

submitted to *Combinatorics, Probability and Computing*, October 23, 2013

# ENTROPY REGION AND CONVOLUTION

FRANTIŠEK MATUŠ AND LÁSZLO CSIRMAZ

ABSTRACT. The entropy region is constructed from vectors of random variables by collecting Shannon entropies of all subvectors. Its shape is studied here by means of polymatroidal constructions, notably by convolution. The closure of the region is decomposed into the direct sum of tight and modular parts, reducing the study to the tight part. The relative interior of the reduction belongs to the entropy region. Behavior of the decomposition under selfadhesivity is clarified. Results are specialized to and completed for the region of four random variables. This and computer experiments help to visualize approximations of a symmetrized part of the entropy region. Four-atom conjecture on the minimization of Ingleton score is refuted.

## 1. INTRODUCTION

The entropy function of a random vector  $(\xi_i)_{i \in N}$  with a finite index set  $N$  maps each  $I \subseteq N$  into the Shannon entropy of the subvector  $(\xi_i)_{i \in I}$ . When the vector takes finite number of values, the entropy function can be considered for a point of the Euclidean space  $\mathbb{R}^{\mathcal{P}(N)}$  where  $\mathcal{P}(N)$  is the power set of  $N$ . Such points define the *entropic region*  $\mathbf{H}_N^{\text{ent}}$ . The closure  $cl(\mathbf{H}_N^{\text{ent}})$  of the region is a convex cone [44, Theorem 1] whose relative interior is contained in  $\mathbf{H}_N^{\text{ent}}$  [34, Theorem 1]. This work studies mostly the shape of  $cl(\mathbf{H}_N^{\text{ent}})$ .

Basic properties of the Shannon entropy imply that any entropy function  $h$  from  $\mathbf{H}_N^{\text{ent}}$  is non-decreasing and submodular, and thus the pair  $(N, h)$  is a polymatroid with the ground set  $N$  and rank function  $h$  [15]. The polymatroidal rank functions on  $N$  form the polyhedral cone  $\mathbf{H}_N$  which, consequently, contains  $cl(\mathbf{H}_N^{\text{ent}})$ . A polymatroid, or its rank function, is called *entropic (almost entropic)* if the rank function belongs to  $\mathbf{H}_N^{\text{ent}}$  ( $cl(\mathbf{H}_N^{\text{ent}})$ ).

In this work, the entropy region and its closure are studied by means of standard constructions on polymatroids, recalled in Section 2. The central working tool is the *convolution* of two polymatroidal rank functions. The crucial property is that  $cl(\mathbf{H}_N^{\text{ent}})$  is closed under convolution with modular polymatroids [34, Theorem 2]. This has consequences on principal extensions and their contractions.

In Section 3, the cone  $cl(\mathbf{H}_N^{\text{ent}})$  is decomposed into the direct sum of two cones, see Corollary 2. The first one consists of rank functions which give the same rank to  $N$  and all subsets with one element less. We call them *tight*. The second one is the cone of modular polymatroids, contained in the entropy region. Hence, the decomposition reduces the study of  $cl(\mathbf{H}_N^{\text{ent}})$  to a cone of lesser dimension. It is also closely related to balanced information-theoretic inequalities [7]. The relative interior of the first cone is exhausted by entropic points, see Theorem 2 in Section 4.

---

The work of FM was partially supported by Grant Agency of the Czech Republic under Grant 13-20012S. The work of LCs was partially supported by TAMOP-4.2.2.C-11/1/KONV-2012-0001 and by the Lendulet Program of the Hungarian Academy of Sciences.

The authors would like to express their gratitude to The Chinese University of Hong Kong for kind hospitality during First Workshop on Entropy and Information Inequalities, April 15-17, 2013, and fruitful discussions with participants.

Section 5 recalls the notion of selfadhesivity that describes amalgamation, or pasting, of copies of a polymatroid. It is the main ingredient in the majority of proofs of non-Shannon information-theoretical inequalities. The selfadhesivity is compared with the decomposition into tight and modular polymatroids. An alternative technique for proving inequalities is briefly discussed and related to principal extensions and their contractions.

Starting from Section 6 the set  $N$  is assumed to have exactly four elements. The role of Ingleton inequality in the structure of  $\mathbf{H}_N$  is recalled. The cone  $\mathcal{cl}(\mathbf{H}_N^{\text{ent}})$  is reduced to its subcone  $\mathbf{L}_{ij}$ , cut off by a reversed Ingleton inequality and tightness. Applying polymatroidal constructions, it is shown that  $\mathbf{L}_{ij}$  is mapped by two linear maps to its face  $\mathbf{F}_{ij}$  of dimension 9, see Theorem 4.

Section 7 investigates a symmetrization of  $\mathbf{F}_{ij}$  whose cross-section  $\mathbf{S}_{ij}$  has dimension three. Various numerical optimization techniques were employed to find an inner approximation of  $\mathbf{S}_{ij}$ . An outer approximation is compiled from available non-Shannon type information inequalities. Results are visualized, and indicate that the two approximations are far from each other. In Section 8, the range of Ingleton score studied and related to the cross-section  $\mathbf{S}_{ij}$ , see Theorem 5. In Example 2, a score is presented that refutes Four-atom conjecture [10, 14].

The concept of entropy region matters for several mathematical and engineering disciplines. The inequalities that hold for the points of the region are called information-theoretic, those that do not follow from the polymatroid axioms are frequently called non-Shannon. Main breakthroughs include finding of the first non-Shannon linear inequality by Zhang-Yeung [46] and the relation to group theory by [6]. The cone  $\mathcal{cl}(\mathbf{H}_N^{\text{ent}})$  is not polyhedral [35] and the structure of non-Shannon inequalities seems to be complex [45, 28, 13, 14, 11, 43]. Reviews are in [9, 34] and elsewhere.

In communications networks, the capacity region of multi-source network coding can be expressed in terms of the entropy region, the reader is referred to the thorough review of the network coding in [2]. Non-Shannon inequalities have a direct impact on converse theorems for multi-terminal problems of information theory, see [42]. In cryptography, the inequalities can improve bounds on the information ratios in secret sharing schemes [4, 5, 10].

In probability theory, the implication problem of conditional independence among subvectors of a random vector can be rephrased via the entropy region, see [41]. The guessing numbers of games on directed graphs and entropies of the graphs can be related to the network coding [38, 17] where non-Shannon inequalities provide sharper bounds [1]. Information-theoretic inequalities are under investigation in additive combinatorics [27]. Last but not least, the information-theoretic inequalities are known to be related to Kolmogorov complexity [28], determinantal inequalities and group-theoretic inequalities [9].

## 2. PRELIMINARIES

This section recalls basic facts about polymatroids and related operations. Auxiliary lemmas are worked out to be used later. Introduction to entropy and the entropy region can be found in the textbooks [12, 42], further material on polymatroids is in [26].

The letter  $N$  always denotes a finite set and  $f, g, h$  real functions on the power set  $\mathcal{P}(N)$  of  $N$ , or points in the  $2^{|N|}$ -dimensional Euclidean space  $\mathbb{R}^{\mathcal{P}(N)}$ . Singletons

and elements of  $N$  are not distinguished and the union sign between subsets of  $N$  is often omitted. For example,  $iJ$  abbreviates  $\{i\} \cup J$  where  $i \in N$  and  $J \subseteq N$ .

For  $I \subseteq N$  let  $\delta_I \in \mathbb{R}^{\mathcal{P}(N)}$  have all coordinates equal to 0 but  $\delta_I(I) = 1$ . For  $I, J \subseteq N$  the expression  $f(I) + f(J) - f(I \cup J) - f(I \cap J)$  is interpreted as the standard scalar product of  $\Delta_{I,J} = \delta_I + \delta_J - \delta_{I \cup J} - \delta_{I \cap J}$  with  $f$ . An alternative notation for  $\Delta_{iL,jL}$  is  $\Delta_{ij|L}$  where  $L \subseteq N$  and  $i, j \in N \setminus L$ .

2.1. The pair  $(N, f)$  is a *polymatroid* when  $f(\emptyset) = 0$ ,  $f$  is nondecreasing, thus  $f(I) \leq f(J)$  for  $I \subseteq J \subseteq N$ , and submodular, thus  $\Delta_{I,J} f \geq 0$  for  $I, J \subseteq N$ . Here,  $N$  is the ground set,  $f(N)$  is the rank and  $f$  is the rank function of the polymatroid. The polymatroid is frequently identified with its rank function. The collection of polymatroidal rank functions forms the closed polyhedral cone  $\mathbf{H}_N$  in the nonnegative orthant of  $\mathbb{R}^{\mathcal{P}(N)}$ . Extreme rays of the cone are mostly unknown. For a review of polymatroids the reader is referred to [26].

The polymatroid is a *matroid* [37] if  $f$  takes integer values and  $f(I) \leq |I|$ ,  $I \subseteq N$ . For  $J \subseteq N$  and  $0 \leq m \leq |N \setminus J|$  integer let  $r_m^J(I) = \min\{m, |I \setminus J|\}$ ,  $I \subseteq N$ . Thus,  $r_m^J$  is a matroidal rank function with the set of loops  $J$ ,  $r_m^J(J) = 0$ , and rank  $m$ . If  $J = \emptyset$  it is sometimes omitted in the superindex.

The polymatroid  $f$  is *modular* if  $\Delta_{I,J} f = 0$  for any  $I$  and  $J$  disjoint. This is equivalent to  $f(I) = \sum_{i \in I} f(i)$ ,  $I \subseteq N$ , or to the single of this equalities with  $I = N$ . The modular polymatroids form the polyhedral cone  $\mathbf{H}_N^{\text{mod}}$  whose extreme rays are generated by the matroids  $r_1^{N \setminus i}$ ,  $i \in N$ .

A polymatroid  $(N, f)$  is *linear* if there exist subspaces  $E_i$ ,  $i \in N$ , of a linear space over a field  $\mathbb{F}$  such that if  $I \subseteq N$  then  $f(I)$  equals the dimension of the sum of  $E_i$  over  $i \in I$ . If  $\mathbb{F}$  is finite then  $f \ln |\mathbb{F}|$  is entropic.

2.2. The *contraction* of a polymatroid  $(N, f)$  along  $I \subseteq N$  is  $(N \setminus I, h)$  where  $h(J) = f(J \cup I) - f(J)$ ,  $J \subseteq N \setminus I$ . (Poly)matroids are closed to contractions. The following lemma is known, e.g. implicit in the proof of [30, Lemma 2], but no reference to a proof seems to be available.

**Lemma 1.** *The almost entropic polymatroids are closed to contractions.*

*Proof.* It suffices to show that if  $f$  is equal to the entropy function of a random vector  $(\xi_i)_{i \in N}$ , then the contraction  $h$  of  $f$  along  $I$  is almost entropic. If  $\xi_i$  takes values in a finite set  $X_i$  then  $\xi_I = (\xi_i)_{i \in I}$  ranges in the product of  $X_i$ ,  $i \in I$ . For every element  $x_I$  of the product that is attained with a positive probability, let  $\eta^{x_I}$  be the random vector  $\xi_{N \setminus I}$  conditioned on the event that  $\xi_I = x_I$ . The entropy function of  $\eta^{x_I}$  is denoted by  $g_{x_I}$ . By an easy calculation, the contraction  $h$  equals the convex combination of the entropy functions  $g_{x_I}$  weighted by the probabilities of the events  $\xi_I = x_I$ . Since  $cl(\mathbf{H}_N^{\text{ent}})$  is convex [44, Theorem 1],  $h$  is almost entropic.  $\square$

2.3. When  $f$  and  $g$  are polymatroidal rank functions on the same ground set  $N$ , their *convolution*  $f * g$  is defined as

$$f * g(I) = \min_{J \subseteq I} \{ f(J) + g(I \setminus J) \}, \quad I \subseteq N.$$

If both  $f$  and  $g$  are modular, then  $f * g$  is also modular, assigning the values  $\min\{f(i), g(i)\}$  to the singletons  $i$  of  $N$ . By [26, Theorem 2.5],  $(N, f * g)$  is a polymatroid whenever  $g$  alone is modular. The following simple assertion may help to build intuition for later proofs.

**Lemma 2.** *Let  $f, g$  be two polymatroids on  $N$  where  $g$  is modular, and  $i \in N$ . If  $f(j) \leq g(j)$  for all  $j \in N \setminus i$  then*

$$\begin{aligned} f * g(I) &= f(I), \\ f * g(iI) &= \min \{ f(I) + g(i), f(iI) \}, \end{aligned} \quad I \subseteq N \setminus i.$$

*If, additionally,  $f(i) \leq g(i)$  then  $f * g = f$ .*

*Proof.* By submodularity of  $f$ , for  $J \subseteq I \subseteq N \setminus i$

$$f(I) + g(\emptyset) \leq f(J) + f(I \setminus J) \leq f(J) + \sum_{j \in I \setminus J} f(j) \leq f(J) + g(I \setminus J)$$

using that  $f(j) \leq g(j)$ ,  $j \in N \setminus i$ , and modularity of  $g$ . This proves that  $f * g(I)$  equals  $f(I)$ . Similarly,

$$f(iI) + g(\emptyset) \leq f(iJ) + f(I \setminus J) \leq f(iJ) + g(I \setminus J)$$

and

$$f(I) + g(i) \leq f(J) + f(I \setminus J) + g(i) \leq f(J) + g(iI \setminus J).$$

Hence,  $f * g(iI)$  is the smaller of the numbers  $f(iI)$  and  $f(I) + g(i)$ .  $\square$

In a notable instance of the convolution, the difference between  $f * g$  and  $f$  is at most at a singleton. This will be used in Theorem 4 to shift almost entropic points.

**Corollary 1.** *Let  $(N, f)$  be a polymatroid,  $i \in N$ , and*

$$\max_{j \in N \setminus i} [f(ij) - f(j)] \leq t \leq f(i).$$

*Let  $(N, g)$  be a modular polymatroid with  $g(i) = t$  and  $f(j) \leq g(j)$ ,  $j \in N \setminus i$ . Then  $f * g$  is equal to  $f$  up to  $f * g(i) = t$ .*

*Proof.* The assumption  $t \leq f(i)$  implies  $f * g(i) = t$ . By Lemma 2,  $f * g$  is equal to  $f$  on the subsets of  $N \setminus i$ . Let  $I \subseteq N \setminus i$  and  $I$  contain some  $j$ . By submodularity and the lower bound on  $t$ ,

$$f(iI) \leq f(I) + f(ij) - f(j) \leq f(I) + t = f(I) + g(i).$$

It follows by Lemma 2 that  $f * g(iI) = f(iI)$ .  $\square$

Since the operation  $*$  is commutative and associative, the convolution of a polymatroid  $f$  with a modular polymatroid  $g$  can be computed iteratively by Lemma 2. It suffices to write  $g$  as the multiple convolution of modular polymatroids  $g_i$ ,  $i \in N$ , such that  $g_i(i) = g(i)$  and  $g_i(j) = r$ ,  $j \in N \setminus i$ , where  $r$  is larger than the values of  $f$  and  $g$  on all singletons.

2.4. The last subsection of this section defines a one-element parallel extension of a polymatroid. This turns into a principal extension when modified by a convolution. Then, the added element is contracted. The polymatroid obtained by these three constructions is employed later in Sections 5 and 6.

Two points  $i, j$  of a polymatroid  $(N, f)$  are *parallel* if  $f(iJ) = f(jJ)$  for every  $J \subseteq N$ . Given any  $i \in N$ , it is always possible to extend  $f$  to  $\mathbb{R}^{\mathcal{P}(0 \cup N)}$ , where  $0 \notin N$ , such that  $i$  and  $0$  are parallel in the extension. More generally, for  $L \subseteq N$  the *extension of  $f$  by 0 parallel to  $L$*  is the polymatroid  $(0 \cup N, h)$  given by  $h(J) = f(J)$  and  $h(0 \cup J) = f(L \cup J)$  where  $J \subseteq N$ . If  $f$  is the entropy function of  $(\xi_i)_{i \in N}$  then  $h$  is entropic as well, completing the random vector by the variable  $\xi_0 = (\xi_i)_{i \in L}$ .

This parallel extension is convolved with the modular polymatroid  $(0 \cup N, g)$  having  $g(0) = t \leq f(L)$  and  $g(i) \geq f(i)$ ,  $i \in N$ , to arrive at the *principal extension*  $f_{L,t}$  of  $f$  on the subset  $L$  with the value  $t$  [26]. By Lemma 2,

$$f_{L,t}(0 \cup I) = \min \{f(I) + t, f(L \cup I)\}, \quad I \subseteq N.$$

In turn, the principal extension  $f_{L,t}$  is contracted by 0 to get the polymatroid on  $N$  with the rank function

$$(1) \quad f_{L,t}^*(I) = \min \{f(I), f(L \cup I) - t\}, \quad I \subseteq N.$$

**Lemma 3.** *If  $(N, f)$  is almost entropic,  $L \subseteq N$  and  $0 \leq t \leq f(L)$  then so is  $(N, f_{L,t}^*)$ .*

*Proof.* If  $f$  is entropic then  $f_{L,t}^* \in cl(\mathbf{H}_N^{\text{ent}})$  by [34, Theorem 2] and Lemma 1 which implies the assertion.  $\square$

Under some assumptions, it is possible to find all minima in (1). Recall that the  $f$ -closure  $cl(I)$  of  $I \subseteq N$  consists of those  $i \in N$  that satisfy  $f(iI) = f(I)$ . By monotonicity and submodularity,  $f(cl(I)) = f(I)$ .

**Lemma 4.** *If  $(N, f)$  is a polymatroid,  $L \subseteq N$  and  $0 \leq t \leq f(L)$  such that*

$$(2) \quad t \leq \min_{I \subseteq N, L \not\subseteq cl(I)} \max_{\ell \in L \setminus cl(I)} [f(\ell \cup I) - f(I)]$$

then

$$f_{L,t}^*(I) = \begin{cases} f(I) - t, & \text{when } L \subseteq cl(I), \\ f(I), & \text{otherwise,} \end{cases} \quad I \subseteq N.$$

*Proof.* The assumption  $0 \leq t \leq f(L)$  is needed to derive (1). If  $L \subseteq cl(I)$  then  $f(L \cup I) \leq f(L \cup cl(I)) = f(cl(I)) = f(I)$ . Hence, the inequality is tight and the minimum in (1) equals  $f(I) - t$ . Otherwise, by the assumption (2),  $t \leq f(\ell \cup I) - f(I)$  for some  $\ell \in L \setminus cl(I)$ . Since  $f(L \cup I) - t \geq f(\ell \cup I) - t \geq f(I)$  the minimum in (1) equals  $f(I)$ .  $\square$

*Remark 1.* The special instance of Lemma 4 is used in the proof of Theorem 4 to shift almost entropic points. There,  $L$  equals a singleton  $k$  contained in  $cl(N \setminus k)$ . In such a case, (2) is a consequence of

$$(3) \quad t \leq \min_{j \in N \setminus k} [f(N \setminus j) - f(N \setminus jk)]$$

because each maximum  $f(kI) - f(I)$  in (2) dominates the right-hand side of (3) by submodularity.

In another special instance  $L = N$  of Lemma 4, the polymatroid  $f_{L,t}^*$  is called the *truncation* of  $f$  by  $t$ , or to  $f(N) - t$ . It was applied e.g. in [8] to investigate linear polymatroids.

### 3. DECOMPOSITION INTO TIGHT AND MODULAR POLYMATROIDS

The cone  $\mathbf{H}_N$  of polymatroidal rank functions  $h$  decomposes into the direct sum of the cone  $\mathbf{H}_N^{\text{ti}}$  of *tight* rank functions and the cone  $\mathbf{H}_N^{\text{mod}}$  of modular functions. Here,  $h$  is tight if  $h(N) = h(N \setminus i)$ ,  $i \in N$ . The decomposition can be written as  $h = h^{\text{ti}} + h^{\text{m}}$  where

$$\begin{aligned} h^{\text{ti}}(I) &= h(I) - \sum_{i \in I} [h(N) - h(N \setminus i)], \\ h^{\text{m}}(I) &= \sum_{i \in I} [h(N) - h(N \setminus i)], \end{aligned} \quad I \subseteq N.$$

It is unique because the linear spaces  $\mathbf{H}_N^{\text{ti}} - \mathbf{H}_N^{\text{ti}}$  and  $\mathbf{H}_N^{\text{mod}} - \mathbf{H}_N^{\text{mod}}$  intersect at the origin. In symbols,  $\mathbf{H}_N = \mathbf{H}_N^{\text{ti}} \oplus \mathbf{H}_N^{\text{mod}}$ .

**Theorem 1.** *If  $h \in \text{cl}(\mathbf{H}_N^{\text{ent}})$  then  $h^{\text{ti}}$  is almost entropic.*

*Proof.* Let  $(N, h)$  be a polymatroid,  $N'$  be a disjoint copy of  $N$  and  $i \mapsto i'$  a bijection between them. The polymatroid  $(N, h)$  extends to  $(N \cup N', f)$  by

$$f(I \cup J') = h(I \cup J), \quad I, J \subseteq N,$$

where  $J' = \{j' : j \in J\}$ . Thus, each  $i'$  is parallel to  $i$ . Let  $(N \cup N', g)$  be a modular polymatroid. Then, for  $I \subseteq N$

$$f * g(I \cup N') = \min_{J \subseteq I, K \subseteq N} [h(J \cup K) + g(I \setminus J) + g(N' \setminus K')].$$

By monotonicity of  $g$ , the bracket does not grow when  $K$  is replaced by  $K \cup J$ . Hence, the minimization can be restricted to the situations when  $J = K \cap I$ , and

$$f * g(I \cup N') = \min_{K \subseteq N} [h(K) + g(I \setminus K) + g(N' \setminus K')].$$

If  $g(i) + g(i') = h(i)$  for  $i \in N$  then

$$f * g(I \cup N') = \min_{K \subseteq N} [h(K) + \sum_{i \in I \setminus K} h(i) + g(N' \setminus (I' \cup K'))].$$

By submodularity of  $h$ , this minimization can be restricted to  $K \supseteq I$ , thus

$$(4) \quad f * g(I \cup N') = \min_{I \subseteq K \subseteq N} [h(K) + g(N' \setminus K')], \quad I \subseteq N.$$

In the case when

$$g(i') = h^{\text{m}}(i) = h(N) - h(N \setminus i) \leq h(i), \quad i \in N,$$

$h$  is decomposed to  $h^{\text{ti}} + h^{\text{m}}$  and the minimum in (4) is equal to  $h^{\text{ti}}(I) + h^{\text{m}}(N)$ . It is attained for  $K = I$ . It follows that  $h^{\text{ti}}(I) = f * g(I \cup N') - f * g(N')$ . Hence,  $h^{\text{ti}}$  is the contraction of  $f * g$  along  $N'$ .

If  $h \in \text{cl}(\mathbf{H}_N^{\text{ent}})$  then  $f$  is almost entropic. The convolution theorem [34, Theorem 2] implies that  $f * g \in \text{cl}(\mathbf{H}_N^{\text{ent}})$ . By Lemma 1,  $h^{\text{ti}} \in \text{cl}(\mathbf{H}_N^{\text{ent}})$ .  $\square$

The closure of the entropic region decomposes analogously to  $\mathbf{H}_N$ . As a consequence, the intersection  $\text{cl}(\mathbf{H}_N^{\text{ent}}) \cap \mathbf{H}_N^{\text{ti}}$  is equal to

$$\text{cl}(\mathbf{H}_N^{\text{ent}})^{\text{ti}} = \{f^{\text{ti}} : f \in \text{cl}(\mathbf{H}_N^{\text{ent}})\}.$$

**Corollary 2.**  $\text{cl}(\mathbf{H}_N^{\text{ent}}) = [\text{cl}(\mathbf{H}_N^{\text{ent}}) \cap \mathbf{H}_N^{\text{ti}}] \oplus \mathbf{H}_N^{\text{mod}}$ .

*Proof.* Theorem 1 and  $\mathbf{H}_N = \mathbf{H}_N^{\text{ti}} \oplus \mathbf{H}_N^{\text{mod}}$  imply the inclusion  $\subseteq$ . The reverse one follows from the facts that  $\text{cl}(\mathbf{H}_N^{\text{ent}})$  is a convex cone and  $\mathbf{H}_N^{\text{mod}} \subseteq \mathbf{H}_N^{\text{ent}}$  [34, Lemma 2].  $\square$

It is open whether  $\mathbf{H}_N^{\text{ent}}$  equals  $[\mathbf{H}_N^{\text{ent}} \cap \mathbf{H}_N^{\text{ti}}] \oplus \mathbf{H}_N^{\text{mod}}$ .

In the remaining part of this section it is shown that Corollary 2 is equivalent to [7, Theorem 1] on balanced inequalities.

Any nonempty closed convex cone  $K$  in a Euclidean space is expressible as intersection of homogeneous closed halfspaces. This is reflected in the notion of the polar cone  $K^\circ$  of  $K$  that consists of the outer normal vectors to  $K$  at the origin,

$$K^\circ = \{(\vartheta_I)_{I \subseteq N} \in \mathbb{R}^{\mathcal{P}(N)} : \sum_{I \subseteq N} \vartheta_I h(I) \leq 0 \text{ for all } h \in K\},$$

see [39, Section 14]. For example, the polar of  $\mathbf{H}_N^{\text{mod}}$  can be defined by the inequalities  $\sum_{I \ni i} \vartheta_I \leq 0$ ,  $i \in N$ ; substituting  $r_i^{N \setminus i}$ ,  $i \in N$ , for  $h$ . The polars of  $\mathbf{H}_N^{\text{ent}}$  and  $cl(\mathbf{H}_N^{\text{ent}})$  coincide and are defined by the very linear information-theoretic inequalities.

By [39, Corollary 16.4.2], Corollary 2 is equivalent to

$$(5) \quad (\mathbf{H}_N^{\text{ent}})^\circ = (cl(\mathbf{H}_N^{\text{ent}})^{\text{ti}})^\circ \cap (\mathbf{H}_N^{\text{mod}})^\circ.$$

It was used tacitly also that  $(\mathbf{H}_N^{\text{ent}})^{\circ\circ} = cl(\mathbf{H}_N^{\text{ent}})$ , and that  $cl(\mathbf{H}_N^{\text{ent}})^{\text{ti}}$ ,  $\mathbf{H}_N^{\text{mod}}$  and their sum are closed. The polar of  $cl(\mathbf{H}_N^{\text{ent}})^{\text{ti}}$  consists of the vectors  $(\vartheta_I)_{I \subseteq N}$  satisfying  $\sum_{I \subseteq N} \vartheta_I h^{\text{ti}}(I) \leq 0$ ,  $h \in cl(\mathbf{H}_N^{\text{ent}})$ , which rewrites to

$$(6) \quad \sum_{I \subseteq N} \vartheta_I h(I) - \sum_{i \in N} [h(N) - h(N \setminus i)] \sum_{I \ni i} \vartheta_I \leq 0, \quad h \in \mathbf{H}_N^{\text{ent}}.$$

In turn, (5) can be rephrased as [7, Theorem 1]: *Given  $(\vartheta_I)_{I \subseteq N}$ , the inequality  $\sum_{I \subseteq N} \vartheta_I h(I) \leq 0$  holds for all  $h \in \mathbf{H}_N^{\text{ent}}$  if and only if (6) is valid and  $\sum_{I \ni i} \vartheta_I \leq 0$ ,  $i \in N$ .*

Writing,  $\tau_I = \vartheta_I$  when  $|I| < |N| - 1$ ,  $\tau_{N \setminus i} = \vartheta_{N \setminus i} + \sum_{I \ni i} \vartheta_I$ ,  $i \in N$ , and  $\tau_N = \vartheta_N - \sum_{I \subseteq N} |I| \vartheta_I$  the inequality in (6) rewrites to  $\sum_{I \subseteq N} \tau_I h(I) \leq 0$ . This one is *balanced* in the sense  $\sum_{I \ni i} \tau_I = 0$ ,  $i \in N$ . Thus, (6) expresses all balanced information-theoretic inequalities.

#### 4. ENTROPY REGION: REGULAR FACES OF $cl(\mathbf{H}_N^{\text{ent}})$

As mentioned earlier, the relative interior of  $cl(\mathbf{H}_N^{\text{ent}})$  belongs to the entropy region  $\mathbf{H}_N^{\text{ent}}$ . Thus,  $\mathbf{H}_N^{\text{ent}}$  and  $cl(\mathbf{H}_N^{\text{ent}})$  differ only on the relative boundary of the latter. This section proves a stronger relation between them, motivated by the decomposition in Corollary 2.

**Theorem 2.**  $ri(cl(\mathbf{H}_N^{\text{ent}})^{\text{ti}}) \oplus \mathbf{H}_N^{\text{mod}} \subseteq \mathbf{H}_N^{\text{ent}}$ .

The proof presented below is based on an auxiliary lemma. At the end of the section, faces of  $cl(\mathbf{H}_N^{\text{ent}})$  are discussed.

**Lemma 5.** *The cone  $cl(\mathbf{H}_N^{\text{ent}})^{\text{ti}}$  contains a dense set of entropic points.*

A proof resorts to polymatroids constructed from groups. Recall that a polymatroid  $(N, f)$  is *group-generated* if there exists a finite group  $G$  with subgroups  $G_i$ ,  $i \in N$ , such that  $f(I) = \ln |G| / |G_I|$  for  $I \subseteq N$ . Here,  $G_I$  abbreviates  $\bigcap_{i \in I} G_i$ . Such a polymatroid is always entropic. In fact, the group  $G$  is endowed with the uniform probability measure and the polymatroid equals the entropy function of  $(\xi_i)_{i \in N}$  where  $\xi_i$  is the factormapping of  $G$  on the family  $G/G_i$  of left cosets of  $G_i$ . The divisions of the group-generated polymatroidal rank functions by positive integers are dense in  $cl(\mathbf{H}_N^{\text{ent}})$  [6, Theorem 4.1].

*Proof of Lemma 5.* Given  $\varepsilon > 0$  and  $g \in cl(\mathbf{H}_N^{\text{ent}})^{\text{ti}}$  there exists, by Theorem 1, a random vector whose entropy function  $h$  satisfies

$$\|h - g\|_\infty \triangleq \max_{I \subseteq N} |h(I) - g(I)| \leq \varepsilon.$$

Since  $g$  is tight

$$\begin{aligned} h^m(N) &= h^m(N) - g^m(N) \\ &\leq \sum_{i \in N} |h(N) - g(N)| + |h(N \setminus i) - g(N \setminus i)| \leq 2\varepsilon |N|. \end{aligned}$$

It is possible to assume that the vector sits on a finite set endowed with the uniform probability measure. By [34, Remark 11], there exists a group  $G$ , group-generated polymatroid  $f$  and integer  $\ell \geq 1$  such that  $\|\frac{1}{\ell}f - h\|_\infty \leq \varepsilon$ . Therefore,

$$\frac{1}{\ell}f^m(N) \leq |\frac{1}{\ell}f^m(N) - h^m(N)| + h^m(N) \leq 4\varepsilon|N|.$$

Let  $(\xi_i)_{i \in N}$  be the corresponding random vector of factormappings of  $G$  onto  $G/G_i$  whose entropy function equals  $f$ . If  $I \subseteq N$  then  $\xi_I \triangleq (\xi_i)_{i \in I}$  takes  $|G/G_I|$  values, each one with the same probability and  $f(I) = \ln |G/G_I|$ . Therefore, for every  $j \in N$  there exists a random variable  $\eta_j$  defined on  $G$  such that it is constant on each coset of  $G/G_N$ , takes  $|G_{N \setminus j}/G_N|$  values and  $(\xi_{N \setminus j}, \eta_j)$  takes  $|G/G_N| = |G/G_{N \setminus j}| |G_{N \setminus j}/G_N|$  values. Necessarily,  $\eta_j$  is a function of  $\xi_N$ , its entropy is  $\ln |G_{N \setminus j}/G_N|$ ,  $\eta_j$  is stochastically independent of  $\xi_{N \setminus j}$ , and they together determine  $\xi_N$ . Let  $h'$  denote the entropy function of  $(\zeta_i)_{i \in N}$  where  $\zeta_i = (\xi_i, \eta_N)$  and  $\eta_N = (\eta_j)_{j \in N}$ . By construction,  $h'(N \setminus i)$  is the entropy of  $(\xi_{N \setminus i}, \eta_N)$ ,  $i \in N$ , and  $h'(N)$  is the entropy of  $(\xi_N, \eta_N)$ . Hence,  $h'$  is a tight entropy function. For  $I \subseteq N$

$$f(I) \leq h'(I) \leq f(I) + \sum_{j \in N} \ln |G_{N \setminus j}/G_N| = f(I) + f^m(N).$$

It follows that  $\|\frac{1}{\ell}h' - \frac{1}{\ell}f\|_\infty \leq \frac{1}{\ell}f^m(N) \leq 4\varepsilon|N|$ . Hence,

$$\|\frac{1}{\ell}h' - g\|_\infty \leq \|\frac{1}{\ell}h' - \frac{1}{\ell}f\|_\infty + \|\frac{1}{\ell}f - h\|_\infty + \|h - g\|_\infty \leq 4\varepsilon|N| + 2\varepsilon.$$

By [34, Lemma 4], the tight polymatroid  $\frac{1}{\ell}h' + \delta r_1$  is entropic for any  $\delta > 0$ . Thus,  $\|(\frac{1}{\ell}h' + \varepsilon r_1) - g\|_\infty \leq 4\varepsilon|N| + 3\varepsilon$  where  $\varepsilon$  can be arbitrarily small.  $\square$

*Remark 2.* It is of separate interest that  $cl(\mathbf{H}_N^{\text{ent}})^{\text{ti}}$  contains a dense set of points in the form  $\frac{1}{m}h''$  where  $h''$  is group-generated and  $m \geq 1$  integer. In fact, the tight entropy function  $h'$  from the previous proof need not be group-generated but arises from random variables defined on  $G$  with the uniform probability measure. Then, by [34, Remark 11],  $h'$  can be arbitrarily well approximated by  $\frac{1}{m}h''$  with  $h''$  group-generated. Since  $h'$  is tight the construction of that remark provides  $h''$  tight as well. Thus, to a given  $g \in cl(\mathbf{H}_N^{\text{ent}})^{\text{ti}}$  it is possible to construct  $\frac{1}{\ell m}h''$  arbitrarily close, as in the above proof.

*Proof of Theorem 2.* Since,  $\mathbf{H}_N^{\text{mod}} \subseteq \mathbf{H}_N^{\text{ent}}$  [34, Lemma 2] and  $\mathbf{H}_N^{\text{ent}}$  is closed to sums it suffices to prove that  $ri(cl(\mathbf{H}_N^{\text{ent}})^{\text{ti}}) \subseteq \mathbf{H}_N^{\text{ent}}$ . Argumentation is analogous to that in the proof of [34, Theorem 1]. By [34, Lemma 3], the matroidal rank functions  $r_1^J$  with  $J \subseteq N$  and  $|J| < |N| - 1$  are linearly independent. Since they are tight and their nonnegative combinations are entropic they span a polyhedral cone contained  $\mathbf{H}_N^{\text{ti}} \cap \mathbf{H}_N^{\text{ent}}$  whose dimension  $2^{|N|} - |N| - 1$  is the same as that of  $cl(\mathbf{H}_N^{\text{ent}})^{\text{ti}}$ . Therefore, if  $\varepsilon > 0$  then the set  $B_\varepsilon$  of polymatroids  $\sum_{J: |J| < |N| - 1} \alpha_J r_1^J$ , where  $0 < \alpha_J < \varepsilon$ , is open in the linear space  $\mathbf{H}_N^{\text{ti}} - \mathbf{H}_N^{\text{ti}}$  and the shifts of these sets provide a base for the relative topology.

Hence, if  $g$  belongs to the relative interior of  $cl(\mathbf{H}_N^{\text{ent}})^{\text{ti}}$  then it belongs to such a shift contained in the relative interior. It follows that  $g - B_\varepsilon$  is a subset of the relative interior for  $\varepsilon > 0$  sufficiently small. Since  $g - B_\varepsilon$  is a relatively open subset of  $cl(\mathbf{H}_N^{\text{ent}})^{\text{ti}}$  it contains an entropic polymatroid  $h$ , by Lemma 5. This implies that  $g$  can be written as  $h + \sum_{J: |J| < |N| - 1} \alpha_J r_1^J$  where all  $\alpha_J$  are nonnegative, and thus is entropic.  $\square$

A convex subset  $F$  of a convex set  $K$  is a *face* if it contains every line segment of  $K$  which has an interior point in  $F$ . A face of a convex cone is a convex cone.



Let us call a face  $F$  of  $cl(\mathbf{H}_N^{\text{ent}})$  *regular* if all relative interior points of  $F$  are entropic, thus  $ri(F) \subseteq \mathbf{H}_N^{\text{ent}}$ . The trivial face  $F = cl(\mathbf{H}_N^{\text{ent}})$  is regular. Since the cones  $cl(\mathbf{H}_N^{\text{ent}})^{\text{ti}}$  and  $\mathbf{H}_N^{\text{mod}}$  are defined by imposing tightness in monotonicity and submodularity, they are faces of  $cl(\mathbf{H}_N^{\text{ent}})$ , by Theorem 1. Each face of  $\mathbf{H}_N^{\text{mod}}$  is a face of  $cl(\mathbf{H}_N^{\text{ent}})$ , and is regular because  $\mathbf{H}_N^{\text{mod}} \subseteq \mathbf{H}_N^{\text{ent}}$ . By Theorem 2, the face  $F = cl(\mathbf{H}_N^{\text{ent}})^{\text{ti}}$  is regular.

### 5. SELFADHESIVITY AND TIGHTNESS

This section recalls the notion of selfadhesivity and explores its relation to the decomposition  $h = h^{\text{ti}} + h^{\text{m}}$  of polymatroids. The role of selfadhesivity in proving information-theoretic inequalities is briefly discussed and compared to an alternative technique by [28].

Two polymatroids  $(N, h)$  and  $(M, g)$  are *adhesive* [33], or *adhere*, if a polymatroid  $(N \cup M, f)$  exists such that  $f(I) = h(I)$  for  $I \subseteq N$ ,  $f(J) = g(J)$  for  $J \subseteq M$ , and  $f(N) + f(M) = f(N \cup M) + f(N \cap M)$ . Thus, the rank function  $f$  is a common extension of  $h$  and  $g$ , and the last equality expresses the adherence. A polymatroid  $(N, h)$  is *selfadhesive at*  $O \subseteq N$  if it adheres with the  $\pi$ -copy  $(\pi(N), h_\pi)$  defined by a bijection  $\pi: N \rightarrow \pi(N)$  such that  $O = N \cap \pi(N)$ ,  $\pi(i) = i$  for  $i \in O$ , and  $h_\pi(\pi(I)) = h(I)$ ,  $I \subseteq N$ . A polymatroid is *selfadhesive* if it is selfadhesive at each  $O \subseteq N$ .

The rank functions of selfadhesive polymatroids on  $N$  form the polyhedral cone  $\mathbf{H}_N^{\text{sa}}$  [33]. This cone decomposes similarly to  $\mathbf{H}_N = \mathbf{H}_N^{\text{ti}} \oplus \mathbf{H}_N^{\text{mod}}$ .

**Theorem 3.** *If  $h \in \mathbf{H}_N^{\text{sa}}$  then  $h^{\text{ti}}$  is selfadhesive.*

*Proof.* Let a polymatroid  $(N, h)$  adhere with a  $\pi$ -copy at  $O = N \cap \pi(N)$  and  $\hat{N} = N \cup \pi(N)$ . Thus, there exists an adhesive extension  $(\hat{N}, \hat{h})$ . This extension is further extended to  $(\hat{N} \cup \hat{N}', f)$ , doubling each element of  $\hat{N}$  by a parallel one in  $\hat{N}'$ , disjoint with  $\hat{N}$ . Similarly to the proof of Theorem 1, a modular polymatroid  $(\hat{N} \cup \hat{N}', g)$  is constructed below such that the contraction of  $f * g$  along  $\hat{N}'$  witnesses that  $(N, h^{\text{ti}})$  is selfadhesive at  $O$ .

The modular rank function  $g$  is defined by

$$\begin{aligned} g(i) &= g(\pi(i)) = h(i) + h(N \setminus i) - h(N), \\ g(i') &= g(\pi(i')) = h(N) - h(N \setminus i), \end{aligned} \quad i \in N.$$

Since  $g(i) + g(i') = \hat{h}(i)$  and  $g(\pi(i)) + g(\pi(i')) = \hat{h}(\pi(i))$ , an analogue of (4) takes the form

$$(7) \quad f * g(I \cup \hat{N}') = \min_{I \subseteq K \subseteq \hat{N}} [\hat{h}(K) + g(\hat{N}' \setminus K)], \quad I \subseteq \hat{N},$$

arguing as in the proof of Theorem 1.

If  $i \in N \setminus O$  then

$$\begin{aligned} \hat{h}^{\text{m}}(i) &= \hat{h}(\hat{N}) - \hat{h}(\hat{N} \setminus i) = 2h(N) - h(O) - [h(N \setminus i) + h(N) - h(O)] \\ &= h(N) - h(N \setminus i) = h^{\text{m}}(i) = g(i') \end{aligned}$$

because  $\hat{h}$  is an adhesive extension of  $h$  and  $h_\pi$ . Analogously, if  $i \in N \setminus O$  then  $\hat{h}^m(\pi(i)) = h^m(i) = g(\pi(i)')$ . Therefore, the bracket in (7) rewrites to

$$\begin{aligned} & \hat{h}^{\text{ti}}(K) + \hat{h}^m(K) + \hat{h}^m(\hat{N} \setminus (O \cup K)) + g(O' \setminus K') \\ &= \hat{h}^{\text{ti}}(K) + \hat{h}^m(\hat{N} \setminus (O \setminus K)) + h^m(O \setminus K). \end{aligned}$$

Hence, the minimization in (7) can be further restricted to  $K \subseteq I \cup O$  and

$$f * g(I \cup \hat{N}') = \hat{h}^m(\hat{N} \setminus (I \cup O)) + \min_{I \subseteq K \subseteq I \cup O} [\hat{h}(K) + h^m(O \setminus K)], \quad I \subseteq \hat{N}.$$

The above minimum can be found in special cases. First,

$$f * g(\hat{N} \cup \hat{N}') = \hat{h}(\hat{N}) = 2h(N) - h(O)$$

using that  $\hat{h}$  extends adhesively  $h$  and its  $\pi$ -copy. Second,

$$f * g(I \cup \hat{N}') = h^{\text{ti}}(I) + h^m(N) + h^m(N \setminus O), \quad I \subseteq N,$$

using that  $\hat{h}(K) + h^m(O \setminus K) = h^{\text{ti}}(K) + h^m(I \cup O)$ . Third,

$$f * g(\pi(I) \cup \hat{N}') = h^{\text{ti}}(I) + h^m(N) + h^m(N \setminus O), \quad I \subseteq N,$$

by symmetry. It follows that the contraction of  $f * g$  along  $\hat{N}'$  extends  $h^{\text{ti}}$  and its  $\pi$ -copy. The rank of the contraction is

$$[2h(N) - h(O)] - [h^m(N) + h^m(N \setminus O)] = 2h^{\text{ti}}(N) - h^{\text{ti}}(O)$$

whence the extension is selfadhesive.  $\square$

**Corollary 3.**  $\mathbf{H}_N^{\text{sa}} = [\mathbf{H}_N^{\text{sa}} \cap \mathbf{H}_N^{\text{ti}}] \oplus \mathbf{H}_N^{\text{mod}}$ .

*Proof.* The inclusion  $\subseteq$  follows from Theorem 3. Since the modular polymatroids have selfadhesive modular extensions and  $\mathbf{H}_N^{\text{sa}}$  is a convex cone the opposite inclusion holds as well.  $\square$

The convex cone  $\text{cl}(\mathbf{H}_N^{\text{ent}})$  is not polyhedral [35], thus its polar cone is not finitely generated. There are infinite sets of linear information-theoretic inequalities [45, 28, 13, 14, 11], some of them rigorously proved and hundreds of them generated in computer experiments. The experiments are based on the fact that the entropic polymatroids are selfadhesive,  $\mathbf{H}_N^{\text{ent}} \subseteq \mathbf{H}_N^{\text{sa}}$ , and iterations of the idea. None of the experiments seems to have taken into account the possible reduction by imposing the tightness, cf. Corollary 3.

A linear information-theoretic inequality

$$\sum_{I \subseteq N} \vartheta_I h(I) \leq 0 \text{ for all } h \in \mathbf{H}_N^{\text{ent}}$$

is of non-Shannon type if  $(\vartheta_I)_{I \subseteq N} \in (\mathbf{H}_N^{\text{ent}})^\circ$  is not in  $\mathbf{H}_N^\circ$ . There are two techniques for proving non-Shannon-type inequalities: either by selfadhesivity, as implicit in the original proof of Zhang-Yeung inequality [44], or alternatively by a lemma of Ahlswede and Körner [12], as proposed in [28]. Recently it was found that the two techniques have the same power [22]. Actually, the original lemma from [12] is not needed and only the following version on extensions suffices for proofs of [28, 22].

**Lemma 6.** *If  $(N, h)$  is almost entropic,  $i \in N$  and  $i' \notin N$  then the polymatroid has an almost entropic extension  $(i' \cup N, g)$  such that*

$$\begin{aligned} g(i' \cup N \setminus i) &= g(N \setminus i), \\ g(i' \cup I) - g(i') &= g(i \cup I) - g(i), \quad I \subseteq N \setminus i. \end{aligned}$$

*Proof.* The assumption implies that there exists an almost entropic and adhesive extension  $(i' \cup N, f)$  of  $(N, h)$  and its copy at  $N \setminus i$ . Let  $g$  denote the contraction  $f_{L,t}^*$  of the principal extension  $f_{L,t}$  of  $f$  on the singleton  $L = i'$  with  $t = h(N) - h(N \setminus i)$ . By Lemma 3,  $g$  is almost entropic. The value  $t$  is at most  $h(i) = f(L)$  whence (1) applies and takes the form

$$g(I) = \min \{f(I), f(i' \cup I) - h(N) + h(N \setminus i)\}, \quad I \subseteq i' \cup N.$$

If  $I \subseteq N \setminus i$  then, using the properties of  $f$  and submodularity,

$$\begin{aligned} g(I) &= \min\{h(I), h(i \cup I) - h(N) + h(N \setminus i)\} = h(I), \\ g(i \cup I) &= \min\{h(i \cup I), f(i' \cup i \cup I) - f(i' \cup N) + f(N)\} = h(i \cup I), \\ g(i' \cup I) &= h(i \cup I) - h(N) + h(N \setminus i). \end{aligned}$$

The first and second equation show that  $g$  is an extension of  $h$ . This and the last one imply  $g(i' \cup N \setminus i) = g(N \setminus i)$  and  $g(i' \cup I) - g(i') = g(i \cup I) - g(i)$ .  $\square$

The main ingredient in the above proof is a contraction of a principal extension, which relies on convolution. This indicates that selfadhesivity, convolution and other constructions on polymatroids seem to be powerful enough to rephrase all existing approaches to proofs of the linear information-theoretic inequalities.

## 6. ENTROPY REGION OF FOUR VARIABLES

This section presents more special applications of polymatroidal constructions and consequences of Theorems 1 and 2 when the ground set  $N$  has four elements. Results on reduction of  $cl(\mathbf{H}_N^{\text{ent}})$  will be used later when minimizing Ingleton score. It is assumed that the four elements  $i, j, k, l$  of  $N$  are always different. In the notation for cones the subscript  $N$  is omitted, for example  $\mathbf{H} = \mathbf{H}_N$ .

When studying the entropic functions of four variables the crucial role is played by the expression

$$h(ik) + h(jk) + h(il) + h(jl) + h(kl) - h(ij) - h(k) - h(l) - h(ikl) - h(jkl)$$

where  $h \in \mathbf{H}$ . It is interpreted also as the scalar product  $\square_{ij}h$  of

$$\square_{ij} = \delta_{ik} + \delta_{il} + \delta_{jk} + \delta_{jl} + \delta_{kl} - \delta_{ij} - \delta_k - \delta_l - \delta_{ikl} - \delta_{jkl}$$

with  $h$ . The inequality  $\square_{ij}h \geq 0$  holds when  $h$  is linear, see the works of Ingleton [20, 21]. Let  $\mathbf{H}^\square$  denote the polyhedral cone of the functions  $h \in \mathbf{H}$  that satisfy the six instances of the Ingleton inequality obtained by the permutation symmetry. By [36, Lemma 3],  $\mathbf{H}^\square$  has dimension 15 and is generated by linear polymatroidal rank functions. Therefore, the functions from  $ri(\mathbf{H}^\square)$  are entropic due to [34, Theorem 1].

By [36, Lemma 4], any  $h \in \mathbf{H} \setminus \mathbf{H}^\square$  violates exactly one of the six Ingleton inequalities. Let  $\mathbf{H}_{(ij)}^\square$  denote the cone of functions  $h \in \mathbf{H}$  with  $\square_{ij}h \leq 0$ . It follows that  $\mathbf{H}$  is union of  $\mathbf{H}^\square$  with the six cones  $\mathbf{H}_{(ij)}^\square, \dots, \mathbf{H}_{(kl)}^\square$ . Focusing primarily on  $cl(\mathbf{H}^{\text{ent}})$ , it contains  $\mathbf{H}^\square$  and is contained in the union. By symmetry, it remains to study  $cl(\mathbf{H}^{\text{ent}}) \cap \mathbf{H}_{(ij)}^\square$ .

Let  $\mathbf{L}_{ij}$  denote the cone  $cl(\mathbf{H}^{\text{ent}})^{\text{ti}} \cap \mathbf{H}_{(ij)}^\square$  of tight and almost entropic polymatroids  $h$  that satisfy the reversed Ingleton inequality  $\square_{ij}h \leq 0$ .

**Corollary 4.**  $cl(\mathbf{H}^{\text{ent}}) \cap \mathbf{H}_{(ij)}^\square = \mathbf{L}_{ij} \oplus \mathbf{H}^{\text{mod}}$ .

*Proof.* Since the expression  $\square_{ij}h$  is balanced,  $\square_{ij}h = \square_{ij}h^{\text{ti}}$  and  $\mathbf{H}^{\text{mod}}$  is contained in  $\mathbf{H}_{(ij)}^{\square}$ . This and Corollary 2 imply the decomposition.  $\square$

The study of  $cl(\mathbf{H}^{\text{ent}})$  thus reduces to that of  $\mathbf{L}_{ij}$ . This cone is contained in  $\mathbf{H}_{(ij)}^{\square} \cap \mathbf{H}^{\text{ti}}$  which is known to be the conic hull of 11 linearly independent polymatroidal rank functions [36, Lemma 6.1]. The most notable one

$$(8) \quad \bar{r}_{ij}(K) = \begin{cases} 3, & K \in \{ik, jk, il, jl, kl\}, \\ \min\{4, 2|K|\}, & \text{otherwise} \end{cases}$$

is not almost entropic by Zhang-Yeung inequality [44]. The remaining ones are matroidal

$$(9) \quad r_1^{\emptyset}, \quad r_3^{\emptyset}, \quad r_1^i, r_1^j, \quad r_2^k, r_2^l, \quad r_1^{ik}, r_1^{jk}, r_1^{il}, r_1^{jl}$$

where the matroids are uniform up to loops. Recall that the subindex denotes the rank and the superindex the set of loops. By the proof of [36, Lemma 6.1], every  $g \in \mathbf{H}_{(ij)}^{\square} \cap \mathbf{H}^{\text{ti}}$  is the unique conic combination of the rank functions from (8) and (9),

$$(10) \quad \begin{aligned} g = & -(\square_{ij}g)\bar{r}_{ij} + (\Delta_{ij|\emptyset}g)r_1 + (\Delta_{kl|ij}g)r_3 \\ & + (\Delta_{kl|i}g)r_1^i + (\Delta_{kl|j}g)r_1^j + (\Delta_{ij|k}g)r_2^k + (\Delta_{ij|l}g)r_2^l \\ & + (\Delta_{jl|k}g)r_1^{ik} + (\Delta_{il|k}g)r_1^{jk} + (\Delta_{jk|l}g)r_1^{il} + (\Delta_{ik|l}g)r_1^{jl} \end{aligned}$$

identifying explicitly the coordinate functionals.

Since the matroids in (9) are linear and there exists an entropic point violating Ingleton inequality, the dimension of  $\mathbf{L}_{ij}$  is 11, the same as that of  $cl(\mathbf{H}^{\text{ent}})^{\text{ti}}$  or  $\mathbf{H}_{(ij)}^{\square} \cap \mathbf{H}^{\text{ti}}$ . Theorem 2 has the following consequence.

**Corollary 5.**  $ri(\mathbf{L}_{ij}) \subseteq \mathbf{H}^{\text{ent}}$ .

The remaining part of this section focuses on faces of  $\mathbf{L}_{ij}$ . The face  $\mathbf{F}_{ij}$  given by the equalities  $\Delta_{ij|\emptyset}g = 0$  and  $\Delta_{kl|ij}g = 0$  plays a special role later, in particular when optimizing Ingleton score.

Let  $A_{i,j}$  and  $B_{i,j,k}$  be the linear mappings defined on  $\mathbb{R}^{\mathcal{P}(N)}$  by

$$A_{i,j}g = g + (\Delta_{ij|\emptyset}g)(r_1^i - r_1) \quad \text{and} \quad B_{i,j,k}g = g + (\Delta_{kl|ij}g)(r_2^k - r_3).$$

**Lemma 7.** *The mappings  $A_{i,j}$  and  $B_{i,j,k}$  commute, leave invariant the hyperplanes given by  $\Delta_{ij|\emptyset}g = 0$  and  $\Delta_{kl|ij}g = 0$ , respectively,  $A_{i,j}$  maps into the first hyperplane,  $B_{i,j,k}$  into the second one, and*

$$\square_{ij}g = \square_{ij}(A_{i,j}g) = \square_{ij}(B_{i,j,k}g), \quad g \in \mathbb{R}^{\mathcal{P}(N)}.$$

A simple proof is omitted, for example

$$(11) \quad A_{i,j}B_{i,j,k}g = B_{i,j,k}A_{i,j}g = g + (\Delta_{ij|\emptyset}g)(r_1^i - r_1) + (\Delta_{kl|ij}g)(r_2^k - r_3).$$

Both  $A_{i,j}$  and  $B_{i,j,k}$  change at most two coordinates in (10).

**Theorem 4.**  $A_{i,j}B_{i,j,k}\mathbf{L}_{ij} = \mathbf{F}_{ij}$ .

*Proof.* The hyperplanes given by  $\Delta_{ij|\emptyset}g = 0$  and  $\Delta_{kl|ij}g = 0$  peel out two facets of  $\mathbf{H}_{(ij)}^{\square} \cap \mathbf{H}^{\text{ti}}$ , due to (10). By Lemma 7,  $A_{i,j}B_{i,j,k}$  maps  $\mathbf{H}_{(ij)}^{\square} \cap \mathbf{H}^{\text{ti}}$  onto the intersection of the two facets. Since  $\mathbf{L}_{ij}$  is equal to  $\mathbf{H}_{(ij)}^{\square} \cap \mathbf{H}^{\text{ti}} \cap cl(\mathbf{H}^{\text{ent}})$  it suffices to prove that both  $A_{i,j}$  and  $B_{i,j,k}$  map  $\mathbf{L}_{ij}$  into  $cl(\mathbf{H}^{\text{ent}})$ .

By the identity

$$\square_{ij} = \Delta_{ij|k} + \Delta_{ik|l} + \Delta_{kl|j} - \Delta_{ik|jl},$$

if  $f \in \mathbf{H}_{(ij)}^\square$ , then  $\Delta_{ik|j} f \geq \Delta_{ij|k} f$ , thus  $f(ij) - f(j) \geq f(ik) - f(k)$ . By symmetry,  $f(ij) - f(j) \geq f(il) - f(l)$ . In turn, Corollary 1 can be applied to  $t = h(ij) - h(j)$ , and provides  $h \in \mathbf{H}$  that coincides with  $f$  except at  $i$  where  $h(i) = f(ij) - f(j)$ . Similarly, the rank functions  $r_1^i$  and  $r_1$  differ only at  $i$  and  $r_1^i(i) - r_1(i) = -1$ . It follows that  $h = A_{i,j} f$ . If additionally  $f \in \mathcal{cl}(\mathbf{H}^{\text{ent}})$  then  $h$ , being the convolution of  $f$  with a modular polymatroid, is almost entropic. Therefore,  $f \in \mathbf{L}_{ij}$  implies  $A_{i,j} f \in \mathcal{cl}(\mathbf{H}^{\text{ent}})$ .

By the identity

$$\square_{ij} = \Delta_{ij|k} + \Delta_{ik|l} + \Delta_{kl|ij} - \Delta_{ik|jl},$$

if  $f \in \mathbf{H}_{(ij)}^\square$  then  $\Delta_{ik|jl} f \geq \Delta_{kl|ij} f$ . Additionally, if  $f$  is tight this inequality rewrites to  $f(ij) \geq f(jl)$ . By symmetry,  $f(ij) \geq f(il)$ . It follows that (3) is valid for  $t = f(N) - f(ij)$ . By Remark 1 and  $t \leq f(k)$ , Lemma 4 is applied with  $L = k$  and provides  $h = f_{k,t}^*$ . This rank function differs from  $f$  by  $t$  on the sets  $I \subseteq N$  with  $k \in \mathcal{cl}(I)$ . By the identity

$$\square_{ij} = \Delta_{ij|k} + \Delta_{ij|l} + \Delta_{kl|ij} - \Delta_{ij|kl},$$

$\Delta_{ij|kl} f \geq \Delta_{kl|ij} f$ . Since  $f$  is tight the inequality rewrites to  $f(ij) \geq f(kl)$ . By symmetry,  $f(ij)$  is maximal among all  $f(J)$  with  $|J| = 2$ . Therefore, if  $t > 0$  then  $k \in \mathcal{cl}(I)$  is equivalent to  $I \ni k$  or  $I = N \setminus k$ . These are exactly the cases when  $r_2^k$  and  $r_3$  differ, and  $r_2^k(I) - r_3(I) = -1$ . It follows from  $t = \Delta_{kl|ij} f$  that  $h = B_{ij,k} f$ . If, additionally,  $f \in \mathcal{cl}(\mathbf{H}^{\text{ent}})$  then  $h$  is almost entropic. Therefore,  $f \in \mathbf{L}_{ij}$  implies  $B_{ij,k} f \in \mathbf{L}_{ij}$ .  $\square$

*Remark 3.* Let  $\mathbf{E}_{ij}$  be the face of  $\mathbf{L}_{ij}$  given by the equalities

$$\Delta_{ij|k} g = 0, \Delta_{ij|l} g = 0, \Delta_{kl|i} g = 0, \Delta_{kl|j} g = 0 \text{ and } \Delta_{kl|ij} g = 0.$$

In [31, Example 2], four random variables are constructed such that their entropy function  $g$  satisfies the above five constraints,  $\square_{ij} g < 0$ , each of  $\Delta_{ij|\emptyset} g$ ,  $\Delta_{j|l} g$ ,  $\Delta_{il|k} g$ ,  $\Delta_{jk|l} g$  and  $\Delta_{ik|l} g$  is positive, and  $g$  is not tight. The lack of tightness makes  $g$  to be outside  $\mathbf{L}_{ij}$ . Nevertheless, Theorem 1 implies that  $g^{\text{ti}} \in \mathcal{cl}(\mathbf{H}^{\text{ent}})$  whence  $g^{\text{ti}}$  belongs to the face  $\mathbf{E}_{ij}$ , even more, it belongs to its relative interior. At the same time, [32, Theorem 4.1] implies that no point of  $\text{ri}(\mathbf{E}_{ij})$  is entropic. This phenomenon can be equivalently rephrased in terms of conditional information inequalities, studied recently in [23, 24, 25].

## 7. SYMMETRIZATION OF $\mathbf{F}_{ij}$

As before, the ground set  $N$  is assumed to have four elements  $i, j, k, l$ , which are always different. In the previous section the study of  $\mathcal{cl}(\mathbf{H}_N^{\text{ent}})$  was reduced to that of  $\mathbf{L}_{ij}$ , and a particular face  $\mathbf{F}_{ij}$  of the latter was identified. Here, a symmetrization of  $\mathbf{F}_{ij}$  is described and its cross-section visualized, owing to numerical computer experiments.

The expression  $\square_{ij}$  and the cones  $\mathbf{L}_{ij}$  and  $\mathbf{F}_{ij}$  enjoy natural symmetries. Namely, if a permutation  $\pi$  on  $N$  stabilizes the two-element set  $ij$  then  $\square_{ij} h = \square_{ij} h_\pi$ ,  $h \in \mathbf{H}$ . Hence  $\mathbf{L}_{ij}$  and  $\mathbf{F}_{ij}$  are closed to the action  $h \mapsto h_\pi$ .

Let  $C_{ij}$  be the linear mapping on  $\mathbb{R}^{\mathcal{P}(N)}$  given by

$$C_{ij}h \triangleq |G_{ij}|^{-1} \sum_{\pi \in G_{ij}} h_{\pi}$$

where  $G_{ij}$  denotes the stabilizer of  $ij$ , consisting of four permutations. By the decomposition (10), if  $h \in \mathbf{H}_{(ij)}^D \cap \mathbf{H}^{\text{ti}}$  then

$$\begin{aligned} C_{ij}h = & -(\square_{ij} h) \bar{r}_{ij} + (\Delta_{ij|\emptyset} h) r_1^{\emptyset} + (\Delta_{kl|ij} h) r_3^{\emptyset} \\ & + \frac{1}{2} [\Delta_{kl|i} h + \Delta_{kl|j} h] [r_1^j + r_1^i] \\ & + \frac{1}{2} [\Delta_{ij|k} h + \Delta_{ij|l} h] [r_2^l + r_2^k] \\ & + \frac{1}{4} [\Delta_{jl|k} h + \Delta_{il|k} h + \Delta_{jk|l} h + \Delta_{ik|l} h] [r_1^{ik} + r_1^{jk} + r_1^{il} + r_1^{jl}]. \end{aligned}$$

It follows that  $C_{ij}\mathbf{L}_{ij}$  has dimension 6 and  $C_{ij}\mathbf{F}_{ij}$  is a face of dimension 4. The cross-section

$$\mathbf{S}_{ij} \triangleq \{h \in C_{ij}\mathbf{F}_{ij} : h(N) = 1\}$$

is three-dimensional. By (10), for  $h \in \mathbf{S}_{ij}$

$$\begin{aligned} 1 = h(N) = & [-4\square_{ij} h] + [\Delta_{kl|i} h + \Delta_{kl|j} h] + [2\Delta_{ij|k} h + 2\Delta_{ij|l} h] \\ & + [\Delta_{jl|k} h + \Delta_{il|k} h + \Delta_{jk|l} h + \Delta_{ik|l} h]. \end{aligned}$$

Denoting by  $\bar{\alpha}_h, \bar{\beta}_h, \bar{\gamma}_h$  and  $\bar{\delta}_h$  the above brackets, respectively, any function  $h \in \mathbf{S}_{ij}$  takes the form

$$h = \bar{\alpha}_h \frac{1}{4} \bar{r}_{ij} + \bar{\beta}_h \frac{1}{2} [r_1^j + r_1^i] + \bar{\gamma}_h \frac{1}{4} [r_2^l + r_2^k] + \bar{\delta}_h \frac{1}{4} [r_1^{ik} + r_1^{jk} + r_1^{il} + r_1^{jl}].$$

Here,  $\bar{\alpha}_h, \bar{\beta}_h, \bar{\gamma}_h$  and  $\bar{\delta}_h$  are nonnegative and sum to one. Further,  $\alpha = \frac{1}{4} \bar{r}_{ij}$ ,  $\beta = \frac{1}{2} [r_1^j + r_1^i]$ ,  $\gamma = \frac{1}{4} [r_2^l + r_2^k]$  and  $\delta = \frac{1}{4} [r_1^{ik} + r_1^{jk} + r_1^{il} + r_1^{jl}]$  are linearly independent polymatroidal rank functions from  $\mathbf{S}_{ij}$ . It follows that  $\mathbf{S}_{ij}$  is a closed convex subset of the tetrahedron with the vertices  $\alpha, \beta, \gamma$  and  $\delta$ . Since the points  $h$  having  $\bar{\alpha}_h = 0$  are almost entropic and  $\bar{r}_{ij}$  is not,  $\mathbf{S}_{ij}$  contains the triangle  $\beta\gamma\delta$  but not the vertex  $\alpha$ .

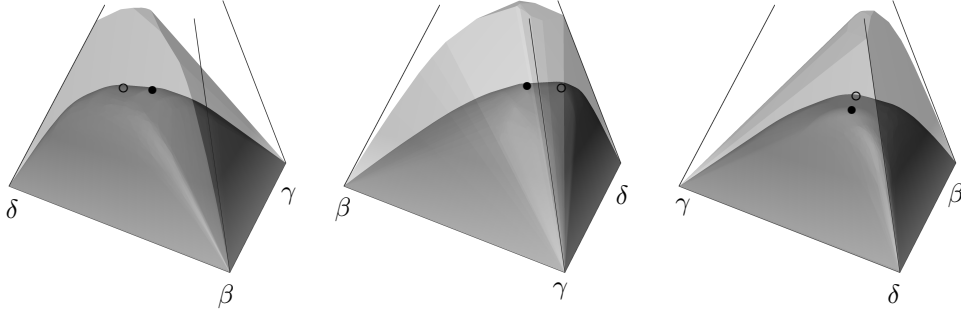


FIGURE 1. Inner and outer approximations of  $\mathbf{S}_{ij}$ .

Computer experiments were run to visualize  $\mathbf{S}_{ij}$ . Involved random variables were limited to take at most 11 values. Various maximization procedures were run numerically over the distributions of four tuples of random variables. The corresponding entropy functions  $f$  were transformed to  $g = C_{ij}A_{i,j}B_{ij,k}f^{\text{ti}}$  and then to  $h = g/g(N)$ , which is the convex combination

$$h = \bar{\alpha}_h \alpha + \bar{\beta}_h \beta + \bar{\gamma}_h \gamma + \bar{\delta}_h \delta \in \mathbf{S}_{ij}.$$

The procedures maximized  $\bar{\alpha}_h$  in various directions, over the distributions. Different methods and strategies were employed, including also randomized search. In this way, over 5 million points from  $\mathcal{S}_{ij}$  have been generated.

In Figure 1, the convex hull of these points is depicted as a dark gray region from three different perspectives. In the images, the vertex  $\alpha$  is missing and the straight lines are the incomplete edges of the tetrahedron incident to  $\alpha$ . The dark gray region is spanned by about 2200 extreme points. The projections of the extreme points from  $\alpha$  to  $\beta\gamma\delta$  do not exhaust the triangle uniformly, see Figure 2. This explains the lack of smoothness of the dark gray region. The two extreme points of the dark gray region depicted in Figure 1 are discussed in Section 8.

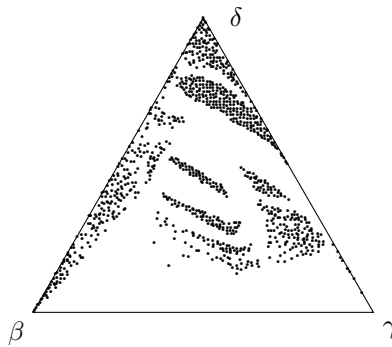


FIGURE 2. Extreme points of the dark gray region, projected to  $\beta\gamma\delta$ .

The light gray region in Figure 1 visualizes an outer approximation of  $\mathcal{S}_{ij}$  which was constructed from hundreds of known non-Shannon information inequalities, mostly from those of [14]. Details are omitted. The gap between the approximations is large.

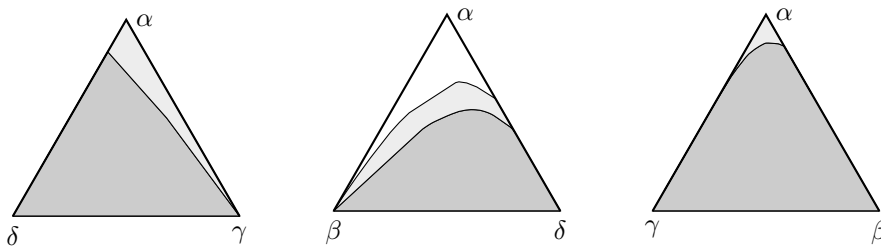


FIGURE 3. Projections of the approximations of  $\mathcal{S}_{ij}$  to triangles.

Figure 3 shows the dark and light gray regions when projected from the vertex  $\beta/\gamma/\delta$  to the opposite triangle of the tetrahedron. By the non-Shannon inequalities (13) discussed in the next section, the only almost entropic points on the edges  $\alpha\beta$  and  $\alpha\gamma$  are  $\beta$  and  $\gamma$ . The analogous statement for the edge  $\alpha\delta$  is open.

### 8. INGLETON SCORE

As before, the ground set  $N$  has four elements and  $ij$  is a two-element subset of  $N$ . The *Ingleton score* of a polymatroidal rank function  $h \neq 0$  is defined as

$\mathbb{I}_{ij}(h) \triangleq \square_{ij} h/h(N)$  [14, Definition 3]. The number

$$\mathbb{I}^* \triangleq \inf \{ \mathbb{I}_{ij}(h) : 0 \neq h \in \mathbf{H}^{\text{ent}} \}$$

is referred to as the *infimal* Ingleton score. This is likely the most interesting number related to the entropy region of four variables. By symmetry,  $\mathbb{I}^*$  does not depend on  $ij$ . This section presents an alternative way of minimization and a new upper bound on this number in Example 2.

First, the minimization is reduced to a three dimensional body.

**Theorem 5.**  $\mathbb{I}^* = \min_{\mathbf{S}_{ij}} \mathbb{I}_{ij}$ .

*Proof.* Since the score is constant along rays and  $\mathbb{I}^*$  is negative

$$\mathbb{I}^* = \min \{ \mathbb{I}_{ij}(h) : h(N) = 1, \square_{ij} h \leq 0 \text{ and } h \in \text{cl}(\mathbf{H}^{\text{ent}}) \},$$

minimizing over a compact. If  $h \in \mathbf{H}_{(ij)}^{\square}$  then  $\mathbb{I}_{ij}(h) \geq \mathbb{I}_{ij}(h^{\text{ti}})$  for  $h^{\text{ti}} \neq 0$ , and  $\mathbb{I}_{ij}(h) = 0$  for  $h^{\text{ti}} = 0 \neq h$ . Hence,

$$(12) \quad \mathbb{I}^* = \min \{ \mathbb{I}_{ij}(h) : h(N) = 1 \text{ and } h \in \mathbf{L}_{ij} \}.$$

Recall that  $\mathbf{L}_{ij} = \text{cl}(\mathbf{H}^{\text{ent}})^{\text{ti}} \cap \mathbf{H}_{(ij)}^{\square}$  is the cone of tight almost entropic rank functions  $h$  with  $\square_{ij} h \leq 0$ . (By Corollary 5, the above minimization can be expressed by special entropy functions.)

If  $g \in \mathbf{L}_{ij}$  then (11) and tightness of  $g$  provide

$$A_{i,j} B_{i,j,k} g(N) = g(N) - \Delta_{kl|ij} g = g(ij).$$

By Lemma 7,  $\mathbb{I}_{ij}(g) \geq \mathbb{I}_{ij}(A_{i,j} B_{i,j,k} g)$  when  $g(ij) > 0$ . If  $g(ij) = 0$  then  $\square_{ij} g = g(k) + g(l) - g(kl)$  which is possible only if  $\mathbb{I}_{ij}(g)$  vanishes. Hence, Theorem 4 implies that the minimization restricts to  $\mathbf{F}_{ij}$ . The assertion follows by symmetrization.  $\square$

The three dimensional body  $\mathbf{S}_{ij}$  is enclosed in the tetrahedron  $\alpha\beta\gamma\delta$  and  $-4\mathbb{I}_{ij}(h)$  is the weight  $\bar{\alpha}_h$  of  $h \in \mathbf{S}_{ij}$  at the vertex  $\alpha$  when  $h$  is written as the unique convex combination of the vertices. Thus, points of  $\mathbf{S}_{ij}$  with the heaviest weight at  $\alpha$  are the minimizers in Theorem 5. It should be also mentioned that is not clear which part of  $\mathbf{S}_{ij}$  is exhausted by the very entropic points.

Lower bounds on  $\mathbb{I}^*$  can be obtained by relaxing  $\mathbf{L}_{ij}$  in (12). The simplest relaxation is to  $\mathbf{H}^{\text{ti}} \cap \mathbf{H}_{(ij)}^{\square}$  because this cone has only one extreme ray allowing for negative scores, namely the one generated by  $\bar{r}_{ij}$ . Therefore, the infimal score  $\mathbb{I}^*$  is lower bounded by  $\mathbb{I}_{ij}(\bar{r}_{ij}) = -\frac{1}{4}$ . With little more work, the bound  $-\frac{1}{6}$  can be obtained by Zhang-Yeung inequality. Better lower bounds are reported in [14], based on further non-Shannon-type inequalities.

Upper bounds on the infimal Ingleton score arise from entropic polymatroids that violate the Ingleton inequality. There are many examples at disposal [40, 31, 32, 46, 19, 16, 18, 29, 3, 43]. The following one has attracted a special attention.

*Example 1.* Let  $\xi_i$  and  $\xi_j$  be exchangeable and 0-1 valued, and  $\xi_i = 1$  with the probability  $\frac{1}{2}$ . Let further  $\xi_k = \min\{\xi_i, \xi_j\}$  and  $\xi_l = \max\{\xi_i, \xi_j\}$ , see [31, Example 1] or [10]. If  $0 \leq p \leq \frac{1}{2}$  denotes the probability of  $\xi_i \xi_j = 00$  and  $h_p$  is the entropy function of  $\xi_i \xi_j \xi_k \xi_l$  then

$$\mathbb{I}_{ij}(h_p) = \frac{\Delta_{ij|\emptyset} h_p - \Delta_{kl|\emptyset} h_p}{h_p(N)}$$



using the identity  $\mathbb{D}_{ij} = \Delta_{kl|i} + \Delta_{kl|j} + \Delta_{ij|\emptyset} - \Delta_{kl|\emptyset}$ . Let  $\kappa(u) = -u \ln u$ ,  $u > 0$ , and  $\kappa(0) = 0$ . The numerator is

$$\begin{aligned} & 2 \ln 2 - 2\kappa(p) - 2\kappa\left(\frac{1}{2} - p\right) - \left[2\kappa(1-p) - \kappa(1-2p)\right] \\ & = (2p+1) \ln 2 - 2\kappa(p) - 2\kappa(1-p) \end{aligned}$$

and the denominator is  $2\kappa(p) + 2\kappa\left(\frac{1}{2} - p\right)$ . The function  $p \mapsto \mathbb{I}_{ij}(h_p)$  is strictly convex and has a unique global minimizer  $p^*$  in the interval  $[0, \frac{1}{2}]$ . Approximately,  $p^* \doteq 0.350457$  and  $\mathbb{I}_{ij}(h_{p^*}) \doteq -0.089373$ .

The guess that  $\mathbb{I}^*$  be equal to  $\mathbb{I}_{ij}(h_{p^*})$  goes back to [10] but the formulation [10, Conjecture 4.1] had a wrong numerical value. The same surmise appeared later in [14] as *Four-atom conjecture*, referring to the four possible values of  $\xi_i \xi_j \xi_k \xi_l$ . The minimization was considered also in [29, 3] that report no score below  $\mathbb{I}_{ij}(h_{p^*})$ . However, the computer experiments discussed in Section 7 found an entropic point that can be transformed to an almost entropic point witnessing failure of Four-atom conjecture.

*Example 2.* Let each of four variables in  $\xi_i \xi_j \xi_k \xi_l$  take values in  $\{0, 1, 2, 3\}$  and  $p, q, r, s, t$  be nonnegative such that  $p + q + r + s + t = \frac{1}{8}$ . The table below lists 40 different configurations of the random vector. Each configuration in any column is attained with the probability given by the label of that column. The remaining configurations have zero probabilities. The corresponding entropy function is denoted by  $f$ .

$p$	$q$	$r$	$s$	$t$
0000	0210	0011	0010	0001
0101	0321	0120	0121	0100
1010	1100	1002	1000	1012
1212	1332	1230	1232	1210
2121	2001	2103	2101	2123
2323	2233	2331	2333	2321
3232	3012	3213	3212	3233
3333	3123	3322	3323	3332

By inspection of the table, in each column any variable takes each value twice. Hence,  $f(i)$ ,  $f(j)$ ,  $f(k)$  and  $f(l)$  are equal to  $2 \ln 2$ . In each column,  $\xi_i \xi_l$  and  $\xi_j \xi_k$  are in the configurations 00, 33, 01, 10, 12, 21, 23, 32. Hence,  $f(il)$  and  $f(jk)$  are equal to  $3 \ln 2$ . In each column but the second/third one,  $\xi_i \xi_j$  and  $\xi_k \xi_l$  are in the configurations 00, 33, 01, 10, 12, 21, 23, 32, otherwise in 11, 22, 02, 20, 13, 31, 03, 30. Hence,

$$\begin{aligned} f(ij) &= 8\kappa(q) + 8\kappa(p+r+s+t) \\ f(kl) &= 8\kappa(r) + 8\kappa(p+q+s+t). \end{aligned}$$

In the first and fifth/forth column,  $\xi_i \xi_k$  and  $\xi_j \xi_l$  are in the configurations 00, 11, 22, 33, each one attained twice, otherwise in 01, 10, 02, 20, 13, 31, 23, 32. Hence,

$$\begin{aligned} f(ik) &= 4\kappa(2p+2t) + 8\kappa(q+r+s) \\ f(jl) &= 4\kappa(2p+2s) + 8\kappa(q+r+t). \end{aligned}$$

Analogous considerations provide

$$\begin{aligned} f(ikl) &= 8\kappa(p+t) + 8\kappa(q+s) + 8\kappa(r) \\ f(jkl) &= 8\kappa(p+s) + 8\kappa(q+t) + 8\kappa(r) \\ f(ijk) &= 8\kappa(p+t) + 8\kappa(r+s) + 8\kappa(q) \\ f(ijl) &= 8\kappa(p+s) + 8\kappa(r+t) + 8\kappa(q). \end{aligned}$$

Since the 40 configurations of the table are all different

$$f(ijkl) = 8\kappa(p) + 8\kappa(q) + 8\kappa(r) + 8\kappa(s) + 8\kappa(t).$$

The choice

$$p = 0.09524, q = 0.02494, r = 0.00160 \text{ and } s = t = 0.00161,$$

where  $r$  is close to  $s = t$ , gives  $\mathbb{I}_{ij}(f) \doteq -0.078277$ . This is yet bigger than the value  $-0.089373$  from Example 1. However,  $\mathbb{I}_{ij}(f^{\text{ti}}) \doteq -0.0912597$ , refuting Four-atom conjecture. Even better, if  $g$  denotes  $A_{i,j}B_{i,j,k}f^{\text{ti}}$  then  $\square_{ij}g = \square_{ij}f$  by Lemma 7, and

$$\begin{aligned} g(N) &= f^{\text{ti}}(N) - \Delta_{kl|ij} f^{\text{ti}} = 2f^{\text{ti}}(N) + f^{\text{ti}}(ij) - f^{\text{ti}}(ijk) - f^{\text{ti}}(ijl) \\ &= f(ij) + f(ikl) + f(jkl) - 2f(N) < f^{\text{ti}}(N) \end{aligned}$$

by (11). Hence, the score  $\mathbb{I}_{ij}(g)$  is approximately  $-0.09243$ , currently the best upper bound on the infimal Ingleton score.

Figure 1 features also two extreme points of the dark gray region. The circle depicts the point  $C_{ij}A_{i,j}h_{p^*}$  where  $h_{p^*}$  was described in Example 1. The bullet depicts  $C_{ij}A_{i,j}B_{i,j,k}f^{\text{ti}}$  where  $f$  is the entropic point from Example 2.

Figure 4 shows the intersections of the dark and light gray regions, approximating  $\mathbf{S}_{ij}$ , with the triangles  $\alpha\beta\gamma$ ,  $\alpha\gamma\delta$  and  $\alpha\delta\beta$ , two more exceptional points of  $\mathbf{S}_{ij}$ , and the role of Zhang-Yeung inequality.

The symmetrized Zhang-Yeung inequality

$$2\square_{ij}h + [\Delta_{ik|l}h + \Delta_{il|k}h + \Delta_{kl|i}h] + [\Delta_{jk|l}h + \Delta_{jl|k}h + \Delta_{kl|j}h] \geq 0,$$

valid for  $h \in \mathbf{H}^{\text{ent}}$ , rewrites to  $\bar{\beta}_h + \bar{\delta}_h \geq \frac{1}{2}\bar{\alpha}_h$ . The plane defined by the equality here is indicated in Figure 4 by the three dashed segments.

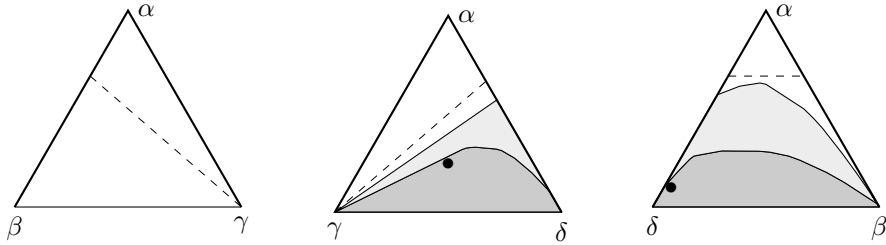


FIGURE 4. Intersections of the approximations of  $\mathbf{S}_{ij}$  with triangles.

By [14, Theorem 10], if  $s \geq 0$  is integer then

$$\begin{aligned} (2^s - 1)\square_{ij}h + \Delta_{kl|i}h + s2^{s-1}[\Delta_{ik|l}h + \Delta_{il|k}h] \\ + ((s-2)2^{s-1} + 1)[\Delta_{jk|l}h + \Delta_{jl|k}h] \geq 0, \quad h \in \mathbf{H}^{\text{ent}}. \end{aligned}$$

This inequality and its instance with  $i$  and  $j$  interchanged sum to

$$(13) \quad \bar{\beta}_h + [(s-1)2^s + 1]\bar{\delta}_h \geq \frac{1}{2}(2^s - 1)\bar{\alpha}_h, \quad h \in \mathbf{H}^{\text{ent}}.$$

Hence, the triangle  $\alpha\beta\gamma$  contains no almost entropic points except those on the edge  $\beta\gamma$ .

The bullet inside the triangle  $\alpha\gamma\delta$  depicts the entropy function  $f_{1/2}$  from Example 1, see also [31, Example 1]. The bullet inside the triangle  $\alpha\beta\delta$  shows the almost entropic point  $C_{ij}A_{i,j}g^{\text{ti}}$  where  $g$  is the entropy function discussed in Remark 3, see also [31, Example 2].

#### REFERENCES

- [1] R. Baber, D. Christofides, A.N. Dang, S. Riis, and E.R. Vaughan (2013) Multiple unicasts, graph guessing games, and non-Shannon inequalities. *NetCod 2013* 1–6.
- [2] R. Bassoli, H. Marques, J. Rodriguez, K.W. Shum and R. Tafazolli (2013) Network Coding Theory: A Survey. *IEEE Communications Surveys & Tutorials*. IEEE Vol. PP (99), 1–29.
- [3] N. Boston and T-T. Nan (2012) Large violations of the Ingleton inequality. *50th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, 1588–1593.
- [4] A. Beimel and I. Orlov (2009) Secret sharing and non-Shannon information inequalities. In: *TCC 2009 Lecture Notes in Computer Science* **5444** 539–557.
- [5] A. Beimel, N. Livne, and C. Padró (2008) Matroids can be far from ideal secret sharing. In: *TCC 2008 Lecture Notes in Computer Science* **4948** 194–212.
- [6] T.H. Chan and R.W. Yeung (2002) On a relation between information inequalities and group theory. *IEEE Trans. Information Theory* **48** 1992–1995.
- [7] T.H. Chan (2003) Balanced information inequalities. *IEEE Trans. Information Theory* **49** 3261–3267.
- [8] T.H. Chan, A. Grant and D. Pflüger (2011) Truncation technique for characterizing linear polymatroids. *IEEE Trans. Information Theory* **57** 6364–6378.
- [9] T.H.Chan (2011) Recent progresses in characterising information inequalities. *Entropy* **13**(2) 379–401.
- [10] L. Csirmaz (1996) The dealer’s random bits in perfect secret sharing schemes. *Studia Sci. Math. Hungar.* **32**, 429–437.
- [11] L. Csirmaz (2013) Book inequalities. (<http://eprints.renyi.hu/70>)
- [12] I. Csiszar and J. Körner (2011) *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge University Press, Cambridge.
- [13] R. Dougherty, C. Freiling, and K. Zeger (2006) Six New Non-Shannon Information Inequalities. *Proceedings IEEE ISIT 2006*, Seattle, Washington, 233–236.
- [14] R. Dougherty, C. Freiling, and K. Zeger (2011) Non-Shannon information inequalities in four random variables. (arXiv:1104.3602v1).
- [15] S. Fujishige (1978) Polymatroidal dependence structure of a set of random variables. *Information and Control* **39** 55–72.
- [16] D. Fong, S. Shadbakht, and B. Hassibi (2008) On the entropy region and the Ingleton inequality. *Mathematical Theory of Networks and Systems (MTNS)*.
- [17] M. Gadouleau and S. Riis (2011) Graph-theoretical constructions for graph entropy and network coding based communications. *IEEE Transactions on Information Theory* **57** 6703–6717.
- [18] B. Hassibi and S. Shadbakht (2007) On a construction of entropic vectors using lattice-generated distributions. *Proceedings IEEE ISIT 2007*, 501–505.
- [19] D. Hammer, A. Romashchenko, A. Shen, and N. Vereshchagin (2000) Inequalities for Shannon entropy and Kolmogorov complexity. *J. of Computer and Systems Sciences* **60** 442–464.
- [20] A.W. Ingleton (1971) Conditions for representability and transversality of matroids. *Proc. Fr. Br. Conf. 1970*, Springer Lecture Notes **211**, Springer-Verlag, Berlin, 62–67.
- [21] A.W. Ingleton (1971) Representation of matroids. *Combinatorial Mathematics and its Applications* (D.J.A. Welsh, ed.), Academic Press, London, 149–167.
- [22] T. Kaced (2013) Equivalence of two proof techniques for non-Shannon-type inequalities. (arXiv:1302.2994).

- [23] T. Kaced and A. Romashchenko (2011) On essentially conditional information inequalities. *Proceedings IEEE ISIT 2011*, 1935–1939.
- [24] T. Kaced and A. Romashchenko (2012) On the non-robustness of essentially conditional information inequalities. *Proceedings IEEE ITW 2012*, 262–266.
- [25] T. Kaced and A. Romashchenko (2013) Conditional information inequalities for entropic and almost entropic points (submitted to IEEE TIT, arXiv:1207.5742v2).
- [26] L. Lovász (1982) Submodular functions and convexity. *Mathematical Programming – The State of the Art* (A. Bachem, M. Grötschel and B. Korte, eds.), Springer-Verlag, Berlin, 234–257.
- [27] M. Madiman, A.W. Marcus and P. Tetali (2010) Information-theoretic inequalities in additive combinatorics. *IEEE ITW 2010* 1–4.
- [28] K. Makarychev, Yu. Makarychev, A. Romashchenko and N. Vereshchagin (2002) A new class of non-Shannon-type inequalities for entropies. *Communications in Information and Systems* **2** 147–166.
- [29] W. Mao, M. Thill and B. Hassibi (2012) On the Ingleton-violating finite groups and group network codes. (arXiv:1202.5599)
- [30] F. Matúš (1994) Probabilistic conditional independence structures and matroid theory: background. *Int. Journal of General Systems* **22** 185–196.
- [31] F. Matúš (1995) Conditional independences among four random variables II. *Combinatorics, Probab. and Comput.* **4** 407–417.
- [32] F. Matúš (1999) Conditional independences among four random variables III: final conclusion. *Combinatorics, Probab. and Comput.* **8** 269–276.
- [33] F. Matúš (2006) Adhesivity of polymatroids. *Discrete Mathematics* **307** 2464–2477.
- [34] F. Matúš (2007) Two constructions on limits of entropy functions. *IEEE Trans. Inform. Theory* **53** 320–330.
- [35] F. Matúš (2007) Infinitely many information inequalities. *Proceedings IEEE ISIT 2007*, Nice, France, 41–44.
- [36] F. Matúš and M. Studený (1995) Conditional independences among four random variables I. *Combinatorics, Probab. and Comput.* **4** 269–278.
- [37] J.G. Oxley (1992) *Matroid Theory*. Oxford University Press, Oxford.
- [38] S. Riis (2007) Information flows, graphs and their guessing numbers. *The Electronic Journal of Combinatorics* **14** 1–17.
- [39] R.T. Rockafellar (1970) *Convex Analysis*. Princeton University Press, Princeton.
- [40] M. Studený (1990) personal communication.
- [41] M. Studený (2005) *Probabilistic Conditional Independence Structures*. Springer, New York.
- [42] R.W. Yeung (2002) *A First Course in Information Theory*. Kluwer Academic/Plenum Publishers, New York.
- [43] J. MacLaren Walsh and S. Weber (2010) Relationships among bounds for the region of entropic vectors in four variables. *2010 Allerton Conference on Communication, Control, and Computing*, 1319–1326.
- [44] Z. Zhang and R.W. Yeung (1997) A non-Shannon-type conditional inequality of information quantities. *IEEE Trans. Information Theory* **43** 1982–1986.
- [45] Z. Zhang (2003) On a new non-Shannon-type information inequality. *Communications in Information and Systems* **3** 47–60.
- [46] Z. Zhang and R.W. Yeung (1998) On characterization of entropy function via information inequalities. *IEEE Trans. Information Theory* **44** 1440–1452.

FRANTIŠEK MATUŠ IS WITH INSTITUTE OF INFORMATION THEORY AND AUTOMATION, ACADEMY OF SCIENCES OF THE CZECH REPUBLIC, POD VODÁRENSKOU VĚŽÍ 4, 182 08 PRAGUE, CZECH REPUBLIC (E-MAIL: MATUS@UTIA.CAS.CZ).

LÁSZLO CSIRMAZ IS WITH CENTRAL EUROPEAN UNIVERSITY, NÁDOR UTCA 9, H-1051, BUDAPEST, HUNGARY, UNIVERSITY OF DEBRECEN, AND RENYI INSTITUTE OF MATHEMATICS (E-MAIL: CSIRMAZ@RENYI.HU).