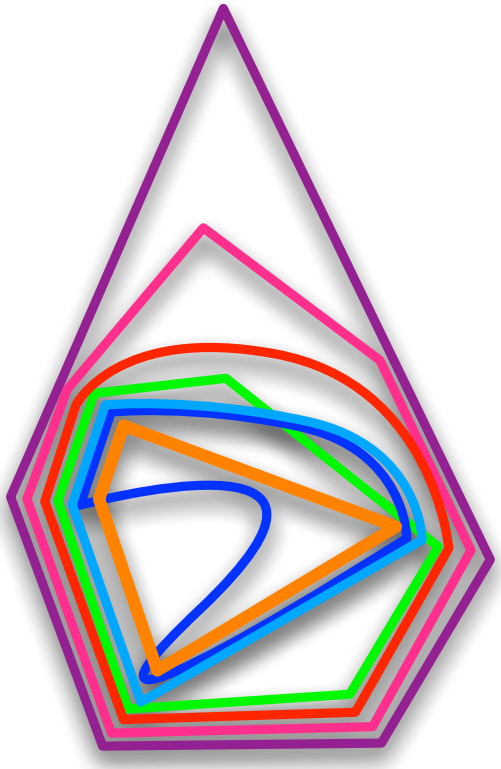


(Roughly) [5] intersect $\bar{\Gamma}_N^*$ with

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

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

Entropic Vector Region: What is known/unknown? – From Outside



Γ_N **Shannon Outer Bound**
$\mathcal{Z}_N$ **Non-Shannon Outer Bound**
$\bar{\Gamma}^*_N$ **Region of Entropic Vectors**
$\mathcal{S}_N$ **Subspace Ranks Bound**
$\mathcal{M}^q_N$ **GF(q)-Representable Matroid Bound**
Φ_4 **binary entropic vectors**
$\mathrm{conv}(\Phi_4)$ **convex hull**

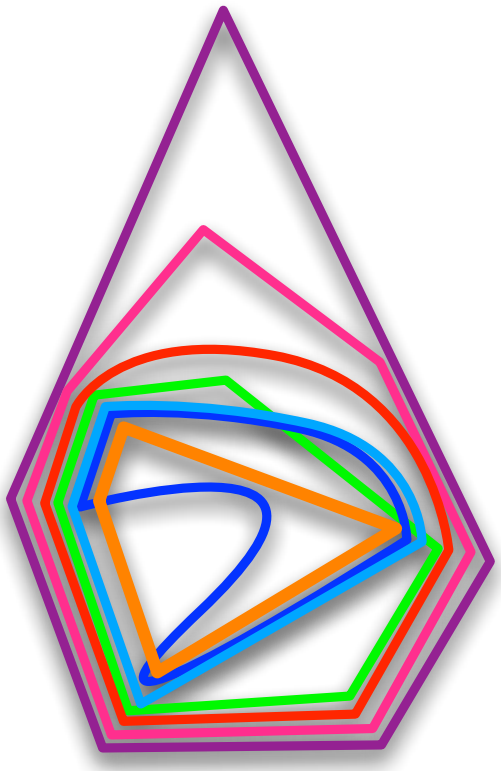
- **Shannon Outer Bound**: Γ_N: entropy is submodular: $I(\boldsymbol{X}_{\mathcal{A}}; \boldsymbol{X}_{\mathcal{B}}|\boldsymbol{X}_{\mathcal{C}}) \geq 0 \quad \forall \mathcal{A}, \mathcal{B}, \mathcal{C}$. A subset of these, the *elemental inequalities* $\Rightarrow$ the rest

$$I(X_i; X_j|\boldsymbol{X}_{\mathcal{K}}) \geq 0, \quad H(X_i|\boldsymbol{X}_{\mathcal{N}\setminus\{i\}}) \geq 0 \quad (1)$$

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

- **Non-Shannon Outer**:[6, 7, 8, 9, 10, 11, 12]
  Zhang & Yeung, 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!
- **Infinite Number of *Linear* Information Inequalities**: [11] Matus a sequence of these inequalities & a curve in $\bar{\Gamma}^*_4 \Rightarrow \bar{\Gamma}^*_N$ is non-polyhedral convex cone for $N \geq 4$.

Entropic Vector Region: What is known/unknown? – From Inside (Linear)

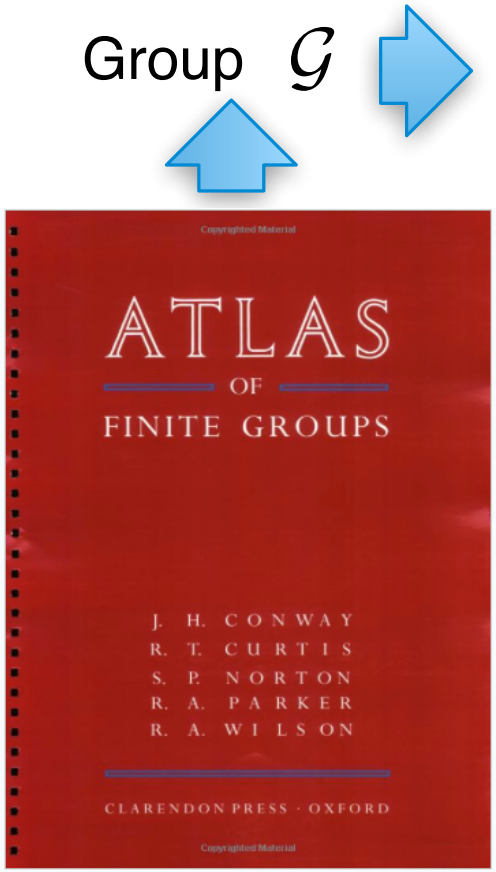


Γ_N Shannon Outer Bound
$\mathcal{Z}_N$ Non-Shannon Outer Bound
$\bar{\Gamma}^*_N$ Region of Entropic Vectors
$\mathcal{S}_N$ Subspace Ranks Bound
$\mathcal{M}^q_N$ GF(q)-Representable Matroid Bound
Φ_4 binary entropic vectors
$\mathrm{conv}(\Phi_4)$ convex hull

- **Matroids**: (rank function) Submodular, non-dec. $r : 2^{\mathcal{N}} \to \mathbb{Z}$, $r(A) \leq |A|$. $(\Gamma_N \cap \mathbb{Z}$ w/ $h(\{i\}) \leq 1)$.
- $(\mathbb{F}_q$-**)Representable Matroid** $r(\mathcal{A}) := \mathrm{rank}(\mathbb{A}_{:,\mathcal{A}})$, $\mathbb{A}$ $(\in \mathbb{F}_q^{r(\mathcal{N})\times N})$. $\propto$ *entropic:* [13, 14, 15]
  $\boldsymbol{X} = \boldsymbol{U}\mathbb{A}$, $\boldsymbol{U} \sim \mathbb{F}_q^{\mathrm{rank}(\mathcal{N})} \Rightarrow h(\boldsymbol{X}_{\mathcal{A}}) = r(A)\log_2(q)$
  take conic hull to get inner bound
- characterize: "forbidden minors" for $q \in \{2, 3, 4\}$. [16] cardinality & integer req. $\Rightarrow$ small $\subsetneq$ for $\bar{\Gamma}^*_4$.
- **Ingleton's Inequality**: [17] $h(\cdot) = r(\cdot)$ is repr. matroid, $\Rightarrow \mathrm{Ingleton}_{kl} \triangleq$
  $$I(i;j|k) + I(i;j|l) + I(k;l) - I(i;j) \geq 0$$
  (not an info. ineq.) derive w/ common inf. [18]
- **Subspace Bounds**: Group variables. (vector linear) Conic hull for 4-rvs $=$*Ingleton Inner bnd* [19, 20, 21].
  $$\mathcal{S}_4 \triangleq \Gamma_4 \cap_{ij\subset\{1,2,3,4\}} \mathrm{Ingleton}_{ij} \geq 0$$
  DFZ obtained $\mathcal{S}_5$ w/ extra ineq.s Unknown ≥ 6.

Entropic Vector Region: What is known/unknown? – From Inside (Groups)

Only provably exhaustive inner bound is due to Chan [22]:



Group $\mathcal{G}$ → Subgroups $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_N$

Group Characterizable Entropic Vector

$$h_{\mathcal{A}} \triangleq \log \frac{|\mathcal{G}|}{\left|\bigcap_{i\in\mathcal{A}} \mathcal{G}_i\right|}$$

collect over all groups/ subgroups

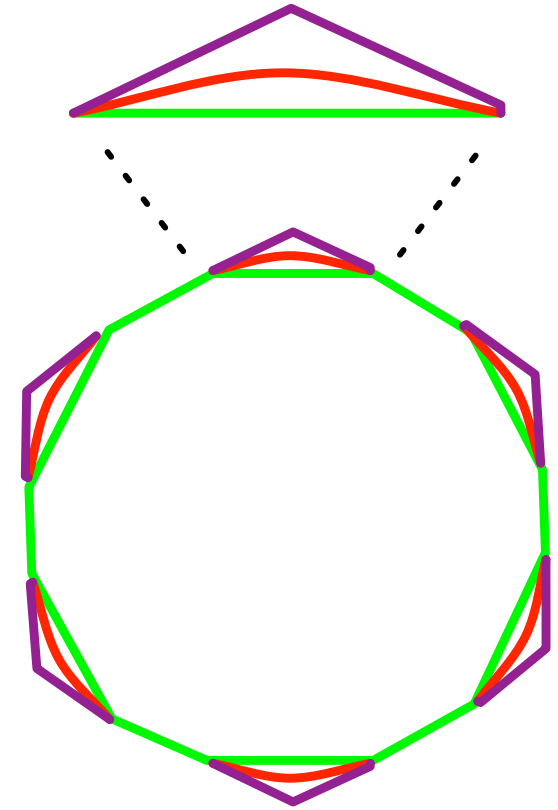
$$\Lambda_N \rightarrow \bar{\Gamma}^*_N = \overline{\mathrm{conic}(\Lambda_N)}$$

Structure of the Unknown Part of Γ_4

$\mathcal{P}_{ij} \triangleq \Gamma_4 \cap \text{Ingleton}_{ij} \leq 0$

$\bar{\mathcal{P}}^*_{ij} \triangleq \bar{\Gamma}^*_4 \cap \text{Ingleton}_{ij} \leq 0$

$\mathcal{S}^{=}_{ij} \triangleq \mathcal{S}_4 \cap \text{Ingleton}_{ij} = 0$



Γ_4 **Shannon Outer Bound**

$\bar{\Gamma}^*_4$ **Region of Entropic Vectors**

$\mathcal{S}_4$ **Ingleton Inner Bound**

Matus 1995 Conditional Independence Relations [23]

- faces of Γ_4 have entropic point in rel. int.
- $h \in \Gamma_4$, $\text{Ingleton}_{ij} < 0 \implies \text{Ingleton}_{kl} \geq 0$
- $\bar{\Gamma}^*_4 = \mathcal{S}_4 \cup_{ij} \bar{\mathcal{P}}^*_{ij}$
- 6 symm. $\mathcal{P}_{ij}$. Each w/ 1 Ingl. vio. extr. ray.

Walsh EII 2013 [24]:

- Even though 15 dim., $\forall h_{\mathcal{A}} \in \text{Ingleton}_{ij}$

$$\text{proj}_{\backslash h_{\mathcal{A}}} \mathcal{P}_{ij} = \text{proj}_{\backslash h_{\mathcal{A}}} \bar{\mathcal{P}}^*_{ij} = \text{proj}_{\backslash h_{\mathcal{A}}} \mathcal{S}^{=}_{ij}$$

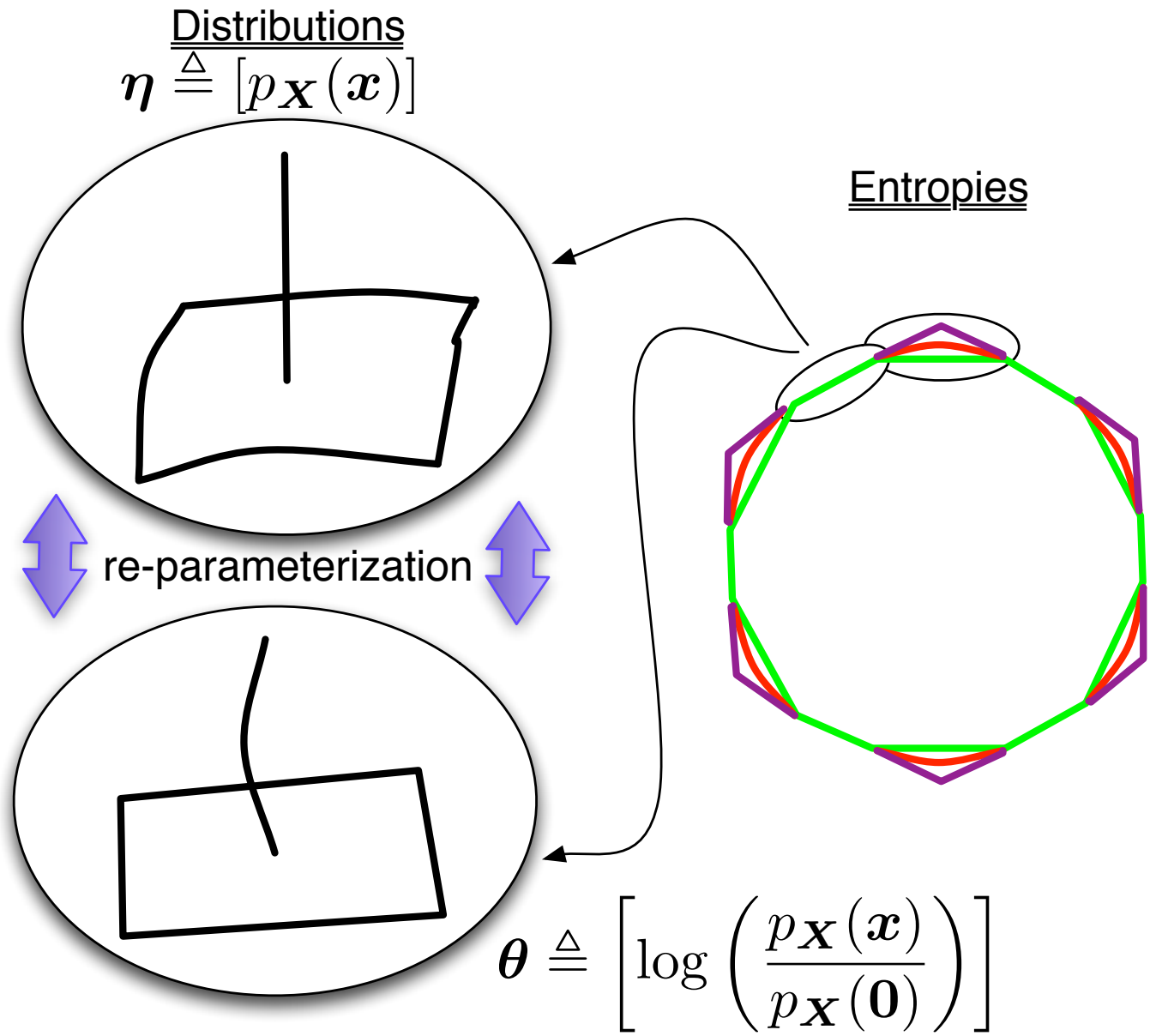
- *Only one necessary nonlinear inform. ineq.!*
- $=$ forms: $(-h_{\mathcal{A}}) \in \text{Ingleton}_{ij}$, $\bar{\mathcal{P}}^*_{ij} = \mathcal{P}_{ij} \cap$

$$h_{\mathcal{A}} \leq g^{up}_{\mathcal{A}}(\mathbf{h}^o_{\backslash \mathcal{A}}) \triangleq \max_{\mathbf{h} \in \bar{\mathcal{P}}^*_{ij}, \mathbf{h}_{\backslash \mathcal{A}} = \mathbf{h}^o_{\backslash \mathcal{A}}} h_{\mathcal{A}} \qquad (2)$$

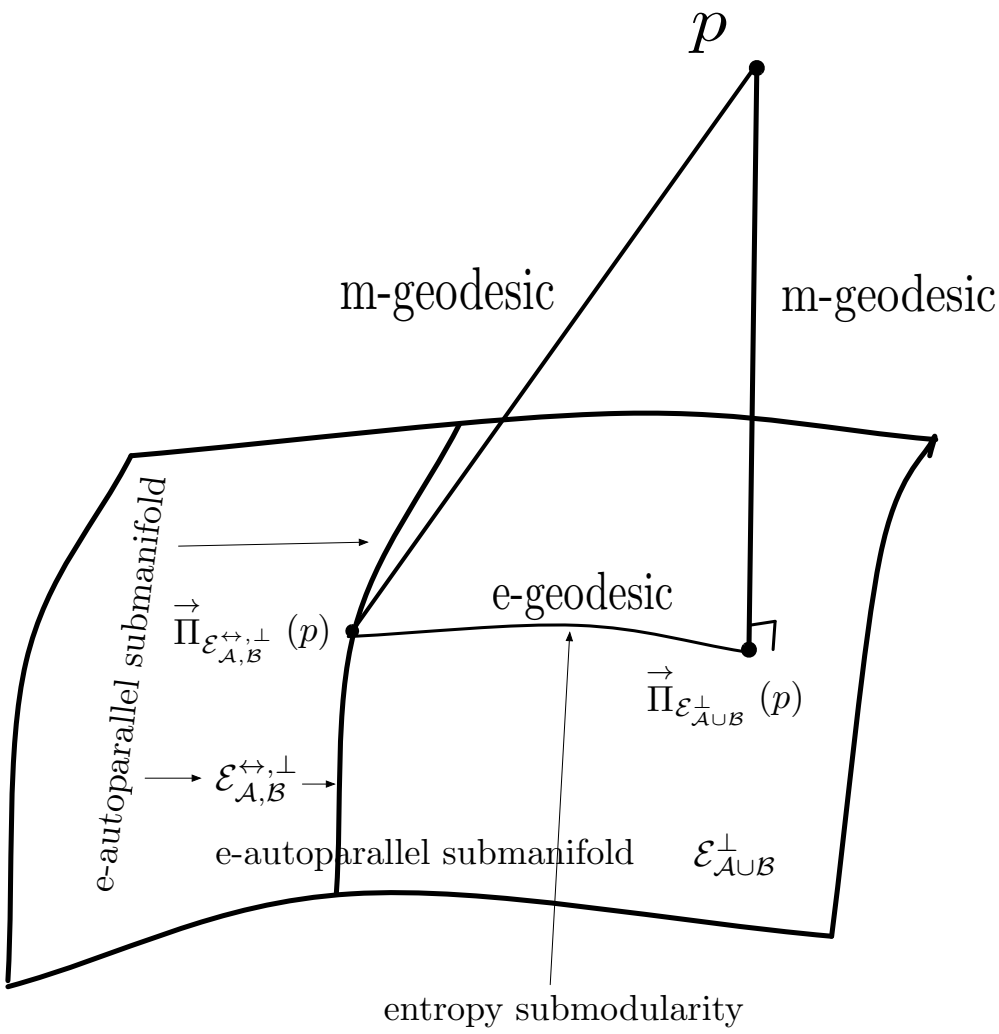
and if $(+h_{\mathcal{A}}) \in \text{Ingleton}_{ij}$, $\bar{\mathcal{P}}^*_{ij} = \mathcal{P}_{ij} \cap$

$$h_{\mathcal{A}} \geq g^{dn}_{\mathcal{A}}(\mathbf{h}^o_{\backslash \mathcal{A}}) \triangleq \min_{\mathbf{h} \in \bar{\mathcal{P}}^*_{ij}, \mathbf{h}_{\backslash \mathcal{A}} = \mathbf{h}^o_{\backslash \mathcal{A}}} h_{\mathcal{A}} \qquad (3)$$

Overall Aim of the Paper/Talk



Information Geometric Properties of Distributions on Shannon Facets



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

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

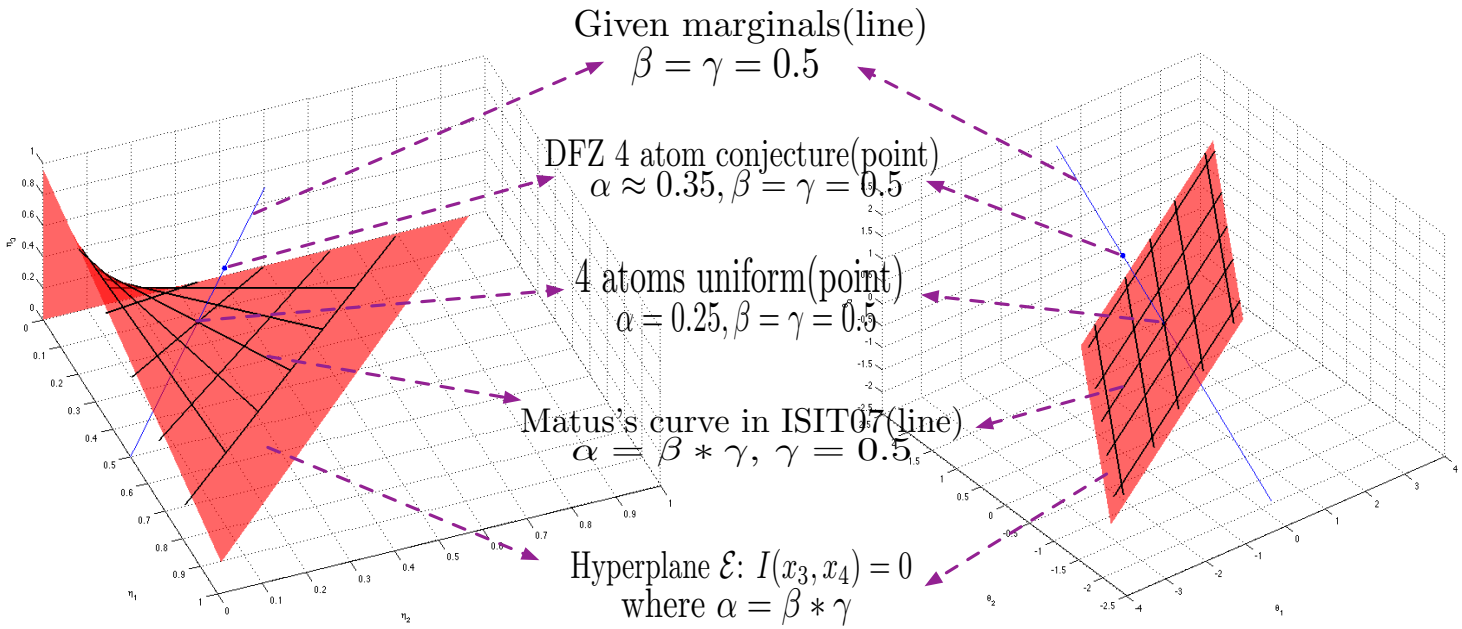
$$\mathcal{E}^{\perp}_{\mathcal{A},\mathcal{B}} = \left\{ \boldsymbol{\theta} \,\middle|\, p_{\boldsymbol{X}} = p_{\boldsymbol{X}_{\mathcal{A}\cup\mathcal{B}}} p_{\boldsymbol{X}_{(\mathcal{A}\cup\mathcal{B})^c}} \right\}$$

- 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 $p_{\boldsymbol{X}_{\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}}}$!
- $\Longrightarrow$ 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

Interesting 4-atom Distribution Support

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

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



The whole 3D space

$p(0000) = \alpha$

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

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

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

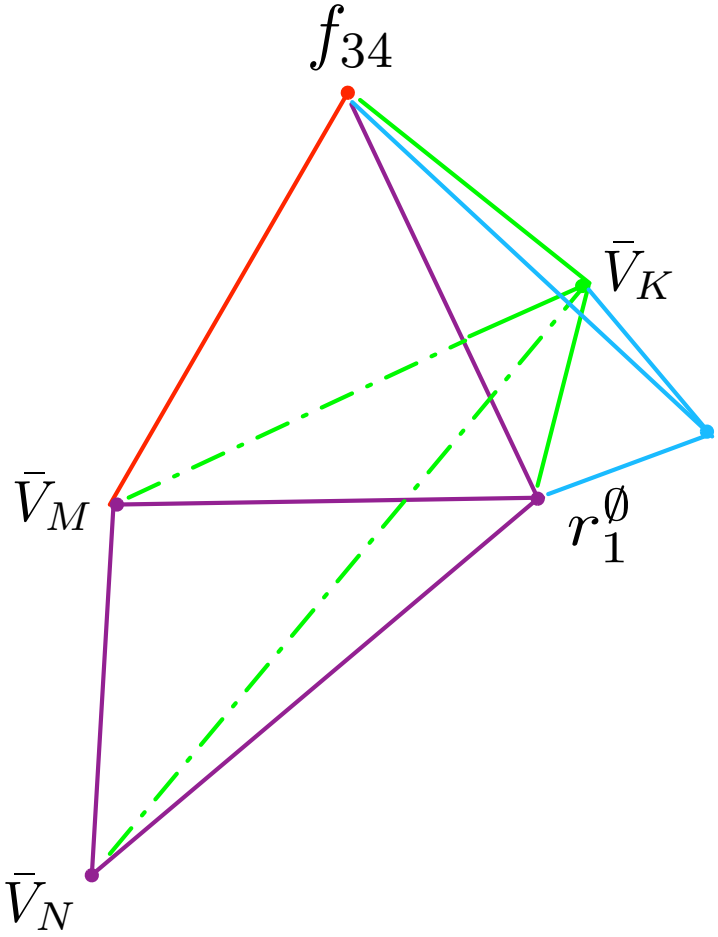
Marginal distribution of x_3 and x_4

$p(x_3 = 0) = \gamma$

$p(x_4 = 0) = \beta$

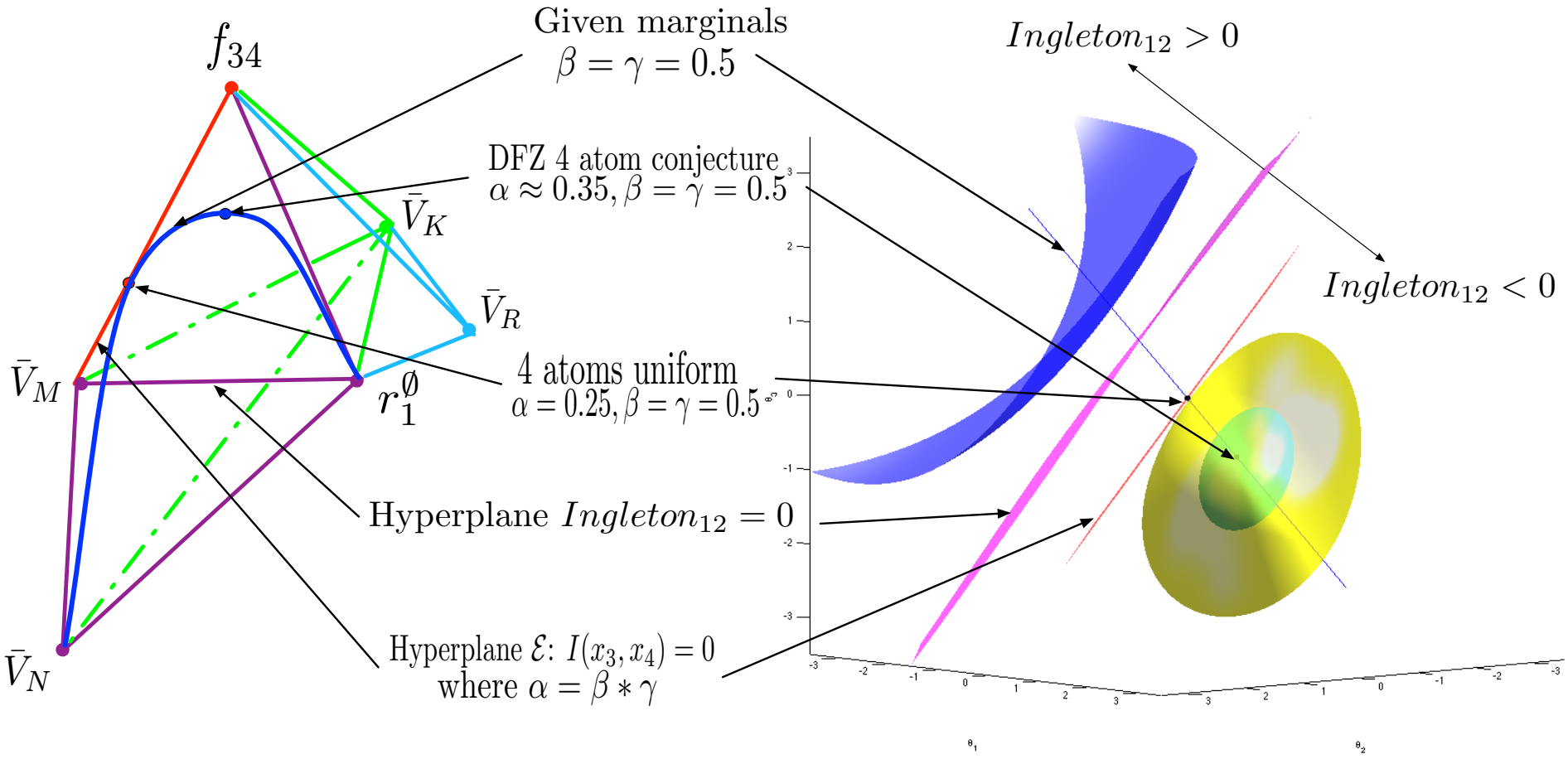
Structure of $\bar{\Gamma}_4^*$: Matúš Notation for $\mathcal{P}_{ij}$

$$\bar{V}_P = (r_1^{\emptyset}, r_1^3, r_1^4, r_1^{13}, r_1^{14}, r_1^{23}, r_1^{24}, r_1^{123}, r_1^{124}, r_1^{134}, r_1^{234}, r_2^1, r_2^2, r_3^{\emptyset}, f_{34})$$
$$= (\bar{V}_M, \bar{V}_K, \bar{V}_R, r_1^{\emptyset}, f_{34})$$



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

Information Geometry of Ingleton Violation for 4-atom Support



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

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

References

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

[2] 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)*, Jun. 2007, pp. 116 – 120.

[3] ——, "An Implicit Characterization of the Achievable Rate Region for Acyclic Multisource Multisink Network Coding," *IEEE Trans. on Information Theory*, vol. 58, no. 9, pp. 5625–5639, Sep. 2012.

[4] T. Chan and A. Grant, "Dualities between entropy functions and network codes," in *Fourth Workshop on Network Coding, Theory and Applications (NetCod)*, January 2008.

[5] R. W. Yeung and Zhen Zhang, "Distributed source coding for satellite communications," *IEEE Trans. on Information Theory*, vol. 45, no. 4, pp. 1111–1120, 1999.

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

[7] 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.

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

[9] 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.

[10] 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.

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

[12] 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

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

[14] Congduan Li, John MacLaren Walsh, Steven Weber, "A computational approach for determining rate regions and codes using entropic vector bounds," in *50th Annual Allerton Conference on Communication, Control and Computing*, Oct. 2012. [Online]. Available: http://www.ece.drexel.edu/walsh/Allerton2012LJW.pdf

[15] Congduan Li, Jayant Apte, John MacLaren Walsh, Steven Weber, "A new computational approach for determining rate regions and optimal codes for coded networks," in *The 2013 IEEE International Symposium on Network Coding (NetCod 2013)*, Jun. 2013. [Online]. Available: http://www.ece.drexel.edu/walsh/LiNetCod13.pdf

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

[17] 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.

[18] 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.

[19] A. S. N. V. D. Hammer, A. Romashchenko, "Inequalities for Shannon Entropy and Kolmogorov Complexity," *Journal of Computer and System Sciences*, no. 60, pp. 442–464, 2000.

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

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

[22] T. Chan and R. Yeung, "On a relation between information inequalities and group theory," *IEEE Trans. on Information Theory*, vol. 48, no. 7, pp. 1992 – 1995, Jul. 2002.

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

[24] J. M. Walsh, "Information Geometry, Polyhedral Computation, and Entropic Vectors," in *First Workshop on Entropy and Information Inequalities*, The Chinese University of Hong Kong, Apr. 2013. [Online]. Available: http://www.youtube.com/watch?v=gYkS73jyvPU&feature=youtu.be