Entropy Region and Convolution

František Matúš 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 self-adhesivity is clarified. Results are specialized and extended to the region constructed from four tuples of random variables. This and computer experiments help to visualize approximations of a symmetrized part of the entropy region. The four-atom conjecture on the minimal Ingleton score is refuted.**

***Index Terms*— Entropy region, entropy function, information-theoretic inequality, non-Shannon inequality, polymatroid, matroid, convolution, selfadhesivity, Ingleton inequality, Zhang-Yeung inequality, four-atom conjecture, Ingleton score.**

I. Introduction

The entropy function of a vector $(\xi_i)_{i\in N}$ of discrete random variables, indexed by a finite set N, maps each subset I of N to the Shannon entropy of the subvector $(\xi_i)_{i\in I}$. This 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, provided the entropies are finite. Instead it is assumed throughout that the vector takes finite number of values. The collection of such points, over all vectors of discrete random variables, defines the *entropic region* $\boldsymbol{H}_N^{\mathsf{ent}}$. The closure $cl(\boldsymbol{H}_N^{\mathsf{ent}})$ of the region is a convex cone [48, Th. 1] whose relative interior is contained in $\boldsymbol{H}_N^{\mathsf{ent}}$ [37, Th. 1]. This work studies mostly the shape of $cl(\boldsymbol{H}_N^{\mathsf{ent}})$.

Basic properties of the Shannon entropy imply that any entropy function h from $\boldsymbol{H}_N^{\mathsf{ent}}$ is non-decreasing and submodular, and thus the pair (N,h) is a polymatroid with the ground set N and rank function h [17]. The polymatroidal rank functions on N form the polyhedral cone $\boldsymbol{H}_N$ which, consequently, contains $cl(\boldsymbol{H}_N^{\mathsf{ent}})$. A polymatroid, or its rank function, is called *entropic (almost entropic)* if the rank function belongs to $\boldsymbol{H}_N^{\mathsf{ent}}$ ($cl(\boldsymbol{H}_N^{\mathsf{ent}})$).

In this work, the entropy region and its closure are studied by means of standard constructions on polymatroids, recalled in Section II. The central working tool is the *convolution* of two polymatroidal rank functions. The crucial property is that $cl(\boldsymbol{H}_N^{\mathsf{ent}})$ is closed under convolution with modular polymatroids [37, Th. 2]. This has consequences on principal extensions and their contractions.

In Section III, the cone $cl(\boldsymbol{H}_N^{\mathsf{ent}})$ is decomposed into the direct sum of two cones, see Corollary 3. 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. The decomposition reduces the study of $cl(\boldsymbol{H}_N^{\mathsf{ent}})$ to a cone of lesser dimension. It is also closely related to balanced information-theoretic inequalities [8]. The relative interior of the first cone is exhausted by entropic points, see Theorem 2 in Section IV.

Section V 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-theoretic 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 VI the set N is assumed to have four elements. The role of Ingleton inequality in the structure of $\boldsymbol{H}_N$ is recalled. To describe the cone $cl(\boldsymbol{H}_N^{\mathsf{ent}})$ it suffices to pass to its subcone $\boldsymbol{L}_{ij}$, cut off by a reversed Ingleton inequality and tightness. Applying polymatroidal constructions, it is shown that $\boldsymbol{L}_{ij}$ is mapped by two linear maps to its face $\boldsymbol{F}_{ij}$ of dimension 9, see Theorem 4.

Section VII investigates a symmetrization of $\boldsymbol{F}_{ij}$. Its cross-section $\boldsymbol{S}_{ij}$ has dimension three. Various numerical optimization techniques were employed to find an inner approximation of $\boldsymbol{S}_{ij}$. An outer approximation is compiled from available non-Shannon type information inequalities. The two approximations are visualized, and it turns out that they are yet far from each other. In Section VIII, the range of Ingleton score is studied and related to the cross-section $\boldsymbol{S}_{ij}$, see Theorem 5. In Example 2, a score is presented that refutes the four-atom conjecture [11], [16].

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 on a fixed ground set are frequently called non-Shannon type information inequalities. Main breakthroughs include finding

Manuscript received July 21, 2015; accepted August 5, 2016. Date of publication August 19, 2016; date of current version October 18, 2016. F. Matúš was supported by the Grant Agency of the Czech Republic under Grant 13-20012S. L. Csirmaz was supported in part by the Research Program under Grant TAMOP-4.2.2.C-11/1/KONV-2012-0001 and in part by the Lendulet Program of the Hungarian Academy of Sciences.

F. Matúš is with the Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, 182 08 Prague, Czech Republic (e-mail: matus@utia.cas.cz).

L. Csirmaz is with Central European University, 1051 Budapest, Hungary, also with the University of Debrecen, 4032 Debrecen, Hungary, and also with the Renyi Institute of Mathematics, 1053 Budapest, Hungary (e-mail: csirmaz@renyi.hu).

Communicated by C. Nair, Associate Editor for Shannon Theory.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2016.2601598

of the first non-Shannon linear inequality [50] and the relation to group theory [7]. The cone $cl(\boldsymbol{H}_N^{\mathsf{ent}})$ is not polyhedral [38] and the structure of non-Shannon inequalities seems to be complex [12], [15], [16], [30], [47], [49]. Reviews are in [10] and [37] 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 [46]. In cryptography, the inequalities can improve bounds on the information ratios in secret sharing schemes [5], [6], [11].

In probability theory, the implication problem of conditional independence among subvectors of a random vector can be rephrased via the entropy region [45]. The guessing numbers of games on directed graphs and entropies of the graphs can be related to the network coding [19], [42], where non-Shannon inequalities provide sharper bounds [1]. Information-theoretic inequalities are under investigation in additive combinatorics [29] and in connection to entropy power inequalities [31]. They started to matter in matroid theory [39]. Last but not least, information-theoretic inequalities are known to be related to Kolmogorov complexity [30], determinantal inequalities and group-theoretic inequalities [10].

The above overview witnesses a broad scope of potential applications of results on the entropy region and its closure.

II. 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 monographs [14], [46]; further material on polymatroids is in [28].

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 δ_I denote the point of $\mathbb{R}^{\mathcal{P}(N)}$ having all the 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 $i, j \in N \setminus L$.

A. Polymatroids

The pair (N, f) is a *polymatroid* when $f(\emptyset) = 0$, f is non-decreasing, thus $f(I) \leqslant f(J)$ for $I\subseteq J\subseteq N$, and submodular, thus $\Delta_{I,J}\, f \geqslant 0$ 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 $\boldsymbol{H}_N$ 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 [28].

The polymatroid is a *matroid* [41] if f takes integer values and $f(I) \leqslant |I|$, $I \subseteq N$. For $J \subseteq N$ and $0 \leqslant m \leqslant |N \setminus J|$ 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. When $J = \emptyset$ the superindex is sometimes omitted.

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 $\boldsymbol{H}_N^{\mathsf{mod}}$ 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, thus is in $\boldsymbol{H}_N^{\mathsf{ent}}$.

The *contraction* of a polymatroid (N, f) along $I \subseteq N$ sits on $N \setminus I$ and has the rank function $J \mapsto f(J \cup I) - f(I)$, $J \subseteq N \setminus I$. (Poly)matroids are closed under contractions. The following lemma is known, e.g. implicit in the proof of [33, Lemma 2], but no reference to a proof seems to be available.

Lemma 1: 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 ξ_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 η^{x_I} be the random vector $\xi_{N\setminus I}$ conditioned on the event that $\xi_I = x_I$. The entropy function of η^{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 $cl(\boldsymbol{H}_N^{\mathsf{ent}})$ is convex [48, Th. 1], h is almost entropic. ∎

B. Convolution

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

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

see [28]. 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 [28, Th. 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 $f(j) \leqslant g(j)$ for all $j \in N \setminus i$ then

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

If, additionally, $f(i) \leqslant g(i)$ then $f{*}g = f$.

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

$$\begin{aligned} f(I) + g(\emptyset) &\leqslant f(J) + f(I \setminus J) \\ &\leqslant f(J) + \textstyle\sum_{j\in I\setminus J} f(j) \leqslant f(J) + g(I \setminus J) \end{aligned}$$

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

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

and

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

Hence, $f{*}g\,(iI)$ is equal to the smaller of the numbers $f(iI)$ and $f(I)+g(i)$. ■

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

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

Let (N,g) be a modular polymatroid such that $g(i)=t$ and $f(j)\leqslant g(j)$, $j\in N\setminus i$. Then $f{*}g$ takes the same values as f with the exception $f{*}g(i)=t$.

Proof: The assumption $t\leqslant f(i)$ 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,

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

It follows by Lemma 2 that $f{*}g(iI)=f(iI)$. ■

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.

C. Principal Extension

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 in these steps is employed later in Sections V and VI.

Two points $i,j\in N$ of a polymatroid (N,f) are *parallel* if $f(iJ)=f(jJ)$ for $J\subseteq N$. This happens if and only if $f(i)=f(ij)=f(j)$. 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 $g(0)=t\leqslant f(L)$ and $g(i)\geqslant f(i)$, $i\in N$, to arrive at the *principal extension $f_{L,t}$ of f on the subset L with the value t* [28]. By Lemma 2,

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

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

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

Lemma 3: If a polymatroid (N,f) is almost entropic, $L\subseteq N$ and $0\leqslant t\leqslant f(L)$ then so is $(N,f^*_{L,t})$.

Proof: If f is entropic then $f_{L,t}$ is almost entropic since $\mathit{cl}(\boldsymbol{H}_N^{\mathsf{ent}})$ is closed under convolutions [37, Th. 2]. This combined with Lemma 1 implies the assertion. ■

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

Lemma 4: If (N,f) is a polymatroid, $L\subseteq N$ and $t\leqslant f(L)$ such that

$$0\leqslant t\leqslant \min_{I\subseteq N,\ L\not\subseteq \mathit{cl}(I)}\ \max_{\ell\in L\setminus \mathit{cl}(I)} [f(\ell\cup I)-f(I)] \tag{2}$$

then

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

Proof: The inequalities $0\leqslant t\leqslant f(L)$ are needed to derive (1). If $L\subseteq \mathit{cl}(I)$ then

$$f(L\cup I)\leqslant f(L\cup \mathit{cl}(I))=f(\mathit{cl}(I))=f(I).$$

Hence, the inequality is tight and the minimum in (1) equals $f(I)-t$. Otherwise, by the assumption (2), $t\leqslant f(\ell\cup I)-f(I)$ for some $\ell\in L\setminus \mathit{cl}(I)$. Since $f(L\cup I)-t\geqslant f(\ell\cup I)-t\geqslant f(I)$ the minimum in (1) equals $f(I)$. ■

Remark 1: A special instance of Lemma 4 is used in the proof of Theorem 4 to shift almost entropic points inside $\mathit{cl}(\boldsymbol{H}_N^{\mathsf{ent}})$. There, L equals a singleton k contained in $\mathit{cl}(N\setminus k)$. In such a case, (2) is a consequence of

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

because each maximal $f(\ell\cup I)-f(I)$ 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 [9] to investigate linear polymatroids.

III. Tight and Modular Polymatroids

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

$$h^{\mathsf{ti}}(I)=h(I)-\textstyle\sum_{i\in I}[h(N)-h(N\setminus i)],$$
$$h^{\mathsf{m}}(I)=\textstyle\sum_{i\in I}[h(N)-h(N\setminus i)],\quad I\subseteq N.$$

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

Theorem 1: If (N,h) is a polymatroid then the tight component h^{ti} can be constructed from h by parallel extensions, convolution and contraction.

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 $I \subseteq N$

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

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

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

If $g(i) + g(i') = h(i)$ for $i \in N$ then $f{*}g(I \cup N')$ is equal to

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

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

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

In the case when

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

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

Corollary 2: If $h \in cl(\boldsymbol{H}_N^{\mathsf{ent}})$ then h^{ti} is almost entropic.

Proof: Keeping the notation of the previous proof, the assumption implies that f is almost entropic. The convolution theorem [37, Th. 2] implies that $f{*}g \in cl(\boldsymbol{H}_N^{\mathsf{ent}})$. By Lemma 1, $h^{\mathsf{ti}} \in cl(\boldsymbol{H}_N^{\mathsf{ent}})$. ∎

The closure of the entropic region decomposes analogously to $\boldsymbol{H}_N$.

Corollary 3: $cl(\boldsymbol{H}_N^{\mathsf{ent}}) = [cl(\boldsymbol{H}_N^{\mathsf{ent}}) \cap \boldsymbol{H}_N^{\mathsf{ti}}] \oplus \boldsymbol{H}_N^{\mathsf{mod}}$.

Proof: Corollary 2 and $\boldsymbol{H}_N = \boldsymbol{H}_N^{\mathsf{ti}} \oplus \boldsymbol{H}_N^{\mathsf{mod}}$ imply the inclusion $\subseteq$. The reverse one holds since $cl(\boldsymbol{H}_N^{\mathsf{ent}})$ is a convex cone and $\boldsymbol{H}_N^{\mathsf{mod}} \subseteq \boldsymbol{H}_N^{\mathsf{ent}}$ [37, Lemma 2]. ∎

As a consequence, $cl(\boldsymbol{H}_N^{\mathsf{ent}}) \cap \boldsymbol{H}_N^{\mathsf{ti}}$ equals

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

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

In the remaining part of the section it is shown that Corollary 3 is equivalent to [8, Th. 1] on balanced inequalities.

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

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

see [43, Sec. 14]. For example, the polar of $\boldsymbol{H}_N^{\mathsf{mod}}$ can be defined by the inequalities $\sum_{I\ni i} \vartheta_I \leqslant 0$, $i \in N$; substituting the matroids $r_1^{N\setminus i}$, $i \in N$, for h. The polars of $\boldsymbol{H}_N^{\mathsf{ent}}$ and $cl(\boldsymbol{H}_N^{\mathsf{ent}})$ coincide and are defined by the very linear information-theoretic inequalities.

By [43, Corollary 16.4.2], Corollary 3 is equivalent to

$$(\boldsymbol{H}_N^{\mathsf{ent}})^\circ = (cl(\boldsymbol{H}_N^{\mathsf{ent}})^{\mathsf{ti}})^\circ \cap (\boldsymbol{H}_N^{\mathsf{mod}})^\circ. \tag{5}$$

It was used tacitly also that $(\boldsymbol{H}_N^{\mathsf{ent}})^{\circ\circ}$ coincides with $cl(\boldsymbol{H}_N^{\mathsf{ent}})$, and that $cl(\boldsymbol{H}_N^{\mathsf{ent}})^{\mathsf{ti}}$, $\boldsymbol{H}_N^{\mathsf{mod}}$ and their sum are closed. The polar of $cl(\boldsymbol{H}_N^{\mathsf{ent}})^{\mathsf{ti}}$ consists of the vectors $(\vartheta_I)_{I\subseteq N}$ satisfying $\sum_{I\subseteq N} \vartheta_I h^{\mathsf{ti}}(I) \leqslant 0$ for $h \in cl(\boldsymbol{H}_N^{\mathsf{ent}})$. This rewrites to

$$\textstyle\sum_{I\subseteq N} \vartheta_I h(I) - \sum_{i\in N}[h(N) - h(N \setminus i)]\sum_{I\ni i}\vartheta_I \leqslant 0. \tag{6}$$

In turn, eq. (5) can be rephrased as follows. *Given* $(\vartheta_I)_{I\subseteq N}$, *the inequality* $\sum_{I\subseteq N} \vartheta_I h(I) \leqslant 0$ *holds for all* $h \in \boldsymbol{H}_N^{\mathsf{ent}}$ *if and only if* (6) *is valid and* $\sum_{I\ni i} \vartheta_I \leqslant 0$, $i \in N$. This was formulated earlier in [8, Th. 1].

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

IV. Entropy Region: Regular Faces of $cl(\boldsymbol{H}_N^{\mathsf{ent}})$

As mentioned earlier, the relative interior of $cl(\boldsymbol{H}_N^{\mathsf{ent}})$ belongs to the entropy region $\boldsymbol{H}_N^{\mathsf{ent}}$. Thus, $\boldsymbol{H}_N^{\mathsf{ent}}$ and $cl(\boldsymbol{H}_N^{\mathsf{ent}})$ 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: $ri(cl(\boldsymbol{H}_N^{\mathsf{ent}})^{\mathsf{ti}}) \oplus \boldsymbol{H}_N^{\mathsf{mod}} \subseteq \boldsymbol{H}_N^{\mathsf{ent}}$.

The proof presented below is based on an auxiliary lemma.

Lemma 5: The cone $cl(\boldsymbol{H}_N^{\mathsf{ent}})^{\mathsf{ti}}$ 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 ξ_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 $cl(\boldsymbol{H}_N^{\mathsf{ent}})$ [7, Th. 4.1].

Proof of Lemma 5: Let $\|h\|_\infty \triangleq \max_{I\subseteq N} |h(I)|$. Given ε positive and $g \in cl(\boldsymbol{H}_N^{\mathsf{ent}})^{\mathsf{ti}}$ there exists a random vector whose entropy function h satisfies $\|h^{\mathsf{ti}} - g\|_\infty < \varepsilon$. By Corollary 2, h^{ti} is almost entropic whence $\|h - g\|_\infty \leqslant \varepsilon$ for some h entropic. Since g is tight,

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

It can be assumed that the random vector sits on a finite set endowed with the uniform probability measure. By [37, Remark 11], there exists a group G, a group-generated polymatroid f and an integer $\ell \geqslant 1$ such that $\|\frac{1}{\ell} f - h\|_\infty \leqslant \varepsilon$. Therefore,

$$\tfrac{1}{\ell} f^{\mathsf{m}}(N) \leqslant |\tfrac{1}{\ell} f^{\mathsf{m}}(N) - h^{\mathsf{m}}(N)| + h^{\mathsf{m}}(N) \leqslant 4\varepsilon|N|.$$

Let $(\xi_i)_{i\in N}$ be the corresponding random vector of factor-mappings 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 η_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, η_j is a function of ξ_N, its entropy is $\ln|G_{N\setminus j}/G_N|$, η_j is stochastically independent of $\xi_{N\setminus j}$, and they together determine ξ_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 (ξ_N, η_N). Hence, h' is a tight entropy function. For $I \subseteq N$

$$\begin{aligned} f(I) \leqslant h'(I) &\leqslant f(I) + \textstyle\sum_{j\in N} \ln|G_{N\setminus j}/G_N| \\ &= f(I) + f^{\mathsf{m}}(N). \end{aligned}$$

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

$$\begin{aligned} \|\tfrac{1}{\ell}h' - g\|_\infty &\leqslant \|\tfrac{1}{\ell}h' - \tfrac{1}{\ell}f\|_\infty + \|\tfrac{1}{\ell}f - h\|_\infty + \|h - g\|_\infty \\ &\leqslant 4\varepsilon|N| + 2\varepsilon. \end{aligned}$$

By [37, Lemma 4], the tight polymatroid $\frac{1}{\ell}h' + \delta r_1$ is entropic for any $\delta > 0$. Thus, $\|(\frac{1}{\ell}h' + \varepsilon r_1) - g\|_\infty \leqslant 4\varepsilon|N| + 3\varepsilon$ where ε can be arbitrarily small. ∎

Remark 2: It is of some interest that $cl(\boldsymbol{H}_N^{\mathsf{ent}})^{\mathsf{ti}}$ contains a dense set of points in the form $\frac{1}{m}h''$ where m is integer and h'' is group-generated. 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 [37, 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 $g \in cl(\boldsymbol{H}_N^{\mathsf{ent}})^{\mathsf{ti}}$ it is possible to construct $\frac{1}{\ell m}h''$ arbitrarily close, as in the above proof.

Proof of Theorem 2: Since, the cone $\boldsymbol{H}_N^{\mathsf{mod}}$ is contained in $\boldsymbol{H}_N^{\mathsf{ent}}$ [37, Lemma 2] and $\boldsymbol{H}_N^{\mathsf{ent}}$ is closed to sums, it suffices to prove that $ri(cl(\boldsymbol{H}_N^{\mathsf{ent}})^{\mathsf{ti}}) \subseteq \boldsymbol{H}_N^{\mathsf{ent}}$. The argumentation is analogous to that in the proof of [37, Th. 1]. By [37, 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 in $\boldsymbol{H}_N^{\mathsf{ti}} \cap \boldsymbol{H}_N^{\mathsf{ent}}$ whose dimension $2^{|N|} - |N| - 1$ is the same as that of $cl(\boldsymbol{H}_N^{\mathsf{ent}})^{\mathsf{ti}}$. Therefore, if $\epsilon > 0$ then the set B_ϵ of polymatroids $\sum_{J:\,|J|<|N|-1} \alpha_J r_1^J$, where $0 < \alpha_J < \epsilon$, is open in the linear space $\boldsymbol{H}_N^{\mathsf{ti}} - \boldsymbol{H}_N^{\mathsf{ti}}$ and the shifts of these sets provide a base for the relative topology.

Hence, if g belongs to the relative interior of $cl(\boldsymbol{H}_N^{\mathsf{ent}})^{\mathsf{ti}}$ then it belongs to such a shift contained in the relative interior. It follows that $g - B_\epsilon$ is a subset of the relative interior for $\epsilon > 0$ sufficiently small. Since $g - B_\epsilon$ is a relatively open subset of $cl(\boldsymbol{H}_N^{\mathsf{ent}})^{\mathsf{ti}}$ 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 α_J are nonnegative, and thus is entropic. ∎

A convex subset F of a convex set K is a *face* if every line segment in K with an interior point in F belongs to F. A face of a convex cone is a convex cone.

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

V. Selfadhesivity and Tightness

This section recalls the notion of selfadhesivity and explores its relation to the decomposition $h = h^{\mathsf{ti}} + h^{\mathsf{m}}$ of polymatroids. Then the role of selfadhesivity in proving information-theoretic inequalities is briefly discussed and compared to an alternative technique from [30].

Two polymatroids (N, h) and (M, g) are *adhesive* [36], or *adhere*, if there exists a polymatroid $(N \cup M, f)$ 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 a π-copy $(\pi(N), h_\pi)$ of itself. This is defined by a bijection $\pi : N \to \pi(N)$ such that such that $O = N \cap \pi(N)$, $\pi(i) = i$ for $i \in O$, and $h_\pi(\pi(I)) = h(I)$ for $I \subseteq N$. 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 $\boldsymbol{H}_N^{\mathsf{sa}}$ [36]. This cone decomposes similarly to $\boldsymbol{H}_N = \boldsymbol{H}_N^{\mathsf{ti}} \oplus \boldsymbol{H}_N^{\mathsf{mod}}$.

Theorem 3: If $h \in \boldsymbol{H}_N^{\mathsf{sa}}$ then h^{ti} is selfadhesive.

Proof: Let a polymatroid (N, h) adhere with a π-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^{ti}) is selfadhesive at O.

The modular rank function g is defined by

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

Since

$$g(i) + g(i') = \hat{h}(i) \text{ and } g(\pi(i)) + g(\pi(i)') = \hat{h}(\pi(i)),$$

an analogue of (4) takes the form

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

arguing as in the proof of Theorem 1.

If $i \in N \setminus O$ then

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

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

$$\hat{h}^{\mathsf{ti}}(K) + \hat{h}^{\mathsf{m}}(K) + \hat{h}^{\mathsf{m}}(\hat{N} \setminus (O \cup K)) + g(O' \setminus K') \\ = \hat{h}^{\mathsf{ti}}(K) + \hat{h}^{\mathsf{m}}(\hat{N} \setminus (O \setminus K)) + h^{\mathsf{m}}(O \setminus K).$$

Hence, the minimization in (7) can be further restricted to $K \subseteq I \cup O$, and if $I \subseteq \hat{N}$ then $f{*}g(I \cup \hat{N}')$ is equal to

$$\hat{h}^{\mathsf{m}}(\hat{N} \setminus (I \cup O)) + \min_{I \subseteq K \subseteq I \cup O} \big[\hat{h}(K) + h^{\mathsf{m}}(O \setminus K)\big].$$

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 π-copy. Second,

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

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

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

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

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

whence the extension is selfadhesive. ■

Corollary 4: $\boldsymbol{H}_N^{\mathsf{sa}} = [\boldsymbol{H}_N^{\mathsf{sa}} \cap \boldsymbol{H}_N^{\mathsf{ti}}] \oplus \boldsymbol{H}_N^{\mathsf{mod}}$.

Proof: The inclusion $\subseteq$ follows from Theorem 3. Since the modular polymatroids have selfadhesive modular extensions and $\boldsymbol{H}_N^{\mathsf{sa}}$ is a convex cone, the opposite inclusion holds as well. ■

The convex cone $cl(\boldsymbol{H}_N^{\mathsf{ent}})$ is not polyhedral [38], thus its polar cone is not finitely generated. There are infinite sets of linear information- theoretic inequalities [12], [15], [16], [30], [49]. Hundreds of them have been generated in computer experiments based on the fact that the entropic polymatroids are selfadhesive, $\boldsymbol{H}_N^{\mathsf{ent}} \subseteq \boldsymbol{H}_N^{\mathsf{sa}}$, and iterations of the idea. None of the experiments seems to have taken into account the possible reduction by imposing the tightness, cf. Corollary 4.

A linear information-theoretic inequality

$$\textstyle\sum_{I \subseteq N} \vartheta_I h(I) \leqslant 0 \text{ for all } h \in \boldsymbol{H}_N^{\mathsf{ent}}$$

is of non-Shannon type if $(\vartheta_I)_{I \subseteq N} \in (\boldsymbol{H}_N^{\mathsf{ent}})^\circ$ is not in $\boldsymbol{H}_N^\circ$. There are two techniques for proving non-Shannon-type inequalities: either by selfadhesivity, as implicit in the original proof of Zhang-Yeung inequality [48], or alternatively by a lemma of Csiszar and Körner [14], as proposed in [30]. Recently it was found that the two techniques have the same power [24]. Actually, the original lemma from [14] is not needed and only the following version on extensions suffices for proofs of [24] and [30].

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

$$g(i' \cup N \setminus i) = g(N \setminus i), \\ g(i' \cup I) - g(i') = g(i \cup I) - g(i), \quad I \subseteq N \setminus i.$$

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 the value $t = h(N) - h(N \setminus i)$. By Lemma 3, g is almost entropic. The value t is at most $h(i) = f(L)$ whence (1) applies and takes the form

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

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

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

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 $g(i' \cup I) - g(i') = g(i \cup I) - g(i)$. ■

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.

VI. Entropy Region of Four Variables

This section presents applications of polymatroidal constructions and consequences of above results in the situation when the ground set N has four elements. It partially describes shape of the cone $cl(\boldsymbol{H}_N^{\mathsf{ent}})$ which is 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 $\boldsymbol{H} = \boldsymbol{H}_N$.

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

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

It is interpreted also as the scalar product $\square_{ij} h$ of

$$\square_{ij} = \delta_{ik} + \delta_{i\ell} + \delta_{jk} + \delta_{j\ell} + \delta_{k\ell} - \delta_{ij} \\ - \delta_k - \delta_\ell - \delta_{ik\ell} - \delta_{jk\ell}$$

with h. The inequality $\square_{ij} h \geqslant 0$ holds when h is linear, see the works of Ingleton [22], [23]. Let $\boldsymbol{H}^{\square}$ denote the polyhedral cone of the functions $h \in \boldsymbol{H}$ that satisfy the six instances of the Ingleton inequality obtained by the permutation symmetry. By [40, Lemma 3], $\boldsymbol{H}^{\square}$ has dimension 15 and is generated by linear polymatroidal rank functions. Therefore, the functions from $ri(\boldsymbol{H}^{\square})$ are entropic due to [37, Th. 1].

By [40, Lemma 4], any $h \in \boldsymbol{H} \setminus \boldsymbol{H}^{\square}$ violates exactly one of the six Ingleton inequalities. Let $\boldsymbol{H}^{\square}_{(ij)}$ denote the cone of functions $h \in \boldsymbol{H}$ with $\square_{ij} h \leqslant 0$. It follows that $\boldsymbol{H}$ is union of $\boldsymbol{H}^{\square}$ with the six cones $\boldsymbol{H}^{\square}_{(ij)}, \ldots, \boldsymbol{H}^{\square}_{(kl)}$. Focusing primarily on the cone $cl(\boldsymbol{H}^{\mathsf{ent}})$, it contains $\boldsymbol{H}^{\square}$ and is contained in the union. By symmetry, it remains to study a single intersection, say $cl(\boldsymbol{H}^{\mathsf{ent}}) \cap \boldsymbol{H}^{\square}_{(ij)}$.

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

Corollary 5: $cl(\boldsymbol{H}^{\mathsf{ent}}) \cap \boldsymbol{H}^{\square}_{(ij)} = \boldsymbol{L}_{ij} \oplus \boldsymbol{H}^{\mathsf{mod}}$.

Proof: Since the expression $\square_{ij} h$ is balanced, $\square_{ij} h$ equals $\square_{ij} h^{\mathsf{ti}}$ and $\boldsymbol{H}^{\mathsf{mod}}$ is contained in $\boldsymbol{H}^{\square}_{(ij)}$. These observations and Corollary 3 imply the decomposition. ∎

The study of $cl(\boldsymbol{H}^{\mathsf{ent}})$ thus reduces to that of $\boldsymbol{L}_{ij}$. This cone is contained in $\boldsymbol{H}^{\square}_{(ij)} \cap \boldsymbol{H}^{\mathsf{ti}}$ which is known to be the conic hull of 11 linearly independent polymatroidal rank functions [40, Lemma 6.1]. The most notable one

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

is not almost entropic by Zhang-Yeung inequality [48]. The remaining ten rank functions are matroidal

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

where all 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 [40, Lemma 6.1], every $g \in \boldsymbol{H}^{\square}_{(ij)} \cap \boldsymbol{H}^{\mathsf{ti}}$ can be written uniquely as

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

This is a conic combination of the rank functions from (8) and (9), identifying explicitly the coordinate functionals.

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

Corollary 6: $ri(\boldsymbol{L}_{ij}) \subseteq \boldsymbol{H}^{\mathsf{ent}}$.

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

Let $A_{i,j}$ and $B_{ij,k}$ be the linear mappings defined by

$$\begin{aligned} A_{i,j} g &= g + (\Delta_{ij|\emptyset}\, g)(r_1^{i} - r_1) \\ B_{ij,k} g &= g + (\Delta_{kl|ij}\, g)(r_2^{k} - r_3) \end{aligned}$$

where $g \in \mathbb{R}^{\mathcal{P}(N)}$.

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

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

A simple proof is omitted, up to the computation

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

that is needed below. Both $A_{i,j}$ and $B_{ij,k}$ change at most two coordinates in the conic combination (10).

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

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

By the identity

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

if $f \in \boldsymbol{H}^{\square}_{(ij)}$ then $\Delta_{ik|j} f \geqslant \Delta_{ij|k} f$, thus $f(ij) - f(j) \geqslant f(ik) - f(k)$. By symmetry, $f(ij) - f(j) \geqslant f(il) - f(l)$. In turn, Corollary 1 can be applied to $t = f(ij) - f(j)$, and provides $h \in \boldsymbol{H}$ 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 $f \in cl(\boldsymbol{H}^{\mathsf{ent}})$ then h, being the convolution of f with a modular polymatroid, is almost entropic. Therefore, $f \in \boldsymbol{L}_{ij}$ implies $A_{i,j} f \in cl(\boldsymbol{H}^{\mathsf{ent}})$.

By the identity

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

if $f \in \boldsymbol{H}^{\square}_{(ij)}$ then $\Delta_{ik|jl} f \geqslant \Delta_{kl|ij} f$. Additionally, if f is tight this inequality rewrites to $f(ij) \geqslant f(jl)$. By symmetry, $f(ij) \geqslant f(il)$. It follows that (3) is valid for $t = f(N) - f(ij)$. By Remark 1 and $t \leqslant f(k)$, 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 $k \in cl(I)$. By the identity

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

$\Delta_{ij|kl} f \geqslant \Delta_{kl|ij} f$. Since f is tight the inequality rewrites to $f(ij) \geqslant f(kl)$. By symmetry, $f(ij)$ is maximal among all $f(J)$ with $|J| = 2$. Therefore, if $t > 0$ then $k \in cl(I)$ 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, $f \in cl(\boldsymbol{H}^{\mathsf{ent}})$ then h is almost entropic. Therefore, $f \in \boldsymbol{L}_{ij}$ implies $B_{ij,k} f \in \boldsymbol{L}_{ij}$. ∎

Remark 3: Let $\boldsymbol{E}_{ij}$ be the face of $\boldsymbol{L}_{ij}$ given by the equalities

$$\begin{aligned} &\Delta_{ij|k}\, g = 0, \quad \Delta_{ij|l}\, g = 0, \ \ \Delta_{kl|i}\, g = 0, \\ &\Delta_{kl|j}\, g = 0 \text{ and } \Delta_{kl|ij}\, g = 0. \end{aligned}$$

In [34, Example 2], four random variables are constructed such that their entropy function g satisfies the above five constraints, $\square_{ij} g < 0$, each of $\Delta_{ij|\emptyset}\, g$, $\Delta_{jl|k}\, g$, $\Delta_{il|k}\, g$, $\Delta_{jk|l}\, g$ and $\Delta_{ik|l}\, g$ is positive, and g is not tight. The lack of tightness makes g to be outside $\boldsymbol{L}_{ij}$. Nevertheless, Corollary 2 implies that g^{ti} is almost entropic whence it belongs to the face $\boldsymbol{E}_{ij}$. Even more, it belongs to its relative interior. At the

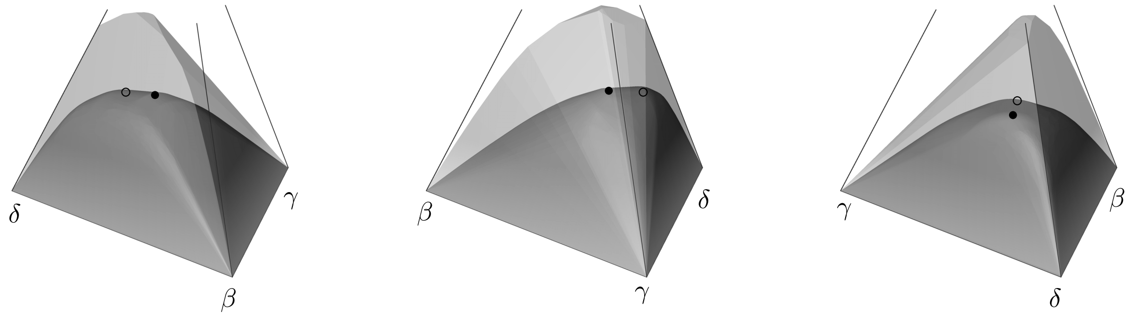

Fig. 1. Inner and outer approximations of $\boldsymbol{S}_{ij}$.

same time, [35, Th. 4.1] implies that no point of $ri(\boldsymbol{E}_{ij})$ is entropic. This phenomenon can be equivalently rephrased in terms of conditional information inequalities, studied recently in [25]–[27].

VII. Symmetrization of $\boldsymbol{F}_{ij}$

In the previous section, assuming $N = \{i, j, k, l\}$, the study of the cone $cl(\boldsymbol{H}_N^{\mathsf{ent}})$ was reduced to that of $\boldsymbol{L}_{ij}$, and a particular face $\boldsymbol{F}_{ij}$ of the latter was identified. Here, a symmetrization of $\boldsymbol{F}_{ij}$ is described and its cross-section visualized, resorting to computer assistance.

The expression $\square_{ij}$ and the cones $\boldsymbol{L}_{ij}$ and $\boldsymbol{F}_{ij}$ enjoy natural symmetries. Namely, if a permutation π on N stabilizes the two-element set ij then $\square_{ij} h = \square_{ij} h_\pi$, $h \in \boldsymbol{H}$. Hence $\boldsymbol{L}_{ij}$ and $\boldsymbol{F}_{ij}$ 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), for $h \in \boldsymbol{H}^{\square}_{(ij)} \cap \boldsymbol{H}^{\mathsf{ti}}$

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

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

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

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

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

Let $\bar{\alpha}_h$, $\bar{\beta}_h$, $\bar{\gamma}_h$ and $\bar{\delta}_h$ denote the above brackets, respectively. They are nonnegative and sum to one. Any function $h \in \boldsymbol{S}_{ij}$ can be written as

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

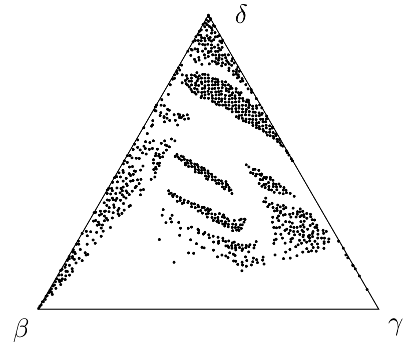

Fig. 2. Extreme points of the dark gray region, projected to $\beta\gamma\delta$.

which is a convex combination of the linearly independent polymatroidal rank functions

$$\begin{aligned} \alpha &= \tfrac{1}{4}\bar{r}_{ij} \quad \beta = \tfrac{1}{2}[r_1^j + r_1^i] \\ \gamma &= \tfrac{1}{4}[r_2^l + r_2^k] \quad \delta = \tfrac{1}{4}[r_1^{ik} + r_1^{jk} + r_1^{il} + r_1^{jl}]. \end{aligned}$$

It follows that $\boldsymbol{S}_{ij}$ is a closed convex subset of the three-dimensional tetrahedron in $\mathbb{R}^4$ with the vertices α, β, γ and δ. Since the points h having $\bar{\alpha}_h = 0$ are almost entropic and $\bar{r}_{ij}$ is not, $\boldsymbol{S}_{ij}$ contains the triangle $\beta\gamma\delta$, but not the vertex α.

Computer experiments were run to visualize $\boldsymbol{S}_{ij}$. The 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 $g = C_{ij} A_{i,j} B_{ij,k} f^{\mathsf{ti}}$ and then to $h = g/g(N)$, which is the convex combination

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

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 $\boldsymbol{S}_{ij}$ 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 α is missing and the straight lines are the incomplete edges of the tetrahedron incident to α. The dark gray region is spanned by about 2200 extreme points. The projections of the extreme points from α 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 VIII.

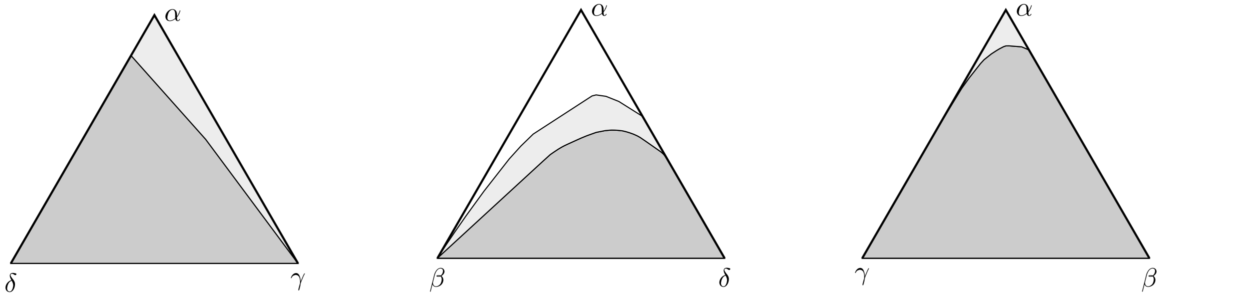

Fig. 3. Projections of the approximations of $\boldsymbol{S}_{ij}$ to triangles.

The light gray region in Figure 1 visualizes an outer approximation of $\boldsymbol{S}_{ij}$ which was constructed from hundreds of known non-Shannon information inequalities, mostly from those of [13] and [16]. Details are omitted. The gap between the approximations is large.

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 β and γ. The analogous statement for the edge $\alpha\delta$ is open.

VIII. 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 $\mathbb{I}_{ij}(h) \triangleq \square_{ij} h / h(N)$ [16, Definition 3]. The number

$$\mathbb{I}^* \triangleq \inf \left\{ \mathbb{I}_{ij}(h) \colon 0 \neq h \in \boldsymbol{H}^{\mathsf{ent}} \right\}$$

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: $\mathbb{I}^* = \min_{\boldsymbol{S}_{ij}} \mathbb{I}_{ij}$.

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

$$\mathbb{I}^* = \min \left\{ \mathbb{I}_{ij}(h) \colon h(N) = 1, \square_{ij} h \leqslant 0 \text{ and } h \in cl(\boldsymbol{H}^{\mathsf{ent}}) \right\},$$

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

$$\mathbb{I}^* = \min \left\{ \mathbb{I}_{ij}(h) \colon h(N) = 1 \text{ and } h \in \boldsymbol{L}_{ij} \right\}. \tag{12}$$

Recall that $\boldsymbol{L}_{ij} = cl(\boldsymbol{H}^{\mathsf{ent}})^{\mathsf{ti}} \cap \boldsymbol{H}^{\square}_{(ij)}$ is the cone of tight almost entropic rank functions h with $\square_{ij} h \leqslant 0$. (By Corollary 6, the above minimization can be expressed by special entropy functions.)

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

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

By Lemma 7, $\mathbb{I}_{ij}(g) \geqslant \mathbb{I}_{ij}(A_{i,j} B_{ij,k} g)$ 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 $\boldsymbol{F}_{ij}$. The assertion follows by symmetrization. ■

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

Lower bounds on $\mathbb{I}^*$ can be obtained by relaxing $\boldsymbol{L}_{ij}$ in (12). The simplest relaxation is to $\boldsymbol{H}^{\mathsf{ti}} \cap \boldsymbol{H}^{\square}_{(ij)}$ 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 a little more work, the bound $-\frac{1}{6}$ can be obtained by Zhang-Yeung inequality. Better lower bounds are reported in [16], 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 [3], [18], [20], [21], [32], [34], [35], [44], [47], [50]. The following one has attracted a special attention.

Example 1: Let ξ_i and ξ_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 [34, Example 1] or [11]. If $0 \leqslant p \leqslant \frac{1}{2}$ 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 $\square_{ij} = \Delta_{kl|i} + \Delta_{kl|j} + \Delta_{ij|\emptyset} - \Delta_{kl|\emptyset}$. Let $\kappa(u) = -u \ln u$, $u > 0$, and $\kappa(0) = 0$. The numerator is

$$\begin{aligned} &2\ln 2 - 2\kappa(p) - 2\kappa(\tfrac{1}{2} - p) - \Big[2\kappa(1-p) - \kappa(1-2p) \Big] \\ &\quad = (2p+1)\ln 2 - 2\kappa(p) - 2\kappa(1-p) \end{aligned}$$

and the denominator is $2\kappa(p) + 2\kappa(\frac{1}{2} - p)$. 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 [11] but the formulation [11, Conjecture 4.1] had a wrong numerical value. The same surmise appeared later in [16] as *the four-atom conjecture*, referring to the four possible values of $\xi_i\xi_j\xi_k\xi_l$. The minimization was considered also in [3] and [32] that report no score below $\mathbb{I}_{ij}(h_{p^*})$. However, the computer

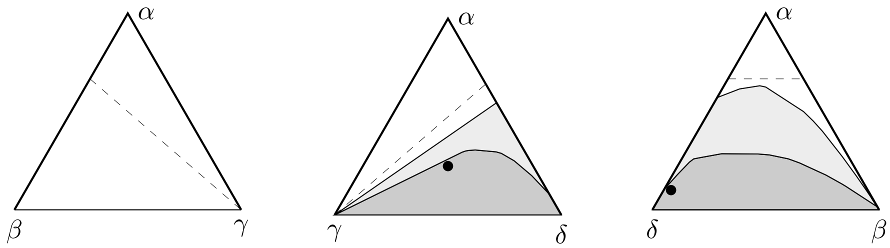

Fig. 4. Intersections of the approximations of $\boldsymbol{S}_{ij}$ with triangles.

experiments discussed in Section VII found an entropic point that can be transformed to an almost entropic point witnessing the failure of the 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,

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

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,

$$f(ik) = 4\kappa(2p+2t) + 8\kappa(q+r+s)$$
$$f(jl) = 4\kappa(2p+2s) + 8\kappa(q+r+t).$$

Analogous considerations provide

$$f(ikl) = 8\kappa(p+t) + 8\kappa(q+s) + 8\kappa(r)$$
$$f(jkl) = 8\kappa(p+s) + 8\kappa(q+t) + 8\kappa(r)$$
$$f(ijk) = 8\kappa(p+t) + 8\kappa(r+s) + 8\kappa(q)$$
$$f(ijl) = 8\kappa(p+s) + 8\kappa(r+t) + 8\kappa(q).$$

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

$$p = 0.09524 \quad q = 0.02494$$
$$r = 0.00160 \quad s = t = 0.00161,$$

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

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

by (11). Hence, the score $\mathbb{I}_{ij}(g)$ is approximately -0.09243, currently the best upper bound on the infimal Ingleton score.[1]

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 $C_{ij}A_{i,j}B_{ij,k}f^{\mathsf{ti}}$ where f is the entropic point from Example 2.

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

The symmetrized Zhang-Yeung inequality

$$\begin{aligned} 2\square_{ij}h &+ \big[\Delta_{ik|l}\,h + \Delta_{il|k}\,h + \Delta_{kl|i}\,h\big] \\ &+ \big[\Delta_{jk|l}\,h + \Delta_{jl|k}\,h + \Delta_{kl|j}\,h\big] \geqslant 0, \end{aligned}$$

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

By [16, Th. 10], if $s \geqslant 0$ is integer then for $h \in \boldsymbol{H}^{\mathsf{ent}}$

$$\begin{aligned} (2^s-1)\square_{ij}h &+ \Delta_{kl|i}\,h + s2^{s-1}\big[\Delta_{ik|l}\,h + \Delta_{il|k}\,h\big] \\ &+ \big((s-2)2^{s-1}+1\big)\big[\Delta_{jk|l}\,h + \Delta_{jl|k}\,h\big] \geqslant 0. \end{aligned}$$

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

$$\bar{\beta}_h + \big[(s-1)2^s+1\big]\bar{\delta}_h \geqslant \tfrac{1}{2}(2^s-1)\bar{\alpha}_h, \quad h \in \boldsymbol{H}^{\mathsf{ent}}. \qquad (13)$$

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 [34, Example 1]. The bullet inside the triangle $\alpha\beta\delta$ shows the almost entropic point $C_{ij}A_{i,j}g^{\mathsf{ti}}$ where g is the entropy function discussed in Remark 3, see also [34, Example 2].

[1] Still better bounds have been reported at the last revision of this work, by experimenting with subgroups of groups [4].

IX. Conclusion

The structure and shape of the entropy region is mostly unknown. There are a couple of general results, including statements such that $cl(\boldsymbol{H}_N^{\mathrm{ent}})$ is a convex cone, its interior consists of entropic points only, or it is selfadhesive (a notion which is behind all known techniques which generate non-Shannon entropy inequalities). Using polymatroidal techniques it was shown here that $cl(\boldsymbol{H}_N^{\mathrm{ent}})$ decomposes into the direct sum of its *modular* and *tight* parts – both smaller dimensional convex cones contained in the boundary of the closure of the entropy region. This fact is closely related to balanced information-theoretic inequalities [8].

While the modular part has a simple structure, the tight part carries over the complexity of the whole entropy region. Section IV proves that the relative interior of the tight part is entropic, namely, this face of $cl(\boldsymbol{H}_N^{\mathrm{ent}})$ is regular. Whether a face of $cl(\boldsymbol{H}_N^{\mathrm{ent}})$ has no entropic points, or is regular, can be expressed in terms of conditional information inequalities, which has been studied extensively in [25]–[27]. Selfadhesivity relativizes to the tight part as was shown in Section V. Actually, the Ahlswede-Körner lemma used by [30] to prove non-Shannon entropy inequalities can be replaced by a consequence of results proved in this section.

Sections VI and VII presented an attempt to visualize the entropy region of four random variables. Applying techniques developed earlier, a 9-dimensional face of $\boldsymbol{H}_4^{\mathrm{ent}}$ has been identified which carries over most of the structure complexity. Its symmetrization, when normalized by the total entropy, is the 3-dimensional body $\boldsymbol{S}_{ij}$, described by 4 homogeneous coordinates. Computer experiments were run to create inner and outer bounds for $\boldsymbol{S}_{ij}$. The gap between the approximations is quite large. The true shape is conjectured to be close to the inner approximation.

The *Ingleton score* measures how much the Ingleton inequality, valid for linear polymatroids, can be violated by entropy functions. This amount turns out to be the height of the body $\boldsymbol{S}_{ij}$ above its base, see Theorem 5. According to the computer simulations, the maximum is numerically attained along a relatively long, almost horizontal ridge, and it is not at the point provided by the four-atom distribution [16]. It is interesting to note that while we know the existence of a distribution with Ingleton score below $\mathbb{I}_{ij}(h_{p^*}) \doteq -0.089373$, no such an *explicit* distribution has been found so far.

Acknowledgment

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.

References

[1] R. Baber, D. Christofides, A. N. Dang, S. Riis, and E. R. Vaughan, “Multiple unicasts, graph guessing games, and non-Shannon inequalities,” in *Proc. NetCod*, Jun. 2013, pp. 1–6.

[2] R. Bassoli, H. Marques, J. Rodriguez, K. W. Shum, and R. Tafazolli, “Network coding theory: A survey,” *IEEE Commun. Surveys Tut.*, vol. 15, no. 5, pp. 1950–1978, 4th Quart., 2013.

[3] N. Boston and T.-T. Nan, “Large violations of the Ingleton inequality,” in *Proc. 50th Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Oct. 2012, pp. 1588–1593.

[4] N. Boston, private communication, 2016.

[5] A. Beimel and I. Orlov, “Secret sharing and non-Shannon information inequalities,” in *Theory of Cryptography* (Lecture Notes in Computer Science), vol. 5444. New York, NY, USA: Springer-Verlag, 2009, pp. 539–557.

[6] A. Beimel, N. Livne, and C. Padró, “Matroids can be far from ideal secret sharing,” in *Theory of Cryptography* (Lecture Notes in Computer Science), vol. 4948. New York, NY, USA: Springer-Verlag, 2008, pp. 194–212.

[7] T. H. Chan and R. W. Yeung, “On a relation between information inequalities and group theory,” *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1992–1995, Jul. 2002.

[8] T. H. Chan, “Balanced information inequalities,” *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3261–3267, Dec. 2003.

[9] T. Chan, A. Grant, and D. Pflüger, “Truncation technique for characterizing linear polymatroids,” *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6364–6378, Oct. 2011.

[10] T. Chan, “Recent progresses in characterising information inequalities,” *Entropy*, vol. 13, no. 2, pp. 379–401, 2011.

[11] L. Csirmaz, “The dealer’s random bits in perfect secret sharing schemes,” *Studia Sci. Math. Hungarica*, vol. 32, nos. 3–4, pp. 429–437, 1996.

[12] L. Csirmaz, “Book inequalities,” *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 6811–6818, Nov. 2014.

[13] L. Csirmaz, “Using multiobjective optimization to map the entropy region,” *Comput. Optim. Appl.*, vol. 63, no. 1, pp. 45–67, 2015.

[14] I. Csiszar and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[15] R. Dougherty, C. Freiling, and K. Zeger, “Six new non-Shannon information inequalities,” in *Proc. IEEE ISIT*, Seattle, WA, USA, Jul. 2006, pp. 233–236.

[16] R. Dougherty, C. F. Freiling, and K. Zeger. (2011). “Non-Shannon information inequalities in four random variables.” [Online]. Available: https://arxiv.org/abs/1104.3602

[17] S. Fujishige, “Polymatroidal dependence structure of a set of random variables,” *Inf. Control*, vol. 39, pp. 55–72, Oct. 1978.

[18] D. Fong, S. Shadbakht, and B. Hassibi, “On the entropy region and the Ingleton inequality,” in *Proc. Math. Theory Netw. Syst. (MTNS)*, 2008, pp. 1–8.

[19] M. Gadouleau and S. Riis, “Graph-theoretical constructions for graph entropy and network coding based communications,” *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6703–6717, Oct. 2011.

[20] B. Hassibi and S. Shadbakht, “On a construction of entropic vectors using lattice-generated distributions,” in *Proc. IEEE ISIT*, Jun. 2007, pp. 501–505.

[21] D. Hammer, A. Romashchenko, A. Shen, and N. Vereshchagin, “Inequalities for Shannon entropy and Kolmogorov complexity,” *J. Comput. Syst. Sci.*, vol. 60, pp. 442–464, Apr. 2000.

[22] A. W. Ingleton, “Conditions for representability and transversality of matroids,” in *Rencontre Franco-Britannique Actes* (Lecture Notes in Mathematics), vol. 211. Berlin, Germany: Springer-Verlag, 1971, pp. 62–66.

[23] A. W. Ingleton, “Representation of matroids,” in *Combinatorial Mathematics and Its Applications*, D. J. A. Welsh, Ed. London, U.K.: Academic, 1971, pp. 149–167.

[24] T. Kaced, “Equivalence of two proof techniques for non-Shannon-type inequalities,” in *Proc. IEEE ISIT*, Jul. 2013, pp. 236–240.

[25] T. Kaced and A. Romashchenko, “On essentially conditional information inequalities,” in *Proc. IEEE ISIT*, Jul./Aug. 2011, pp. 1935–1939.

[26] T. Kaced and A. Romashchenko, “On the non-robustness of essentially conditional information inequalities,” in *Proc. IEEE ITW*, Sep. 2012, pp. 262–266.

[27] T. Kaced and A. Romashchenko, “Conditional information inequalities for entropic and almost entropic points,” *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7149–7167, Nov. 2013.

[28] L. Lovász, “Submodular functions and convexity,” in *Mathematical Programming—The State of the Art*, A. Bachem, M. Grötschel, and B. Korte, Eds. Berlin, Germany: Springer-Verlag, 1982, pp. 235–257.

[29] M. Madiman, A. W. Marcus, and P. Tetali, “Information-theoretic inequalities in additive combinatorics,” in *Proc. IEEE ITW*, Jan. 2010, pp. 1–4.

[30] K. Makarychev, Y. Makarychev, A. Romashchenko, and N. Vereshchagin, “A new class of non-Shannon-type inequalities for entropies,” *Commun. Inf. Syst.*, vol. 2, pp. 147–166, Dec. 2002.

[31] M. Madiman and F. A. Ghassemi, “Combinatorial entropy power inequalities: Preliminary study of the Stam region,” *IEEE Trans. Inf. Theory*, to be published.
[32] W. Mao, M. Thill, and B. Hassibi. (2012). “On the Ingleton-violating finite groups and group network codes.” [Online]. Available: http://arxiv.org/abs/1202.5599
[33] F. Matúš, “Probabilistic conditional independence structures and matroid theory: Background,” *Int. J. General Syst.*, vol. 22, no. 2, pp. 185–196, 1994.
[34] F. Matúš, “Conditional independences among four random variables II,” *Combinat., Probab. Comput.*, vol. 4, pp. 407–417, Dec. 1995.
[35] F. Matúš, “Conditional independences among four random variables III: Final conclusion,” *Combinat., Probab. Comput.*, vol. 8, no. 3, pp. 269–276, 1999.
[36] F. Matúš, “Adhesivity of polymatroids,” *Discrete Math.*, vol. 307, pp. 2464–2477, Oct. 2006.
[37] F. Matúš, “Two constructions on limits of entropy functions,” *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 320–330, Jan. 2007.
[38] F. Matúš, “Infinitely many information inequalities,” in *Proc. IEEE ISIT*, Nice, France, Jun. 2007, pp. 41–44.
[39] F. Matúš, “Classes of matroids closed under the minors and principal extensions,” *Combinatorica*, under final consideration.
[40] F. Matúš and M. I. Studený, “Conditional independences among four random variables I,” *Combinat., Probab. Comput.*, vol. 4, pp. 269–278, Sep. 1995.
[41] J. G. Oxley, *Matroid Theory*. Oxford, U.K.: Oxford Univ. Press, 1992.
[42] S. Riis, “Information flows, graphs and their guessing numbers,” *Electron. J. Combinat.*, vol. 14, pp. 1–17, Jun. 2007.
[43] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ, USA: Princeton Univ. Press, 1970.
[44] M. Studený, personal communication, 1990.
[45] M. Studený, *Probabilistic Conditional Independence Structures*. New York, NY, USA: Springer, 2005.
[46] R. W. Yeung, *A First Course in Information Theory*. New York, NY, USA: Kluwer, 2002.
[47] J. M. Walsh and S. Weber, “Relationships among bounds for the region of entropic vectors in four variables,” in *Proc. Allerton Conf. Commun., Control, Comput.*, 2010, pp. 1319–1326.
[48] Z. Zhang and R. W. Yeung, “A non-Shannon-type conditional inequality of information quantities,” *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1982–1986, Nov. 1997.
[49] Z. Zhang, “On a new non-Shannon type information inequality,” *Commun. Inf. Syst.*, vol. 3, no. 1, pp. 47–60, 2003.
[50] Z. Zhang and R. W. Yeung, “On characterization of entropy function via information inequalities,” *IEEE Trans. Inf. Theory*, vol. 44, no. 4, pp. 1440–1452, Jul. 1998.

František Matúš, biography not available at the time of publication.

Lászlo Csirmaz, biography not available at the time of publication.