

On the Entropy Region and the Ingleton Inequality

D. Fong^{*}, S. Shadbakht[†] and B. Hassibi[‡]

1 Introduction

Let $\mathcal{N} = \{1, \dots, n\}$ and let $X_1, \dots, X_n$ be n jointly distributed, discrete random variables with distribution $p(X_1, \dots, X_n)$. Then for any $\alpha \subseteq \mathcal{N}$, the entropy of $X_\alpha = \{X_i, i \in \alpha\}$ is defined as follows:

$$H(X_\alpha) = - \sum_{x_\alpha} p_{X_\alpha}(x_\alpha) \log p_{X_\alpha}(x_\alpha) \quad (1)$$

The vector of all $2^n - 1$ joint entropies is called an entropy vector. On the other hand a $2^n - 1$ dimensional vector whose entries can be regarded as the joint entropies of n jointly distributed random variables is referred to as entropic. The region of all entropic vectors is denoted by Γ_n^* [2]. It turns out that characterizing this region is the key to the solution of many open problems. Determining the compatibility of independence relations among a number of random variables [3], optimizing information quantities and discovery of new information inequalities for ranks of arbitrary matroids [4, 1, 5, 6, 7, 8] are some examples. Moreover it has been shown that there is a fundamental connection between the capacity of information networks and the entropy region [9, 10] and since the capacity region of many even simple networks such as relay channel are still open, full characterization of entropy region is of central importance in network information theory. However this has proven to be notoriously difficult and all the attempts so far have only resulted in partial characterizations for $n \geq 4$. Nonetheless, for $n \leq 3$ the region has been fully characterized [11]. In fact it is well known that the joint entropies $H(X_\alpha)$ (or H_α for simplicity) satisfy the following inequalities:

^{*}Department of Mathematics, Chinese University of Hong Kong, dfong@caltech.edu

[†]Department of Electrical Engineering, California Institute of Technology, sormeh@caltech.edu

[‡]Department of Electrical Engineering, California Institute of Technology, hassibi@caltech.edu

1. $H_\emptyset = 0$
2. For $\alpha \subseteq \beta$: $H_\alpha \leq H_\beta$
3. For any α, β : $H_{\alpha \cup \beta} + H_{\alpha \cap \beta} \leq H_\alpha + H_\beta$.

These are called the basic inequalities of Shannon information measures and the last one is referred to as the submodularity property. They all follow from the non-negativity of conditional mutual information. Any inequality obtained from positive linear combination of conditional mutual information is called a "Shannon type" inequality. The space of all $2^n - 1$ dimensional vectors satisfying only Shannon-type inequalities is denoted by Γ_n . It has been shown that $\Gamma_2^* = \Gamma_2$ and $\bar{\Gamma}_3^* = \Gamma_3$ where $\bar{\Gamma}_3^*$ denotes the closure of Γ_3^* [11]. However, for $n \geq 4$, in 1998 the first non-Shannon type information inequality was discovered [11] which demonstrated that Γ_4^* is strictly smaller than Γ_4 . Since then many other non-Shannon type inequalities have been discovered [6, 12, 13, 14]. New information inequalities of course result in new (hopefully tighter) outerbounds for Γ_n^* , however achievable regions are obtained through determining innerbounds and there has been less focus on this [1, 15, 16].

A remarkable result of [17] reveals a deep connection between entropy region and group theory.

In the remainder, we first give the relationship between the entropy region, group theory and quasi-uniform distributions. We later use this to fully construct entropy regions of 2 and 3 random variables and give an innerbound for Γ_4^* from quasi-uniform distributions of alphabet size 2. We then show that the well-known Ingleton bound for the entropy region cannot be violated with the use of some classical groups such as the general linear group and finally we prove that the non-trivial Ingleton-violating example of [1] cannot be constructed via a group of order 156.

2 Entropy region and group theory

2.1 group-characterizable entropies, quasi-uniform distributions

Let G denote a finite group with subgroups $G_1, \dots, G_n$. Define $2^n - 1$ dimensional vector g such that $g_\alpha = \log \frac{|G|}{|\cap_{i \in \alpha} G_i|}$ where $g_\alpha, \alpha \subseteq \mathcal{N}$ are the elements of g . Then g is called *group characterizable* and it can be shown that it is in fact entropic. Furthermore if we denote the region of all group-characterizable entropy functions as γ_n then [17]:

$$\overline{\text{con}}(\gamma_n) = \bar{\Gamma}_n^* \quad (2)$$

Therefore any information inequality as $\sum_\alpha a_\alpha H_\alpha \geq 0$ can be translated to its corresponding group inequality $\sum_\alpha a_\alpha \log \frac{|G|}{|\cap_{i \in \alpha} G_i|} \geq 0$.

Along the same path let $p_{X_N}(x_N)$ denote the joint distribution of n random variables $X_1, \dots, X_n$. The distribution is called *quasi-uniform* if all the marginal distributions $p_{X_\alpha}(x_\alpha), \alpha \subseteq \mathcal{N}$ satisfy:

$$p_{X_\alpha}(x_\alpha) = c_\alpha \quad \text{or} \quad 0 \quad (3)$$

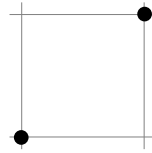


Figure 1: A quasi-uniform distribution of two random variables, alphabet size 2.

for some constant c_α . Computing the entropy of such distributions is straightforward, i.e. $H_\alpha = \log \frac{1}{c_\alpha}$. Let Λ_n denote the space of all entropy vectors obtained from quasi-uniform distributions. It has been shown that

$$\overline{\text{con}}(\Lambda_n) = \bar{\Gamma}_n^* \quad (4)$$

In other words the collection of all group characterizable entropy functions or the set of all quasi-uniform distributions are sufficient to characterize Γ_n^* [18]. However determining all the quasi-uniform distributions is still a formidable task. We have surprisingly found that considering quasi-uniform distributions of alphabet size 2 alone, gives a full characterization of the entropy region of 2 and 3 random variables and an innerbound for Γ_4^* . For alphabet size 2, it is rather easy to obtain all such distributions.

2.2 Quasi-uniform distributions of alphabet size 2

We represent each distribution by a lattice diagram where probabilities are uniformly distributed among all the dots in each diagram. For example, Figure 1 represents the distribution $p(0,0) = p(1,1) = 1/2$, $p(0,1) = p(1,0) = 0$.

For 2 random variables, there are 5 different quasi-uniform distributions of alphabet size 2 (Figure 2). Using the Double Description Method [19] to compute the convex hull, we find that the 5 entropies correspond to the 5 vertices of the 3 dimensional polyhedron described by the following inequalities:

$$\begin{aligned} h_i &\leq 1 & i &= 1, 2 \\ h_i &\leq h_{ij} & i, j &= 1, 2 \\ h_{12} &\leq h_1 + h_2 \end{aligned}$$

Recalling that $\Gamma_2^* = \Gamma_2$, we conclude that the convex cone of the above 5 entropy vectors is in fact Γ_2^* . Determining all the quasi-uniform distributions of alphabet

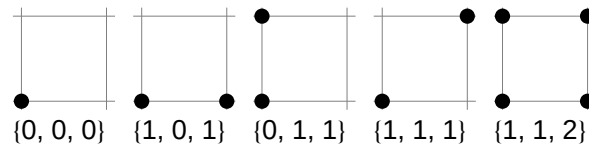


Figure 2: Quasi-uniform distribution of two random variables and their corresponding entropy vectors.

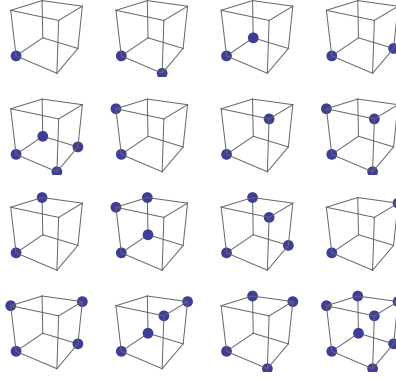


Figure 3: Quasi-uniform distribution of three random variables, alphabet size 2

size 2 for 3 random variables results in 16 entropy vectors (Figure 3). These correspond to all the vertices of a 7 dimensional polytope, whose 12 facets represent Shannon-type inequalities. The 12 inequalities are:

$$\begin{aligned} h_i &\leq 1 \\ h_{ij} &\leq h_i + h_j \\ h_{ij} &\leq h_{ijk} \\ h_i + h_{ijk} &\leq h_{ij} + h_{ik} \end{aligned}$$

for all permutations $\{i, j, k\}$ of $\{1, 2, 3\}$. Now since $\bar{\Gamma}_3^* = \Gamma_3$, the convex cone of the above 16 entropy vectors clearly gives $\bar{\Gamma}_3^*$.

For 4 random variables, a similar investigation gives 67 different quasi-uniform distributions corresponding to a 15 dimensional polytope with 94 facets.

Unlike the case of 2 and 3 random variables, the convex cone of these entropy vectors gives a loose innerbound for Γ_4^* . Specifically this polytope is obtained by the intersection of 6 hyperplanes of the following form and 88 other hyperplanes:

$$h_i + h_j + h_{ijk} + h_{ijl} + h_{kl} \leq h_{ij} + h_{ik} + h_{il} + h_{jk} + h_{jl} \quad (5)$$

for all permutation $\{i, j, k, l\}$ of $\{1, 2, 3, 4\}$. This is called the *Ingleton inequality* [20] and it holds for the ranks of representable matroids. In terms of characterizing the entropy region, it is one of the best known innerbounds [1]. Entropies obtained from abelian groups and normal-subgroups of a general group are known to satisfy the Ingleton inequality [21, 22]. Therefore in an attempt to find tighter innerbounds for the entropy region we need to focus on non-abelian groups.

3 Non-Abelian groups and the Ingleton inequality

Representing the Ingleton inequality in terms of subgroups of a group G we obtain:

$$\frac{|G_i||G_j||G_{kl}||G_{ijk}||G_{ijl}|}{|G_{ij}||G_{ik}||G_{il}||G_{jk}||G_{jl}|} \geq 1 \quad (6)$$

To violate (6), one should think of non-abelian groups. Investigation of some classical groups such as the general linear group interestingly show that even these groups satisfy the Ingleton inequality.

3.1 Some classical groups and the Ingleton inequality

Let $G = GL(n, q)$ and m_1, m_2, m_3, m_4 be 4 positive integers such that $m_1 + m_2 + m_3 + m_4 = n$. A general subgroup of G , say G_i would have the following structure:

$$G_i = \left\{ \begin{pmatrix} A_i & 0 & B_i \\ 0 & I_{m_i} & 0 \\ C_i & 0 & D_i \end{pmatrix} : A_i \in GL\left(\sum_{j=1}^{i-1} m_j, q\right), \begin{pmatrix} A_i & B_i \\ C_i & D_i \end{pmatrix} \in GL(n - m_i, q) \right\} \quad (7)$$

Recall that the order of a general linear group over a finite field with q elements is $(q^n - 1)(q^n - q) \cdots (q^n - q^{n-1})$. A direct computation using this fact, shows that (6) always holds.

We also showed that we cannot violate the Ingleton bound using a similar construction as the above using special linear groups, orthogonal groups and symmetric groups.

3.2 An attempt to construct an Ingleton violating entropy vector

One of the known non-trivial examples which violate the Ingleton inequality is constructed by using a pair of projective planes [1]. The resulting entropy vector is as follows:

$$\begin{aligned} h_1 &= h_2 = h_3 = h_4 = \log_2 13 \\ h_{12} &= \log_2 156, h_{34} = \log_2 78 \\ h_{23} &= h_{24} = h_{13} = h_{14} = \log_2 52 \\ h_{123} &= h_{124} = h_{134} = h_{234} = h_{1234} = \log_2 156 \end{aligned}$$

The fact that $h_{1234} = \log_2 156$, suggests to look for a group of order 156 to find a group-characterization of the entropy vector. Therefore we need to find a group G and four subgroups G_1, G_2, G_3, G_4 such that

$$\begin{aligned} |G| &= 156, |G_1| = |G_2| = |G_3| = |G_4| = 12 \\ |G_{12}| &= 1, |G_{34}| = 2 \\ |G_{23}| &= |G_{24}| = |G_{13}| = |G_{14}| = 3 \\ |G_{123}| &= |G_{124}| = |G_{134}| = |G_{234}| = |G_{1234}| = 1 \end{aligned}$$

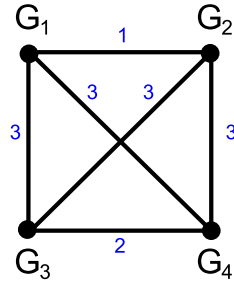


Figure 4: Required intersection order for subgroups of the group of order 156

The corresponding order of the intersection of the subgroups are shown in Figure 4.

If any one of groups among G_1, G_2, G_3, G_4 has only one subgroup of order 3, then it must intersect with two other groups at this subgroup. This implies that those two groups must again intersect at an order of at least 3. However, this violates the fact that G_1, G_2 intersect at only 1 element and G_3, G_4 intersect at only 2 elements. Therefore, by contradiction, each of the groups G_1, G_2, G_3, G_4 must have at least two subgroups of order 3.

G_1, G_2, G_3, G_4 are groups of order 12, and there is only one isomorphic class of group of 12; the alternating group A_4 , that has at least two subgroups of order 3. This implies that each of the four groups must be isomorphic to A_4 . This gives us more information about the group G that we are looking for:

- G contains four Sylow-2 subgroups
- G contains four Sylow-3 subgroups
- G contains no element of order 4

The first two claims are obvious while the last one can be seen directly from the Third Sylow Theorem [23] that all Sylow-2 subgroups are isomorphic and they must be isomorphic to $C_2 \times C_2$.

We then check the above three properties against all the 18 isomorphic classes of groups of order 156 using the GAP computational discrete algebra system. Among the 18 groups, 17 of them violate some of the above properties. For the remaining one, all of its order 12 subgroups are isomorphic to $C_6 \times C_2$, not A_4 . Therefore, we conclude that there does not exist a group of order 156 which satisfy our requirement for entropies.

4 Conclusion

We studied the connection between the entropy region, quasi-uniform distributions and group theory. We provided new constructions of the entropy region for $n = 2, 3$ random variables using binary quasi-uniform distributions and showed that it gives a strict inner bound for $n \geq 4$. In particular, for $n \geq 4$ our construction satisfies the Ingleton inequality. We also presented some (unsuccessful) group-theoretic approaches to violate the Ingleton bound.

Bibliography

- [1] Z. ZHANG AND R. YEUNG, *On Characterization of Entropy Function via Information Inequalities*, IEEE Trans. on Information Theory, 44 (1998), pp. 1440–1452.
- [2] R.W. YEUNG, *A Framework for Linear Information Inequalities*, IEEE Trans. on Information Theory, 43 (1997), pp. 1924–1934.
- [3] F. MATUS, *Conditional independences among four random variables III: final conclusion*, Combinatorics, Probability and Computing, 8 (1999), pp. 269–276.
- [4] F. MATUS, *Infinitely many information inequalities*, IEEE International Symposium on Information Theory (ISIT), 2007, pp. 41–44.
- [5] S. FUJISHIJE, *Polymatroidal Dependence Structure of a Set of Random Variables*, Information and Control, 39(1978), pp. 55–72.
- [6] F. MATUS, *Two constructions on limits of entropy functions*, IEEE Trans. on Information Theory, 53(2007), pp. 320–330.
- [7] D. HAMMER, A. E. ROMASHCHENKO, A. SHEN, N. K. VERESHCHAGIN, *Inequalities for Shannon entropies and Kolmogorov complexity*, IEEE conf. on computational complexity, Ulm, Germany, 1997, pp. 13–23.
- [8] R. DOUGHERTY AND C. FREILING AND K. ZEGER, *Matroids, networks and non-shannon information inequalities*, IEEE Trans. on Information Theory, 53(2007), pp. 1949–1969.
- [9] B. HASSIBI AND S. SHADBAKHT, *Normalized entropy vectors, network information theory and convex optimization*, Information theory workshop, Bergen, Norway, 2007.
- [10] X. YAN, R.W. YEUNG, Z. ZHANG, *The capacity region for multi-source multi-sink network coding*, IEEE International Symposium on Information Theory (ISIT), 2007, pp. 116–120.
- [11] ZHEN ZHANG AND RAYMOND YEUNG, *A Non-Shannon-Type Conditional Inequality of Information Quantities*, IEEE Trans. on Information Theory, 43(1997), pp. 1982–1986.

- [12] R. DOUGHERTY, C. FREILING, K. ZEGER, *Six New Non-Shannon Information Inequalities*, IEEE International Symposium on Information Theory (ISIT), 2006, pp. 233–236.
- [13] K. MAKARYCHEV, Y. MAKARYCHEV, A. ROMASHCHENKO AND N. VERESHCHAGIN, *A New Class of Non-Shannon-Type Inequalities for Entropies*, Communications In Information and Systems, 2(2002), pp. 147–166.
- [14] Z. ZHANG, *On a new non-Shannon type information inequality*, Communications In Inf. and Systems, 3(2003), pp. 47–60.
- [15] F. MATUS AND M. STUDENY, *Conditional independences among four random variables I*, Combin., Prob. Comput., 4(1995), pp. 269–278.
- [16] B. HASSIBI AND S. SHADBAKHT, *On a construction of entropic vectors using lattice-generated distributions*, IEEE International Symposium on Information Theory (ISIT), 2007, pp. 501–505.
- [17] T.H. CHAN AND R.W. YEUNG, *On a Relation between Information Inequalities and Group Theory*, IEEE Trans. on Information Theory, 48(2002), pp. 1992–1995.
- [18] T.H. CHAN, *A Combinatorial Approach to Information Inequalities*, Communications In Information and Systems, 1(2001), pp. 241–254.
- [19] FUKUDA, KOMEI, *Double Description Method Revisited*, Lecture Notes In Computer Science, 1120(1995), pp. 91.
- [20] A. W. INGLETON, *Representation of matroids*, Combinatorial mathematics and its applications, D. Welsh, Ed. London: Academic Press, 1971, pp. 149–167.
- [21] HUA LI, EDWIN K. P. CHONG, *On connection between group homomorphism and the Ingleton inequality*, IEEE International Symposium on Information Theory (ISIT), 2007, pp. 1996–1999.
- [22] T. H. CHAN, *Group characterizable entropy functions*, IEEE Int. Symp. on Inf. Theory (ISIT), 2007, pp. 506–510.
- [23] MICHAEL ARTIN, *Algebra*, Prentice Hall, 1991.