Proving Information Inequalities by Gaussian Elimination

Laigang Guo
School of Mathematical Sciences
Beijing Normal University
Beijing, China.
Email: lgguo@bnu.edu.cn

Raymond W. Yeung
Institute of Network Coding
The Chinese University of Hong Kong
Hong Kong, China.
Email: whyeung@ie.cuhk.edu.hk

Xiao-Shan Gao
Key Laboratory of Mathematics Mechanization
Chinese Academy of Sciences
Beijing, China.
Email: xgao@mmrc.iss.ac.cn

2024 IEEE International Symposium on Information Theory (ISIT) | 979-8-3503-8284-6/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISIT57864.2024.10619371

***Abstract*—The proof of information inequalities under linear constraints on the information measures is an important problem in information theory. For this purpose, ITIP and other variant algorithms have been developed and implemented, which are all based on solving a linear program (LP). Building on our recent work [13], we develop in this paper an enhanced approach for solving this problem.**

I. Introduction

A framework for linear information inequalities was introduced in [1]. Based on this framework, the problem of verifying Shannon-type inequalities can be formulated as a linear program (LP), and a software package based on MATLAB called Information Theoretic Inequality Prover (ITIP) was developed [2]. Subsequently, different variations of ITIP have been developed [3]–[7].

The LP-based approach is in general not computationally efficient because it does not take advantage of the special structure of the underlying LP. To tackle this issue, we developed in [13] a set of algorithms that can be implemented by symbolic computation. Based on these algorithms, we devised procedures for reducing the original LP to the minimal size, which can be solved easily. These procedures are computationally more efficient than solving the original LP directly. In this paper, we develop a different symbolic approach which not only make the reduction from the original LP to the minimal size more efficient, but also in many cases can prove the information inequality without solving any LP. The reader is referred to [8, Chs. 13-15] for the background and to [14] for the proofs omitted here.

II. Information inequality preliminaries

Throughout this paper, all random variables are discrete. Unless otherwise specified, all information expressions involve some or all of the random variables $X_1, X_2, \ldots, X_n$. The value of n will be specified when necessary. Denote the set $\{1, 2, \ldots, n\}$ by $\mathcal{N}_n$, the set $\{0, 1, 2, \ldots\}$ by $\mathbb{N}_{\geq 0}$ and the set $\{1, 2, \ldots\}$ by $\mathbb{N}_{>0}$.

Theorem II.1. [1] *Any Shannon's information measure can be expressed as a conic combination of the following two elemental forms of Shannon's information measures:*

i) $H(X_i|X_{\mathcal{N}_n-\{i\}})$

ii) $I(X_i; X_j|X_K)$*, where* $i \neq j$ *and* $K \subseteq \mathcal{N}_n - \{i, j\}$.

The nonnegativity of the two elemental forms of Shannon's information measures forms a proper but equivalent subset of the set of basic inequalities. The inequalities in this smaller set are called the *elemental inequalities*. In [1], the minimality of the elemental inequalities is also proved. The total number of elemental inequalities is equal to $u \triangleq n + \binom{n}{2}2^{n-1}$.

Shannon's information measures, with conditional mutual information being the general form, can be expressed as a linear combination of joint entropies. For the random variables $X_1, X_2, \ldots, X_n$, there are a total of $2^n - 1$ joint entropies. By regarding the joint entropies as variables, the basic (elemental) inequalities become linear inequality constraints in $\mathbb{R}^{2^n-1}$. By the same token, the linear equality constraints on Shannon's information measures imposed by the problem under discussion become linear equality constraints in $\mathbb{R}^{2^n-1}$. This way, the problem of verifying a (linear) Shannon-type inequality can be formulated as a linear program (LP), which is described next.

Let $\mathbf{h}$ be the column (2^n-1)-vector of the joint entropies of $X_1, X_2, \ldots, X_n$. The set of elemental inequalities can be written as $\mathbf{Gh} \geq 0$, where $\mathbf{G}$ is an $u \times (2^n - 1)$ matrix and $\mathbf{Gh} \geq 0$ means all the components of $\mathbf{Gh}$ are nonnegative. Likewise, the constraints on the joint entropies can be written as $\mathbf{Qh} = 0$. When there is no constraint on the joint entropies, $\mathbf{Q}$ is assumed to contain zero rows. The following theorem enables a Shannon-type inequality to be verified by solving an LP.

Theorem II.2. [1] $\mathbf{b}^\top\mathbf{h} \geq 0$ *is a Shannon-type inequality under the constraint* $\mathbf{Qh} = 0$ *if and only if the minimum of the problem*

$$\textit{Minimize } \mathbf{b}^\top\mathbf{h} \textit{, subject to } \mathbf{Gh} \geq 0 \textit{ and } \mathbf{Qh} = 0$$

is zero.

III. Algorithms for homogeneous linear inequlities

In this section, we will develop new algorithms for proving information inequalities. For details, one can refer to [14].

Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$, and let $\mathbb{R}_h[\mathbf{x}]$ be the set of all homogeneous linear polynomials in $\mathbf{x}$ with real coefficients. In this paper, unless otherwise specified, we assume that all polynomials are linear and homogeneous, all inequality sets have the form $S_f = \{f_i \geq 0, i \in \mathcal{N}_m\}$, with $f_i \not\equiv 0$ and $f_i \in \mathbb{R}_h[\mathbf{x}]$, and all equality sets have the form $E_{\tilde{f}} = \{\tilde{f}_i = 0, i \in \mathcal{N}_{\tilde{m}}\}$ with $\tilde{f}_i \not\equiv 0$ and $\tilde{f}_i \in \mathbb{R}_h[\mathbf{x}]$.

For a given set of polynomials $P_f = \{f_i, i \in \mathcal{N}_m\}$ and the corresponding set of inequalities $S_f = \{f_i \geq 0, i \in \mathcal{N}_m\}$, and a given set of polynomials $P_{\tilde{f}} = \{\tilde{f}_i, i \in \mathcal{N}_{\tilde{m}}\}$ and the corresponding set of equalities $E_{\tilde{f}} = \{\tilde{f}_i = 0, i \in \mathcal{N}_{\tilde{m}}\}$, where f_i and $\tilde{f}_i$ are polynomials in $\mathbf{x}$, we write $S_f = \mathcal{R}(P_f)$, $P_f = \mathcal{R}^{-1}(S_f)$, $E_{\tilde{f}} = \widetilde{\mathcal{R}}(P_{\tilde{f}})$ and $P_{\tilde{f}} = \widetilde{\mathcal{R}}^{-1}(E_{\tilde{f}})$.

Definition III.1. *Let $S_f = \{f_i \geq 0, i \in \mathcal{N}_m\}$ and $S_{f'} = \{f'_i \geq 0, i \in \mathcal{N}_{m'}\}$ be two inequality sets, and $E_{\tilde{f}}$ and $E_{\tilde{f}'}$ be two equality sets. We write $S_{f'} \subseteq S_f$ if $\mathcal{R}^{-1}(S_{f'}) \subseteq \mathcal{R}^{-1}(S_f)$, and $E_{\tilde{f}'} \subseteq E_{\tilde{f}}$ if $\widetilde{\mathcal{R}}^{-1}(E_{\tilde{f}'}) \subseteq \widetilde{\mathcal{R}}^{-1}(E_{\tilde{f}})$. Furthermore, we write $(f_i \geq 0) \in S_f$ to mean that the inequality $f_i \geq 0$ is in S_f.*

Definition III.2. *Let $\mathbb{R}_{>0}$ and $\mathbb{R}_{\geq 0}$ be the sets of positive and nonnegative real numbers, respectively. A linear polynomial F in $\mathbf{x}$ is called a positive (nonnegative) linear combination of polynomials f_j in $\mathbf{x}$, $j = 1, \ldots, m$, if $F = \sum_{j=1}^m r_j f_j$ with $r_j \in \mathbb{R}_{>0}$ ($r_j \in \mathbb{R}_{\geq 0}$). A nonnegative linear combination is also called a conic combination.*

Definition III.3. *The inequalities $f_1 \geq 0, f_2 \geq 0, \ldots, f_m \geq 0$ imply the inequality $f \geq 0$ if the following holds:*

For all $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x}$ satisfying $f_1 \geq 0, f_2 \geq 0, \ldots, f_m \geq 0$ implies $\mathbf{x}$ satisfies $f \geq 0$.

Definition III.4. *Given a set of inequalities $S_f = \{f_i \geq 0, i \in \mathcal{N}_m\}$, for $i \in \mathcal{N}_m$, $f_i \geq 0$ is called a redundant inequality if $f_i \geq 0$ is implied by the inequalities $f_j \geq 0$, where $j \in \mathcal{N}_m \backslash \{i\}$.*

Definition III.5. *Let $S_f = \{f_i(\mathbf{x}) \geq 0, i \in \mathcal{N}_m\}$ be an inequality set. If $f_k(\mathbf{x}) = 0$ for all solutions $\mathbf{x}$ of S_f, then $f_k(\mathbf{x}) = 0$ is called an implied equality of S_f. The inequality set S_f is called a pure inequality set if S_f has no implied equalities.*

Lemma III.1. [13] *Let $S_f = \{f_i(\mathbf{x}) \geq 0, i \in \mathcal{N}_m\}$ be an inequality set. Then $f_k = 0$ is an implied equality of S_f if and only if*

$$f_k(\mathbf{x}) \equiv \sum_{i=1, i\neq k}^{m} p_i f_i(\mathbf{x}), \tag{1}$$

where $p_i \leq 0$ for all $i \in \mathcal{N}_m \backslash \{k\}$.

Lemma III.2. [12] *Given $h_1, \ldots, h_m, h_0 \in \mathbb{R}_h[\mathbf{x}]$, $h_1 \geq 0, ..., h_m \geq 0$ imply $h_0 \geq 0$ if and only if h_0 is a conic combination of $h_1, \ldots, h_m$.*

The following proposition is well known (see for example [9, Chapter 1]).

Proposition III.1. *Under the variable order $x_1 \succ x_2 \succ \cdots \succ x_n$, the linear equation system $E_{\tilde{f}} = \{\tilde{f}_i = 0,\ i \in \mathcal{N}_{\widetilde{m}}\}$ can be reduced by the Gauss-Jordan elimination to the unique form*

$$\widetilde{E} = \{x_{k_i} - U_i = 0, i \in \mathcal{N}_{\widetilde{n}}\}, \tag{2}$$

where $\widetilde{n}$ is the rank of the linear system $E_{\tilde{f}}$, $k_1 < k_2 < \cdots < k_{\widetilde{n}}$, x_{k_i} is the leading term of $x_{k_i} - U_i$, and U_i is a linear function in $\{x_j,\ \text{for } k_i < j \leq n,\ \ j \neq k_l,\ i < l \leq \tilde{n}\}$.

Among $x_1, x_2, \ldots, x_n$, the variable x_{k_i}, $i \in \mathcal{N}_{\tilde{n}}$ are called the pivot variables, and the rest are called the free variables.

We call the equality set $\widetilde{E}$ the reduced row echelon form (RREF) of $E_{\tilde{f}}$. Likewise, we call the polynomial set $\widetilde{\mathcal{R}}^{-1}(\widetilde{E})$ the RREF of $\widetilde{\mathcal{R}}^{-1}(E_{\tilde{f}})$. We say applying the Gauss-Jordan elimination to $\widetilde{\mathcal{R}}^{-1}(E_{\tilde{f}})$ to mean finding $\widetilde{\mathcal{R}}^{-1}(\widetilde{E})$ by Proposition III.1.

Algorithm 1 Dimension reduction

Input: S_f, $E_{\tilde{f}}$.
Output: The remainder set R_f.
1: Compute $\widetilde{E}$ for $E_{\tilde{f}}$ by Proposition III.1.
2: Substitute x_{k_i} by U_i in $\mathcal{R}^{-1}(S_f)$ to obtain the set R.
3: Let $R_f = R \backslash \{0\}$.
4: **return** $\mathcal{R}(R_f)$.

We say reducing S_f by $E_{\tilde{f}}$ to mean using Algorithm 1 to find $\mathcal{R}(R_f)$. We also say reducing P_f by $E_{\tilde{f}}$ to mean using Algorithm 1 to find R_f, called *the remainder set* (or remainder if R_f is a singleton).

Definition III.6. *Let $f \in \mathbb{R}_h[\mathbf{x}]$ and $x_1 \succ x_2 \succ \cdots x_n$ be a fixed variable order. The variable set of f, denoted by $V(f)$, is the set containing all the variables of f. The variable sequence of f, denoted by $\mathcal{V}(f)$, is the sequence containing all the variables of f in the given order. The coefficient sequence of f, denoted by $\mathcal{C}(f)$, is the sequence containing the coefficients corresponding to the variables in $\mathcal{V}(f)$. We adopt the convention that $\mathcal{C}(f) = [0]$ and $V(f) = \emptyset$ for $f \equiv 0$.*

Definition III.7. *Let $P_f = \{f_i,\ i \in \mathcal{N}_m\}$, where $f_i \in \mathbb{R}_h[\mathbf{x}]$. The variable set of P_f, denoted by $V(P_f)$, is the set containing all the variables of f_i's, i.e., $V(P_f) = \cup_{i \in \mathcal{N}_m} V(f_i)$.*

Example III.1. *Let $P_f = \{f_1, f_2\}$, where $f_1 = x_1 + x_2$, $f_2 = x_1 - x_3$. Then, we have*

$$V(f_1) = \{x_1, x_2\},\ \mathcal{V}(f_1) = [x_1, x_2],\ \mathcal{C}(f_1) = [1, 1],$$
$$V(f_2) = \{x_1, x_3\},\ \text{and } V(P_f) = \{x_1, x_2, x_3\}.$$

Observe that for any polynomial $f(\mathbf{x})$, the following equality holds:

$$\{\mathbf{x} : f(\mathbf{x}) \geq 0\} = \mathrm{Proj}_{\mathbf{x}}\{(\mathbf{x}, a) : f(\mathbf{x}) - a = 0, a \geq 0\}.$$

Note that on the RHS, a new variable a is introduced. Motivated by this observation, in the sequel we will say that an inequality $f(\mathbf{x}) \geq 0$ is equivalent to the semi-algebraic set $\{f(\mathbf{x}) - a = 0,\ a \geq 0\}$. Also, $\{f_i(\mathbf{x}) \geq 0, i \in \mathcal{N}_m\}$ is equivalent to $\{f_i(\mathbf{x}) - a_i = 0,\ a_i \geq 0, i \in \mathcal{N}_m\}$.

Definition III.8. *Let $H = \{h_i, i \in \mathcal{N}_m\}$ be a set of polynomials, where $h_i \in \mathbb{R}_h[\mathbf{b}]$ and $\mathbf{b} = (x_1, \ldots, x_n, a_1, \ldots, a_m)^T$. Under the variable order $x_1 \succ \cdots \succ x_n \succ a_1 \succ \cdots \succ a_m$, we can obtain the RREF of H, denoted by $\widetilde{H}$. Let $\widetilde{H} = H_1 \cup H_2$, where*

$V(h) \cap \{x_1, x_2, \ldots, x_n\} \neq \emptyset$ for every $h \in H_1$, and
$V(h) \cap \{x_1, x_2, \ldots, x_n\} = \emptyset$ and
$V(h) \cap \{a_1, a_2, \ldots, a_m\} \neq \emptyset$ for every $h \in H_2$.

H_1 is called the partial RREF of H in $\mathbf{x}$ and $\mathbf{a}$, and H_2 is called the partial RREF of H in $\mathbf{a}$.

Let $F_0 \in \mathbb{R}_h[\mathbf{x}]$ and $S_f = \{f_i \geq 0, i \in \mathcal{N}_m\}$, where $f_i \in \mathbb{R}_h[\mathbf{x}]$. In the rest of this section, we discuss how to solve the following problem.

Problem III.1. *Prove $F_0 \geq 0$ subject to S_f.*

We first give a method implemented by the following algorithm for reducing Problem III.1 to another LP.

Algorithm 2 LP reduction

Input: Problem III.1
Output: A reduced LP.
1: Let $G_i = f_i - a_i$, $i \in \mathcal{N}_m$, where a_i's are assumed to satisfy $a_i \geq 0,\ i \in \mathcal{N}_m$.
2: Fix the variable order $x_1 \succ x_2 \succ \cdots \succ x_n \succ a_1 \succ \cdots \succ a_m$.
3: Apply the Gauss-Jordan elimination to $\{G_i, i \in \mathcal{N}_m\}$ and obtain the RREF.
4: Let J_0 be the partial RREF of $\{G_i, i \in \mathcal{N}_m\}$ in $\mathbf{x}$ and $\mathbf{a}$, and J_1 be the partial RREF of $\{G_i, i \in \mathcal{N}_m\}$ in $\mathbf{a}$.
5: Reduce F_0 by J_0 to obtain F.
6: The Problem III.1 is equivalent to
Problem III.2. *Prove $F \geq 0$ subject to $\widetilde{\mathcal{R}}(J_1)$ and $a_i \geq 0,\ i \in \mathcal{N}_m$.*
7: **return** Problem III.2.

Remark III.1. *In Algorithm 2, if Problem III.1 can be solved, then F needs to satisfy $V(F) \cap \{x_1, \ldots, x_n\} = \emptyset$. If there exist $x_i \in V(F)$, then x_i is a free variable in Problem III.2, and Problem III.2 cannot be solved. Thus Problem III.1 cannot be solved. For example, we consider the problem*

***P1**: Prove $x_1 + x_3 \geq 0$ subject to $x_1 \geq 0$ and $x_2 \geq 0$.*

Running Algorithm 2, the above problem becomes

***P2**: Prove $a_1 + x_3 \geq 0$ subject to $a_1 \geq 0$.*

*Obviously, **P2** cannot be proved since x_3 is a free variable.*

Let $\boldsymbol{a} = (a_1, \ldots, a_m)^T$, $F \in \mathbb{R}_h[\boldsymbol{a}]$, $f_i \in \mathbb{R}_h[\boldsymbol{a}]$ for $i \in \mathcal{N}_{\widetilde{m}}$, $S_a = \{a_i \geq 0,\ i \in \mathcal{N}_m\}$, and $E_a = \{f_i = 0,\ i \in \mathcal{N}_{\widetilde{m}}\}$. Based on the discussion above, we only need to consider the case that F satisfies $V(F) \cap \{x_1, \ldots, x_n\} = \emptyset$.

To facilitate the discussion, we restate Problem III.2 in a general form:

Problem III.3. *Prove $F \geq 0$ subject to E_a and S_a.*

We say that a problem as given in Problem III.3 is "solvable" if $F \geq 0$ is implied by E_a and S_a.

Definition III.9. *Let $E_{\widetilde{f}} = \{\widetilde{f}_i = 0, i \in \mathcal{N}_{\widetilde{m}}\}$ and $E_{f'} = \{f'_i = 0, i \in \mathcal{N}_{m'}\}$ be two equality sets, where $\tilde{f}_i,\ f'_i \in \mathbb{R}_h[\mathbf{x}]$. If the solution sets of $E_{f'}$ and $E_{\widetilde{f}}$ are the same, then we say that $E_{\widetilde{f}}$ and $E_{f'}$ are equivalent.*[1]

Definition III.10. *Let $h_i \in \mathbb{R}_h[\boldsymbol{a}]$, $i = 1, 2$, where $\mathbf{a} = (a_1, \ldots, a_m)^T$ and let $E_{\tilde{f}} = \{\tilde{f}_i = 0, i \in \mathcal{N}_{\widetilde{m}}\}$ be an equality set, where $\tilde{f}_i \in \mathbb{R}_h[\boldsymbol{a}]$ for all $i \in \mathcal{N}_{\widetilde{m}}$. We say h_1 can be transformed to h_2 by $E_{\tilde{f}}$ if $h_1 \equiv h_2 + h_3$, where $h_3 \equiv \sum_{i=1}^{m'} q_i f'_i$, $q_i \in \mathbb{R}$ and $E_{f'} = \{f'_i = 0, i \in \mathcal{N}_{m'}\}$ is an equivalent set of $E_{\tilde{f}}$.*

Theorem III.1 ([14]). *Problem III.3 is solvable if and only if F can be transformed into a conic combination of $a_i, i \in \mathcal{N}_m$ by E_a.*

Definition III.11. *Let $E_a = \{f_i = 0,\ i \in \mathcal{N}_{\widetilde{m}}\}$, where f_i is a polynomial in $\mathbf{a}$, be an equality set. We say eliminating a variable a_i from E_a to mean solving for a_i in some $f_i = 0$ with $a_i \in V(f_i)$ to obtain $a_i = A_i$ and then substituting $a_i = A_i$ into E_a to obtain $E_A = subs(a_i = A_i, E_a) \backslash \{0 = 0\}$.*

Let F be a polynomial in $\mathbf{a}$. We say eliminating a_i from F by E_a to mean eliminating a_i from E_a to obtain $a_i = A_i$ and E_A, and then substituting $a_i = A_i$ into F to obtain $F_1 = subs(a_i = A_i, F)$.

The notions of redundant inequality and implied equality in Definitions III.4 and III.5, respectively can be applied in the more general setting in Problem III.3. Specifically, $a_i = 0,\ i \in \mathcal{N}_m$ is an implied equality if $-a_i \geq 0$ is provable subject to E_a and S_a. Also, by eliminating a_i for some $i \in \mathcal{N}_m$ from E_a to obtain $a_i = A_i$ and E_A, $a_i \geq 0$ is a redundant inequality if $A_i \geq 0$ is provable subject to E_A and $S_a \backslash \{a_i \geq 0\}$.

Definition III.12. *Let f be a polynomial in $\mathbf{a} = \{a_1, a_2, \ldots, a_m\}$. Let $\widetilde{m} \leq m$ and $j_1, j_2, \ldots, j_{\widetilde{m}}$ be distinct elements of $\{1, 2, \ldots, m\}$. If $f = \sum_{i=1}^{\widetilde{m}} p_i a_{j_i}$ or $f = -\sum_{i=1}^{\widetilde{m}} p_i a_{j_i}$ with $p_i > 0$, then f is called a Type I linear combination of a_{j_i}. If $f = \sum_{i=1}^{\widetilde{m}-1} p_i a_{j_i} - p_{\widetilde{m}} a_{j_{\widetilde{m}}}$ or $f = -\sum_{i=1}^{\widetilde{m}-1} p_i a_{j_i} + p_{\widetilde{m}} a_{j_{\widetilde{m}}}$ with $p_i > 0$, then f is called a Type II linear combination of a_{j_i}, and let $single(f) = a_{j_{\widetilde{m}}}$.*

Definition III.13. *In Problem III.3, if $(f = 0) \in E_a$ and*

1) *if f is Type I, then $a_i = 0$ for $a_i \in V(f)$ are called trivially implied equalities;*

[1]With a slight of abuse of terminology, the solution set of $E_{\tilde{f}}$ refers to the set $\{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : \tilde{f}_i = 0,\ i \in \mathcal{N}_{\widetilde{m}}\}$.

2) *if f is Type II, then $single(f) \geq 0$ is called a trivially redundant inequality.*

Example III.2. *Let $E_a = \{f_i = 0,\ i \in \mathcal{N}_4\}$, where $f_1 = a_1 + a_2$, $f_2 = -a_1 - a_2$, $f_3 = a_4 - a_5 - a_6$, and $f_4 = a_7 + a_8 - 2a_9$. Then f_1 and f_2 are Type I, f_3 and f_4 are Type II, $single(f_3) = a_4$, and $single(f_4) = a_9$. It can readily be checked that $a_1 = 0$ and $a_2 = 0$ are trivially implied equalities, and $a_4 \geq 0$ and $a_9 \geq 0$ are trivially redundant inequalities.*

In the rest of the paper, we denote the ith element of a sequence B by $B[i]$. We also denote the ith element of a set S of polynomials in $\mathbf{x}$ by $S[i]$, where the elements in S are assumed to be sorted in lexicographic order. For example, $x_1 + 2x_2 \succ x_2 + x_5$ and $x_3 + x_5 \succ x_3 + x_6$.

Definition III.14. *For a set S, let $|S|$ be the number of elements involved in S.*

Now we develop an algorithm to remove all trivially implied equalities and trivially redundant inequalities in the constraints in Problem III.3. To facilitate the discussion, we use $subs(\cdot,\cdot)$ to denote eleminating one or more variables from a set of polynomials by substitution.

Algorithm 3 Preprocessing Problem III.3

Input: Problem III.3.
Output: A reduced LP for Problem III.3.

Let $E_1 := \widetilde{\mathcal{R}}^{-1}(E_a)$, $S_1 := \mathcal{R}^{-1}(S_a)$, $F_1 := F$, $i_1 := 1$.
while $i_1 = 1$ **do**
  Let $i_1 := 0$.
  for i from 1 to $|E_1|$ **do**
    Let $f := E_1[i]$.
    if f is Type I **then**
      // In this case, all equalities in $\widetilde{\mathcal{R}}(V(f))$ are trivially implied equalities.
      $E_1 := subs(\widetilde{\mathcal{R}}(V(f)), E_1)\backslash\{0\}$.
      $S_1 := S_1\backslash V(f)$.
      $F_1 := subs(\widetilde{\mathcal{R}}(V(f)), F_1)$.
      $i_1 := 1$.
    end if
    if f is Type II **then**
      // In this case, the inequality $single(f) \geq 0$ is a trivially redundant inequality.
      $E_1 := subs(single(f) = solve(f, single(f)), E_1)\backslash\{0\}$.
      $S_1 := S_1\backslash\{single(f)\}$.
      $F_1 := subs(single(f) = solve(f, single(f)), F_1)$.
      $i_1 := 1$.
    end if
  end for
end while
return A reduced LP:
  Problem III.4. *Prove $F_1 \geq 0$ subject to $\widetilde{\mathcal{R}}(E_1)$ and $\mathcal{R}(S_1)$.*

Algorithm 3 removes all the trivially implied equalities and trivially redundant inequalities from Problem III.3. Toward solving Problem III.3, we first apply Algorithm 3 to reduce it to Problem III.4. The next algorithm is a heuristic that attempts to solve this problem. If unsuccesful, the algorithms in [14, Appendix A] will be applied to further reduce the LP into a smaller one that contains no implied equality and redundant inequality.

Algorithm 4 Heuristic search for a conic combination

Input: Problem III.4.
Output: SUCCESSFUL, or UNSUCCESSFUL and a reduced LP.

Let $J := E_1$, $J_2 := \emptyset$.
Let $\mathcal{V}(F_1) = [a_{i_1}, \ldots, a_{i_{n_3}}]$ and $\mathcal{C}(F_1) = [p_1, \ldots, p_{n_3}]$, where $1 \leq n_3 \leq m$ and the coefficient p_j corresponds to the variable a_{i_j} for all $j \in \mathcal{N}_{n_3}$.
while (there exists $p_j < 0$ for some $j \in \mathcal{N}_{n_3}$) $\wedge$ ($|J| > 0$) $\wedge$ ($a_{i_j} \in V(f)$ for some $f \in J$) **do**
  Solve a_{i_j} from $f = 0$ to yield $a_{i_j} = A_{i_j}$ such that A_{i_j} is a polynomial in $V(f)\backslash\{a_{i_j}\}$.
  $F_1 := F_1 - p_j(a_{i_j} - A_{i_j})$.
  $J := subs(a_{i_j} = A_{i_j}, J)\backslash\{0\}$.
  $J_2 := subs(a_{i_j} = A_{i_j}, J_2) \cup \{a_{i_j} - A_{i_j}\}$.
  Update $\mathcal{V}(F_1)$ and $\mathcal{C}(F_1)$.
end while
if there does not exist a negative element in $\mathcal{C}(F_1)$ **then**
  // $F_1 \geq 0$ is obviously implied by $\mathcal{R}(S_1)$.
  Return 'SUCCESSFUL'.
else
  // Need to solve
    Problem III.5. *Prove $F_1 \geq 0$ subject to $\widetilde{R}(J \cup J_2)$ and $\mathcal{R}(S_1)$.*
  // Instead of reducing F_1 by $J \cup J_2$ directly, since J_2 is already in row echelon form after the WHILE loop, we can simplify the computation as follows.
  Reduce F_1 and J_2 by J to obtain the remainder F_2 and the remainder set $\widetilde{J}_2$, respectively, and also the RREF of J denoted by $\widetilde{J}$.
  Let $\widetilde{\mathcal{E}}_1 = \widetilde{J} \cup \widetilde{J}_2$, which is an RREF of $\widetilde{\mathcal{R}}^{-1}(E_a)$.
  // Problem III.5 is reduced to
    Problem III.6. *Prove $F_2 \geq 0$ subject to $\widetilde{R}(\widetilde{\mathcal{E}}_1)$ and $\mathcal{R}(S_1)$.*
  Apply the algorithms in [14, Appendix A] to Problem III.6 to obtain a reduction of Problem III.4:
    Problem III.7. *Prove $F_3 \geq 0$ subject to $\widetilde{\mathcal{R}}(\widetilde{\mathcal{E}}_2)$ and $\mathcal{R}(V(\{F_3\} \cup \widetilde{\mathcal{E}}_2))$.*
  // Problem III.7 contains no implied equalities and redundant inequalities. Thus we only need to consider the inequality constraints $\mathcal{R}(V(\{F_3\} \cup \widetilde{\mathcal{E}}_2))$ instead of $\mathcal{R}(S_1)$, where $|V(\{F_3\} \cup \widetilde{\mathcal{E}}_2)| \leq |S_1|$.
  Return 'UNSUCCESSFUL' and Problem III.7.
end if

In [14], we also give an example to show that Algorithm 4 is not always successful even though the problem is solvable. In general, different decisions made in the algorithm can lead to different outcomes.

Assume that Algorithm 4 outputs 'UNSUCCESSFUL' and Problem III.7, which is a reduction of Problem III.4. We now present the following algorithm for solving this problem.

Algorithm 5 Solving Problem III.7

Input: Problem III.7.
Output: The statement "Problem III.7 is solvable" is TRUE or FALSE.

Assume that $\widetilde{\mathcal{E}}_2$ has the form $\widetilde{\mathcal{E}}_2 = \{a_{k_l} - A_{k_l}, l \in \mathcal{N}_r\}$, where

r is the rank of $\widetilde{\mathcal{E}}_2$, and A_{k_l}'s are linear combinations of the free variables $a_{k_{r+1}}, \ldots, a_{k_t}$, where $t = |V(\widetilde{\mathcal{E}}_2)| \leq m$.

2: Let $F_4 \equiv F_3 + \sum_{l=1}^{r} p_l(a_{k_l} - A_{k_l})$, where $p_l, 1 \leq l \leq r$ are to be determined. Since F_3 and A_{k_l}'s are in terms of the free variables, we can rewrite F_4 as $F_4 \equiv \sum_{l=1}^{r} p_l a_{k_l} + \sum_{l=r+1}^{t} P_l a_{k_l}$, where P_l's are linear combinations of p_l's.

3: // By Theorem III.1, Problem III.7 can be proved if and only if F_4 can be expressed as a conic combination of a_i's.

4: Solve the following LP:

Problem III.8. $\min(0)$ such that $p_l \geq 0,\ l \in \mathcal{N}_r$ and $P_l \geq 0,\ l \in \mathcal{N}_t \backslash \mathcal{N}_r$.

5: **if** Problem III.8 can be solved **then**

6: Declare that "Problem III.7 can be solved" is 'TRUE'.

7: **else**

8: Declare that "Problem III.7 can be solved" is 'FALSE'.

9: **end if**

10: **return** The argument "the Problem III.7 can be solved" is TRUE or FALSE.

IV. Procedure for proving information inequality

In this section, we present a procedure for proving information inequalities under the constraint of an inequality set and/or equality set. They are designed in the spirit of Theorem II.2. To simplify the discussion, $H(X_1, X_2, \ldots, X_n)$ will be denoted by $h_{1,2,\ldots,n}$, so on and so forth. For a joint entropy $t = h_{i_1,i_2,\ldots,i_n}$, the set $L(t) = \{i_1, i_2, ..., i_n\}$ is called the *subscript set* of t. The following defines an order among the joint entropies.

Definition IV.1. *Let $t_1 = h_{i_1,i_2,\ldots,i_{n_1}}$ and $t_2 = h_{j_1,j_2,\ldots,j_{n_2}}$ be two joint entropies. We write $t_1 \succ t_2$ if one of the following conditions is satisfied:*

1) $|L(t_1)| > |L(t_2)|$,
2) $|L(t_1)| = |L(t_2)|$, $i_l = j_l$ *for* $l = 1, \ldots, k-1$ *and* $i_k > j_k$.

Next, we present our procedure.

Input:

Objective information inequality: $\bar{F} \geq 0$.

Elemental information inequalities: $\bar{C}_i \geq 0,\ i = 1, \ldots, m_1$.

Additional constraints: $\bar{C}_j \geq 0,\ j = m_1 + 1, \ldots, m_2$; $\bar{C}_k = 0,\ k = m_2 + 1, \ldots, m_3$.

// Here, $\bar{F}$, $\bar{C}_i$, $\bar{C}_j$, and $\bar{C}_k$ are linear combination of Shannon's information measures.

Output: A proof of $\bar{F} \geq 0$ if it is implied by the elemental inequalities and the additional constriants.

Step 1. Transform $\bar{F}$ and $\bar{C}_i$, $i \in \mathcal{N}_{m_3}$ to homogeneous linear polynomials $\widetilde{F}$ and $\widetilde{C}_i$, $i \in \mathcal{N}_{m_3}$ in joint entropies.

Step 2. Fix the joint entropies' order $h_{1,2,\ldots,n} \succ \cdots \succ h_1$. Apply Algorithm 1 to reduce the inequality set $\{\widetilde{C}_i \geq 0, i \in \mathcal{N}_{m_2}\}$ by the equality set $\{\widetilde{C}_i = 0, i \in \mathcal{N}_{m_3} \backslash \mathcal{N}_{m_2}\}$ to obtain the reduced inequality set $\{C_i \geq 0, i \in \mathcal{N}_m\}$.

Step 3. Reduce $\widetilde{F}$ by the equality set $\{\widetilde{C}_i = 0, i \in \mathcal{N}_{m_3} \backslash \mathcal{N}_{m_2}\}$ to obtain F_5.

// We need to solve

Problem IV.1. *Prove $F_5 \geq 0$ under the constraints $C_i \geq 0,\ i \in \mathcal{N}_m$.*

Step 4. Under the variable order $h_{1,2,\ldots,n} \succ \cdots \succ h_1 \succ a_1 \succ \cdots \succ a_m$, apply Algorithm 2 to Problem IV.1 to obtain

Problem III.2$(*)$. Prove $F \geq 0$ subject to $\widetilde{\mathcal{R}}(J_1)$ and $a_i \geq 0,\ i \in \mathcal{N}_m$, where $J_1 = \{f_i,\ i \in \mathcal{N}_{m_4}\}$.

Step 5. Apply Algorithm 3 and Algorithm 4 successively to the above problem. If Algorithm 4 outputs 'SUCCESSFUL', then the objective function $\bar{F} \geq 0$ is proved. Otherwise, the following reduced LP is obtained:

Problem III.7$(*)$. Prove $F_3 \geq 0$ subject to $\widetilde{\mathcal{R}}(\widetilde{\mathcal{E}}_2)$ and $\mathcal{R}(V(\{F_3\} \cup \widetilde{\mathcal{E}}_2))$, where $\widetilde{\mathcal{E}}_2 = \{\tilde{f}_i,\ i \in \mathcal{N}_{m_5}\}$.

// Note that $m_5 \leq m_4$ and $|\mathcal{R}(V(\{F_3\} \cup \widetilde{\mathcal{E}}_2))| \leq m$.

Step 6. Apply Algorithm 5 to the above problem. If Algorithm 5 outputs 'TRUE', then the objective function $\bar{F} \geq 0$ is proved. Otherwise, declare 'Not Provable'.

Next, we give an example to show the effectiveness of our procedure.

Example IV.1. $I(X_i; X_4) = 0,\ i = 1, 2, 3$ *and* $H(X_4|X_i, X_j) = 0, 1 \leq i < j \leq 3 \Rightarrow H(X_i) \geq H(X_4)$.

The inequality above can be proved by our procedure. The details can be found in [14]. Table I shows the advantage of our procedure by comparing it with the Direct LP method and our previous work [13].

TABLE I

	Number of variables	Number of equality constraints	Number of Inequality constraints
Direct LP method	15	6	28
LP in [13]	2	0	6
LP in this work	no LP needs to be solved		

In [14], we also apply our procedure to tackle the problem studied by Tian [11] regarding a conjecture on the rate region for (4,3,3) exact-repair regeneration codes [10] and show a significant reduction in the complexity of the problem compared with our previous work [13].

V. Concluding remarks

Since different elimination choices of variables in Algorithm 4 can lead to different results, our heuristic method may not necessarily succeed. Nevertheless, if the first attempt is unsuccessful, we can repeat the attempt with different elimination choices of variables for a certain maximum number of times. Even if Algorithm 4 cannot solve the problem directly, it can still reduce the problem to the minimal LP in a shorter time and with less memory compared with our previous work [13]. The reader is referred to [14] for the details, where we also included a detailed discussion on the advantage and effectiveness of our procedure for solving more elaborate problems such as Tian's problem [11].

REFERENCES

[1] R. W. Yeung, "A framework for linear information inequalities," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1924-1934, Nov. 1997.

[2] R. W. Yeung and Y.-O. Yan (1996), Information Theoretic Inequality Prover (ITIP), MATLAB Program Software Package. [Online]. Available: http://home.ie.cuhk.edu.hk/ ITIP

[3] R. Pulikkoonattu and S. Diggavi (2006), Xitip, ITIP-Based C Program Software Package. [Online]. Available: http://xitip.epfl.ch

[4] L. Csirmaz (2016), A MINimal Information Theoretic Inequality Prover (Minitip). [Online]. Available: https://github.com/lcsirmaz/minitip

[5] C. T. Li (2020), Python Symbolic Information Theoretic Inequality Prover (psitip). [Online]. Available: https://github.com/cheuktingli/

[6] N. Rathenakar, S. Diggavi, T. Gläβle, E. Perron, R. Pulikkoonattu, R. W. Yeung, and Y.-O. Yan (2020), Online X-Information Theoretic Inequalities Prover (oXitip). [Online]. Available: http://www.oxitip.com

[7] S.-W. Ho, L. Ling, C. W. Tan, and R. W. Yeung, "Proving and disproving information inequalities: Theory and scalable algorithms," *IEEE Trans. Inf. Theory*, vol. 66, no. 9, pp. 5522-536, Sep. 2020.

[8] R. W. Yeung, *Information Theory and Network Coding*. New York, NY, USA: Springer, 2008.

[9] D. C. Lay, *Linear Algebra and Its Applications*, 5th Edition. Pearson, 2016.

[10] A. G. Dimakis, P. B. Godfrey, Y. Wu, M. J. Wainwright, and K. Ramchandran, "Network coding for distributed storage systems," *IEEE Trans. Inform. Theory*, vol. 56, pp. 4539-4551, Sept 2010.

[11] C. Tian, "Characterizing the rate region of the (4, 3, 3) exact-repair regenerating codes," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 5, pp: 967-975, 2014.

[12] A. Ben-Tal, A. Nemirovski, *Lecture notes: optimization III*, New York, NY, USA: Springer, 2022.

[13] L. Guo, R. W. Yeung and X. -S. Gao, "Proving Information Inequalities and Identities with Symbolic Computation," IEEE Transactions on Information Theory, vol. 69, no. 8, pp. 4799-4811, 2023.

[14] L. Guo, R. W. Yeung and X. -S. Gao, "Proving Information Inequalities by Gaussian Elimination," arXiv preprint, arXiv:2401.14916, 2024.