Linear Complexity Entropy Regions

Yirui Liu and John MacLaren Walsh
Department of Electrical and Computer Engineering
Drexel University
Philadelphia, PA 19104, USA
Email: {yl926, jmw96}@drexel.edu

***Abstract*—A great many problems in network information theory, both fundamental and applied, involve determining a minimal set of inequalities linking the Shannon entropies of a certain collection of subsets of random variables. In principle this minimal set of inequalities could be determined from the entropy region, whose dimensions are all the subsets of random variables, by projecting out dimensions corresponding to the irrelevant subsets. As a general solution technique, however, this method is plagued both by the incompletely known nature of the entropy region as well as the exponential complexity of its bounds. Even worse, for four or more random variables, it is known that the set of linear information inequalities necessary to completely describe the entropy region must be uncountably infinite. A natural question arises then, if there are certain nontrivial collections of subsets where the inequalities linking only these subsets is both completely known, and have inequality descriptions that are linear in the number of random variables. This paper answers this question in the affirmative. A decomposition expressing the collection of inequalities linking a larger collection of subsets from that of smaller collections of subsets is first proven. This decomposition is then used to provide systems of subsets for which it both exhaustively determines the complete list of inequalities, which is linear in the number of variables.**

***Index Terms*—information inequalities, complexity of entropy region**

I. Introduction

For two sets A and B, we will denote by $A^B = \{\boldsymbol{x}\,|\boldsymbol{x} : B \to A\}$. For two subsets $\mathscr{P}_1 \subseteq \mathbb{R}^{A_1}$ and $\mathscr{P}_2 \subseteq \mathbb{R}^{A_2}$ of Euclidean space, define $\mathscr{P}_1 \bowtie \mathscr{P}_2$ as,

$$\mathscr{P}_1 \bowtie \mathscr{P}_2 = \left\{ \boldsymbol{f} \in \mathbb{R}^{A_1 \cup A_2} \middle| \begin{array}{l} \exists\ \boldsymbol{g}_1 \in \mathscr{P}_1, \boldsymbol{g}_2 \in \mathscr{P}_2 \text{ s.t :} \\ \boldsymbol{f}(a) = \boldsymbol{g}_1(a), \forall a \in A_1 \\ \boldsymbol{f}(b) = \boldsymbol{g}_2(b), \forall b \in A_2 \end{array} \right\} \tag{1}$$

Let the ground set $E \subseteq \mathbb{N}$ index a set of $|E|$ finite discrete random variables (RVs), and let,

$$\Delta_E = \left\{ p_E \middle| \begin{array}{l} \mathcal{X} \subseteq \mathbb{N}, p_E : \mathcal{X}^E \to [0,1], \\ \sum_{\boldsymbol{x}_E \in \mathcal{X}^E} p_E(\boldsymbol{x}_E) = 1 \end{array} \right\} \tag{2}$$

be a set of valid joint probability mass functions (PMFs) for these RVs paired with their ranges. For any such joint PMF $p_E \in \Delta_E$ and any subset $A \subset E$ define the generalized marginalization operator $M : \Delta_E \times 2^E \to \bigcup_{A \subseteq E} \Delta_A$ with,

$$M(p_E, A) = p_A,\ \ A \neq \varnothing \tag{3}$$

where,

$$p_A(\boldsymbol{x}_A) = \sum_{\substack{\boldsymbol{x}_E \in \mathcal{X}^E \text{ s.t.} \\ \boldsymbol{x}_E(e) = \boldsymbol{x}_A(e)\ \forall e \in A}} p_E(\boldsymbol{x}_E) \tag{4}$$

Given two PMFs p_{E_1} and p_{E_2} such that $E_1 \cap E_2 \neq \varnothing$, we write

$$p_{E_1} \overset{M}{=} p_{E_2} \tag{5}$$

if and only if

$$M(p_{E_1}, E_1 \cap E_2) = M(p_{E_2}, E_1 \cap E_2) \tag{6}$$

Define the function $\hat{\boldsymbol{h}}_E : \Delta_E \to \mathbb{R}^{2^E}$ such that for any $A \subseteq E$,

$$\hat{\boldsymbol{h}}_E(p_E) : A \mapsto -\sum_{\boldsymbol{x}_A \in \mathcal{X}^A} p_A(\boldsymbol{x}_A) \log_2 (p_A(\boldsymbol{x}_A)) \tag{7}$$

with again $M(p_E, A) = p_A$. Here, based on preference, we can either think of $\hat{\boldsymbol{h}}_E(p_E)$ as a vector whose elements are indexed by all subsets 2^E, or equivalently as a function $\hat{\boldsymbol{h}}_E(p_E) : 2^E \to \mathbb{R}$ assigning a number to to each such subset of E. Either way, the value associated with the subset A is the joint Shannon entropy of that subset of the random variables.

The entropy region for the ground set E will be the image of the set of all joint PMFs Δ_E under this this map $\hat{\boldsymbol{h}}_E$

$$\Gamma_E^* = \hat{\boldsymbol{h}}_E(\Delta_E) \tag{8}$$

denoted as $\boldsymbol{h} = [\boldsymbol{h}(A)|A \subseteq E]$ an entropic vector in Γ_E^* and let,

$$\Xi = \hat{\boldsymbol{h}}_E^{-1}(\boldsymbol{h}) \tag{9}$$

be the collection of PMFs that is associated with $\boldsymbol{h}$.

Γ_E^* was first introduced in [1], motivated by the desire to develop a new framework to proof information inequalities. By definition, Γ_E^* gives a complete characterization of all information inequalities and thus is closely related to some important problems in probability theory and information theory. The attempt to fully characterize Γ_E^* started from the so-called basic information inequalities [2], which are implied from the non-negativity of the conditional mutual informations and are of the form:

$$\boldsymbol{h}(A) + \boldsymbol{h}(B) \geqslant \boldsymbol{h}(A \cup B) + \boldsymbol{h}(A \cap B), \forall A, B \subseteq E \tag{10}$$

Each vector $\boldsymbol{h} \in \Gamma_E^*$ must satisfies all the inequalities of the form (10), which means, when Γ_E is defined as,

$$\Gamma_E = \{\boldsymbol{h} \in \mathbb{R}^{2^E} | \boldsymbol{h} \text{ satisfies (10)}, \forall A, B \subseteq E\} \tag{11}$$

we have an outer bound $\Gamma^*_E \subseteq \Gamma_E$. $\overline{\Gamma^*_E}$ denotes the closure of Γ^*_E, the natural question of whether or not $\overline{\Gamma^*_E} \overset{?}{=} \Gamma_E$ motivates the finding of non-Shannon type inequalities [3]–[5], which are defined as the information inequalities that can not be implied from basic inequalities in (10). The closure of the entropy region $\overline{\Gamma^*_E}$ was proved to be a closed convex cone [3], and for $|E| \leqslant 3$ the cone is polyhedral and $\overline{\Gamma^*_E} = \Gamma_E$. The equality doesn't hold, meaning $\overline{\Gamma^*_E} \neq \Gamma_E$, for $|E| \geqslant 4$, moreover $\overline{\Gamma^*_E}$ is not polyhedral when $|E| \geqslant 4$, and its inequality and extreme ray descriptions are incompletely known [6], [7].

Significant progresses have been achieved on difference aspects of the topic [8]–[10] [11]–[13], However, much less has been written considering the instance when we only need the list of all fundamental inequalities linking those entropies of only certain subsets. Furthermore, it is only this restriction to only a certain collection of subsets that is often required in the multiterminal information theory applications. Hence, in this article we are aiming to study the restriction of the entropy region to only those dimensions associated with a certain given collection of subsets.

To describe this concept properly, consider a collection of subsets $\mathcal{F} \subseteq 2^E$, and define the projection,

$$\text{proj}_{\mathcal{F}}\ \overline{\Gamma^*_E} = \left\{ h \in \mathbb{R}^{\mathcal{F}} \middle| \exists \boldsymbol{h} \in \overline{\Gamma^*_E},\ h(A) = \boldsymbol{h}(A)\ \forall A \in \mathcal{F} \right\} \tag{12}$$

Again, most applied problems, for instance network coding capacity regions [14], depend on Γ^*_E only through its projection $\text{proj}_{\mathcal{F}}\ \Gamma^*_E$ for some smaller collection of subsets $\mathcal{F} \subsetneq 2^E$. Even worse, even the outer bounds such as the Shannon outer bound have both a number of inequalities and dimensions that grow exponentially in $|E|$. *Significantly, while $\overline{\Gamma^*_E}$ is non-polyhedral for $|E| \geqslant 4$, and these outer bounds have complexities that are exponential in $|E|$, there is the possibility that its projection proj$_{\mathcal{F}}$ $\overline{\Gamma^*_E}$ can in some instances be polyhedral and, further, have a number of inequalities and dimensions that grow linearly in $|E|$.* This paper proves that indeed, many non-trivial such instances exist.

II. Preliminaries

To understand the general idea behind the method the paper utilizes to prove these existence of these cases, let $E = E_1 \cup E_2$, and suppose that $\mathcal{F} = 2^{E_1} \cup 2^{E_2}$. In the trivial case that $E_1 \cap E_2 = \varnothing$, it is direct to show that,

$$\text{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\Gamma^*_E}) = \overline{\Gamma^*_{E_1}} \bigotimes \overline{\Gamma^*_{E_2}} \tag{13}$$

The more general case, when there's a collection of subsets $E_1, ..., E_m$ such that $E_i \cap E_j \neq \varnothing$ for some $i, j \in [1, m]$, is more complicated. The main contribution of the paper is summarized as the following theorem.

A. Main Results

Theorem 1. *Let $n \in \mathbb{N}$, consider a collection of $m = 2n$ subsets $E_{2i+1} = \{3i+1, 3i+2, 3i+3\}$, $E_{2i+2} = \{3i+2, 3i+3, 3i+4\}$, $\forall i \in [0, n-1]$. Let $\mathcal{F} = \bigcup_{i=1}^m 2^{E_i}$, then proj$_{\mathcal{F}}(\overline{\Gamma^*_E})$ is a **polyhedral cone** with:*

*1) **explicitly known** inequality descriptions.*

*2) a number of inequalities that grows **linearly** in $|E|$.*

Theorem 1 is proved based on the following two theorems.

Theorem 2. *Let E_1,E_2 two ground sets with such that $|E_1 \cap E_2| = 1$, then*

$$\textit{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\Gamma^*_E}) = \overline{\Gamma^*_{E_1}} \bigotimes \overline{\Gamma^*_{E_2}} \tag{14}$$

Theorem 3. *If $E_1 = \{1, 2, 3\}$, $E_2 = \{2, 3, 4\}$ and $E = E_1 \cup E_2 = \{1, 2, 3, 4\}$, then,*

$$\textit{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\Gamma^*_E}) = \overline{\Gamma^*_{E_1}} \bigotimes \overline{\Gamma^*_{E_2}} \tag{15}$$

B. Outer bounds on proj$_{\cup_{i=1}^m 2^{E_i}}(\overline{\Gamma^*_E})$

A natural outer bound of $\text{proj}_{\cup_{i=1}^m 2^{E_i}}(\overline{\Gamma^*_E})$ would concatenate the inequality descriptions for $\overline{\Gamma^*_{E_i}}, i \in [1, m]$ to get,

$$\text{proj}_{\cup_{i=1}^m 2^{E_i}}(\overline{\Gamma^*_E}) \subseteq \bigotimes_{i=1}^{m} \overline{\Gamma^*_{E_i}} \tag{16}$$

A tighter outer bound forces there to be a way to select the distributions to match on their marginals,

$$\text{proj}_{\cup_{i=1}^m 2^{E_i}}(\overline{\Gamma^*_E}) \subseteq (\bigotimes_{i=1}^{m} \overline{\Gamma^*_{E_i}}) \bigotimes \overline{\mathcal{P}^{=}_{E_1,...,E_m}} \tag{17}$$

where,

$$\overline{\mathcal{P}^{=}_{E_1,...,E_m}} = \left\{ \overline{\boldsymbol{f}} \in \bigotimes_{i=1}^{m} \overline{\Gamma^*_{E_i}} \middle| \begin{array}{l} \exists\ \boldsymbol{h}_i^{(n)} \in \Gamma^*_{E_i}, i \in [1, m]: \\ 1) \lim_{n \to \infty} \boldsymbol{h}_i^{(n)} = \overline{\boldsymbol{h}_i}, i \in [1, m] \\ 2) \overline{\boldsymbol{f}}(A) = \overline{\boldsymbol{h}_i}(A),\ A \subseteq E_i, i \in [1, m] \\ 3) \forall i \in [1, m], \exists\ p_{E_i}^{(n)} \in \hat{\boldsymbol{h}}_{E_i}^{-1}(\boldsymbol{h}_i^{(n)}),\ \text{s.t.} \\ \quad \forall i, j \in [1, m] \text{ if } E_i \cap E_j \neq \varnothing, \\ \quad \text{then } p_{E_i}^{(n)} \overset{M}{=} p_{E_j}^{(n)},\ \forall n \in \mathbb{N} \end{array} \right\} \tag{18}$$

C. Marginal problem and Running Intersection Property

Definition 1. *(The marginal problem on PMF) Consider a collection of marginal PMFs p_{E_1},..., p_{E_m},, the marginal problem wonders if there exists a joint PMF p_E with $E = \bigcup_{i=1}^m E_i$, such that,*

$$M(p_E, E_i) = p_{E_i},\ \forall i \in [1, m] \tag{19}$$

Defined above is the marginal problem on PMF, which has received research interests in different literature [15]–[18]. The problem is trivial when the collection of PMFs are defined on disjoint ground sets, in which case the stochastic product of the marginals give us a valid joint PMF. However, the problem is not trivial anymore when the ground sets are not disjoint. One obvious necessary condition for such joint PMF to exist for a marginal problem is pairwise compatibility, meaning for any pair of E_i and E_j, $M(p_{E_i}, E_i \cap E_j) = M(p_{E_j}, E_i \cap E_j)$ if $E_i \cap E_j \neq \varnothing$.

So the question becomes when does the pairwise compatibility implies a global compatibility. In fact, using the theory of hypergraphs, Beeri et al. [19](see also [20]) established a theorem on this problem stating that,

Theorem 4. *(Beeri et al. [19]) The following two conditions are equivalent:*

1) For any collection of pairwise compatible marginal PMFs $p_{E_1}, ..., p_{E_m}$ over $(E_1, ..., E_m)$, exists a joint PMF p_E, over $E = \bigcup_{i=1}^{m} E_i$ such that,

$$M(p_E, E_i) = p_{E_i},, \ \forall i \in [1, m] \tag{20}$$

*2) The collection of ground sets $E_1, ..., E_m$ satisfy Running Intersection Property (**RIP**).*

with the running intersection property defined as,

Definition 2. *(Running Intersection Property (RIP)) A sequence of subsets E_1, E_2,..., E_m of a finite ground set E satisfies the **Running Intersection Property** if for every $k \in [2, m]$ the intersection of E_k with $\bigcup_{j=1}^{k-1} E_j$ is contained in one of these previous subsets, that is,*

$$\forall k \in [2, m], \exists i \in [1, k-1] \ s.t. \ E_k \cap (\bigcup_{j=1}^{k-1} E_j) \subseteq E_i \tag{21}$$

III. Theorems and proofs

Proof. (sketch of the proof of **Theorem** 2) The containment

$$\text{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\Gamma_E^*}) \subseteq \overline{\Gamma_{E_1}^*} \bigotimes \overline{\Gamma_{E_2}^*} \tag{22}$$

is straightforward, so to prove the theorem is suffices to prove that

$$\overline{\Gamma_{E_1}^*} \bigotimes \overline{\Gamma_{E_2}^*} \subseteq \text{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\Gamma_E^*}) \tag{23}$$

The group characterizable entropy region is helpful to prove this containment - by **Corollary** 4.1 in [21], we know that,

$$\overline{\text{con}}(\Upsilon_E) = \overline{\Gamma_E^*} \tag{24}$$

In fact, using a rational approximation of the conic combination coefficients, one can show that $\overline{\text{con}}(\Upsilon_E) = \overline{\text{ray}}(\Upsilon_E)$. Thus we must show that for any pair of sequences of scaled group characterizable entropy vectors $\alpha_k \boldsymbol{h}_1^k \in \text{ray}(\Upsilon_E)$ and $\alpha_2^k \boldsymbol{h}_2^k \in \text{ray}(\Upsilon_E)$ with $\lim_{k\to\infty} \alpha_1^k \boldsymbol{h}_1^k(A) = \lim_{k\to\infty} \alpha_2^k \boldsymbol{h}_2^k(A), \forall A \subseteq E_1 \cap E_2$, we can construct a sequence of global entropy vectors $\alpha^k \boldsymbol{h}^k$ such that $\lim_{k\to\infty} \alpha^k \boldsymbol{h}^k(A) = \lim_{k\to\infty} \boldsymbol{h}_1^k(A), \ \forall A \subseteq E_1$ and $\lim_{k\to\infty} \alpha^k \boldsymbol{h}^k(A) = \lim_{k\to\infty} \boldsymbol{h}_2^k(A), \ \forall A \subseteq E_2$. Dirichlet's approximation theorem enables one to construct a rational approximation with a common denominator for α_1^k, α_2^k and $\alpha_1^k \boldsymbol{h}_1(A) - \alpha_2^k \boldsymbol{h}_2(A)$, as $\frac{n_1}{q}, \frac{n_2}{q}, \frac{e}{q}$, respectively that enables $|n_1 \boldsymbol{h}_1^k(A) - n_2 \boldsymbol{h}_2^k(A) - e|$ to be made arbitrarily small. Repeat the group structure for $\boldsymbol{h}_i^k$ n_i times for $i \in \{1, 2\}$, and build a map from the cosets for group associated with A for $i = 1$ to those for $i = 2$ that is one to one until running out of cosets for $i = 1$, mapping the remaining cosets to a deterministic extra symbol, drawing 2^e times uniformly from the selected coset for $i = 2$. The subset entropies of this construction, using the coset index from $i = 1$ as the random variable for the overlapping element, provide the required entropies in the limit. ∎

Lemma 1. *Given a collection of subsets E_1,...,E_m with $E = \cup_{i=1}^{m} E_i$ such that for any $i, j \in [1, m]$*

$$|E_i \cap E_j| = \begin{cases} 1, & |i-j| = 1 \\ 0, & |i-j| > 1 \end{cases} \tag{25}$$

then,

$$\textit{proj}_{\cup_{i=1}^{m} 2^{E_i}}(\overline{\Gamma_E^*}) = \bigotimes_{i=1}^{m} \overline{\Gamma_{E_i}^*} \tag{26}$$

Proof. (proof of **Lemma**1) Prove by induction, **Theorem** 2 gives the base case that,

$$\text{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\Gamma_E^*}) = \overline{\Gamma_{E_1}^*} \bigotimes \overline{\Gamma_{E_2}^*} \tag{27}$$

Next for the inductive step, let $E'_k = \cup_{i=1}^{k} E_i$ and assume that

$$\text{proj}_{\cup_{i=1}^{k} 2^{E_i}}(\overline{\Gamma_{E'_k}^*}) = \bigotimes_{i=1}^{k} \overline{\Gamma_{E_i}^*} \tag{28}$$

To prove it holds when adding E_{k+1}, let $E'_{k+1} = E'_k \cup E_{k+1}$, then applying **Theorem** 2 on the pair of sets E_{k+1} and E'_k we have,

$$\text{proj}_{2^{E'_k} \cup 2^{E_{k+1}}}(\overline{\Gamma_{E'_{k+1}}^*}) = \overline{\Gamma_{E'_k}^*} \bigotimes \overline{\Gamma_{E_{k+1}}^*} \tag{29}$$

Now project both sides of (29) down to $\cup_{i=1}^{k+1} 2^{E_i}$, we have

$$\begin{aligned} &\text{proj}_{\cup_{i=1}^{k+1} 2^{E_i}}(\text{proj}_{2^{E'_k} \cup 2^{E_{k+1}}}(\overline{\Gamma_{E'_{k+1}}^*})) \\ =&\text{proj}_{\cup_{i=1}^{k+1} 2^{E_i}}(\overline{\Gamma_{E'_{k+1}}^*}) \\ =&\text{proj}_{\cup_{i=1}^{k+1} 2^{E_i}}(\overline{\Gamma_{E'_k}^*} \bigotimes \overline{\Gamma_{E_{k+1}}^*}) \\ =&\text{proj}_{\cup_{i=1}^{k} 2^{E_i}}(\overline{\Gamma_{E'_k}^*}) \bigotimes \overline{\Gamma_{E_{k+1}}^*} \\ =&(\bigotimes_{i=1}^{k} \overline{\Gamma_{E_i}^*}) \bigotimes \overline{\Gamma_{E_{k+1}}^*} \\ =&\bigotimes_{i=1}^{k+1} \overline{\Gamma_{E_i}^*} \end{aligned} \tag{30}$$

∎

Lemma 2. *The $\bigotimes$ is both associative and commutative.*

Proof. Straightforward. ∎

Proof. (proof of **Theorem** 3) To show that,

$$\text{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\Gamma_E^*}) \subseteq \overline{\Gamma_{E_1}^*} \bigotimes \overline{\Gamma_{E_2}^*} \tag{31}$$

Pick an arbitrary vector $\overline{\boldsymbol{v}} \in \text{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\Gamma_E^*})$, then we can find a vector $\overline{\boldsymbol{h}} \in \overline{\Gamma_E^*}$ such that $\overline{\boldsymbol{v}} = \text{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\boldsymbol{h}})$, which means,

$$\overline{\boldsymbol{v}}(A) = \overline{\boldsymbol{h}}(A), \forall A \subseteq E_1 \text{ or } A \subseteq E_2 \tag{32}$$

By definition of $\overline{\Gamma_E^*}$, $\overline{\boldsymbol{h}}$ is a limit point of Γ_E^*, so we can find a sequence of entropic vectors $\boldsymbol{h}^{(n)} \in \Gamma_E^*$ each of which is associated with a PMF $p_E^{(n)}$. Create from $\boldsymbol{h}^{(n)}$ and $\overline{\boldsymbol{h}}$ the following terms,

$$\overline{\boldsymbol{k}} = [\overline{k}(A)|\overline{k}(A) = \overline{\boldsymbol{h}}(A), \forall A \subseteq E_1] \tag{33}$$

$$\bar{\boldsymbol{l}} = [\bar{l}(A)|\bar{l}(A) = \overline{\boldsymbol{h}}(A), \forall A \subseteq E_2] \tag{34}$$

$$\boldsymbol{k}^{(n)} = [k^{(n)}(A)|k^{(n)}(A) = \boldsymbol{h}^{(n)}(A), \forall A \subseteq E_1], \tag{35}$$

$$\boldsymbol{l}^{(n)} = [l^{(n)}(A)|l^{(n)}(A) = \boldsymbol{h}^{(n)}(A), \forall A \subseteq E_2] \tag{36}$$

It is not hard to see that by the above construction, $\boldsymbol{k}^{(n)}$ and $\boldsymbol{l}^{(n)}$ are sequence of entropic vectors, which implies that $\overline{\boldsymbol{k}} \in \overline{\Gamma^*_{E_1}}$ and $\bar{\boldsymbol{l}} \in \overline{\Gamma^*_{E_2}}$. Now combining (32), (33) and (34) we have $\overline{\boldsymbol{v}} \in \overline{\Gamma^*_{E_1}} \bigotimes \overline{\Gamma^*_{E_2}}$, which then implies that

$$\begin{aligned} &\overline{\boldsymbol{v}} \in \overline{\Gamma^*_{E_1}} \bigotimes \overline{\Gamma^*_{E_2}} \\ \implies &\text{proj}_{2^{E_1} \cup 2^{E_1}}(\overline{\Gamma^*_E}) \subseteq \overline{\Gamma^*_{E_1}} \bigotimes \overline{\Gamma^*_{E_2}} \end{aligned} \tag{37}$$

To show that,

$$\overline{\Gamma^*_{E_1}} \bigotimes \overline{\Gamma^*_{E_2}} \subseteq \text{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\Gamma^*_E}) \tag{38}$$

By assumption we have,

$$\overline{\Gamma^*_E} = \overline{\Gamma^*_{\{1,2,3,4\}}}$$

$$\begin{bmatrix}
1, 1, 1, 0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 2, 1, 1, 1, 1, 1, 2, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1 \\
1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 0, 0, 1, 0, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 0, 1, 1 \\
1, 1, 0, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 2, 1, 1, 1, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1 \\
1, 1, 1, 1, 0, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 0 \\
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 1, 1, 3, 2, 2, 3, 2, 2, 2, 2, 2, 2, 2, 2, 1, 2, 2, 2, 1, 1, 2, 2 \\
1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 0, 1, 0, 1, 2, 2, 3, 3, 2, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 2, 1, 2 \\
1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 3, 2, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 2, 2, 2, 1, 1, 2 \\
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 3, 3, 3, 3, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2 \\
1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 0, 3, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 2, 2, 1, 2, 1 \\
1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 0, 2, 3, 3, 2, 2, 2, 2, 2, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1 \\
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 3, 3, 3, 3, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2 \\
1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 2, 3, 2, 3, 2, 2, 2, 2, 2, 2, 2, 1, 2, 2, 2, 2, 1, 2, 2, 1 \\
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 3, 3, 3, 3, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2 \\
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 3, 3, 3, 3, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2 \\
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 3, 3, 3, 3, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2
\end{bmatrix}$$

Figure 1: 35 extreme rays of $\Gamma^{In}_{\{1,2,3,4\}}$. Each column indicates an extreme ray, each row indicates dimension labelled as, from left to right, $h(\{1\})$, $h(\{2\})$, $h(\{3\})$, $h(\{4\})$, $h(\{1,2\})$,$h(\{1,3\})$, $h(\{2,3\})$, $h(\{1,2,3\})$, $h(\{2,4\})$,$h(\{3,4\})$, $h(\{2,3,4\})$, $h(\{1,4\})$, $h(\{1,2,4\})$, $h(\{1,3,4\})$, and $h(\{1,2,3,4\})$

As shown in Fig.1, let $\Gamma^{In}_{\{1,2,3,4\}}$ be an inner bound of $\overline{\Gamma^*_{\{1,2,3,4\}}}$ constructed by taking the conic hull of all but 6 bad extreme rays of the Shannon outer bound $\Gamma_{\{1,2,3,4\}}$ of $\overline{\Gamma^*_{\{1,2,3,4\}}}$, then we must have that,

$$\text{proj}_{2^{E_1} \cup 2^{E_2}}(\Gamma^{In}_{\{1,2,3,4\}}) \subseteq \text{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\Gamma^*_{\{1,2,3,4\}}}) \tag{39}$$

We actually calculated $\text{proj}_{2^{E_1} \cup 2^{E_2}}(\Gamma^{In}_{\{1,2,3,4\}})$, which ended up to contain 23 extreme rays as shown in Fig.2. We then calculated $\Gamma_{\{1,2,3\}} \bigotimes \Gamma_{\{2,3,4\}}$ and verified that its extreme rays are exactly the same as Fig.2, which means,

$$\text{proj}_{2^{E_1} \cup 2^{E_2}}(\Gamma^{In}_{\{1,2,3,4\}}) = \Gamma_{\{1,2,3\}} \bigotimes \Gamma_{\{2,3,4\}} \tag{40}$$

together with the fact that $\Gamma_{\{1,2,3\}} = \overline{\Gamma^*_{\{1,2,3\}}}$ and $\Gamma_{\{2,3,4\}} = \overline{\Gamma^*_{\{2,3,4\}}}$(Shannon outer bound is tight on $\overline{\Gamma^*_E}$ when $|E| \leqslant 3$), we have,

$$\begin{aligned} &\overline{\Gamma^*_{\{1,2,3\}}} \bigotimes \overline{\Gamma^*_{\{2,3,4\}}} \\ \subseteq &\Gamma_{\{1,2,3\}} \bigotimes \Gamma_{\{2,3,4\}} \\ = &\text{proj}_{2^{E_1} \cup 2^{E_2}}(\Gamma^{In}_{\{1,2,3,4\}}) \\ \subseteq &\text{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\Gamma^*_{\{1,2,3,4\}}}) \end{aligned} \tag{41}$$

$$\begin{bmatrix}
1, 1, 1, 0, 1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 1, 2, 1, 1, 1 \\
1, 0, 1, 0, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1, 1, 1, 1, 1 \\
1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1 \\
1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 1, 1, 1 \\
1, 1, 1, 0, 2, 0, 1, 1, 1, 1, 1, 1, 2, 1, 1, 1, 0, 1, 2, 2, 2, 2, 2 \\
1, 1, 1, 0, 2, 1, 0, 1, 1, 1, 0, 1, 2, 1, 1, 2, 1, 1, 2, 2, 1, 2, 2 \\
1, 1, 1, 0, 2, 1, 1, 0, 1, 1, 1, 1, 2, 1, 2, 2, 1, 1, 2, 2, 2, 2, 2 \\
1, 1, 1, 0, 2, 1, 1, 1, 1, 1, 1, 1, 2, 1, 2, 2, 1, 1, 2, 2, 2, 2, 2 \\
1, 0, 1, 1, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1, 2, 2, 1, 1, 2, 2, 2, 2, 2 \\
1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 1, 1, 2, 1, 2, 2, 1, 1, 2, 2, 2, 2, 1 \\
1, 1, 1, 1, 2, 1, 1, 0, 1, 1, 1, 1, 2, 1, 2, 2, 1, 1, 2, 2, 2, 2, 2
\end{bmatrix}$$

Figure 2: 23 extreme rays of $\text{proj}_{2^{E_1} \cup 2^{E_2}}(\Gamma^{In}_{\{1,2,3,4\}})$. Each column indicates an extreme ray, each row indicates dimension labelled as, from left to right, $h(\{1\})$, $h(\{2\})$, $h(\{3\})$, $h(\{4\})$, $h(\{1,2\})$,$h(\{1,3\})$, $h(\{2,3\})$, $h(\{1,2,3\})$, $h(\{2,4\})$,$h(\{3,4\})$ and $h(\{2,3,4\})$.

which together with (37) implies that,

$$\text{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\Gamma^*_E}) = \overline{\Gamma^*_{E_1}} \bigotimes \overline{\Gamma^*_{E_2}} \tag{42}$$

■

Theorem 5. *Let* $n \in \mathbb{N}$*, consider a collection of* $m = 2n$ *subsets* $E_{2i+1} = \{3i+1, 3i+2, 3i+3\}$, $E_{2i+2} = \{3i+2, 3i+3, 3i+4\}$, $\forall i \in [0, n-1]$. *Let* $\mathcal{F} = \bigcup_{i=1}^{m} 2^{E_i}$*, then,*

$$\textit{proj}_{\mathcal{F}}(\overline{\Gamma^*_E}) = \bigotimes_{i=0}^{n-1} \left(\overline{\Gamma^*_{E_{2i+1}}} \bigotimes \overline{\Gamma^*_{E_{2i+2}}} \right) \tag{43}$$

Proof. (proof of **Theorem** 5) $\forall i \in [0, n-1]$, let $B_i = E_{2i+1} \cup E_{2i+2} = \{3i+1, 3i+2, 3i+3, 3i+4\}$, then we must have that for any $i, j \in [0, n-1]$,

$$|B_i \cap B_j| = \begin{cases} 1, & |i-j| = 1 \\ 0, & |i-j| > 1 \end{cases} \tag{44}$$

So from **Lemma**1 we have,

$$\text{proj}_{\cup_{i=0}^{n-1} 2^{B_i}}(\overline{\Gamma^*_E}) = \bigotimes_{i=0}^{n-1} \overline{\Gamma^*_{B_i}} \tag{45}$$

Project both sides of (45) onto $\mathcal{F}$ we have

$$\begin{aligned} &\text{proj}_{\mathcal{F}}(\text{proj}_{\cup_{i=0}^{n-1} 2^{B_i}}(\overline{\Gamma^*_E})) \\ \overset{T_1}{=} &\text{proj}_{\mathcal{F}}(\overline{\Gamma^*_E}) \\ = &\text{proj}_{\mathcal{F}} \left(\bigotimes_{i=0}^{n-1} \overline{\Gamma^*_{B_i}} \right) \\ = &\bigotimes_{i=0}^{n-1} \text{proj}_{2^{E_{2i+1}} \cup 2^{E_{2i+2}}}(\overline{\Gamma^*_{B_i}}) \\ \overset{T_2}{=} &\bigotimes_{i=0}^{n-1} (\overline{\Gamma^*_{E_{2i+1}}} \bigotimes \overline{\Gamma^*_{E_{2i+2}}}) \end{aligned} \tag{46}$$

where T_1 holds because $\mathcal{F} \subseteq \cup_{i=0}^{n-1} 2^{B_i}$, T_2 holds because $\forall i \in [0, n-1]$

$$|E_{2i+1}| = 3 \tag{47}$$

$$|E_{2i+2}| = 3 \tag{48}$$

$$|E_{2i+1} \cap E_{2i+2}| = 2 \tag{49}$$

$$|B_i| = |E_{2i+1} \cup E_{2i+2}| = 4 \tag{50}$$

So from **Theorem** 3 we know that $\text{proj}_{2^{E_{2i+1}}\cup 2^{E_{2i+2}}}(\overline{\Gamma^*_{B_i}}) = \overline{\Gamma^*_{E_{2i+1}}} \bigotimes \overline{\Gamma^*_{E_{2i+2}}}$. ∎

Now we have everything we need to prove **Theorem** 1.

Proof. (proof of **Theorem** 1) From **Theorem** 5 we know that

$$\text{proj}_{\mathcal{F}}(\overline{\Gamma^*_E}) = \bigotimes_{i=0}^{n-1} \left(\overline{\Gamma^*_{E_{2i+1}}} \bigotimes \overline{\Gamma^*_{E_{2i+2}}}\right) \tag{51}$$

given that for each $i \in [0, n-1]$, both E_{2i+1} and E_{2i+2} have cardinality 3, we know that

$$\Gamma_{E_{2i+1}} = \overline{\Gamma^*_{E_{2i+1}}} \tag{52}$$

$$\Gamma_{E_{2i+2}} = \overline{\Gamma^*_{E_{2i+2}}} \tag{53}$$

Plugging the above two equations into (51) we have

$$\text{proj}_{\mathcal{F}}(\overline{\Gamma^*_E}) = \bigotimes_{i=0}^{n-1} \left(\Gamma_{E_{2i+1}} \bigotimes \Gamma_{E_{2i+2}}\right) \tag{54}$$

So either $\Gamma_{E_{2i+1}}$ or $\Gamma_{E_{2i+2}}$ involved in (54) is polyhedral, the **explicit** inequality representation of $\text{proj}_{\mathcal{F}}(\overline{\Gamma^*_E})$ is nothing but the stacking of all inequalities defining $\Gamma_{E_{2i+1}}$ and $\Gamma_{E_{2i+2}}$ together, which gives,

$$\boldsymbol{h}(E_i \backslash \{j\}) \leqslant \boldsymbol{h}(E_i), \forall j \in E_i, \forall i \in [1, m] \tag{55}$$

$$\begin{aligned} \boldsymbol{h}(\{j,k\} \cup A) + \boldsymbol{h}(A) \leqslant \boldsymbol{h}(\{j\} \cup A) + \boldsymbol{h}(\{k\} \cup A), \\ \forall A \subseteq E_i \backslash \{j,k\}, \forall \{j,k\} \subseteq E_i, \forall i \in [1, m] \end{aligned} \tag{56}$$

So the number of these inequalities (including replicas) is,

$$m(3 + \binom{3}{2} 2^{3-2}) = 9m = 6(|E| - 4) + 18 \tag{57}$$

The number of these inequalities after excluding replicas is,

$$\frac{m}{2}\left(2(3 + \binom{3}{2} 2^{3-2}) - 3\right) = \frac{15m}{2} = 5(|E| - 4) + 15 \tag{58}$$

which means that the number of inequalities defining $\text{proj}_{\mathcal{F}}(\overline{\Gamma^*_E})$ grows **linearly** in $|E|$. ∎

IV. Discussion and Future Work

This manuscript proved several instances of systems of subsets $\mathcal{F}$, for which $\text{proj}_{\mathcal{F}}(\overline{\Gamma^*_E})$ is equal to the pasting of a collection of small entropy regions together, including several in which E grows arbitrarily large, and, further, several for which, utilizing the known inequalities for the entropy region on two and three random variables, a polyhedral description of this projection of the entropy region can be provided. Additionally, it is shown that that not only may this set be polyhedral, but also that it may have a number of inequalities that grows only linearly in $|E|$, which provides a stark contrast to the behavior of, for example, the commonly used Shannon outer bound Γ_E which requires a number of inequalities that grows exponentially. A companion manuscript [] reviews how key problems often expressed in terms of $\overline{\Gamma^*_E}$ or its bounds, only depend on it through $\text{proj}_{\mathcal{F}}(\overline{\Gamma^*_E})$, so this difference in description complexity can be highly useful in applications. Current and future work is cataloging further collections of systems of subsets $\mathcal{F}$ enabling this type of decomposition.

References

[1] R. W. Yeung, "A framework for linear information inequalities," *IEEE Transactions on Information Theory*, vol. 43, no. 6, pp. 1924–1934, 1997.

[2] Z. Zhang and R. W. Yeung, "On characterization of entropy function via information inequalities," *IEEE Transactions on Information Theory*, vol. 44, no. 4, pp. 1440–1452, 1998.

[3] Z. Zhang and R. W. Yeung, "A non-shannon-type conditional inequality of information quantities," *IEEE Transactions on Information Theory*, vol. 43, no. 6, pp. 1982–1986, 1997.

[4] R. Dougherty, C. Freiling, and K. Zeger, "Non-shannon information inequalities in four random variables," *arXiv preprint arXiv:1104.3602*, 2011.

[5] K. Makarychev, Y. Makarychev, A. Romashchenko, and N. Vereshchagin, "A new class of non-shannon-type inequalities for entropies," *Communications in Information and Systems*, vol. 2, no. 2, pp. 147–166, 2002.

[6] F. Matus, "Infinitely many information inequalities," in *2007 IEEE International Symposium on Information Theory*, pp. 41–44, 2007.

[7] F. Matus, "Two constructions on limits of entropy functions," *IEEE Transactions on Information Theory*, vol. 53, no. 1, pp. 320–330, 2007.

[8] F. Matúš and L. Csirmaz, "Entropy region and convolution," *IEEE Transactions on Information Theory*, vol. 62, no. 11, pp. 6007–6018, 2016.

[9] T. H. Chan, "Group characterizable entropy functions," in *2007 IEEE International Symposium on Information Theory*, pp. 506–510, June 2007.

[10] K. Zhang and C. Tian, "On the symmetry reduction of information inequalities," *IEEE Transactions on Communications*, vol. 66, no. 6, pp. 2396–2408, 2017.

[11] R. Yeung, *A First Course in Information Theory*. No. v. 1 in A First Course in Information Theory, Springer US, 2002.

[12] R. Dougherty, C. Freiling, and K. Zeger, "Networks, matroids, and non-shannon information inequalities," *IEEE Transactions on Information Theory*, vol. 53, no. 6, pp. 1949–1969, 2007.

[13] T. Kaced, "Equivalence of two proof techniques for non-shannon-type inequalities," in *2013 IEEE International Symposium on Information Theory*, pp. 236–240, IEEE, 2013.

[14] R. W. Yeung, *Information Theory and Network Coding*. Springer Publishing Company, Incorporated, 1 ed., 2008.

[15] V. Strassen *et al.*, "The existence of probability measures with given marginals," *Annals of Mathematical Statistics*, vol. 36, no. 2, pp. 423–439, 1965.

[16] R. Jiroušek, "Solution of the marginal problem and decomposable distributions," *Kybernetika*, vol. 27, no. 5, pp. 403–412, 1991.

[17] T. Fritz and R. Chaves, "Entropic inequalities and marginal problems," *IEEE transactions on information theory*, vol. 59, no. 2, pp. 803–817, 2012.

[18] E. Miranda and M. Zaffalon, "Compatibility, desirability, and the running intersection property," *Artificial Intelligence*, vol. 283, p. 103274, 2020.

[19] C. Beeri, R. Fagin, D. Maier, and M. Yannakakis, "On the desirability of acyclic database schemes," *J. ACM*, vol. 30, p. 479–513, July 1983.

[20] F. Malvestuto, "Existence of extensions and product extensions for discrete probability distributions," *Discrete Mathematics*, vol. 69, no. 1, pp. 61 – 77, 1988.

[21] T. H. Chan and R. W. Yeung, "On a relation between information inequalities and group theory," *IEEE Transactions on Information Theory*, vol. 48, pp. 1992–1995, July 2002.