OPEN ACCESS

2nd International Electronic Conference on Entropy and Its Applications
15-30 November 2015

www.sciforum.net/conference/ecea-2

Conference Proceedings Paper – Entropy

Only One Nonlinear Non-Shannon Inequality is Necessary for Four Variables

Yunshu Liu * and John MacLaren Walsh

ECE Department, Drexel University, Philadelphia, PA 19104, USA

* Author to whom correspondence should be addressed; E-Mail: yunshu.liu@drexel.edu.

Published: 13 November 2015

Abstract: The region of entropic vectors $\overline{\Gamma}^*_N$ has been shown to be at the core of determining fundamental limits for network coding, distributed storage, conditional independence relations, and information theory. Characterizing this region is a problem that lies at the intersection of probability theory, group theory, and convex optimization. A 2^N-1 dimensional vector is said to be entropic if each of its entries can be regarded as the joint entropy of a particular subset of N discrete random variables. While the explicit characterization of the region of entropic vectors $\overline{\Gamma}^*_N$ is unknown for $N \geqslant 4$, here we prove that only one form of nonlinear non-shannon inequality is necessary to fully characterize $\overline{\Gamma}^*_4$. We identify this inequality in terms of a function that is the solution to an optimization problem. We also give some symmetry and convexity properties of this function which rely on the structure of the region of entropic vectors and Ingleton inequalities. This result shows that inner and outer bounds to the region of entropic vectors can be created by upper and lower bounding the function that is the answer to this optimization problem.

Keywords: Information Theory; Entropic Vector; Non-Shannon Inequality

---

1. Introduction

The region of entropic vectors $\overline{\Gamma}^*_N$ has been shown to be a key quantity in determining fundamental limits in several contexts in network coding [1], distributed storage [2], group theory [3], and information theory [1]. $\overline{\Gamma}^*_N$ is a convex cone that has recently shown to be non-polyhedral [4], but its boundaries remain unknown. In §2, we give some background on the region of entropic vectors and outer bounds for it. In §3, we review inner bounds for this region, then characterize the gap between the Shannon outer bound and the Ingleton inner bound for 4 random variables. Next, in §4, we present our main result in which we argue the complete characterization of $\overline{\Gamma}^*_4$ can be seen as finding a single nonlinear inequality determined by a single nonlinear function. After defining the function as the solution to an optimization problem, we investigate some properties of it.

2. The Region of Entropic Vectors

Consider a set of N discrete random variables $\boldsymbol{X} = (X_1, \dots, X_N)$, $\mathscr{N} = \{1, \dots, N\}$ with joint probability mass function $p_{\boldsymbol{X}}(\boldsymbol{x})$. For every non-empty subset of these random variables $\boldsymbol{X}_{\mathscr{A}} := (X_n \mid n \in \mathscr{A})$, $\mathscr{A} \subset \{1, \dots, N\}$, there is a Shannon entropy $H(\boldsymbol{X}_{\mathscr{A}})$ associated with it, which can be calculated from $p_{\boldsymbol{X}_{\mathscr{A}}}(\boldsymbol{x}_{\mathscr{A}}) = \sum_{\boldsymbol{x}_{\mathscr{N}\setminus\mathscr{A}}} p_{\boldsymbol{X}_{\mathscr{N}}}(\boldsymbol{x})$ via

$$H(\boldsymbol{X}_{\mathscr{A}}) = \sum_{\boldsymbol{x}_{\mathscr{A}}} -p_{\boldsymbol{X}_{\mathscr{A}}}(\boldsymbol{x}_{\mathscr{A}}) \log_2 p_{\boldsymbol{X}_{\mathscr{A}}}(\boldsymbol{x}_{\mathscr{A}}) \tag{1}$$

If we stack these 2^N-1 joint entropies associated with all the non-empty subsets into a vector $\boldsymbol{h} = \boldsymbol{h}(p_{\boldsymbol{X}}) = (H(\boldsymbol{X}_{\mathscr{A}})|\mathscr{A} \subseteq \mathscr{N})$, $\boldsymbol{h}(p_{\boldsymbol{X}})$ is clearly a function of the joint distribution $p_{\boldsymbol{X}}$. A vector $\boldsymbol{h}_? \in \mathbb{R}^{2^N-1}$ is said to be *entropic* if there exist some joint distribution $p_{\boldsymbol{X}}$ such that $\boldsymbol{h}_? = \boldsymbol{h}(p_{\boldsymbol{X}})$. Γ^*_N is then defined as the image of the set $\mathscr{D} = \{p_{\boldsymbol{X}} | p_{\boldsymbol{X}}(\boldsymbol{x}) \geq 0, \sum_{\boldsymbol{x}} p_{\boldsymbol{X}}(\boldsymbol{x}) = 1\}$:

$$\Gamma^*_N = \boldsymbol{h}(\mathscr{D}) \subsetneq \mathbb{R}^{2^N-1} \tag{2}$$

The closure of this set $\overline{\Gamma}^*_N$ is a convex cone [1], but surprisingly little else is known about the boundaries of it for $N \geq 4$.

With the convention that $h_\emptyset = 0$, entropy is *sub-modular* [1,5], meaning that

$$h_{\mathscr{A}} + h_{\mathscr{B}} \geq h_{\mathscr{A}\cap\mathscr{B}} + h_{\mathscr{A}\cup\mathscr{B}} \quad \forall \mathscr{A}, \mathscr{B} \subseteq \mathscr{N}, \tag{3}$$

and is also *non-decreasing* and *non-negative*, meaning that

$$h_{\mathscr{A}} \geq h_{\mathscr{B}} \geq 0 \quad \forall \mathscr{B} \subseteq \mathscr{A} \subseteq \mathscr{N}. \tag{4}$$

The inequalities (3) and (4) together are known as the *polymatroidal axioms* [1][5], a function satisfy them is called the *rank function* of a *polymatroid*. If in addition to obeying the polymatroidal axioms (3) and (4), a set function r also satisfies

$$r_{\mathscr{A}} \leq |\mathscr{A}|, \quad r_{\mathscr{A}} \in \mathbb{Z} \quad \forall \mathscr{A} \subseteq \mathscr{N} \tag{5}$$

then it is called the rank function of a *matroid* on the ground set $\mathscr{N}$.

Since an entropic vector must obey the polymatroidal axioms, the set of all valid rank functions of polymatroids forms a natural outer bound for Γ^*_N and is known as the Shannon outer bound Γ_N[1,5]:

$$\Gamma_N = \left\{ \boldsymbol{h} \;\middle|\; \begin{array}{c} \boldsymbol{h} \in \mathbb{R}^{2^N-1} \\ h_{\mathscr{A}} + h_{\mathscr{B}} \geq h_{\mathscr{A}\cap\mathscr{B}} + h_{\mathscr{A}\cup\mathscr{B}} \;\forall \mathscr{A}, \mathscr{B} \subseteq \mathscr{N} \\ h_{\mathscr{P}} \geq h_{\mathscr{Q}} \geq 0 \quad \forall \mathscr{Q} \subseteq \mathscr{P} \subseteq \mathscr{N} \end{array} \right\} \tag{6}$$

Γ_N is a polyhedron, we have $\Gamma_2 = \Gamma_2^*$ and $\Gamma_3 = \overline{\Gamma}_3^*$, for $N \geq 4$, however, $\Gamma_N \neq \overline{\Gamma}_N^*$. Zhang and Yeung first showed this in [5] by proving a new inequality among 4 variables

$$2I(C;D) \leq I(A;B) + I(A;C,D) + 3I(C;D|A) + I(C;D|B) \tag{7}$$

which held for entropies, but is not implied by the polymatroidal axioms. They called it a *non-Shannon type* inequality to distinguish it from inequalities implied by Γ_N. In the next few years, a few authors have generated new non-Shannon type inequalities [6–8]. Then Matúš in [4] showed that $\overline{\Gamma}_N^*$ is not a polyhedron for $N \geq 4$. The proof was carried out by constructing several sequences of non-Shannon inequalities, including

$$s[I(A;B|C)+I(A;B|D)+I(C;D)-I(A;B)]+I(B;C|A)+\frac{s(s+1)}{2}[I(A;C|B)+I(A;B|C)] \geq 0 \tag{8}$$

Notice that (8) is the same as Zhang-Yeung inequality (7) when $s = 1$. Additionally, the infinite sequence of inequalities was used with a curve constructed from a particular form of distributions to prove $\overline{\Gamma}_N^*$ is not a polyhedron. Despite this characterization, even $\overline{\Gamma}_4^*$ is still not fully understand. Since then, many authors has been investigating the properties of $\overline{\Gamma}_N^*$ with the hope of ultimately fully characterizing the region [6,9–14].

3. Structure of $\overline{\Gamma}_4^*$: the gap between Ingleton inner bound $\mathscr{S}_4$ and Shannon outer bound Γ_4

Let's first introduce some basics in linear polymatroids and the Ingeton inner bound. Fix a $N' > N$, and partition the set $\{1, \ldots, N'\}$ into N disjoint sets $\mathscr{T}_1, \ldots, \mathscr{T}_N$. Let $\boldsymbol{U}$ be a length r row vector whose elements are i.i.d. uniform over $GF(q)$, and let $\mathbf{G}$ be a particular $r \times N'$ deterministic matrix with elements in $GF(q)$. Consider the N' dimensional vector

$$\boldsymbol{Y} = \boldsymbol{U}\mathbf{G}, \text{ and define } \boldsymbol{X}_i = \boldsymbol{Y}_{\mathscr{T}_i},\ i \in \{1, \ldots, N\}.$$

The subset entropies of the random variables $\{\boldsymbol{X}_i\}$ obey

$$H(\boldsymbol{X}_{\mathscr{A}}) = r(\mathscr{A}) \log_2(q) = \text{rank}\left([\mathbf{G}_{\mathscr{T}_i} | i \in \mathscr{A}]\right) \log_2(q). \tag{9}$$

A set function $r(\cdot)$ created in such a manner is called a linear polymatriod or a subspace rank functions. It obeys the polymatroidal axioms, and is additionally proportional to an integer valued vector, however it need not obey the cardinality constraint therefore it is not necessarily the rank function of a matroid.

Such a construction is clearly related to a representable matroid on a larger ground set[15]. Indeed, the subspace rank function vector is merely formed by taking some of the elements from the $2^{N'}$-1 representable matroid rank function vector associated with $\mathbf{G}$. That is, rank function vectors created via (9) are projections of rank function vectors of representable matroids.

Rank functions capable of being represented in the manner for some N', q and $\mathbf{G}$, are called subspace ranks in some contexts [16–18], while other papers effectively define a collection of vector random variables created in this manner a subspace arrangement [19].

Define $\mathscr{S}_N$ to be the conic hull of all subspace ranks for N subspaces. It is known that $\mathscr{S}_N$ is an inner bound for $\overline{\Gamma}_N^*$[16], which we call subspace inner bound. So far $\mathscr{S}_N$ is only known for $N \leq 5$ ([18,19]).

More specifically, $\mathscr{S}_2 = \overline{\Gamma}^*_2 = \Gamma_2$, $\mathscr{S}_3 = \overline{\Gamma}^*_3 = \Gamma_3$. As with most entropy vector sets, things start to get interesting at $N = 4$ variables (subspaces). For $N = 4$, $\mathscr{S}_4$ is given by the Shannon type inequalities (i.e. the polymatroidal axioms) together with six additional inequalities known as *Ingleton*'s inequality [16,17,20] which states that for $N = 4$ random variables

$$Ingleton_{ij} = I(X_k; X_l|X_i) + I(X_k; X_l|X_j) + I(X_i; X_j) - I(X_k; X_l) \geq 0 \tag{10}$$

Thus, $\mathscr{S}_4$ is usually called the Ingleton inner bound. We know Γ_4 is generated by 28 elemental Shannon type information inequalities[1]. As for $\mathscr{S}_4$, in addition to the the 28 Shannon type information inequalities, we also need six Ingleton's inequalities (10), thus $\mathscr{S}_4 \subsetneq \Gamma_4$.

In [17] it is stated that Γ_4 is the disjoint union of $\mathscr{S}_4$ and six cones $\{h \in \Gamma_4 | Ingleton_{ij} < 0\}$. The six cones $G_4^{ij} = \{h \in \Gamma_4 | Ingleton_{ij} \leq 0\}$ are symmetric due to the permutation of inequalities $Ingleton_{ij}$, so it sufficient to study only one of the cones. Furthermore, [17] gave the extreme rays of G_4^{ij} in Lemma 1 by using the following functions.
For $\mathscr{N} = \{1, 2, 3, 4\}$, with $\mathscr{I} \subseteq \mathscr{N}$ and $0 \leq t \leq |\mathscr{N} \backslash \mathscr{I}|$, define

$$r_t^{\mathscr{I}}(\mathscr{J}) = \min\{t, |\mathscr{J} \backslash \mathscr{I}|\} \; with \; \mathscr{J} \subseteq \mathscr{I}$$

$$g_i^{(2)}(\mathscr{J}) = \begin{cases} 2 & \text{if } \mathscr{J} = i \\ min\{2, |\mathscr{J}|\} & \text{if } \mathscr{J} \neq i \end{cases}$$

$$g_i^{(3)}(\mathscr{J}) = \begin{cases} |\mathscr{J}| & \text{if } i \notin \mathscr{J} \\ min\{3, |\mathscr{J}| + 1\} & \text{if } i \in \mathscr{J} \end{cases}$$

$$f_{ij}(\mathscr{K}) = \begin{cases} 3 & \text{if } \mathscr{K} \in \{ik, jk, il, jl, kl\} \\ min\{4, 2|\mathscr{K}|\} & \text{otherwise} \end{cases}$$

Lemma 1. *(Matúš)[17] The cone $G_4^{ij} = \{h \in \Gamma_4 | Ingleton_{ij} \leq 0\}$, $i, j \in \mathscr{N}$ distinct is the convex hull of 15 extreme rays. They are generated by the 15 linearly independent functions f_{ij}, r_1^{ijk}, r_1^{ijl}, r_1^{ikl}, r_1^{jkl}, $r_1^{\emptyset}$, $r_3^{\emptyset}$, r_1^{i}, r_1^{j}, r_1^{ik}, r_1^{jk}, r_1^{il}, r_1^{jl}, r_2^{k}, r_2^{l}, where $kl = \mathscr{N} \backslash ij$.*

Note that among the 15 extreme rays of G_4^{ij}, 14 extreme rays r_1^{ijk}, r_1^{ijl}, r_1^{ikl}, r_1^{jkl}, $r_1^{\emptyset}$, $r_3^{\emptyset}$, r_1^{i}, r_1^{j}, r_1^{ik}, r_1^{jk}, r_1^{il}, r_1^{jl}, r_2^{k}, r_2^{l} are also extreme rays of $\mathscr{S}_4$ and thus entropic, which leaves f_{ij} the only extreme ray in G_4^{ij} that is not entropic. It is easily verified that $\overline{\Gamma}^*_4$ is known as long as we know the structure of six cones $\overline{\Gamma}^*_4 \cap G_4^{ij}$. Due to symmetry, we only need to focus on one of the six cones $\overline{\Gamma}^*_4 \cap G_4^{34}$, thus we define $P_4^{34} = \overline{\Gamma}^*_4 \cap G_4^{34}$ and study the properties of P_4^{34} in the rest of the paper.

4. Understanding the structure of P_4^{34} by projection

4.1. Derivation of a single nonlinear function

One way to propose the problem of characterizing the entropy region is by the following optimization problem

$$\gamma(\boldsymbol{a}) = \min_{h \in \Gamma^*_N} \sum_{\mathscr{A} \subseteq \mathscr{N}} a_{\mathscr{A}} h_{\mathscr{A}} \tag{11}$$

where $a_{\mathscr{A}} \in \mathscr{R}$ and $\boldsymbol{a} = [a_{\mathscr{A}} | \mathscr{A} \subseteq \mathscr{N}]$. The resulting system of inequalities $\{\boldsymbol{a}^T \mathbf{h} \geq \gamma(\boldsymbol{a}) | \quad \forall \boldsymbol{a} \in \mathbb{R}^{2^N - 1}\}$, has each inequality linear in $\mathbf{h}$, and the minimal, non-redundant, subset of these inequalities is uncountably infinite due to the non-polyhedral nature of $\bar{\Gamma}^*_N$. Hence, while solving the program in principle provides a characterization to the region of entropic vectors, the resulting characterization with uncountably infinite cardinality is likely to be very difficult to use.

By studying the conditions on the solution to 11, in [3], the authors defined the notion of a *quasi-uniform* distribution and made the following connection between Γ^*_n and Λ_n (the space of entropy vectors generated by quasi-uniform distributions).

Theorem 1. *(Chan)[3] The closure of the cone of Λ_n is the closure of Γ^*_n : $\overline{con(\Lambda_n)} = \overline{\Gamma}^*_n$*

From Theorem 1, we know finding all entropic vectors associated with quasi-uniform distribution are sufficient to characterize the entropy region, however, determining all quasi-uniform distributions is a hard combinatorial problem, while taking their conic hull and reaching a nonlinear inequality description of the resulting non-polyhedral set appears even harder, perhaps impossible. Thus new methods to simplify the optimization problem should be explored. Our main result in the next theorem shows that in order to characterize $\overline{\Gamma}^*_4$, we can simplify the optimization problem (11) by utilizing extra structure of P^{34}_4.

Theorem 2 (Only one non-Shannon inequality is necessary)**.** *To determine the structure of $\overline{\Gamma}^*_4$, it suffices to find a single nonlinear inequality. In particular, select any $h_{\mathscr{A}} \in Ingleton_{ij}$. The region P^{ij}_4 is equivalently defined as:*

$$P^{ij}_4 = \left\{ \boldsymbol{h} \in \mathbb{R}^{15} \left| \begin{array}{c} A\boldsymbol{h}_{\setminus \mathscr{A}} \leqslant \boldsymbol{b} \quad (= G^{ij}_4 \textit{ project out } h_{\mathscr{A}}) \\ h_{\mathscr{A}} \geqslant g^{low}_{\mathscr{A}}(h_{\setminus \mathscr{A}}) \\ h_{\mathscr{A}} \leqslant g^{up}_{\mathscr{A}}(h_{\setminus \mathscr{A}}) \end{array} \right. \right\} \tag{12}$$

where $\boldsymbol{h}_{\setminus \mathscr{A}}$ is the 14 dimensional vector excluding $h_{\mathscr{A}}$,

$$g^{low}_{\mathscr{A}}(\boldsymbol{h}_{\setminus \mathscr{A}}) = \min_{[h_{\mathscr{A}} \ \boldsymbol{h}^T_{\setminus \mathscr{A}}]^T \in P^{ij}_4} h_{\mathscr{A}}, \tag{13}$$

$$g^{up}_{\mathscr{A}}(\boldsymbol{h}_{\setminus \mathscr{A}}) = \max_{[h_{\mathscr{A}} \ \boldsymbol{h}^T_{\setminus \mathscr{A}}]^T \in P^{ij}_4} h_{\mathscr{A}}. \tag{14}$$

Furthermore, if the coefficient of $h_{\mathscr{A}}$ in $Ingleton_{ij}$ is positive, $h_{\mathscr{A}} \leqslant g^{up}_{\mathscr{A}}(h_{\setminus \mathscr{A}})$ is the inequality $Ingleton_{ij} \leqslant 0$. Similarly, if the coefficient of $h_{\mathscr{A}}$ in $Ingleton_{ij}$ is negative, $h_{\mathscr{A}} \geqslant g^{low}_{\mathscr{A}}(h_{\setminus \mathscr{A}})$ is the inequality $Ingleton_{ij} \leqslant 0$.

Proof: We know G^{34}_4 is a 15 dimensional polyhedral cone. Inside this cone, some of the points are entropic, some are not, that is to say, $P^{34}_4 \subsetneq G^{34}_4$. From Lemma 1 we obtain the 15 extreme rays of G^{34}_4: f_{34}, r^{134}_1, r^{234}_1, r^{123}_1, r^{124}_1, $r^{\emptyset}_1$, $r^{\emptyset}_3$, r^3_1, r^4_1, r^{13}_1, r^{14}_1, r^{23}_1, r^{24}_1, r^1_2, r^2_2, where each of these extreme rays are 15 dimensional, corresponding to the 15 joint entropy $h_{\mathscr{A}}$ for $\mathscr{A} \subset \mathscr{N}$. The elements of these extreme rays are listed in Fig. 1. As shown in Fig. 1 with the green rows, if we project out h_{123} from these 15 extreme rays, the only ray which is not entropic, f_{34}, falls into the conic hull of the other 14 entropic extreme rays, that is to say, $\pi_{\setminus h_{123}} P^{34}_4 = \pi_{\setminus h_{123}} G^{34}_4$. Furthermore, one can easily verify that the

Figure 1. The extreme rays of G_4^{34}. The top row is the ray f_{34}, and all of its coefficients except in the red column (corresponding to h_{123}) are the sum of the entries in the green rows. Hence $\pi_{\backslash h_{123}} G_4^{34}$ is entirely entropic.

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

same statement holds if we drop any one of the 10 joint entropies $h_{\mathscr{A}} \in Ingleton_{34}$ by summing other extreme ray rows to get all but the dropped dimension. This then implies that for $h_{\mathscr{A}} \in Ingleton_{34}$ the projected polyhedron $\pi_{\backslash h_{\mathscr{A}}} G_4^{34}$ (from which the dimension $h_{\mathscr{A}}$ is dropped) is entirely entropic, and hence $\pi_{\backslash h_{\mathscr{A}}} P_4^{34} = \pi_{\backslash h_{\mathscr{A}}} G_4^{34}$. Hence, for some $h_{\mathscr{A}}$ with a non-zero coefficient in $Ingleton_{34}$, given any point $\boldsymbol{h}_{\backslash \mathscr{A}} \in \pi_{\backslash h_{\mathscr{A}}} G_4^{34} (= \pi_{\backslash h_{\mathscr{A}}} P_4^{34})$, the problem of determining whether or not $[h_{\mathscr{A}} \boldsymbol{h}_{\backslash \mathscr{A}}^T]^T$ is an entropic vector in P_4^{34} is equivalent to determining if $h_{\mathscr{A}}$ is compatible with the specified $\boldsymbol{h}_{\backslash \mathscr{A}}$, as P_4^{34} is convex. The set of such compatible $h_{\mathscr{A}}$s must be an interval $[g^{low}(\boldsymbol{h}_{\backslash \mathscr{A}}), g^{up}(\boldsymbol{h}_{\backslash \mathscr{A}})]$ with functions defined via (13) and (14). This concludes the proof of (12).

To see why one of the two inequalities in (14),(13) is just the Ingleton inequality $Ingleton_{34}$, observe that for the case of dropping out h_{123}, the only lower bound for h_{123} in G_4^{34} is given by $Ingleton_{34} \leqslant 0$ (all other inequalities have positive coefficients for this variable in the non-redundant inequality description of G_4^{34} depicted in Fig. 2). Thus, if $\mathbf{h} \in P_4^{34}$, then $\mathbf{h} \in G_4^{34}$, and

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

Furthermore, $\{Ingleton_{34}{=}0 \cap G_4^{34}\} = \{Ingleton_{34}{=}0 \cap P_4^{34}\}$ since all $\{Ingleton_{34} = 0\}$ rays of the outer bound G_4^{34} are entropic, and there is only one ray with a non-zero $Ingleton_{34}$, so the extreme rays of $\{Ingleton_{34}{=}0 \cap G_4^{34}\}$ are all entropic. This means that for any $\mathbf{h}_{\backslash 123} \in \pi_{\backslash 123} G_4^{34}$, the minimum for h_{123} specified by $Ingleton_{34}$ is attainable, and hence $g_{123}^{low}(\mathbf{h}_{\backslash 123}) = -h_1 - h_2 + h_{12} + h_{13} + h_{23} + h_{14} + h_{24} - h_{124} - h_{34}$.

Thus, the problem of determining $\overline{\Gamma}_4^*$ is equivalent to determining a single nonlinear function $g_{123}^{up}(\mathbf{h}_{\backslash 123})$. A parallel proof applied for other $h_{\mathscr{A}}$ with a non-zero coefficient in $Ingleton_{ij}$ yields the remaining conclusions. ■

Figure 2. The coefficients of the non-redundant inequalities in G_4^{34}. Note that in each column where $Ingleton_{34}$ has a non-zero coefficient, it is the only coefficient with its sign.

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

From Theorem 2, we ten nonlinear inequalities (depending on which $\mathscr{A}$ with $h_{\mathscr{A}}$ appearing in $Ingleton_{ij}$ is selected), any single one of which completely determines P_4^{ij}, and thus, with its six permutations, determine $\overline{\Gamma}_4^*$. This theorem largely simplifies the optimization problem of determining $\overline{\Gamma}_4^*$, in that we only need to work on maximizing or minimizing a single entropy $h_{\mathscr{A}}$ given any $\mathbf{h}_{\setminus\mathscr{A}}$ in the polyhedral cone $\pi_{\setminus h_{\mathscr{A}}} G_4^{ij}$, which is entirely entropic.

4.2. Properties of $g_{\mathscr{A}}^{up}(\mathbf{h}_{\setminus\mathscr{A}})$ and $g_{\mathscr{A}}^{low}(\mathbf{h}_{\setminus\mathscr{A}})$

Based on the analysis in the above section, once we know any one of the ten nonlinear functions, $g_1^{up}(\mathbf{h}_{\setminus 1})$, $g_2^{up}(\mathbf{h}_{\setminus 2})$, $g_{34}^{up}(\mathbf{h}_{\setminus 34})$, $g_{123}^{up}(\mathbf{h}_{\setminus 123})$, $g_{124}^{up}(\mathbf{h}_{\setminus 124})$, $g_{12}^{low}(\mathbf{h}_{\setminus 12})$, $g_{13}^{low}(\mathbf{h}_{\setminus 13})$, $g_{14}^{low}(\mathbf{h}_{\setminus 14})$, $g_{23}^{low}(\mathbf{h}_{\setminus 23})$, and $g_{24}^{low}(\mathbf{h}_{\setminus 24})$ we know P_4^{34} and hence $\overline{\Gamma}_4^*$.

In this section, we investigate the properties of these functions, including the properties of a single nonlinear function, as well as the relationship between different nonlinear functions. The first result is the convexity of $-g_{\mathscr{A}}^{up}(\mathbf{h}_{\setminus\mathscr{A}})$ and $g_{\mathscr{A}}^{low}(\mathbf{h}_{\setminus\mathscr{A}})$.

Lemma 2. *The following functions corresponding to P_4^{34} are convex:*

$$-g_1^{up}(\mathbf{h}_{\setminus 1}),\ -g_2^{up}(\mathbf{h}_{\setminus 2}),\ -g_{34}^{up}(\mathbf{h}_{\setminus 34}), -g_{123}^{up}(\mathbf{h}_{\setminus 123}),\ -g_{124}^{up}(\mathbf{h}_{\setminus 124})$$
$$g_{12}^{low}(\mathbf{h}_{\setminus 12}),\ g_{13}^{low}(\mathbf{h}_{\setminus 13}),\ g_{14}^{low}(\mathbf{h}_{\setminus 14}),\ g_{23}^{low}(\mathbf{h}_{\setminus 23}),\ g_{24}^{low}(\mathbf{h}_{\setminus 24})$$

Proof: Without loss of generality, we investigate the convexity of $g_1^{up}(\mathbf{h}_{\setminus 1})$. Let $\boldsymbol{h}^a = [h_1^a \quad \boldsymbol{h}_{\setminus 1}^a]^T$, $\boldsymbol{h}^b = [h_1^b \quad \boldsymbol{h}_{\setminus 1}^b]^T$ be any two entropic vectors in the pyramid P_4^{34}. Since $\overline{\Gamma}_4^*$ is a convex set, P_4^{34} is also convex. Thus for $\forall\ 0 \leqslant \lambda \leqslant 1$, we have $\lambda\boldsymbol{h}^a + (1-\lambda)\boldsymbol{h}^b \in P_4^{34}$. According to Theorem 2, we have

$$\lambda h_1^a + (1-\lambda)h_1^b \leqslant g_1^{up}(\lambda h_{\setminus 1}^a + (1-\lambda)h_{\setminus 1}^b) \tag{15}$$

Furthermore, for some $\boldsymbol{h}^a$ and $\boldsymbol{h}^b$ to make g_1^{up} tight, besides (15), the following two conditions also hold:

$$h_1^a \leqslant \boldsymbol{h}_1^a = g_1^{up}(\boldsymbol{h}_{\setminus 1}^a) \quad h_1^b \leqslant \boldsymbol{h}_1^b = g_1^{up}(\boldsymbol{h}_{\setminus 1}^b) \tag{16}$$

Combining (15) and (16), we get

$$\lambda h_1^a + (1-\lambda)h_1^b \leqslant \lambda \boldsymbol{h}_1^a + (1-\lambda)\boldsymbol{h}_1^b =$$
$$\lambda g_1^{up}(\boldsymbol{h}_{\setminus 1}^a) + (1-\lambda)g_1^{up}(\boldsymbol{h}_{\setminus 1}^b) \leqslant g_1^{up}(\lambda \boldsymbol{h}_{\setminus 1}^a + (1-\lambda)\boldsymbol{h}_{\setminus 1}^b)$$

Thus $g_1^{up}(h_{\setminus 1})$ is a concave function. Similarly we can prove the convexity of other functions listed above. ■

Next we would like to study the symmetry properties of $g_1^{up}(h_{\setminus 1})$. From the form of Ingleton inequality (10), we notice that for a given distribution, if we swap the position of X_i and X_j, the value calculated from $Ingleton_{ij}$ remain unchanged, same properties hold if we exchange X_k and X_l. However, for a given distribution which has its entropic vector $\boldsymbol{h}^a$ tight on g_1^{up} (thus $h_1^a = g_1^{up}(\boldsymbol{h}_{\setminus 1}^a)$), due to symmetry, the entropic vector $\boldsymbol{h}^b$ corresponding to the distribution swapping X_i and X_j(and/or swap X_k and X_l) will still be on the boundary and satisfy $h_1^b = g_1^{up}(\boldsymbol{h}_{\setminus 1}^b)$. Based on this fact, and that $g_1^{up}(\mathbf{h}_{\setminus 1})$ is a concave function, we get the following theorem.

Theorem 3. *Suppose we have a distribution p_X with corresponding entropic vector $\mathbf{h}^a$ tight on g_1^{up}, and denote $\mathbf{h}^b$ the entropic vector from swapping X_3 and X_4 in $p_{\mathbf{X}}$, then $g_1^{up}(\boldsymbol{h}_{\setminus 1}^a) = g_1^{up}(\boldsymbol{h}_{\setminus 1}^b)$ and*

$$\max_{\lambda \in [0,\ 1]} g_1^{up}(\lambda \boldsymbol{h}_{\setminus 1}^a + (1-\lambda)\boldsymbol{h}_{\setminus 1}^b) = g_1^{up}(\frac{1}{2}\boldsymbol{h}_{\setminus 1}^a + \frac{1}{2}\boldsymbol{h}_{\setminus 1}^b) \tag{17}$$

thus the maximum of g_1^{up} along $\lambda \boldsymbol{h}_{\setminus 1}^a + (1-\lambda)\boldsymbol{h}_{\setminus 1}^b$ must be obtained at entropic vectors satisfying $h_3 = h_4$, $h_{13} = h_{14}$, $h_{23} = h_{24}$ and $h_{123} = h_{124}$.

Proof: First we need to point out the symmetry between $\boldsymbol{h}^a$ and $\boldsymbol{h}^b$ caused by the exchange of X_3 and X_4. For

$$\boldsymbol{h}^a = [h_1^a\ h_2^a\ h_{12}^a\ h_3^a\ h_{13}^a\ h_{23}^a\ h_{123}^a h_4^a h_{14}^a\ h_{24}^a\ h_{124}^a\ h_{34}^a\ h_{134}^a\ h_{234}^a\ h_{1234}^a] \tag{18}$$

it can be easily verified that

$$\boldsymbol{h}^b = [h_1^a\ h_2^a\ h_{12}^a\ h_4^a\ h_{14}^a\ h_{24}^a\ h_{124}^a h_3^a h_{13}^a\ h_{23}^a\ h_{123}^a\ h_{34}^a\ h_{134}^a\ h_{234}^a\ h_{1234}^a] \tag{19}$$

Since both $\boldsymbol{h}^a$ and $\boldsymbol{h}^b$ are tight on g_1^{up},

$$h_1^a = g_1^{up}(\boldsymbol{h}_{\setminus 1}^a) \quad h_1^b = g_1^{up}(\boldsymbol{h}_{\setminus 1}^b)$$

Thus $\boldsymbol{h}_1^a = \boldsymbol{h}_1^b$ implies $g_1^{up}(\boldsymbol{h}_{\setminus 1}^a) = g_1^{up}(\boldsymbol{h}_{\setminus 1}^b)$, which also guarantee $g_1^{up}(\lambda \boldsymbol{h}_{\setminus 1}^a + (1-\lambda)\boldsymbol{h}_{\setminus 1}^b) = g_1^{up}((1-\lambda)\boldsymbol{h}_{\setminus 1}^a + \lambda \boldsymbol{h}_{\setminus 1}^b)$.

Now we proof (17) by contradiction, suppose $\exists \lambda' \in [0,\ 1], \lambda' \neq \frac{1}{2}$ such that

$$g_1^{up}(\lambda' \boldsymbol{h}_{\setminus 1}^a + (1-\lambda')\boldsymbol{h}_{\setminus 1}^b) > g_1^{up}(\frac{1}{2}\boldsymbol{h}_{\setminus 1}^a + \frac{1}{2}\boldsymbol{h}_{\setminus 1}^b) \tag{20}$$

Since $g_1^{up}(h_{\backslash 1})$ is a concave function,

$$\begin{aligned}&g_1^{up}(\lambda'\boldsymbol{h}^a_{\backslash 1}+(1-\lambda')\boldsymbol{h}^b_{\backslash 1})\\&=g_1^{up}((1-\lambda')\boldsymbol{h}^a_{\backslash 1}+\lambda'\boldsymbol{h}^b_{\backslash 1})\\&=\frac{1}{2}g_1^{up}(\lambda'\boldsymbol{h}^a_{\backslash 1}+(1-\lambda')\boldsymbol{h}^b_{\backslash 1})+\frac{1}{2}g_1^{up}((1-\lambda')\boldsymbol{h}^a_{\backslash 1}+\lambda'\boldsymbol{h}^b_{\backslash 1})\\&\leqslant g_1^{up}(\frac{1}{2}[\lambda'\boldsymbol{h}^a_{\backslash 1}+(1-\lambda')\boldsymbol{h}^b_{\backslash 1}]+\frac{1}{2}[(1-\lambda')\boldsymbol{h}^a_{\backslash 1}+\lambda'\boldsymbol{h}^b_{\backslash 1}])\\&=g_1^{up}(\frac{1}{2}\boldsymbol{h}^a_{\backslash 1}+\frac{1}{2}\boldsymbol{h}^b_{\backslash 1})\end{aligned}$$

which contradicts the assumption, and proves (17). Because of the symmetry between $\boldsymbol{h}^a$ in (18) and $\boldsymbol{h}^b$ in (19), entropic vector $\frac{1}{2}\boldsymbol{h}^a_{\backslash 1}+\frac{1}{2}\boldsymbol{h}^b_{\backslash 1}$ will have the properties that $h_3=h_4$, $h_{13}=h_{14}$, $h_{23}=h_{24}$ and $h_{123}=h_{124}$. ■

Next we are going to investigate the relationship between g_1^{up} and g_2^{up} by swapping X_1 and X_2 OR swapping both X_1, X_2 and X_3, X_4. For a distribution p_X with corresponding entropic vector $\boldsymbol{h}^a$ tight on g_1^{up}, we denote $\boldsymbol{h}^c$ the entropic vector from swapping X_1 and X_2 in p_X, $\boldsymbol{h}^d$ be entropic vector from swapping both X_1, X_2 and X_3, X_4. For

$$\boldsymbol{h}^a=[h^a_1\ h^a_2\ h^a_{12}\ h^a_3\ h^a_{13}\ h^a_{23}\ h^a_{123}h^a_4h^a_{14}\ h^a_{24}\ h^a_{124}\ h^a_{34}\ h^a_{134}\ h^a_{234}\ h^a_{1234}]$$

it can be easily verified that

$$\boldsymbol{h}^c=[h^a_2\ h^a_1\ h^a_{12}\ h^a_3\ h^a_{23}\ h^a_{13}\ h^a_{123}h^a_4h^a_{24}\ h^a_{14}\ h^a_{124}\ h^a_{34}\ h^a_{234}\ h^a_{134}\ h^a_{1234}] \tag{21}$$

$$\boldsymbol{h}^d=[h^a_2\ h^a_1\ h^a_{12}\ h^a_4\ h^a_{24}\ h^a_{14}\ h^a_{124}h^a_3h^a_{23}\ h^a_{13}\ h^a_{123}\ h^a_{34}\ h^a_{234}\ h^a_{134}\ h^a_{1234}] \tag{22}$$

Thus from $h^a_1=h^c_2=h^d_2$ we get

$$g_1^{up}(\boldsymbol{h}^a_{\backslash 1})=g_2^{up}(\boldsymbol{h}^c_{\backslash 2})=g_2^{up}(\boldsymbol{h}^d_{\backslash 2}) \tag{23}$$

which leads to the following theorem:

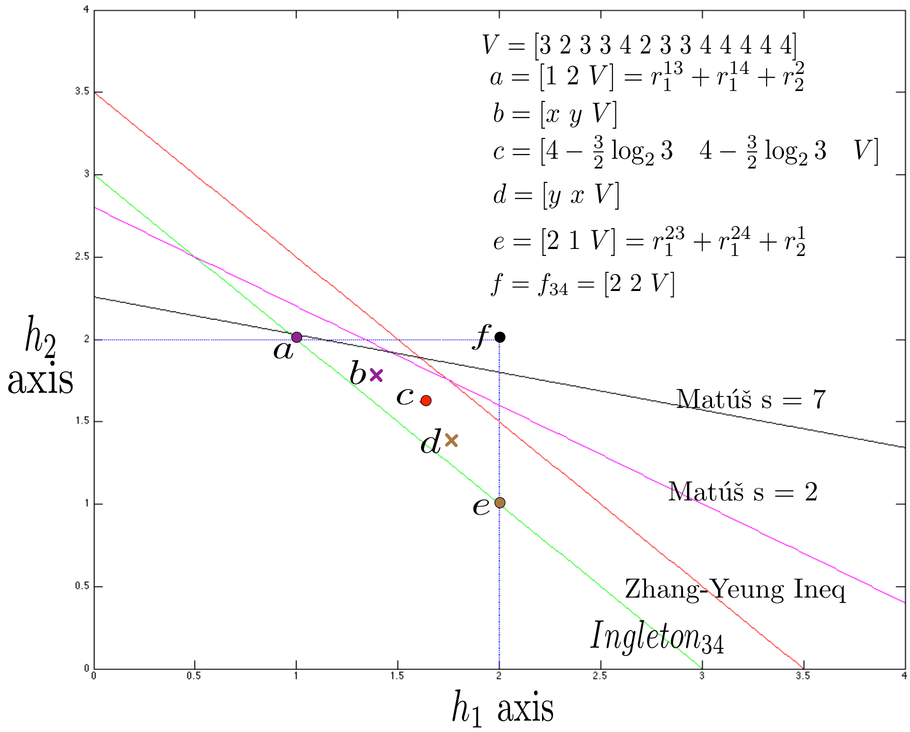
Theorem 4. *Suppose we have a distribution p_X with corresponding entropic vector $\boldsymbol{h}^a$ tight on g_1^{up}, we denote by $\boldsymbol{h}^c$ the entropic vector from swapping X_1 and X_2 in p_X, and $\boldsymbol{h}^d$ the entropic vector from permuting both X_1, X_2 and X_3,X_4. Then*

$$g_1^{up}(\boldsymbol{h}^a_{\backslash 1})=g_2^{up}(\boldsymbol{h}^c_{\backslash 2})=g_2^{up}(\boldsymbol{h}^d_{\backslash 2}) \tag{24}$$

Furthermore, if the entropic vector $\boldsymbol{h}^e$ associated with some distribution p_X satisfies $h_{13}=h_{23}$, $h_{14}=h_{24}$ and $h_{134}=h_{234}$, then $g_1^{up}(\boldsymbol{h}^e_{\backslash 1})=g_2^{up}(\boldsymbol{h}^e_{\backslash 1})$; if the entropic vector h^{fE} associated with some distribution p_X satisfies $h_3=h_4$, $h_{13}=h_{24}$, $h_{14}=h_{23}$, $h_{123}=h_{124}$ and $h_{134}=h_{234}$, then $g_1^{up}(h^{fE}_{\backslash 1})=g_2^{up}(h^{fE}_{\backslash 1})$.

Example 1: In order to explain Theorem 4, we consider the example such that we fix the last 13 dimension of entropic vector to $V=[3\ 2\ 3\ 3\ 4\ 2\ 3\ 3\ 4\ 4\ 4\ 4\ 4]$ and only consider the first two dimensions

Figure 3. Entropic vector hyperplane with only h_1 and h_2 coordinate not fixed



h_1 and h_2, which is shown in Figure 3. Since Γ_4^* is a 15 dimensional convex cone, if we fixed 13 dimensional to V, only h_1 and h_2 should be considered, thus we can easily plot the constrained region for visualization.

In Figure 3, f is the one of the 6 bad extreme rays(extreme rays of Γ_4 that are not entropic). The rectangle formed by connecting (0,0), (2,0), (0,2) and f is the mapping of Shannon outer bound Γ_4 onto this plane. The green line connecting a and e is the projection of $Ingleton_{34}$ onto the plane. Notice we also plot inequality (7) and (8) for some values of s in the figure for the comparison between Ingleton inner bound, Shannon outer bound and non-Shannon outer bound. The red dot point c is the entropic vector of the binary distribution with only four outcomes: (0000)(0110)(1010)(1111), each of the outcomes occur with probability $\frac{1}{4}$, and following from the convention of [21], we call it the 4 atom uniform point.

Since we already know $a = [1\ 2\ V]$ and $e = [2\ 1\ V]$ must lie on the boundary of P_4^{34}, thus $g_1^{up}([2\ V]) = g_2^{up}([2\ V])$ and $g_1^{up}([1\ V]) = g_2^{up}([1\ V])$. More generally, for any entropic vector $b = [x\ y\ V]$ on the boundary, we have $g_1^{up}([x\ V]) = g_2^{up}([x\ V])$ and $g_1^{up}([y\ V]) = g_2^{up}([y\ V])$. Thus we can say that when we constrain the last 13 dimension of entropic vector to $V = [3\ 2\ 3\ 3\ 4\ 2\ 3\ 3\ 4\ 4\ 4\ 4\ 4]$, the two function g_1^{up} and g_2^{up} always give us the same value, that is to say they are identical when fixed in this hyperplane.

5. Conclusions

In this paper, we proved that the problem of characterizing the region of entropic vectors was equivalent to finding a single non-linear inequality solving one of ten interchangeable optimization problems. Additionally, we investigated some symmetry and convexity based properties of the functions that are the solutions to these optimizations problem. Our future work is focused on calculating upper and lower bounds for these nonlinear functions.

Acknowledgements

The authors gratefully acknowledge NSF support of this research through grants CCF-1053702.

Conflicts of Interest

"The authors declare no conflict of interest".

References

1. Raymond W. Yeung. Information Theory and Network Coding. *Springer* **2008**.
2. Tian, C. Rate region of the (4, 3, 3) exact-repair regenerating codes. *IEEE International Symposium on Information Theory (ISIT)* **2013**, pp. 1426–1430.
3. Chan, T.; Yeung, R. On a relation between information inequalities and group theory. *IEEE Trans. on Information Theory* **2002**, *48*, 1992 – 1995.
4. František Matúš. Infinitely Many Information Inequalities. *IEEE Int. Symp. Information Theory (ISIT)* **2007**, pp. 41 – 44.
5. Zhen Zhang and Raymond W. Yeung. On Characterization of Entropy Function via Information Inequalities. *IEEE Trans. on Information Theory* **1998**, *44*.
6. Kaced, T. Equivalence of Two Proof Techniques for Non-Shannon-type Inequalities. *IEEE International Symposium on Information Theory (ISIT)* **2013**.
7. Makarychev, K.; Makarychev, Y.; Romashchenko, A.; Vereshchagin, N. A new class of non-Shannon-type inequalities for entropies. *Communication in Information and Systems* **2002**, *2*, 147–166.
8. Dougherty, R.; Freiling, C.; Zeger, K. Six new non-Shannon information inequalities. *IEEE International Symposium on Information Theory (ISIT)* **2006**, pp. 233–236.
9. Hassibi, B.; Shadbakht, S. On a Construction of Entropic Vectors Using Lattice-Generated Distributions. *IEEE International Symposium on Information Theory (ISIT)* **2007**, pp. 501 – 505.
10. Weidong Xu, Jia Wang, Jun Sun. A projection method for derivation of non-Shannon-type information inequalities. *IEEE International Symposium on Information Theory (ISIT)* **2008**, pp. 2116 – 2120.
11. J. M. Walsh and S. Weber. A Recursive Construction of the Set of Binary Entropy Vectors and Related Inner Bounds for the Entropy Region. *IEEE Trans. Inform. Theory* **2011**, *57*, 6356–6363.
12. F. Matúš and L. Csirmaz. Entropy region and convolution. arXiv:1310.5957v1.

13. Yunshu Liu and J. M. Walsh. Bounding the Entropic Region via Information Geometry. *IEEE Information Theory Workshop* **2013**, pp. 577–581.
14. T. Chan and A. Grant. Non-linear Information Inequalities. *Entropy* **2008**.
15. James Oxley. Matroid Theory, 2nd. Ed. *Oxford University Press* **2011**.
16. D. Hammer, A. Romashschenko, A. Shen, N. Vereshchagin. Inequalities for Shannon Entropy and Kolmogorov Complexity. *Journal of Computer and System Sciences* **2000**, *60*, 442–464.
17. F. Matúš and M. Studený. Conditional Independences among Four Random Variables I. *Combinatorics, Probability and Computing* **1995**, *4*, 269–278.
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* **2011**, *188*, 152–161.
20. A. W. Ingleton. Representation of Matroids. *Combinatorial Mathematics and its Applications* **1971**, pp. 149–167.
21. Randall Dougherty, Chris Freiling, Kenneth Zeger. Non-Shannon Information Inequalities in Four Random Variables. arXiv:1104.3602v1.