How to Use Undiscovered Information Inequalities: Direct Applications of the Copy Lemma

Emirhan Gürpınar, Andrei Romashchenko

To cite this version:

HAL Id: lirmm-01990325
https://hal-lirmm.ccsd.cnrs.fr/lirmm-01990325
Submitted on 29 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Abstract—We discuss linear programming techniques that help to deduce corollaries of non classic inequalities for Shannon’s entropy. These applications involve implicitly some (known or unknown) non-classic universal inequalities for Shannon’s entropy, though we do not derive these inequalities explicitly. To reduce the computational complexity of these problems, we extensively use symmetry considerations.

We present two examples of usage of these techniques: we provide a reduced size formal inference of the best known bound for the Ingleton score (originally proven by Dougherty et al. with explicitly derived non Shannon type inequalities), and improve the lower bound for the optimal information ratio of the secret sharing scheme for an access structure based on the Vámos matroid.

I. INTRODUCTION

We can associate with an $n$-tuple of jointly distributed random variables its entropy profile that consists of the $2^n - 1$ values of Shannon’s entropy of each sub-tuple of the given tuple. We say that a point in $\mathbb{R}^{2^n - 1}$ is entropic if it represents the entropy profile of some distribution. The entropic points satisfy different information inequalities — the constraints that characterize the range of admissible entropies for jointly distributed variables. The most known and well understood information inequalities are so-called Shannon-type inequalities that are defined as linear combinations of several instances of the basic inequality $I(A : B; C) \geq 0$, where $A, B, C$ are any subsets (possible empty) of random variables.

In 1998 Z. Zhang and R.W. Yeung discovered the first example of an (unconditional) non Shannon-type information inequality, which was a linear inequality for entropies of a quadruple of random variables that cannot be represented as a combination of basic inequalities. Now many other non Shannon-type information inequalities are known (including several infinite families of non Shannon-type information inequality), see, e.g., [6], [11], [17].

The works on non-Shannon type information inequalities are focused on the following fundamental question: can we prove some particular property for Shannon’s entropy (usually expressed as an equality or an inequality), assuming some specific constraints (which are also given as equalities and inequalities for Shannon’s entropy, including Shannon type and non Shannon type universal information inequalities). Not surprisingly, these arguments typically use the technique of linear programming: at the first stage, to derive new information inequalities (and to prove that they are indeed new) and then, at the second stage, to apply the derived inequalities to some specific problem (e.g., in secret sharing, network coding, and so on). Many applications of this type involve heavy computations: the number of information inequalities in $n$ random variables grows exponentially with $n$. It seems that in several problems the progress is stopped since the computational complexity of the corresponding linear programs is too high for the modern computers. So the question arises: can we reduce the computational complexity of the relevant problems of linear programming? In this paper we show that in some cases new progress can be made with a reasonable combination of previously known techniques.

Recently Farràs et al., [26], observed that the two stages of the scheme mentioned above (deriving new information inequalities and applying them) can be merged in a unique problem of linear programming: in a specific application of information inequalities, instead of enumerating explicitly all known non-Shannon information inequalities, we can include in a linear program the constraints from which those inequalities are (or can be) derived. The constraints used in [26] were instances of the Ahlswede–Körner lemma, [7]. As a tool of inference of non Shannon type inequalities, the Ahlswede–Körner lemma is essentially equivalent to the Copy Lemma from [17] “cloning” one single random variable, see [20]. In this paper we use a technique similar to [20], but employing the constraints obtained from the Copy Lemma. This permits us to extend the technique: we apply the Copy Lemma to make “clones” of pairs of jointly distributed random variables (instead of copies of individual random variables used in [26]).

To reduce the dimension of the corresponding problem of linear programming, we use symmetries of the problems under consideration. Our usage of symmetries is similar to the ideas proposed in [24], where the authors suggested to use symmetries to reduce the complexity of computing the outer bounds on network coding capacity. For a study of the general problem of Shannon and non Shannon type inequalities for symmetric entropy points we refer the reader to [25].

To illustrate the power of this technique, we discuss two examples. In the first one we give a formal inference of the best known bound for the Ingleton score (originally proven in [12]), with only four auxiliary random variable and three applications of the Copy Lemma. The obtained computer-assistant proof is very fast, and the “formally checkable”
version of the proof can be reduced to a combination of 167 equalities and inequalities with rational coefficients. In our second example we improve the known lower bound for the optimal information ratio of the secret sharing schemes associated with the Vámos matroid (for the access structure denoted V₀).

Though the techniques used in this paper are rather simple, their combination proves to be fruitful. We believe that this approach can be used in other problems of information theory.

**Notation:**

We denote \([n] = \{1, \ldots, n\}\). Given a tuple of jointly distributed random variable \((X₁, \ldots, Xₙ)\) and a set of indices \(V = \{i₁, \ldots, iₖ\} \subset [n]\) we denote by \(H(X_V)\) the Shannon entropy \(H(X₁, \ldots, Xₖ)\). We fix an arbitrary order on non-empty subsets of the set of indices \([n]\) and assign to each distribution \(X = (X₁, \ldots, Xₙ)\) its entropy profile, i.e., the vector of \(2^n - 1\) entropies \(H(X_V)\) (for all non-empty \(V \subset [n]\)).

We say that a point in the space \(V_n = \mathbb{R}^{2^n-1}\) is entropic if it represents the entropy profile of some distribution. Following \(\text{[4]}\), we denote by \(\Gamma^n\) the set of all entropic points. We denote by \(\Gamma^n\) the topological close \(\Gamma^n\); its elements are called almost entropic.

We use the standard abbreviations for linear combinations of coordinates in the entropy profile:

\[
\begin{align*}
H(X_V; X_W) &:= H(X_{V \cup W}) - H(X_W) \\
I(X_V; X_W) &:= H(X_V) + H(X_W) - H(X_{V \cup W}) \\
I(X_V; X_W | X_U) &:= H(X_{V \cup W}) + H(X_{U \cup W}) - H(X_{V \cup W | U}) - H(X_U).
\end{align*}
\]

We extend this notation to almost entropic points.

II. **Copy Lemma**

The following lemma was used (somewhat implicitly) in the very first proof of a non Shannon type inequality in \(\text{[4]}\), see also \(\text{[13]}\,\text{Lemma 14.8}\). The general version of this lemma appeared in \(\text{[9]}\) and \(\text{[17]}\). In \(\text{[17]}\) it was presented explicitly, with a complete and detailed proof.

**Lemma 1 (Copy Lemma):** For every tuples of jointly distributed random variables \((X, Y, Z)\) there exists a distribution \((A, B, C, C')\) such that

- the distribution \((A, B, C)\) coincides with the distribution of \((X, Y, Z)\); the distribution of \((A, C')\) coincides with the distribution of \((X, Z)\), and
- \(I(C'; B; C | A) = 0\).

In what follows we abuse the notation and identify \((A, B, C)\) with \((X, Y, Z)\). That is, we say that the initial distribution \((X, Y, Z)\) can be extended\(^1\) to the distribution \((X, Y, Z, Z')\), where \((X, Z)\) and \((X, Z')\) have isomorphic joint distributions, and \(I(Z'; Y; Z|X) = 0\). Following \(\text{[17]}\), we say that \(Z'\) is a \(Y\)-copy of \(Z\) over \(X\), denoted by \(Z' := Y\text{-copy}(Z|X)\).

Note that Lemma\(^1\) remains valid if \(X, Y, Z\) are not individual variables but tuples of variables.

F. Matuš and L. Csirmaz proposed several extensions and generalizations of the Copy Lemma (polymatroid convolution, book extension, maximum entropy extension, see, e.g., \(\text{[12]}\,\text{, [21]}\,\text{, [23]}\)). However, to the best of our knowledge, the “classic” version of the Copy Lemma is strong enough for all known proofs of non Shannon type inequalities.

All known proofs of (unconditional) non Shannon type information inequalities can be presented in the following style: we start with a distribution \((X₁, \ldots, Xₙ)\); using the Copy Lemma we supplement to this distribution several new variable \((Y₁, \ldots, Yₖ)\); then we enumerate the basic Shannon’s inequalities for the joint distribution \((X₁, \ldots, Xₙ, Y₁, \ldots, Yₖ)\) and show that a combination of these inequalities (together with the constraints from the Copy Lemma) gives the desired new inequality for \((X₁, \ldots, Xₙ)\) (the entropy terms with new variable \(Y_j\) should be canceled). We refer the reader to \(\text{[17]}\) for many instances of this argument.

In what follows we use a very similar scheme of proof. The difference is that we superimpose constraints (equalities for the coordinates of the entropy profile of \((X₁, \ldots, Xₙ)\)) that are specific for the given specific problem. That is, in each application we prove a conditional non Shannon type inequality (which is valid only assuming some linear constraints for the entropy quantities). Each of these arguments can be theoretically translated in a more conventional style: we can at first prove some new unconditional non Shannon type inequalities, and then combine them with the given constraints and deduce the desired corollary. However, we do not need to reveal and formulate explicitly these unconditional inequalities hidden behind the proof.

III. **Symmetries**

In this section we discuss a simple formalism of symmetry considerations that helps to reduce the dimension of a linear programs associated with problems concerning information inequalities (assuming that the problem has a particular symmetric structure). A similar formalism (action of permutations on multidimensional distributions and their entropy profiles, information inequalities for distributions with symmetric entropy profiles) was systematically studied in \(\text{[25]}\).

Let \(X^{(n)} = (X₁, \ldots, Xₙ)\) be a joint distribution of \(n\) random variables. We denote by \(\hat{H}(X)\) its entropy profile. Let \(\pi : [n] \rightarrow [n]\) be a permutation of indices (an \(n\)-permutation). This permutation induces a natural transformation of the distribution \(X\):

\[X = (X₁, \ldots, Xₙ) \xrightarrow{\pi} \hat{X} := (X₁^{(\pi(1))}, \ldots, Xₙ^{(\pi(n))})\]

and therefore a transformation of the entropy profile \(\hat{H}(X) \xrightarrow{\pi} \hat{H}(\hat{X})\).

**Example 1:** If \(\pi\) is the transposition of indices 1 and 2, then the mapping defined above exchanges the entropy variables but tuples of variables.

\(^1\)Formally speaking, such an “extension” may require a subdivision of the initial probabilistic space, so technically \((A, B, C)\) and \((X, Y, Z)\) are defined on different probabilistic spaces.
values $H(X_1)$ and $H(X_2)$, the values $H(X_1, X_3)$ and $H(X_2, X_3)$, etc., and does not change the values $H(X_1, X_2)$, $H(X_1, X_2, X_3)$, etc.

Thus, every permutation of indices $\pi$ induces a transformation of the set of all entropic points $\Gamma^*_n$. This transformation of the entropic points can be naturally extended to a linear transformation of the entire space $V_n$ (defined by a suitable permutation of the coordinate axis). In what follows we denote this transformation by $T_\pi$. (Note that $\pi$ is a permutation on $n$ elements, while $T_\pi$ is a transformation of the space of dimension $2^n - 1$.) In the usual algebraic terminology, we can say that we defined a representation of the symmetric group (the group of all $n$-permutations) by linear transformations of the space $V_n$.

We say that a point $x \in V_n$ is $\pi$-invariant if $T_\pi(x) = x$. Similarly, we say that a set of points $S \subseteq V_n$ is $\pi$-invariant, if $T_\pi(x)$ is a permutation of points in $S$. For a subgroup $G$ of $n$-permutations, we say that a point (or a subset) in $V_n$ is $G$-invariant, if it is $\pi$-invariant for every $\pi \in G$.

We denote the set of $G$-invariant points in $V_n$ by $\text{Inv}(G)$. In what follows we typically study linear and affine spaces that are $G$-invariant for some specific group of permutations $G$. Note that every $G$-invariant linear space in $V_n$ contains the set $\text{Inv}(G)$ as a subspace.

The sets of all entropic and all almost entropic points ($\Gamma^*_n$ and $\Gamma^*_n$ respectfully) are obviously $\pi$-invariant for every $n$-permutation $\pi$. For every group $G$ of $n$-permutations, the intersections of $\Gamma^*_n$ or $\Gamma^*_n$ with any $G$-invariant set gives another $G$-invariant set. In what follows we discuss several simple examples of this type.

**Example 2:** Let $L_{ent}$ be the set of points in $\Gamma^*_n$ such that

\[
\begin{align*}
H(X_1|X_2, X_3) &= 0, \\
H(X_1|X_2, X_4) &= 0, \\
H(X_2|X_1, X_3) &= 0, \\
H(X_2|X_1, X_4) &= 0, \\
H(X_1, X_2, X_3, X_4) &= 1.
\end{align*}
\]

Since $L_{ent}$ is defined by affine conditions, it can be represented as the intersection of $\Gamma^*_n$ with an affine subspace $L$ (of co-dimension 5) in the entire space $V_4 = \mathbb{R}^{15}$. It is easy to see that this subspace $L$ is invariant with respect to transpositions $\pi_{12}$ and $\pi_{13}$ ($\pi_{12}$ exchanges the indices 1 and 2, and $\pi_{13}$ exchanges the indices 1 and 3). Therefore, this subspace is $G$-invariant, where $G$ is the subgroup of the symmetric group generated by the transpositions $\pi_{12}$ and $\pi_{34}$ (this subgroup consists of four permutations).

We say that a linear function $L: \mathbb{R}^{2^n-1} \to \mathbb{R}$ is $\pi$-invariant, if $L(T_\pi(x)) = L(x)$ for all $x \in \mathbb{R}^{2^n-1}$. Similarly, for a subgroup of $n$-permutations $G$, we say that a linear function is $G$-invariant, if it is $\pi$-invariant for every $\pi \in G$.

**Example 3:** Ingleton’s quantity $\text{Ing}(X_1, X_2, X_3, X_4)$ defined as

\[I(X_1; X_2|X_3) + I(X_1; X_2|X_4) + I(X_3; X_4) - I(X_1; X_2)\]
can be extended to a linear function on the space $V_4 = \mathbb{R}^{15}$. This function is invariant with respect to the transposition of indices 1 and 2, to the transposition of indices 3 and 4, and with respect to the group (of size four) generated by these two transpositions.

**Claim 1:** Let $G = \{\pi_1, \ldots, \pi_k\}$ be a subgroup of $n$-permutations, and let $x$ be a point in $V_n$. Then the center of mass of the points

\[T_{\pi_1}(x), \ldots, T_{\pi_k}(x)\]

is $G$-invariant.

**Proof:** It is easy to see that the set of points (1) is $\pi$-invariant for each $\pi \in G$. Therefore, the center of mass of this set is also $G$-invariant.

**Lemma 2:** Let $G = \{\pi_1, \ldots, \pi_k\}$ be a subgroup of $n$-permutations, $L$ be a $G$-invariant convex set in $V_n$, and $F: V_n \to \mathbb{R}$ be a $G$-invariant linear function. Then

\[\inf_{x \in L} F(x) = \inf_{y \in L \cap \text{Inv}(G)} F(y)\]

**Proof:** It is enough to prove that for every $x \in L$ there exists a $y \in L \cap \text{Inv}(G)$ such that

\[F(x) = F(y)\]

To this end we take the points

\[T_{\pi_1}(x), \ldots, T_{\pi_k}(x)\]

define $y$ as the center of mass of these $k$ points. Since $L$ is $G$-invariant, each point $T_{\pi_i}(x)$ belongs to $L$. From convexity of $L$ it follows that $y$ also belongs to $L$, and Claim 1 implies that $y$ belongs to $\text{Inv}(G)$. Thus, the constructed point belongs to $L \cap \text{Inv}(G)$.

On the other hand, from $G$-invariance of $F$ it follows that $F(T_{\pi_1}(x)) = \ldots = F(T_{\pi_k}(x)).$ (2)

Since $F$ is linear, we conclude that $F$ at the center of mass of $T_{\pi_i}(x)$ is equal to the same values as the points in (2), and the lemma is proven.

**Example 4:** Assume we are looking for the minimum of $\text{Ing}(X_1, X_2, X_3, X_4)$ (see Example 3) on the set of almost entropic points (points in $\Gamma^*_3$) that satisfy the linear constraints

\[
\begin{align*}
H(X_1|X_2, X_3) &= 0, \\
H(X_1|X_2, X_4) &= 0, \\
H(X_2|X_1, X_3) &= 0, \\
H(X_2|X_1, X_4) &= 0, \\
H(X_1, X_2, X_3, X_4) &= 1.
\end{align*}
\]

(see Example 3). We define $G$ as the group of permutations generated by two transpositions $\pi_{12}$ and $\pi_{34}$ that exchange the indices (1,2) and (3,4) respectively.

We claim that the required extremal value of $\text{Ing}(X_1, X_2, X_3, X_4)$ is achieved at some $G$-invariant point. Indeed, the given constraints define in $V_4$ a $G$-invariant affine subspace. The intersection of this subspace with the convex cone $\Gamma^*_3$ is a $G$-invariant convex set. On the other hand, Ingleton’s quantity $\text{Ing}(X_1, X_2, X_3, X_4)$ is a $G$-invariant linear function. Therefore, we can apply Lemma 2 and conclude that desired extremum is achieved at some point in $\text{Inv}(G)$. 


IV. THE STANDARD BENCHMARK: BOUNDS FOR INGLETON’S QUANTITY

It is widely believed that one of the most interesting questions on the structure of \( \Gamma^* \) is the question on the Ingleton’s quantity: what is the worst possible violation of Ingleton’s inequality

\[
\text{Ing}(A, B, C, D) \geq 0. \tag{3}
\]

This inequality is true for \textit{linearly representable} entropic points but for some other entropic points it is violated (see \textbf{[3]}). The question is how far below 0 can go the Ingleton score

\[
\text{Ing}(A, B, C, D) \geq 0, \text{ (here we use the terminology from [17])}
\]

From the Shannon type inequalities it follows only that this score is greater than \(-1/4\). The first non Shannon type inequality from \textbf{[4]} implies that this score is greater than \(-1/6\). The best bound was discovered by Dougherty \textit{et al.}

\textit{Theorem 1 (originally proven in [17]):} For every quadruple of jointly distributed random variables \((A, B, C, D)\)

\[
\text{Ing}(A, B, C, D) \geq -3/19.
\]

We show that this bound can be derived with only three applications of the Copy Lemma (with four new variables), with the help of symmetry considerations from the previous section. This approach gives a pretty simple way to confirm the ratio with the help of symmetry considerations from the previous applications of the Copy Lemma (with four new variables), \textbf{[4, 1]}. The first non Shannon type inequality from \textbf{[4]} implies that this score is greater than \(-1/6\). The best bound was discovered by Dougherty \textit{et al.}

\[\text{Ing}(A, B, C, D) \geq -3/19.\]

\textbf{Sketch of the proof:} Our goal it to find the minimal value of the Ingleton score \textbf{[3]} for entropic points. It is equivalent to the minimal value of \(\text{Ing}(A, B, X, Y)\) for almost entropic points satisfying the normalization condition \(H(A, B, C, D) = 1\).

The objective function \(\text{Ing}(A, B, C, D)\) and the normalization constraint are invariant with respect to the transpositions \(\pi_{A,B}\) and \(\pi_{C,D}\) that exchange the variables \(A, B\) and \(C, D\) respectively. Therefore, we can apply Lemma \textbf{[2]} and conclude that the required minimal value can be found in the space of points that are \(\pi_{A,B}\)- and \(\pi_{C,D}\)-invariant.

The invariance means that we can restrict ourselves to the points in \(\Gamma^*_4\) such that

\[
\begin{align*}
H(A) &= H(B), \quad H(A, C, D) = H(B, C, D), \\
H(A, C) &= H(B, C), \quad H(A, D) = H(B, D), \\
H(C) &= H(D), \quad H(A, B, C) = H(A, B, D), \\
H(A, C) &= H(A, D), \quad H(B, C) = H(B, D).
\end{align*}
\tag{4}
\]

The crucial part of the proof is, of course, the Copy Lemma. We use the trick \textbf{[4]} from \textbf{[17]} proof of Ineq. 33, p. 19 [and define

\[
\{ (R,S) : = \emptyset\text{-copy}(B,D|A,C), \\
T := (D,R)\text{-copy}(C|A,B,S), \\
\tag{5}
\]

Technically, this means that we take the conditions (equalities) from the Copy Lemma (Lemma \textbf{[1]} that determine the entropies of the introduced variables and the properties of conditional independence for each instance of the Copy Lemma.

At last, we add to the normalization condition \(H(A, B, C, D) = 1\) and the constraints \textbf{[4]} and \textbf{[5]} the Shannon type inequalities for \((A, B, C, D, R, S, T, U)\). Thus, we obtain a linear program for the \(2^8 - 1 = 255\) variables (corresponding to the coordinates of the entropy profile of \((A, B, C, D, R, S, T, U)\)) with the objective function

\[
\text{Ing}(A, B, C, D) \rightarrow \min
\]

This problem is easily for the standard linear programming solvers (we made experiments with the solvers \textbf{[27]–[29]}). We can extract from the dual solution of this linear program a rational combination of inequalities that gives the desired bound (without floating point arithmetics nor rounding). We present this “formally checkable” proof in Appendix.

V. SECRET SHARING SCHEMES ON THE VÁMOS MATROID

The notion of a \textit{secret sharing scheme} was introduced by Shamir \textbf{[2]} and Blakley \textbf{[1]}. Nowadays, secret sharing is an important component of many cryptographic protocols. We refer the reader to \textbf{[13]} for an excellent survey of secret sharing and its applications.

The aim of \textit{perfect secret sharing} is to distribute a \textit{secret value} among a set of parties in such a way that only the \textit{qualified sets} of parties can recover the secret value; the non-qualified sets should have no information about the secret. The family of qualified sets of parties is called the access structure of the scheme. This family of sets must be monotone (a superset of a qualified set is also qualified).

The \textit{information ratio} of a secret sharing schemes is the maximal ratio between the size (the entropy) of a share and the size (the entropy) of the secret. The optimization of this parameter is the central problem of secret sharing. This question was extensively studied for several families of access structures. The main technique for proving lower bounds on complexity of secret sharing schemes is nowadays the linear programming applied to the constraints formulated in terms of classic and non-classic information inequalities, see \textbf{[19]}.

In what follows we apply our technique (the Copy Lemma combined with symmetry considerations and linear programming) to one of the classic access structures — the access structure defined on the Vámos matroid (more technically, one of two access structures that can be defined on the Vámos matroid). We skip the general motivation (the Vámos matroid is particularly interesting as a simple example of a non-representable matroid, see \textbf{[13]} for details) and formulate directly the problem of the information ratio for this access
structure in terms of Shannon’s entropy of the involved random variables.

We define an access structure with parties \( \{1, \ldots, 7\} \). The minimal qualified sets are the 3-sets \( \{1, 2, 3\} \), \( \{1, 4, 5\} \) and all 4-sets not containing them, with three exceptions \( \{2, 3, 4, 5\}, \{2, 3, 6, 7\}, \{4, 5, 4, 7\} \). We deal with jointly distributed random variables \( (S_0, \ldots, S_7) \), where \( S_0 \) is the secret and \( S_1, \ldots, S_7 \) are the shares given to the seven parties involved in the scheme. We require that

\( i \) every minimal qualified set \( V \) listed above
\( H(S_0|S_V) = 0; \)
\( ii \) if \( V \subseteq \{1, \ldots, 7\} \) includes no minimal qualified set, then

\( H(S_0|S_V) = H(S_0). \)

Under these constraints, we are looking for the extremal value

\[
\max\{H(S_1), \ldots, H(S_7)\} \rightarrow \inf
\]

Since the set of almost-entropic points is a closed convex cone, we can add the normalization \( H(S_0) = 1 \) and rewrite the objective function as

\[
\max\{H(S_1), \ldots, H(S_7)\} \rightarrow \inf
\]

It is known that the required minimum is not greater than \( 4/3, \)

The history of improvements of the lower bound for this problem is presented in the table:

<table>
<thead>
<tr>
<th>year</th>
<th>bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td>( &gt; 1 )</td>
</tr>
<tr>
<td>2006</td>
<td>( \geq 1 + \Omega(1/\sqrt{k}) ) for a secret of size ( k )</td>
</tr>
<tr>
<td>2008</td>
<td>( \geq 11/10 )</td>
</tr>
<tr>
<td>2011</td>
<td>( \geq 9/8 = 1.125 )</td>
</tr>
<tr>
<td>2013</td>
<td>( \geq 67/59 \approx 1.135593 )</td>
</tr>
<tr>
<td>2018</td>
<td>( \geq 33/29 \approx 1.137931 )</td>
</tr>
<tr>
<td>this paper</td>
<td>( \geq 561/491 \approx 1.142566 )</td>
</tr>
</tbody>
</table>

Notice that the objective function, the normalization condition \( H(S_0) = 1 \), and the linear constraints in \( i \) and \( ii \) above are invariant with respect to the permutation that

\( • \) swap the indices 2 and 3,

\( • \) swap the indices 4 and 5,

\( • \) swap the indices 6 and 7,

\( • \) swap the pairs \( (2, 3) \) and \( (4, 5) \).

Therefore we can apply Lemma \( \ref{lemma:copy-lemma} \) and reduce the problem to those almost entropic points that satisfy the symmetries defined above.

We denote

\( V := (S_0, S_1) \) and \( W := (S_6, S_7) \)

define four new variables by applying twice the Copy Lemma:

let \( (V', W') \) be an \( (S_0, S_1, S_6, S_7) \)-copy of \( (V, W) \)
over \( (S_2, S_3, S_4, S_5) \);

let \( (V'', W'') \) be an \( (S_0, S_1, S_6, S_7, V', W') \)-copy of \( (V, W) \)
over \( (S_2, S_3, S_4, S_5) \).

**Remark 1:** Each instance of the Copy Lemma includes the property of conditional independence. In the construction described above we have two applications of the Copy Lemma, and therefore two independence conditions:

\[
I(V', W'; S_0, S_1, S_6, S_7|S_2, S_3, S_4, S_5) = 0
\]

\[
I(V'', W''; S_0, S_1, S_6, S_7, V', W'|S_2, S_3, S_4, S_5) = 0.
\]

These two conditions can be merged in one (more symmetric) constraint

\[
H(S_0, S_1, S_6, S_7, V', W', V''|S_2, S_3, S_4, S_5) =
H(S_0, S_1, S_6, S_7|S_2, S_3, S_4, S_5)
+ H(V', W'|S_2, S_3, S_4, S_5)
+ H(V'', W''|S_2, S_3, S_4, S_5).
\]

Thus, we obtain a linear program with the following constraints:

- the conditions \( i \) and \( ii \) that define the access structure of the secret sharing scheme,
- the normalization \( H(S_0) = 1 \),
- the equalities that follow from the symmetry of the access structure,
- the conditions of the two instances of the Copy Lemma,
- and the Shannon type inequalities for twelve random variables \( (S_0, \ldots, S_7, V', V'', W', W'') \) (i.e., for \( 212 - 1 \) coordinates of their entropy profile).

The goal is to optimize the objective function \( \ref{eq:optimized-objective} \).

**Remark 2:** We do not need to include in the linear program the variables \( V \) and \( W \) and their entropies — these two variables are used only as a notation in the Copy Lemma. With a help of a computer (we used the solvers \( \ref{solver:linear-programming-solvers} \)–\( \ref{solver:linear-programming-solvers} \)) we find the optimal solution of this linear program: it is equal to \( 561/491 \). Thus we establish the following statement.

**Theorem 2:** The optimal information rate of the secret sharing scheme for the access structure \( V_0 \) on the Vámos matroid is not less than \( 561/491 = 1.142566 \ldots \)

**Remark 3:** The standard linear programming solvers use the floating point arithmetic, and we should keep in mind the rounding errors. Our linear program contains only constraints with integer coefficients, so the values of the variables for the optimal solution of this linear program (the primal and the dual) can be taken rational. These rational values can be found by exact computing, without rounding. To compute the exact solution of a linear program we use the rational linear programming solver \( \text{QSopt-ex} \), see \( \ref{solver:exact-linear-programming-solver} \), \( \ref{solver:exact-linear-programming-solver} \).

The found exact dual solutions (rational linear combinations of constraints) for this linear program consist of more than \( 10^3 \) equalities and inequalities. The found rational linear combinations of equalities and inequalities provide a formal mathematical proof of Theorem 2, though such a proof hardly can be called “human-readable.”

Computing the exact solution of a linear program is rather time consuming. If we only need a lower bound for the optimal solution, it is enough to find a feasible solution of the dual problem that approximates (well enough) the optimal solution. Such an approximation can be found much faster than the exact rational solution, with more conventional floating point LP solvers.
Remak 4: The linear program constructed in this section would have the same optimal value if we omit the symmetry conditions (note that, unlike the previous section, in the proof of Theorem 2 the Copy Lemma is applied in symmetric settings). However, these symmetry conditions are not useless. In our experiments with exact calculations, the symmetry constraints help to find a rational solution of the dual problem with smaller denominators. Also, in several experiments with floating point solvers, the symmetry constraints make the computation (or approximation) of the optimal solution faster. However, this is not a universal rule: in some experiments the symmetry constraints make the computation only slower — it depends on the used linear programming algorithm and on the chosen scaling of the objective function and of the constraints.

VI. Conclusion

When we derive balanced information inequalities, the technique of the Copy Lemma with “clones” of individual random variables is equivalent to the technique of the Ahlswede–Körner lemma, see [20]. In this paper we used the Copy Lemma in a stronger form, making “clones” of pairs of correlated variables. In the proofs of Theorem 1 and Theorem 2 it was crucial that we can “duplicate” in one shot a pair of correlated random variables, not only a single random variable.

The Ahlswede–Körner lemma used in [26] has a clear intuitive meaning: it can be interpreted in terms of extraction of the common information, see [7] and [6]. It would be interesting to find a structural property of the probability distribution that would give an intuitive explanation of the efficiency of the Copy Lemma with two (or more) copied variables.

References

[27] QSopt-ex: an exact linear programming solver. The authors of the solver are David Applelegate, William Cook, Sanjeeb Dash, and Daniel Espinoza. The code is distributed under GNU GPL v2.1. Website of the project: https://www.math.uwaterloo.ca/~bico/qsopt/ex/ A fork of QSopt-ex on GitHub: https://github.com/jonies/qsopt-ex
[28] GNU Linear Programming Kit, developed by Andrew O. Makhorin. Website of the project: www.gnu.org/software/glpk/
APPENDIX

In what follows we present a list of 167 equalities and inequalities. An integer linear combination of these inequalities implies Theorem I.

Definition of the standard information quantities:

\[ H(A) = \sum_{i} -p_i \log p_i \]

with a factor of 2586.

\[ H(A; B) = \sum_{i} \sum_{j} p_{ij} \log \left( \frac{1}{p_{ij}} \right) \]

with a factor of 948.

\[ H(A; B; C) = \sum_{i} \sum_{j} \sum_{k} p_{ijk} \log \left( \frac{1}{p_{ijk}} \right) \]

with a factor of 2376.

\[ H(A; B; C; D) = \sum_{i} \sum_{j} \sum_{k} \sum_{l} p_{ijkl} \log \left( \frac{1}{p_{ijkl}} \right) \]

with a factor of 3239.

\[ H(A; B; C; D; R; S; T; U) = \sum_{i} \sum_{j} \sum_{k} \sum_{l} \sum_{m} \sum_{n} \sum_{o} p_{ijklmnop} \log \left( \frac{1}{p_{ijklmnop}} \right) \]

with a factor of 1337.

\[ H(A; B; C; D; R; S; T; U) - H(A; B; C; R; S; U) = 0 \]

Shannon type inequalities:

\[ H(A; B; C; D; R; S; T; U) - H(A; B; C; R; S; U) \geq 0 \]

with a factor of 111.

\[ H(A; B; C; D; S; T; U) - H(B; C; D; S; U) \geq 0 \]

with a factor of 549.

\[ H(A; B; C; D; S; R; U) - H(A; B; C; D; S; T) \geq 0 \]

with a factor of 105.

\[ H(A; B; C; D; S; R; T) - H(A; B; C; D; S; T) \geq 0 \]

with a factor of 136.

\[ H(A; B; C; D; S; T; U) - H(A; B; C; D; S; T) \geq 0 \]

with a factor of 849.

\[ H(A; B; C; D; S; T; U) - H(A; B; C; D; S; T) \geq 0 \]

with a factor of 337.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 270.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 208.

\[ H(A; B; C; D; S; T; U) - H(A; B; C; D; S; T) \geq 0 \]

with a factor of 341.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 107.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 191.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 448.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 948.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 107.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 191.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 448.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 948.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 107.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 191.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 448.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 948.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 107.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 191.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 448.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 948.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 107.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 191.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 448.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 948.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 107.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 191.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 448.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 948.

\[ H(A; B; C; D; S; T; U) - H(A; C; D; R; S; U) \geq 0 \]

with a factor of 107.