Proving Information Inequalities by Gaussian Elimination

Laigang Guo, *Member, IEEE,* Raymond W. Yeung, *Fellow, IEEE,* and Xiao-Shan Gao, *Senior Member, IEEE*

***Abstract*—The proof of information inequalities and identities 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 [23], we developed in this paper an enhanced approach for solving this problem.**

***Index Terms*—Entropy, mutual information, information inequality, information identity, machine proving, ITIP.**

I. Introduction

In information theory, we may need to prove various information inequalities and identities that involve Shannon's information measures. For example, such information inequalities and identities play a crucial role in establishing the converse of most coding theorems. However, proving an information inequality or identity involving more than a few random variables can be highly non-trivial.

To tackle this problem, 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 [3]. Subsequently, different variations of ITIP have been developed. Instead of MATLAB, Xitip [4] uses a C-based linear programming solver, and it has been further developed into its web-based version, oXitip [7]. minitip [5] is a C-based version of ITIP that adopts a simplified syntax and has a user-friendly syntax checker. psitip [6] is a Python library that can verify unconstrained/constrained/existential entropy inequalities. It is a computer algebra system where random variables, expressions, and regions are objects that can be manipulated. AITIP [9], [8] is a platform that not only provides analytical proofs for Shannon-type inequalities but also give hints on constructing a smallest counterexample in case the inequality to be verified is not a Shannon-type inequality. There are also some works that use advanced algorithms for linear programming (LP) and polyhedron computing for proving information inequalities. For the details of LP and polyhedron computing methods and their applications in proving information inequalities, the readers are referred to [21], [25], [26], [27], [28], [29].

Using the above LP-based approach, to prove an information identity $f = 0$, two LPs need to be solved, one for proving the inequality $f \geq 0$ and the other for proving the inequality $f \leq 0$. Roughly speaking, the amount of computation for proving an information identity is twice the amount for proving an information inequality. If the underlying random variables exhibit certain Markov or functional dependence structures, there exist more efficient approaches to proving information identities [11][13].

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 [23] 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 specific contributions of this paper are:

1) We develop a heuristic method to prove an information inequality. This heuristic method does not prove an information inequality by directly solving the associated LP, but rather expedites the proof process through polynomial reduction (Gaussian elimination).

2) This heuristic method may not succeed in proving the inequality. If it does not succeed, it can simplify the original LP into a smaller LP.

3) We give several examples that verify the effectiveness of our method.

This paper is organized as follows. In Section II, we review the linear programming method for proving information inequalities. In Section III, we develop the main

This article was presented in part at the 2024 IEEE International Symposium on Information Theory. (Corresponding author: Laigang Guo.)

L. Guo is with the Laboratory of Mathematics and Complex Systems (Ministry of Education), School of Mathematical Sciences, Beijing Normal University, Beijing, China. e-mail: (lgguo@bnu.edu.cn).

R. W. Yeung is with the Institute of Network Coding and the Department of Information Engineering, The Chinese University of Hong Kong, N.T., Hong Kong. e-mail: (whyeung@ie.cuhk.edu.hk).

X.-S. Gao is with the Key Laboratory of Mathematics Mechanization, Institute of Systems Science, AMSS, Chinese Academy of Sciences, and University of Chinese Academy of Sciences, Beijing, China. e-mail: (xgao@mmrc.iss.ac.cn).

algorithms for homogeneous linear inequalities. In Section IV, we present the procedures for proving information inequalities and identities. In Section V, we present two applications that demonstrate the effectiveness of our approach. Conclusion and Discussion are given in Section VI.

II. Information inequality preliminaries

In this section, we present some basic results related to information inequalities and their verification. For a comprehensive discussion on the topic, we refer the reader to [2], [10, Chs. 13-15].

It is well known that all Shannon's information measures, namely entropy, conditional entropy, mutual information, and conditional mutual information are always nonnegative. The nonnegativity of all Shannon's information measures forms a set of inequalities called the *basic inequalities*. The set of basic inequalities, however, is not minimal in the sense that some basic inequalities are implied by the others. For example,

$$H(X|Y) \geq 0 \text{ and } I(X;Y) \geq 0,$$

which are both basic inequalities involving random variables X and Y, imply

$$H(X) = H(X|Y) + I(X;Y) \geq 0,$$

again a basic inequality involving X and Y.

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

$$u = n + \sum_{r=0}^{n-2} \binom{n}{r} \binom{n-r}{2} = n + \binom{n}{2} 2^{n-2}.$$

In this paper, inequalities (identities) involving only Shannon's information measures are referred to as information inequalities (identities). The elemental inequalities are called *unconstrained* information inequalities because they hold for all joint distributions of the random variables. In information theory, we very often deal with information inequalities (identities) that hold under certain constraints on the joint distribution of the random variables. These are called *constrained* information inequalities (identities), and the associated constraints are usually expressible as linear constraints on Shannon's information measures. We will confine our discussion to constrained inequalities of this type.

Example II.1. *The celebrated data processing theorem asserts that for any four random variables* X*,* Y*,* Z *and* T*, if* $X \to Y \to Z \to T$ *forms a Markov chain, then* $I(X;T) \leq I(Y;Z)$*. Here,* $I(X;T) \leq I(Y;Z)$ *is a constrained information inequality under the constraint* $X \to Y \to Z \to T$*, which is equivalent to*

$$\begin{cases} I(X;Z|Y) &= 0 \\ I(X,Y;T|Z) &= 0, \end{cases}$$

or

$$I(X;Z|Y) + I(X,Y;T|Z) = 0$$

owing to the nonnegativity of conditional mutual information. Either way, the Markov chain can be expressed as a set of linear constraint(s) on Shannon's information measures.

Information inequalities (unconstrained or constrained) that are implied by the basic inequalities are called *Shannon-type* inequalities. Most of the information inequalities that are known belong to this type. However, *non-Shannon-type* inequalities do exist, e.g., [12]. See [10, Ch. 15] for a discussion.

Shannon's information measures, with conditional mutual information being the general form, can be expressed as a linear combination of joint entropies by means of following identity:

$$\begin{aligned} I(X_G; X_{G'}|X_{G''}) &= H(X_G, X_{G''}) + H(X_{G',G''}) \\ &\quad - H(X_G, X_{G'}, X_{G''}) - H(X_{G''}). \end{aligned}$$

where $G, G', G'' \subseteq \mathcal{N}_n$. 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* $\mathbf{Q}\mathbf{h} = 0$ *if and only if the minimum of the problem*

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

is zero.

III. Algorithms for Homogeneous Linear Inequlities

In this section, we will develop new algorithms for proving information inequalities and identities. We will start by discussing some notions pertaining to linear inequality sets and linear equality sets. Then we will state some related properties that are necessary for developing these algorithms. For details, one can refer to [21], [22].

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 $E_{\tilde{f}} = \{\tilde{f}_i = 0, i \in \mathcal{N}_{\widetilde{m}}\}$ 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 $P_{\tilde{f}} = \{\tilde{f}_i, i \in \mathcal{N}_{\widetilde{m}}\}$ and the corresponding set of equalities $E_{\tilde{f}} = \{\tilde{f}_i = 0, i \in \mathcal{N}_{\widetilde{m}}\}$, 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:*

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

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. [23] *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. [22] *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* $h_1, \ldots, h_m$.

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. *For a polynomial* F *in* $\mathbf{x}$, *let* $|F|$ *be the number of variables involved in* F.

Definition III.8. *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*

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

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

$$\{\mathbf{x} : f(\mathbf{x}) \geq 0\} = \mathrm{Proj}_{\mathbf{x}}\{(\mathbf{x}, a) : f(\mathbf{x}) - a = 0, a \geq 0\}.^1$$

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 $\{f_i(\mathbf{x}) - a_i = 0, \ a_i \geq 0, i \in \mathcal{N}_m\}$.

The following proposition is well known (see for example [15, 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*

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

[1] $\mathrm{Proj}_{\mathbf{x}} S$ denotes the projection of set S on $\mathbf{x}$.

This article has been accepted for publication in IEEE Transactions on Information Theory. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIT.2025.3531133

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 $\{x_j,\ for\ k_i < j \leq n,\ \ j \neq k_l,\ i < l \leq \tilde{n}\}$, with $k_{\tilde{n}+1} = n + 1$ by convention.

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.

Definition III.9. *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 $\mathbf{a}$.

Algorithm 1 Dimension Reduction

Input: S_f, $E_{\tilde{f}}$.
Output: The remainder set R_f.

Compute $\widetilde{E}$ for $E_{\tilde{f}}$ by Proposition III.1.
Substitute x_{k_i} by U_i in $\mathcal{R}^{-1}(S_f)$ to obtain the set R.
Let $R_f = R \backslash \{0\}$.
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.10. *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.*[2]

Definition III.11. *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 $E_{\tilde{f}}$.*

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

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 Algorithm

Input: Problem III.1
Output: A reduced LP.

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

Theorem III.1. *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.*

Proof. The sufficiency is obvious. We only need to prove the necessity.

Assume that Problem III.3 is solvable. By Proposition III.1, we compute the RREF of E_a, denoted by $\widetilde{E} = \{a_{k_i} - U_i = 0,\ i \in \mathcal{N}_{\widetilde{n}}\}$, and substitute a_{k_i} in F by U_i to obtain F_1. Then we see that

$$F \equiv F_1 + \sum_{i=1}^{\widetilde{n}} q_i(a_{k_i} - U_i),\ q_i \in \mathbb{R}. \tag{3}$$

In other words, F can be transformed to F_1 by E_a. Then we substitute a_{k_i} in S_a by U_i to obtain $S_o = \{g_i \geq 0,\ i \in \mathcal{N}_m\}$, where $g_{k_i} = U_i$ for $i \in \mathcal{N}_{\widetilde{n}}$ and $g_j = a_j$ for $j \in \mathcal{N}_m \backslash \{k_i, i \in \mathcal{N}_{\widetilde{n}}\}$.

Now Problem III.3 is equivalent to

Problem III.4. *Prove $F_1 \geq 0$ subject to S_o.*

By Lemma III.2, Problem III.4 is solvable if and only if F_1 is a conic combination of $g_i,\ i \in \mathcal{N}_m$. Suppose $F_1 \equiv \sum_{i=1}^{m} p_i g_i$ with $p_i \in \mathbb{R}_{\geq 0}$. Then

$$\begin{aligned} F_1 &\equiv \sum_{j=1}^{m} p_j g_j \\ &\equiv \sum_{i=1}^{\widetilde{n}} p_{k_i} U_i + \sum_{j \in \mathcal{N}_m \backslash \{k_i, i \in \mathcal{N}_{\widetilde{n}}\}} p_j a_j \\ &\equiv \sum_{i=1}^{\widetilde{n}} p_{k_i} a_{k_i} - \sum_{i=1}^{\widetilde{n}} p_{k_i}(a_{k_i} - U_i) + \sum_{j \in \mathcal{N}_m \backslash \{k_i, i \in \mathcal{N}_{\widetilde{n}}\}} p_j a_j \\ &\equiv \sum_{i=1}^{m} p_i a_i - \sum_{i=1}^{\widetilde{n}} p_{k_i}(a_{k_i} - U_i). \\ &= \sum_{i=1}^{m} p_i a_i, \end{aligned} \tag{4}$$

where the last step follows from the constraints in $\widetilde{E}$.

So, F_1 can be expressed as a conic combination of a_i's. Then, by (3) and the second last line above, we obtain

$$\begin{aligned} F &\equiv F_1 + \sum_{i=1}^{\widetilde{n}} q_i(a_{k_i} - U_i) \\ &\equiv \sum_{i=1}^{m} p_i a_i - \sum_{i=1}^{\widetilde{n}} p_{k_i}(a_{k_i} - U_i) + \sum_{i=1}^{\widetilde{n}} q_i(a_{k_i} - U_i) \\ &\equiv \sum_{i=1}^{m} p_i a_i + \sum_{i=1}^{\widetilde{n}} (q_i - p_{k_i})(a_{k_i} - U_i) \end{aligned} \tag{5}$$

where $p_i \in \mathbb{R}_{\geq 0}$, $p_{k_i} \in \mathbb{R}_{\geq 0}$ and $q_i \in \mathbb{R}$.

From Definition III.11, we see that F_1 is a conic combination of $g_i,\ i \in \mathcal{N}_m$ if and only if F can be transformed into a conic combination of $a_i,\ i \in \mathcal{N}_m$ by E_a. Hence, following the discussion in the foregoing, the theorem is proved. □

Definition III.12. *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 $S_a \backslash \{a_i \geq 0\}$.

Example III.2. *Let $S_a = \{a_i \geq 0,\ i \in \mathcal{N}_5\}$ and $E_a = \{f_1 = 0, f_2 = 0\}$, where $f_1 = a_1 + a_2$ and $f_2 = a_3 - a_4 - a_5$. Using $f_1 = 0$, $a_1 \geq 0$ and $a_2 \geq 0$, we can obtain that $-a_1 \geq 0$ and $-a_2 \geq 0$. Thus $a_1 = 0$ and $a_2 = 0$ are implied equalities.*

By eliminating a_3 from E_a, we obtain $a_3 = a_4 + a_5$ and $E_A = \{a_1 + a_2 = 0\}$. Since $a_4 + a_5 \geq 0$ is obviously provable subject to E_A and $S_a \backslash \{a_3 \geq 0\}$, we have that $a_3 \geq 0$ is a redundant inequality.

Definition III.13. *Let f be a polynomial in $\mathbf{a} = \{a_1, a_2, \ldots, a_m\}$. Let $\bar{m} \leq m$ and $j_1, j_2, \ldots, j_{\bar{m}}$ be distinct elements of $\{1, 2, \ldots, m\}$. If $f = \sum_{i=1}^{\bar{m}} p_i a_{j_i}$ or $f = -\sum_{i=1}^{\bar{m}} p_i a_{j_i}$ with $p_i > 0$, then f is called a Type I linear combination of a_{j_i}. If $f = \sum_{i=1}^{\bar{m}-1} p_i a_{j_i} - p_{\bar{m}} a_{j_{\bar{m}}}$ or $f = -\sum_{i=1}^{\bar{m}-1} p_i a_{j_i} + p_{\bar{m}} a_{j_{\bar{m}}}$ with $p_i > 0$, then f is called a Type II linear combination of a_{j_i}, and let $single(f) = a_{j_{\bar{m}}}$.*

Definition III.14. *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;*
2) *if f is Type II, then $single(f) \geq 0$ is called a trivially redundant inequality.*

Example III.3. *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 $x_3 + x_5 \succ x_3 + x_6$.

Now we develop an algorithm to remove all trivially implied equalities and trivially redundant inequalities 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. The use of this notation will be illustrated in Example III.4.

This article has been accepted for publication in IEEE Transactions on Information Theory. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIT.2025.3531133

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
    if $i_1 = 1$ **then**
      Terminate the FOR loop.
    end if
  end for
end while
return A reduced LP:
  Problem III.5. *Prove* $F_1 \geq 0$ *subject to* $\widetilde{\mathcal{R}}(E_1)$ *and* $\mathcal{R}(S_1)$.

Example III.4. *We want to prove* $F = a_1 + 2a_2 - a_3 \geq 0$ *subject to* $E_a = \{a_1 + a_2 - a_3 - a_4 - a_5 = 0,\ a_1 + a_4 = 0\}$ *and* $S_a = \{a_i \geq 0,\ i \in \mathcal{N}_5\}$. *Following Algorithm 3, we give the steps in detail.*

Step 1. $F_1 = a_1 + 2a_2 - a_3$, $E_1 = \{a_1 + a_2 - a_3 - a_4 - a_5,\ a_1 + a_4\}$, $S_1 = \{a_1, a_2, a_3, a_4, a_5\}$.

Step 2. Since $a_1 + a_4$ *is Type I, we obtain* $a_1 = 0$ *and* $a_4 = 0$ *from* $a_1 + a_4 = 0$, $a_1 \geq 0$, *and* $a_4 \geq 0$.

Step 3. Obtain $E_1 := subs(a_1 = 0, a_4 = 0, E_1\})\backslash\{0\} = \{a_2 - a_3 - a_5\}$, $F_1 := subs(a_1 = 0, a_4 = 0, F_1) = 2a_2 - a_3$, *and* $S_1 := S_1\backslash\{a_1, a_4\} = \{a_2, a_3, a_5\}$.

Step 4. Now $a_2 - a_3 - a_5$ *is Type II. Then solve* a_2 *from* $a_2 - a_3 - a_5$ *to obtain* $a_2 = a_3 + a_5$.

Step 5. Obtain $F_1 := subs(a_2 = a_3 + a_5, F_1) = a_3 + 2a_5$, $E_1 := subs(a_2 = a_3 + a_5, E_1)\backslash\{0\} = \emptyset$, *and* $S_1 := S_1\backslash\{a_2\} = \{a_3, a_5\}$.

Now the reduced problem is to prove $a_3 + 2a_5 \geq 0$ *subject to* $a_3 \geq 0$ *and* $a_5 \geq 0$, *which is obviously solvable.*

Algorithm 3 removes all the trivially implied equalities and trivially redundant inequalities from Problem III.3. In Appendix A, we will develop two enhancements of Algorithm 3: Algorithm 6 for removing all implied equalities and Algorithm 7 for removing all redundant inequalities.

Toward solving Problem III.3, we first apply Algorithm 3 to reduce it to Problem III.5. The next algorithm is a heuristic that attempts to solve this problem. If unsuccessful, Algorithms 6 and 7 will be applied to further reduce the LP into a smaller one that contains no implied equality and redundant inequality. This will be illustrated in Example III.5.

Algorithm 4 Heuristic search for a conic combination

Input: Problem III.5.
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.6. *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.6 is reduced to
    Problem III.7. *Prove* $F_2 \geq 0$ *subject to* $\widetilde{R}(\widetilde{\mathcal{E}}_1)$ *and* $\mathcal{R}(S_1)$.
  Apply Algorithms 6 and 7 to Problem III.7 to obtain a reduction of Problem III.5:

This article has been accepted for publication in IEEE Transactions on Information Theory. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TIT.2025.3531133

Problem III.8. *Prove $F_3 \geq 0$ subject to $\widetilde{\mathcal{R}}(\widetilde{\mathcal{E}}_2)$ and $\mathcal{R}(V(\{F_3\} \cup \widetilde{\mathcal{E}}_2))$.*

20: // Problem III.8 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 $|V(\{F_3\} \cup \widetilde{\mathcal{E}}_2)| \leq |S_1|$.

21: Return 'UNSUCCESSFUL' and Problem III.8.

22: **end if**

Remark III.2. *In the WHILE loop in line 3 of Algorithm 4, we need to choose one variable a_{i_j} which has a negative coefficient in the objective polynomial F_1, and then solve a_{i_j} from one $f \in J$ that satisfies $a_{i_j} \in V(f)$. Here, a_{i_j} and f are chosen randomly by using the "choose" command under the RandomTools package in MAPLE. This package uses the "Mersenne Twister" pseudorandom number generator (PRNG) by default. In order to achieve repeatability of the experiments, we need to set the random seed explicitly. The "randomize" command under the RandomTools package can be used to set the random seed. In this paper, we use the seed, "randomize(0)".[3] For the information inequality proof problems discussed in this paper, our experiments have shown that the choice of the random seed may affect the results of one or several cycles in Algorithm 4, but it has little effect on the complete run of the algorithm, especially for large information inequality proofs. In addition, because the choices of a_{i_j} and f are random, it is possible that the same sequence of variables and polynomials are chosen in two different attempts. However, the probability of this happening in two large-scale random calculations is very low, and the installation of any mechanism to present this will increase the average computational complexity. Of course, it may be possible to devise mechanisms to search among the possibilities, still in a random manner, but at a lower computational cost to avoid repeating the same choice. This will be left for future work.*

Next, we 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.

Example III.5. *We want to determine whether $F = -\frac{1}{2}a_1 - a_2 + a_3 + a_4 + a_5 - a_6 + a_7 + a_9 \geq 0$ subject to $S_a = \{a_i \geq 0, i \in \mathcal{N}_{12}\}$ and $E_a = \{a_1 + a_2 - a_3 - a_4 = 0, a_1 + a_2 - a_4 + a_9 + a_{10} - a_{11} - a_{12} = 0, a_6 - a_9 - a_{10} + a_{11} + a_{12} = 0, a_5 - 2a_6 = 0, a_7 + a_8 = 0\}$. Following Algorithm 4, we give the steps in detail.*

***Step 1**. Run Algorithm 3 to obtain*

***Problem III.5**$(*)$. Prove $F_1 \geq 0$ subject to $\widetilde{\mathcal{R}}(E_1)$ and $\mathcal{R}(S_1)$, where $F_1 = -\frac{1}{2}a_1 - a_2 + a_3 + a_4 + a_6 + a_9$, $E_1 = \{a_1 + a_2 - a_3 - a_4,\ a_1 + a_2 - a_4 + a_9 + a_{10} - a_{11} - a_{12},\ a_6 - a_9 - a_{10} + a_{11} + a_{12}\}$, and $S_1 = \{a_1, a_2, a_3, a_4, a_6, a_9, a_{10}, a_{11}, a_{12}\}$.*

Here, we use Problem III.5$(*)$ to denote a special instance of Problem III.5. Similar notations will apply.

Let $J := E_1 = \{a_1 + a_2 - a_3 - a_4,\ a_1 + a_2 - a_4 + a_9 + a_{10} - a_{11} - a_{12},\ a_6 - a_9 - a_{10} + a_{11} + a_{12}\}$, $J_2 := \emptyset$.

Referring to Line 4 of Algorithm 4, we discuss two possible cases.

Case 1: Assume that we solve a_2 from $a_1 + a_2 - a_3 - a_4 = 0$.

Step 2. *Solve a_2 from $a_1 + a_2 - a_3 - a_4 = 0$ to obtain $a_2 = -a_1 + a_3 + a_4$.*

$F_1 := subs(a_2 = -a_1 + a_3 + a_4, F_1) = \frac{1}{2}a_1 + a_6 + a_9$.

Then $F \geq 0$ is proved.

Case 2: Assume that we solve a_1 from $a_1 + a_2 - a_3 - a_4 = 0$.

Step 2. *Solve a_1 from $a_1 + a_2 - a_3 - a_4 = 0$ to obtain $a_1 = -a_2 + a_3 + a_4$.*

$F_1 := subs(a_1 = -a_2 + a_3 + a_4, F_1) = -\frac{1}{2}(a_2 - a_3 - a_4) + a_6 + a_9$.

$J := subs(a_1 = -a_2 + a_3 + a_4, J)\backslash\{0\} = \{a_3 + a_9 + a_{10} - a_{11} - a_{12}, a_6 - a_9 - a_{10} + a_{11} + a_{12}\}$.

$J_2 := subs(a_1 = -a_2 + a_3 + a_4, J_2) \cup \{a_1 + a_2 - a_3 - a_4\} = \{a_1 + a_2 - a_3 - a_4\}$.

After executing this step, we observe that $a_2 \notin V(J)$ and the **while** *loop in Algorithm 4 is terminated. However, we have not yet solved the problem. Thus, we need to continue with the remaining steps in Algorithm 4.*

Step 3. *Reduce F_1 and J_2 by J to obtain the remainder $F_2 = -\frac{1}{2}(a_2 - a_4 - 3a_9 - a_{10} + a_{11} + a_{12})$ and the remainder set $\widetilde{J}_2 = \{a_1 + a_2 - a_4 + a_9 + a_{10} - a_{11} - a_{12}\}$, respectively, and also the RREF of J denoted by $\widetilde{J} = \{a_3 + a_9 + a_{10} - a_{11} - a_{12}, a_6 - a_9 - a_{10} + a_{11} + a_{12}\}$.*

Step 4. *Let $\widetilde{\mathcal{E}}_1 = \widetilde{J} \cup \widetilde{J}_2 = \{a_1 + a_2 - a_4 + a_9 + a_{10} - a_{11} - a_{12}, a_3 + a_9 + a_{10} - a_{11} - a_{12}, a_6 - a_9 - a_{10} + a_{11} + a_{12}\}$.*

Now the problem becomes

***Problem III.7**$(*)$. Prove $F_2 \geq 0$ subject to $\widetilde{\mathcal{R}}(\widetilde{\mathcal{E}}_1)$ and $\mathcal{R}(S_1)$, where $F_2 = -\frac{1}{2}(a_2 - a_4 - 3a_9 - a_{10} + a_{11} + a_{12})$, $\widetilde{\mathcal{E}}_1 = \{a_1 + a_2 - a_4 + a_9 + a_{10} - a_{11} - a_{12}, a_3 + a_9 + a_{10} - a_{11} - a_{12}, a_6 - a_9 - a_{10} + a_{11} + a_{12}\}$, and $S_1 = \{a_1, a_2, a_3, a_4, a_6, a_9, a_{10}, a_{11}, a_{12}\}$.*

Step 5. *Run Algorithms 6 and 7 to reduce the problem to*

***Problem III.8**$(*)$. 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 $F_3 = \frac{1}{2}a_1 - a_{10} + a_{11} + a_{12}$ and $\widetilde{\mathcal{E}}_2 = \{a_9 + a_{10} - a_{11} - a_{12}\}$.*

In this step, all the implied equalities and redundant inequalities in Problem III.7$(*)$ are removed. The detailed steps of this reduction from Problem III.7$(*)$ to Problem III.8$(*)$ are given in Appendix A.

Since $F_3 + \widetilde{\mathcal{E}}_2[1] = \frac{1}{2}a_1 + a_9 \geq 0$, the above LP is solvable. Thus, $F \geq 0$ is provable.

In Line 4 of Algorithm 4, we need to solve a_{i_j} from $f = 0$ for some $f \in J$. Different choices of a_{i_j}'s can

[3] For details about the command "randomize" and RandomTools package, one can refer to the MAPLE official website user manual.

lead to different outcomes, which has been shown in Case 1 and Case 2 in Example III.5. Similarly, different choices of $f \in J$ can also lead to different outcomes. For example, following from Example III.5, instead of solving a_1 or a_2 from $a_1 + a_2 - a_3 - a_4 = 0$ in Step 2, one can also solve a_1 or a_2 from $a_1 + a_2 - a_4 + a_9 + a_{10} - a_{11} - a_{12} = 0$. The details are omitted here.

Assume that Algorithm 4 outputs 'UNSUCCESSFUL' and Problem III.8, which is a reduction of Problem III.5. We now present the following algorithm for solving this problem.

Algorithm 5 Solving Problem III.8

Input: Problem III.8.

Output: The statement "Problem III.8 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 $a_{k_{r+1}}, \dots, a_{k_t}$, where $t = |V(\widetilde{\mathcal{E}}_2)| \leq m$.

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.

// By Theorem III.1, Problem III.8 can be proved if and only if F_4 can be expressed as a conic combination of a_i's.

Solve the following LP:

Problem III.9. $\min(0)$ such that $p_l \geq 0,\ l \in \mathcal{N}_r$ and $P_l \geq 0,\ l \in \mathcal{N}_t \backslash \mathcal{N}_r$.

if Problem III.9 can be solved **then**

  Declare that "Problem III.8 can be solved" is 'TRUE'.

else

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

end if

return The argument "the Problem III.8 can be solved" is TRUE or FALSE.

Remark III.3. *From Theorem III.1, 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. We first apply Algorithm 3 to Problem III.3 to obtain Problem III.5, which contains no trivially implied equalities and trivially redundant inequalities. Then we use Algorithm 4, a heuristic method, to try to find a conic combination for F_1. Specifically, we identify a monomial term of F_1 with a negative coefficient and use E_1 to eliminate the corresponding variable in F_1. Then we repeat this operation until it can not be done. There can be two possibilities. If a conic combination for F_1 is obtained, then the problem is solved. If a conic combination for F_1 is not obtained, then the problem may or may not be solvable. Algorithm 5 deals with this case. Here, even if Algorithm 4 can not solve the problem directly, it will reduce Problem III.3 to an equivalent LP but smaller in size, which can be solved effectively by Algorithm 5.*

IV. Procedures for Proving Information Inequality and Identity

In this section, we present two procedures for proving information inequalities and identities 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, $H(X_1, X_2, \dots, X_n)$ will be denoted by $h_{1,2,\dots,n}$, so on and so forth. For a joint entropy $t = h_{i_1,i_2,\dots,i_n}$, 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 $t_1 = h_{i_1,i_2,\dots,i_{n_1}}$ and $t_2 = h_{j_1,j_2,\dots,j_{n_2}}$ 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* $l = 1, \dots, k-1$ *and* $i_k > j_k$.

A. Procedure I: Proving Information Inequalities

Input:
Objective information inequality: $\bar{F} \geq 0$.
Elemental information inequalities: $\bar{C}_i \geq 0,\ i = 1, \dots, m_1$.
Additional constraints: $\bar{C}_j \geq 0,\ j = m_1 + 1, \dots, m_2$; $\bar{C}_k = 0,\ k = m_2 + 1, \dots, m_3$.
// 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 $h_{1,2,\dots,n} \succ \cdots \succ h_1$. 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 $h_{1,2,\dots,n} \succ \cdots \succ h_1 \succ a_1 \succ \cdots \succ a_m$, 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.8$(*)$. 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'.

B. Procedure II: Proving Information Identities

Input:
Objective information identity: $\bar{F} = 0$.
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 combinations of information measures.
Output: A proof of $\bar{F} = 0$ if it is implied by the elemental inequalities and the additional constraints.

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_6.

// We need to solve

Problem IV.2. *Prove $F_6 \geq 0$ 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.2 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 6 to the above problem to obtain a reduced and pure LP:

Problem VI.2$(*)$. Prove F_1 subject to $\widetilde{R}(E_1)$ and $\mathcal{R}(V(\{F_1\} \cup E_1))$.

If $F_1 \equiv 0$, then the objective function $\bar{F} = 0$ 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 $H(X_4|X_i, X_j) = 0, 1 \leq i < j \leq 3 \Rightarrow H(X_i) \geq H(X_4)$.*

This example has been discussed in [23]. Due to the symmetry of this problem, we only need to prove $H(X_1) \geq H(X_4)$. Next we give the proof based on Procedure I.

Step 1. We need to solve the problem:
Prove

$$\widetilde{F} = h_1 - h_4$$

under constraints $\widetilde{C}_i \geq 0,\ i = 1, \ldots, 28,\ \widetilde{C}_i = 0,\ i = 29, \ldots, 34$.

Steps 2-3. After reduction, the above problem becomes: Prove

$$F = h_1 - h_4$$

under the constraints $C_i \geq 0,\ i = 1, \ldots, 18$, where

$$\begin{aligned}
&C_1 = h_4,\ C_2 = h_1 + h_2 - h_{1,2},\\
&C_3 = h_1 + h_2 + h_4 - h_{1,2},\\
&C_4 = h_1 + h_3 - h_{1,3}, C_5 = h_1 + h_3 + h_4 - h_{1,3},\\
&C_6 = h_2 + h_3 - h_{2,3}, C_7 = h_2 + h_3 + h_4 - h_{2,3},\\
&C_8 = -h_1 + h_{1,2} + h_{1,3} - h_{1,2,3},\\
&C_9 = -h_2 + h_{1,2} + h_{2,3} - h_{1,2,3},\\
&C_{10} = -h_3 + h_{1,3} + h_{2,3} - h_{1,2,3},\\
&C_{11} = -h_1 - h_4 + h_{1,2} + h_{1,3} - h_{1,2,3,4},\\
&C_{12} = -h_2 - h_4 + h_{1,2} + h_{2,3} - h_{1,2,3,4},\\
&C_{13} = -h_3 - h_4 + h_{1,3} + h_{2,3} - h_{1,2,3,4},\\
&C_{14} = h_{1,2,3} - h_{1,2,3,4}, C_{15} = -h_{1,2} + h_{1,2,3,4},\\
&C_{16} = -h_{1,3} + h_{1,2,3,4}, C_{17} = -h_{2,3} + h_{1,2,3,4},\\
&C_{18} = -h_{1,2,3} + h_{1,2,3,4}.
\end{aligned} \tag{6}$$

Step 4. Apply Algorithm 2 to the above problem to obtain a reduced problem:

Problem III.2$(*)$. Prove $F \geq 0$ subject to $\widetilde{\mathcal{R}}(J_1)$ and $a_i \geq 0,\ i \in \mathcal{N}_{18}$, where $F = a_6 + a_7 + a_{14}$ and $J_1 = \{f_i,\ i \in \mathcal{N}_9\}$.

Step 5. Since F is a conic combination of $a_i,\ i \in \mathcal{N}_{18}$, $\bar{F} \geq 0$ is provable.

The aforementioned problem is proved without solving any LPs. If we want to further find a reduced minimal problem, then we can apply Algorithm 6 and 7 to Problem III.2 to obtain the following LP that contains no implied equality and redundant inequality:

Prove F_1 subject to $\widetilde{R}(E_1)$ and $\mathcal{R}(V(\{F_1\} \cup E_1))$, where $E_1 = \{\tilde{f}_1, \tilde{f}_2\}$,

$$\begin{aligned}
&F_1 = a_6 - a_{11} + a_{12} + a_{13} + a_{17},\\
&\tilde{f}_1 = a_4 - a_6 + a_{11} - a_{12},\\
&\tilde{f}_2 = a_2 - a_6 + a_{11} - a_{13}.
\end{aligned} \tag{7}$$

Since $F_1 + \tilde{f}_2 = a_2 + a_{12} + a_{17}$, we show that $F_1 \geq 0$. Thus an explicit proof is given.

V. Two Applications

In this section, we will present two applications of our method. The first one is an information inequality proved in [24]. The second one is the example used in [23], in which we can significantly reduce the required computation for solving the LP.

A. Dougherty-Freiling-Zeger's Problem

The information theoretic inequality needs to be proved in [24] is specified by an LP with 8 random variables.

TABLE I

	Number of variables	Number of equality constraints	Number of Inequality constraints
Direct LP method	255	14	1800
ITIP	241	0	1800
LP obtained in [23]	206	0	1673
This work	no LP needs to be solved		

Problem $\mathbf{P_1}$: Prove $I(B;D,X,Z) \le I(W;A,B,C,D)$ under the constraints

$$\begin{array}{l} I(W;A,B,C,D) = I(X;A,B,W), \\ I(W;A,B,C,D) = I(Y;B,C,X), \\ I(W;A,B,C,D) = I(Z;C,D,Y), \\ I(A;B,C,D,Z) = I(B;D,X,Z), \\ I(B;A,D,W,Z) = I(B;D,X,Z), \\ I(C;A,D,W,Z) = I(B;D,X,Z), \\ I(D;A,B,C,Y) = I(B;D,X,Z), \\ I(C;A,W,Y) = I(B;D,X,Z), \\ I(B;A) = 0,\ I(C;A,B) = 0, \\ I(D;A,B,C) = 0,\ I(X;C,D|A,B,W) = 0, \\ I(Y;A,D,W|B,C,X) = 0, \\ I(Z;A,B,W,X|C,D,Y) = 0. \end{array} \tag{8}$$

We now solve **Problem $\mathbf{P_1}$** by using Procedure I.

Input:

Objective information inequality: $\bar{F} = I(W;A,B,C,D) - I(B;D,X,Z) \ge 0$.

Inequality Constraints: the elemental information inequalities generated by random variables A,B,C,D,X,Y,Z,W (totally 1800 inequalities).

Equality Constraints: totally 14 equalities in (8).

Step 1. The variable vector generated from A,B,C,D,X,Y,Z,W has $2^8 - 1$ elements (joint entropies). Transform $\bar{F}$ into the joint entropy form $\widetilde{F}$. Express the elemental information inequalities in terms of the joint entropies to obtain $\widetilde{C}_i,\ i \in \mathcal{N}_{1800}$. Likewise, express the equality constraints in (8) in terms of the joint entropies to obtain $\widetilde{C}_i = 0,\ i \in \mathcal{N}_{1814} \backslash \mathcal{N}_{1800}$.

Step 2. Apply Algorithm 1 to reduce $\{\widetilde{C}_i,\ i \in \mathcal{N}_{1800}\}$ by $\{\widetilde{C}_i, i \in \mathcal{N}_{1814} \backslash \mathcal{N}_{1800}\}$ to obtain $\{C_i \ge 0,\ i \in \mathcal{N}_{1793}\}$.

Step 3. Reduce $\widetilde{F}$ by $\{\widetilde{C}_i, i \in \mathcal{N}_{1814} \backslash \mathcal{N}_{1800}\}$ to obtain

$F_5 = h_{1,2,3} - h_{1,2,3,4} - h_{1,2,3,7} + h_{1,2,3,4,7} + h_{1,2,3,4,5,6,7} + h_8 - h_{1,2,3,4,5,6,7,8}$.

We need to solve

Problem IV.1$(*)$. Prove $F_5 \ge 0$ under the constraints $C_i \ge 0, i \in \mathcal{N}_{1793}$.

Step 4. Apply Algorithm 2 to obtain a reduced LP:

Problem III.2$(*)$. Prove $F \ge 0$ subject to $\widetilde{\mathcal{R}}(J_1)$ and $a_i \ge 0,\ i \in \mathcal{N}_{1793}$, where $J_1 = \{f_i,\ i \in \mathcal{N}_{1559}\}$.

Step 5. Apply Algorithm 3 and Algorithm 4 to the above problem successively. Algorithm 4 outputs 'SUCCESSFUL'. Thus $\bar{F} \ge 0$ is provable.

In other words, we show that F can be transformed to a conic combination of $\{a_i \ge 0,\ i \in \mathcal{N}_{1793}\}$ by $\{f_i = 0, i \in \mathcal{N}_{1559}\}$. This is given by

$$\begin{aligned} F_3 \ &= \tfrac{1}{2}(a_{24} + a_{28} + a_{35} + a_{129} + a_{185} + a_{520} \\ &+ a_{1048} + a_{1053} + a_{1187} + a_{1237} + a_{1556} + a_{1628} \\ &+ a_{1681} + a_{1782}). \end{aligned}$$

Thus $\bar{F} \ge 0$ is provable.

Table I shows the advantage of Procedure I for this example by comparing it with the Direct LP method induced by Theorem II.2, with ITIP, and with the procedure in [23]. Note that the procedure in [23] can reduce this example to the minimum LP to the greatest extent possible. However, we even do not need to solve LP by Procedure I in this work.

B. Tian's Problem

The framework of regenerating codes introduced in the seminal work of Dimakis *et al.* [16] addresses the fundamental tradeoff between the storage and repair bandwidth in erasure-coded distributed storage systems. In [17], a new outer bound on the rate region for $(4,3,3)$ exact-repair regenerating codes was obtained. This outer bound was proved by means of a computational approach built upon the LP framework in [1] for proving Shannon-type inequalities. The LP that needs to be solved, however, is exceedingly large. In order to make the computation manageable, Tian took advantage of the symmetry of the problem and other problem-specific structures to reduce the numbers of variables and constraints in the LP. This outer bound not only can provide a complete characterization of the rate region, but also proves the existence of a non-vanishing gap between the optimal tradeoff of the exact-repair codes and that of the functional-repair codes for the parameter set $(4,3,3)$. It was the first time that a non-trivial information theory problem was solved using this LP framework.[4]

In this work, we apply the results in the previous sections to Tian's problem, and give a simpler proof by our new method. We first give the abstract formulation of the problem.

[4]It was subsequently proved analytically by Sasidharan *et al.* [18] that the same holds for every parameter set.

Definition V.1. *A permutation π on the set $\mathcal{N}_4$ is a one-to-one mapping π: $\mathcal{N}_4 \to \mathcal{N}_4$. The collection of all permutations is denoted as $\prod$.*

In the problem formulation, we consider 16 random variables grouped into the following two sets:

$$\begin{aligned}\mathcal{W} &= \{W_1, W_2, W_3, W_4\},\\ \mathcal{S} &= \{S_{1,2}, S_{1,3}, S_{1,4}, S_{2,1}, S_{2,3}, S_{2,4}, S_{3,1}, S_{3,2}, S_{3,4},\\ &\quad S_{4,1}, S_{4,2}, S_{4,3}\}.\end{aligned}$$

A permutation π on $\mathcal{N}_4$ is applied to map one random variable to another random variable. For example, the permutation $\pi(1,2,3,4) = (2,3,1,4)$ maps the random variable W_1 to W_2. Similarly it maps the random variable $S_{i,j}$ to $S_{\pi(i),\pi(j)}$. When π is applied to a set of random variables, the permutation is applied to every random variable in the set. For example for the aforementioned permutation π, we have $\pi(W_1,\ S_{2,3}) = (W_2,\ S_{3,1})$.

The original problem is

Problem $\mathbf{P_6}$: Prove

$$4\alpha + 6\beta \geq 3B \tag{9}$$

under the constraints

- C1 $H(\pi(\mathcal{A}), \pi(\mathcal{B})) = H(\mathcal{A}, \mathcal{B})$, for any sets $\mathcal{A} \subseteq \mathcal{S}$ and $\mathcal{B} \subseteq \mathcal{W}$ and any permutation $\pi \in \prod$,
- C2 $H(\mathcal{W} \cup \mathcal{S}|\mathcal{A}) = 0$, any $\mathcal{A} \subseteq \mathcal{W} : |\mathcal{A}| = 3$,
- C3 $H(S_{i,j}|W_i) = 0,\ j \in \mathcal{N}_4,\ i \in \mathcal{N}_4\backslash\{j\}$,
- C4 $H(W_j|\{S_{i,j} \in \mathcal{S} : i \in \mathcal{N}_n\backslash\{j\}\}) = 0$, for any $j \in \mathcal{N}_4$,
- C5 $H(\mathcal{W} \cup \mathcal{S}) = B$,
- C6 $H(\mathcal{A}) = B$, for any $\mathcal{A}$ such that $|\mathcal{A} \cap \mathcal{W}| \geq 3$,
- C7 $H(W_i) \leq \alpha$, $W_i \in \mathcal{W}$,
- C8 $H(S_{i,j}) \leq \beta$, $S_{i,j} \in \mathcal{S}$.

For this specific problem, the random variables involved exhibit strong symmetry due to the setup of the problem. To reduce the scale of the problem, Tian proved in [17, Section III-B] that only a subset of the random variables in $\mathcal{W} \cup \mathcal{S}$ is needed for solving **Problem $\mathbf{P_6}$**. A similar idea was also used in [19], [20].

According to Tian's proof in Section III-B of [17], **Problem $\mathbf{P_6}$** can be reduced to the following simpler problem,

Problem $\mathbf{P_7}$: Prove

$$4\alpha + 6\beta \geq 3B \tag{10}$$

under the constraints: C1, C3, C4, C6, C7 and C8 on the 12 random variables in the set

$$\begin{aligned}\mathcal{W}_1 \cup \mathcal{S}_1 &= \{W_1, W_2, W_4\} \cup \{S_{2,1}, S_{3,1}, S_{4,1}, S_{1,2}, S_{3,2},\\ &\quad S_{4,2}, S_{1,4}, S_{2,4}, S_{3,4}\}.\end{aligned}$$

Remark V.1. *In the following computation, in order to simplify the notation, we will use, for example, $h_{1,2,3,4,5,6,7,8,9,10,11,12}$ to represent the joint entropy $H(W_1, W_2, W_4, S_{2,1}, S_{3,1}, S_{4,1}, S_{1,2}, S_{3,2}, S_{4,2}, S_{1,4}, S_{2,4}, S_{3,4})$. Similarly, we will use h_1 to represent $H(W_1)$, $h_{2,5}$ to represent $H(W_2, S_{3,1})$, so on and so forth.*

We now solve **Problem $\mathbf{P_7}$** by using Procedure I.

Input:

Objective information inequality: $\bar{F} = 4\alpha + 6\beta - 3B \geq 0$.
Inequality Constraints: the elemental information inequalities generated by random variables $\mathcal{W}_1 \cup \mathcal{S}_1$ (total 67596 inequalities); C7 and C8 (total 12 inequalities).
Equality Constraints: C1, C3, C4 and C6 (total 22945 equalities).

Step 1. The variable vector generated from $\mathcal{W}_1 \cup \mathcal{S}_1$ has $2^{12} - 1$ elements (joint entropies). Express C7, C8 and the elemental information inequalities in terms of the joint entropies to obtain $\widetilde{C}_i,\ i \in \mathcal{N}_{67608}$. According to conditions C1, C3, C4 and C6, write equality constraints in joint entropy forms: $\widetilde{C}_i = 0,\ i \in \mathcal{N}_{90553}\backslash\mathcal{N}_{67608}$.

Step 2. Apply Algorithm 1 to reduce $\{\widetilde{C}_i,\ i \in \mathcal{N}_{67608}\}$ by $\{\widetilde{C}_i, i \in \mathcal{N}_{90553}\backslash\mathcal{N}_{67608}\}$ to obtain $\{C_i \geq 0,\ i \in \mathcal{N}_{10189}\}$.

Step 3. Reduce $\widetilde{F}$ by $\{\widetilde{C}_i, i \in \mathcal{N}_{90553}\backslash\mathcal{N}_{67608}\}$ to obtain $F_5 = 4\alpha + 6\beta - 3h_{2,3,5,6,7,8,10}$.

We need to solve

Problem IV.1$(*)$. Prove $F_5 \geq 0$ under the constraints $C_i \geq 0, i \in \mathcal{N}_{10189}$.

Step 4. Apply Algorithm 2 to obtain a reduced LP:

Problem III.2$(*)$. Prove $F \geq 0$ subject to $\widetilde{\mathcal{R}}(J_1)$ and $a_i \geq 0,\ i \in \mathcal{N}_{10189}$, where $J_1 = \{f_i,\ i \in \mathcal{N}_{9859}\}$.

Step 5. Apply Algorithm 3 and Algorithm 4 to the above problem successively. Algorithm 4 outputs 'SUCCESSFUL'. Thus $\bar{F} \geq 0$ is provable.

In other words, we show that F can be transformed to a conic combination of $\{a_i,\ i \in \mathcal{N}_{10189}\}$ by $\{f_i = 0, i \in \mathcal{N}_{9859}\}$. This is given by

$$\begin{aligned}F_3 &= 7a_6 + 2a_{85} + 4a_{94} + a_{119} + a_{167} + 3a_{169} + 4a_{211}\\ &\quad + a_{223} + a_{290} + a_{335} + a_{340} + a_{353} + 4a_{450} + 4a_{484}\\ &\quad + a_{519} + a_{667} + a_{727} + 4a_{735} + a_{819} + a_{820} + a_{827}\\ &\quad + a_{829} + a_{859} + a_{868} + 3a_{906} + a_{916} + 4a_{10188}\\ &\quad + 6a_{10189}.\end{aligned} \tag{11}$$

The formulas used above is listed in Appendix B.

Table II shows the advantage of Procedure I for Tian's problem by comparing it with the Direct LP method induced by Theorem II.2, ITIP, Tian's method in [17], and our previous work in [23].

Note that even if Algorithm 4 cannot obtain a conic combination as desired, it can still reduce the problem to the minimal LP in a shorter time and with less memory compared with our previous work [23]. Table III shows the advantage of Procedure I for reducing Tian's problem by comparing it with the procedure in [23]. In Table III, "Time" and "Memory" refer to the time and memory it takes to simplify the original LP to the minimal LP, respectively. The experiment results are obtained by MAPLE running on a desktop PC with an i7-6700 Core, 3.40GHz CPU and 16G memory.

TABLE II

	Number of variables	Number of Equality constraints	Number of Inequality constraints
Direct LP method	4098	22945	67608
ITIP	600	0	67608
Tian's Method	176	0	6152
LP in [23]	101	0	649
This work	no LP needs to be solved		

TABLE III

	Number of variables	Number of Inequality constraints	Time	Memory
LP in [23]	101	649	23000s	900M
LP in this work	101	649	33s	70M

VI. Conclusion and discussion

In this paper, we have developed a heuristic method to prove information inequalities and identities. This method does not prove an information inequalities or identities by directly solving the associated LP, but rather expedites the explicit proof process through polynomial reduction. The method may not succeed in proving the inequality or identity every time. If it does not succeed, it can simplify the original LP into a smaller LP. We have given several examples to verify the effectiveness of our method. It is observed from these examples that the average complexity of our method is polynomial in the dimension of the entropy vector, while the complexity of ITIP and most subsequent works based on linear programming are estimated to be exponential in the dimension of the entropy vector, and the complexity of the method proposed in [22] is roughly between the above two. Experiments have shown that for most problems with equality constraints, we can have not only one non-negative representation, but many non-negative representations. Based on this fact, we can obtain a non-negative representation with very few attempts using Algorithm 4, which is also verified by the experimental results in Section V.

As discussed in Section III, since different elimination choices of variables can lead to different results, our heuristic method (Algorithm 4) 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. Next, we summarize in Table IV some experimental results on the effectiveness of Algorithm 4 for solving various problems. In the table, "TS" denotes the number of times we need to run Algorithm 4 to obtain a successful result, "TSH" denotes the number of times out of one hundred runs[5] of Algorithm 4 that are successful, and "Time" denotes the total time required to repeatedly run Algorithm 4 to obtain a successful result.

[5]Note that the one hundred attempts here may contain replications, but as mentioned in Remark III.2, the probability of this is very low. For details, refer to Remark III.2.

TABLE IV

	TS	TSH	Time
Example III.5	2	52	2s
Example IV.1	1	100	0.2s
Dougherty-Freiling-Zeger's problem	3	32	11s
Tian's problem	12	10	80s

The data given in Table IV is for reference only. The problems listed in the table are all solvable. For problems that are not solvable, we have to use Algorithm 5 to solve an LP which typically has a much smaller size compared with the original problem, and Algorithm 5 will output 'FALSE'. Compared with [23], the method here for obtaining the reduced minimal characterization set is considerably simpler.

To end this paper, we put forth the following conjecture on the effectiveness of Algorithm 4:

Conjecture. If the problem is solvable, then there exists at least one ordering of the variables such that Algorithm 4 outputs 'SUCCESSFUL'.

Appendices

A. *Two enhancements of Algorithm 3*

In this section, we present two algorithms as enhancements of Algorithm 3. We call Problem III.3 a pure LP if it contains no implied equality, and call it a minimal LP if it contains no redundant inequality.

First, we give a general algorithm for reducing Problem III.3 to a pure LP.

Algorithm 6 Pure LP Algorithm

Input: Problem III.3.

Output: A pure LP.

Use Algorithm 3 to reduce Problem III.3 to
Problem III.5. Proving $F_1 \geq 0$ subject to $\widetilde{\mathcal{R}}(E_1)$ and $\mathcal{R}(S_1)$.
for i from 1 to m **do**
  Apply Algorithm 4 to solve the following LP:
    Problem VI.1. *Prove* $-a_i \geq 0$ *subject to* $\widetilde{\mathcal{R}}(E_1)$ *and* $\mathcal{R}(S_1)$.
  if Algorithm 4 outputs 'SUCCESSFUL' **then**
    $E_1 := subs(a_i = 0, E_1)\backslash\{0\}$.
    $S_1 := S_1\backslash\{a_i\}$.
    $F_1 := subs(a_i = 0, F_1)$.
  else
    Algorithm 4 outputs a reduced LP. Apply Algorithm 5 to solve this LP.
    if Algorithm 5 outputs 'TRUE' **then**
      $E_1 := subs(a_i = 0, E_1)\backslash\{0\}$.
      $S_1 := S_1\backslash\{a_i\}$.
      $F_1 := subs(a_i = 0, F_1)$.
    end if
  end if
end for
Reduce F_1 by E_1 to obtain the remainder F_2 and the RREF of E_1, E_2.
return A pure LP:
  Problem VI.2. *Prove* F_2 *subject to* $\widetilde{R}(E_2)$ *and* $\mathcal{R}(V(\{F_2\} \cup E_2))$.

Next, we give a general algorithm for finding a minimal LP from Problem III.3.

Algorithm 7 Minimal LP Algorithm

Input: Problem III.3.
Output: A minimal LP.

Run Algorithm 3 to reduce Problem III.3 to
Problem III.5. Prove $F_1 \geq 0$ subject to $\widetilde{\mathcal{R}}(E_1)$ and $\mathcal{R}(S_1)$.
for i from 1 to m **do**
  Let $\bar{S}_a = S_1\backslash\{a_i\}$.
  if $a_i \in V(f)$ for some $f \in E_1$ **then**
    Solve a_i from $f = 0$ to obtain $a_i = A_i$.
  else
    Let A_i be a_i.
  end if
  $\bar{E}_a := subs(a_i = A_i, E_1)\backslash\{0\}$.
  Run Algorithm 4 to solve the following LP:
    Problem VI.3. *Prove* $A_i \geq 0$ *subject to* $\widetilde{\mathcal{R}}(\bar{E}_a)$ *and* $\mathcal{R}(\bar{S}_a)$.
  if Algorithm 4 outputs 'SUCCESSFUL' **then**
    $E_1 := subs(a_i = A_i, E_1)\backslash\{0\}$.
    $S_1 := S_1\backslash\{a_i\}$.
    $F_1 := subs(a_i = A_i, F_1)$.
  else
    Algorithm 4 outputs a reduced LP. Apply Algorithm 5 to this LP.
    if Algorithm 5 outputs 'TRUE' **then**
      $E_1 := subs(a_i = A_i, E_1)\backslash\{0\}$.
      $S_1 := S_1\backslash\{a_i\}$.
      $F_1 := subs(a_i = A_i, F_1)$.
    end if
  end if
end for
Reduce F_1 by E_1 to obtain the remainder F_2 and the RREF of E_1, E_2.
return A minimal LP:
  Problem VI.4. *Proving* F_2 *subject to* $\widetilde{\mathcal{R}}(E_2)$ *and* $\mathcal{R}(V(\{F_2\} \cup E_2))$.

Next, we give the detailed steps of the reduction from Problem III.7($*$) to Problem III.8($*$).

// We first follow Algorithm 6.

Step 1. Use Algorithm 3 to reduce Problem III.7($*$) to

Problem III.5($*$). Proving $F_1 \geq 0$ subject to $\widetilde{\mathcal{R}}(E_1)$ and $\mathcal{R}(S_1)$, where $F_1 = -\frac{1}{2}(a_2 - a_4 - 3a_9 - a_{10} + a_{11} + a_{12})$, $E_1 = \{a_1 + a_2 - a_4 + a_9 + a_{10} - a_{11} - a_{12}, a_3 + a_9 + a_{10} - a_{11} - a_{12}, a_6 - a_9 - a_{10} + a_{11} + a_{12}\}$, and $S_1 = \{a_1, a_2, a_3, a_4, a_6, a_9, a_{10}, a_{11}, a_{12}\}$.

Step 2. For $i \in \mathcal{N}_{12}\backslash\{5, 7, 8\}$, run Algorithm 4 to solve the following LP:

Problem VI.1($*$). Prove $-a_i \geq 0$ subject to $\widetilde{\mathcal{R}}(E_1)$ and $\mathcal{R}(S_1)$.

Step 3. Algorithm 4 outputs 'SUCCESSFUL' when $i = 3$, then let

$E_1 = subs(a_3 = 0, E_1)\backslash\{0\} = \{a_1 + a_2 - a_4 + a_9 + a_{10} - a_{11} - a_{12}, a_9 + a_{10} - a_{11} - a_{12}, a_6 - a_9 - a_{10} + a_{11} + a_{12}\}$,

$S_1 = S_1\backslash\{a_3\} = \{a_1, a_2, a_4, a_6, a_9, a_{10}, a_{11}, a_{12}\}$,

$F_1 = subs(a_i = 0, F_1) = -\frac{1}{2}(a_2 - a_4 - 3a_9 - a_{10} + a_{11} + a_{12})$.

Step 4. Algorithm 4 outputs 'SUCCESSFUL' when $i = 6$, then let

$E_1 = subs(a_6 = 0, E_1)\backslash\{0\} = \{a_1 + a_2 - a_4 + a_9 + a_{10} - a_{11} - a_{12}, a_9 + a_{10} - a_{11} - a_{12}, -a_9 - a_{10} + a_{11} + a_{12}\}$,

$S_1 = S_1\backslash\{a_6\} = \{a_1, a_2, a_4, a_9, a_{10}, a_{11}, a_{12}\}$,

$F_1 = subs(a_i = 0, F_1) = -\frac{1}{2}(a_2 - a_4 - 3a_9 - a_{10} + a_{11} + a_{12})$.

// For $i \in \mathcal{N}_{12}\backslash\{3, 5, 6, 7, 8\}$, Algorithm 4 outputs 'UNSUCCESSFUL' and Algorithm 5 outputs 'FALSE'.

Step 5. Reduce F_1 by E_1 to obtain

Problem VI.2($*$). Prove F_2 subject to $\widetilde{R}(E_2)$ and $\mathcal{R}(V(\{F_2\} \cup E_2))$,

where $F_2 = -\frac{1}{2}a_2 + \frac{1}{2}a_4 - a_{10} + a_{11} + a_{12}$ and $E_2 = \{a_1 + a_2 - a_4, a_9 + a_{10} - a_{11} - a_{12}\}$.

// Next, we will follow Algorithm 7.

Step 6. Use Algorithm 3 to reduce Problem VI.2($*$) to **Problem III.5**($*$). Proving $F_1 \geq 0$ subject to $\widetilde{\mathcal{R}}(E_1)$ and $\mathcal{R}(S_1)$,

where $F_1 = \frac{1}{2}a_1 - a_{10} + a_{11} + a_{12}$, $E_1 = \{a_9 + a_{10} - a_{11} - a_{12}\}$, and $S_1 = \{a_1, a_9, a_{10}, a_{11}, a_{12}\}$.

// Now we obtain Problem III.8 in Example III.5.

B. *Formulas in* (11)

In this section, we list the formulas used in (11).

$$
\begin{aligned}
&2h_{11} - h_{11,12} = a_6,\\
&h_{6,7,8,9,10,11,12} - h_{2,3,8,9,10,11,12} = a_{85},\\
&2h_{11,12} - h_{3,9,10,11,12} - h_{11} = a_{94},\\
&2h_{3,8,9,11} - h_{2,3,6,9,10,12} - h_{3,9,12} = a_{119},\\
&h_{3,9} + h_{9,12} - h_{3,8,9} - h_{11} = a_{167},\\
&h_{3,9} + h_{11,12} - h_{3,8,9} - h_{11} = a_{169},\\
&h_{3,8,9} + h_{8,9,10,12} - h_{3,5,7,9} - h_{8,10,12} = a_{211},\\
&h_{3,9,12} + h_{6,7,9,10} - h_{1,5,10,12} - h_{6,9,11} = a_{223},\\
&h_{8,11,12} + h_{6,9,11} - h_{6,7,9,10} - h_{9,12} = a_{290},\\
&h_{1,5,10,12} + h_{6,7,9,10,11} - h_{3,8,9,11,12} - h_{6,7,9,10} = a_{335},\\
&h_{2,3,11,12} + h_{3,8,10,11} - h_{2,3,8,10,11} - h_{3,8,9,11} = a_{340},\\
&h_{2,3,11,12} + h_{3,5,6,7,10,12} - h_{2,3,8,11,12} - h_{3,5,6,7,9,10} = a_{353},\\
&h_{3,8,10,12} + h_{3,5,7,9} - h_{2,3,11,12} - h_{8,9,10,12} = a_{450},\\
&h_{3,9,11,12} + h_{8,10,12} - h_{3,8,10,12} - h_{11,12} = a_{484},\\
&h_{6,7,9,10} + h_{8,9,11,12} - h_{6,7,9,10,11} - h_{8,11,12} = a_{519},\\
&h_{2,3,8,11,12} + h_{3,5,7,9,10,11,12} - h_{2,3,8,9,10,11,12}\\
&\quad - h_{3,5,6,7,10,12} = a_{667},\\
&h_{3,8,9,11,12} + h_{6,8,9,11,12} - h_{3,5,6,7,10,12} - h_{8,9,11,12} = a_{727},\\
&h_{3,9,10,11,12} + h_{3,8,10,12} - h_{8,9,10,11,12} - h_{3,9,11,12} = a_{735},\\
&h_{8,9,10,11,12} + h_{3,5,6,7,10,12} - h_{3,5,7,9,10,11,12}\\
&\quad - h_{3,8,10,12} = a_{819},\\
&h_{8,9,10,11,12} + h_{3,5,6,7,10,12} - h_{3,5,7,9,10,11,12}\\
&\quad - h_{3,8,9,11,12} = a_{820},\\
&h_{8,9,10,11,12} + h_{2,3,8,9,10,11,12} - h_{6,7,8,9,10,11,12}\\
&\quad - h_{3,8,10,11} = a_{827},\\
&h_{2,3,6,9,10,12} + h_{2,3,8,10,11} - h_{2,3,8,9,10,11,12}\\
&\quad - h_{2,3,11,12} = a_{829},\\
&h_{3,5,6,7,9,10} + h_{3,8,9,11,12} - h_{3,5,6,7,10,12} - h_{3,8,9,11} = a_{859},\\
&h_{3,5,6,7,10,12} + h_{8,9,10,11,12} - h_{2,3,8,9,10,11,12}\\
&\quad - h_{6,8,9,11,12} = a_{868},\\
&h_{2,3,8,9,10,11,12} + h_{2,3,11,12} - h_{2,3,5,6,7,8,10}\\
&\quad - h_{3,8,10,12} = a_{906},\\
&h_{3,5,7,9,10,11,12} + h_{2,3,8,9,10,11,12} - h_{6,7,8,9,10,11,12}\\
&\quad - h_{3,5,6,7,10,12} = a_{916},\\
&\alpha - h_{3,9} = a_{10188}, \quad \beta - h_{11} = a_{10189}.
\end{aligned}
$$

ACKNOWLEDGMENT

The authors would like to thank Tao Guo, Fangwei Ye, Cheuk Ting Li and Chao Tian for their invaluable comments and also would like to thank the associate editor and the anonymous reviewers for their suggestions, which greatly improved the paper. This work is partially supported by a research grant of the National Natural Science Foundation of China (12301647, 11688101), and the Research Grants Council of the Hong Kong (14207219, CUHK SRFS2223-4S03).

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 C. T. Li, “Machine-Proving of Entropy Inequalities,” IEEE BITS the Information Theory Magazine, vol. 1, no. 1, pp. 12-22, 1 Sept. 2021.
[3] 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
[4] R. Pulikkoonattu and S. Diggavi (2006), Xitip, ITIP-Based C Program Software Package. [Online]. Available: http://xitip.epfl.ch
[5] L. Csirmaz (2016), A MINimal Information Theoretic Inequality Prover (Minitip). [Online]. Available: https://github.com/lcsirmaz/minitip
[6] C. T. Li (2020), Python Symbolic Information Theoretic Inequality Prover (psitip). [Online]. Available: https://github.com/cheuktingli/
[7] 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
[8] 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.
[9] S.-W. Ho, L. Ling, C. W. Tan, and R. W. Yeung (2020), AITIP. [Online]. Available: https://github.com/convexsoft/AITIP
[10] R. W. Yeung, *Information Theory and Network Coding*. New York, NY, USA: Springer, 2008.
[11] R. W. Yeung, T. T. Lee and Z. Ye, ”Information-theoretic characterizations of conditional mutual independence and Markov random fields,” *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1996-2011, July 2002.
[12] Z. Zhang and R. W. Yeung, “On characterization of entropy function via information inequalities,” *IEEE Trans. Inf. Theory*, vol. 44, pp. 1440-1452, July 1998.
[13] T. Chan, S. Thakor, and A. Grant, “Minimal characterization of Shannon-type inequalities under functional dependence and full conditional independence structures,” *IEEE Trans. Inf. Theory*, vol. 65, no. 7, pp. 4041-4051, Jul. 2019.
[14] R. W. Yeung, “A new outlook on Shannon’s information measures,” *IEEE Trans. Inform. Theory*, vol. 37, pp. 466-74, May 1991.
[15] D. C. Lay, *Linear Algebra and Its Applications*, 5th Edition. Pearson, 2016.
[16] 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.
[17] 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.
[18] B. Sasidharan, N. Prakash, M. N. Krishnan, M. Vajha, K. Senthoor, P. V. Kumar, “Outer bounds on the storage-repair bandwidth trade-off of exact-repair regenerating codes,” *International Journal of Information and Coding Theory*, vol. 3, no. 4, pp: 255-298, 2016.
[19] C. Tian, “A note on the rate region of exact-repair regenerating codes”. arXiv:1503.00011, Mar. 2015.
[20] W. Chen, C. Tian, “A New Approach to Compute Information Theoretic Outer Bounds and Its Application to Regenerating Codes”. arXiv:2205.01612, 2022.
[21] A. Schrijver, *Combinatorial optimization: polyhedra and efficiency*. Berlin: Springer, 2003.
[22] A. Ben-Tal, A. Nemirovski, *Lecture notes: optimization III*, New York, NY, USA: Springer, 2022.
[23] L. Guo, R. W. Yeung and X. -S. Gao, ”Proving Information Inequalities and Identities with Symbolic Computation,” *IEEE Trans. Inform. Theory*, vol. 69, no. 8, pp. 4799-4811, Aug. 2023.
[24] R. Dougherty, C. Freiling and K. Zeger, ”Networks, Matroids, and Non-Shannon Information Inequalities,” *IEEE Trans. Inform. Theory*, vol. 53, no. 6, pp. 1949-1969, 2007.

[25] C. Lassez and J.-L. Lassez, "Quantifier elimination for conjunctions of linear constraints via a convex hull algorithm," Symbolic and Numerical Computation for Artificial Intelligence, pp. 103–122, 1992.

[26] J. Apte, J.M. Walsh, "Exploiting symmetry in computing polyhedral bounds on network coding rate regions," in Proceedings of the International Symposium on Network Coding (NetCod), June 2015, pp. 76–80.

[27] C. Tian, "Symmetry, outer bounds, and code constructions: A computer-aided investigation on the fundamental limits of caching," MDPI Entropy, vol. 20, no. 8, pp 603.1-43, Aug. 2018.

[28] W. Xu, J. Wang, and J. Sun, "A projection method for derivation of nonShannon-type information inequalities," in Proc. IEEE ISIT, Jul. 2008, pp. 2116–2120.

[29] C. T. Li, "An Automated Theorem Proving Framework for Information-Theoretic Results," *IEEE Trans. Inform. Theory*, vol. 69, no. 11, pp. 6857-6877, Nov. 2023.

Laigang Guo (Member, IEEE) received the Ph.D. degree in applied mathematics from the University of Chinese Academy of Sciences, Beijing, China, in 2019. He joined the National Center for Mathematics and Interdisciplinary Sciences, Chinese Academy of Sciences and the Institute of Network Coding, The Chinese University of Hong Kong as a postdoctoral fellow in 2019 and 2021, respectively. Currently, he is an assistant professor with the School of Mathematical Sciences, Beijing Normal University. His research interests are symbolic computation methods in information theory and nonlinear systems. He was a recipient of the president award of Chinese Academy of Sciences.

Raymond W. Yeung (Fellow, IEEE) was born in Hong Kong, in 1962. He received the B.S., M.Eng., and Ph.D. degrees in electrical engineering from Cornell University, Ithaca, NY, USA, in 1984, 1985, and 1988, respectively. He was on leave at the Ecole Nationale Supérieure des Télécommunications, Paris, France, in Fall 1986. He was a member of Technical Staff at AT&T Bell Laboratories from 1988 to 1991. He has held visiting positions at Cornell University, Nankai University, Bielefeld University, Copenhagen University, the Tokyo Institute of Technology, the Munich University of Technology, and Columbia University. Since 1991, he has been with The Chinese University of Hong Kong, where he is currently a Choh-Ming Li Professor of information engineering and the Co-Director of the Institute of Network Coding. He was a Consultant in a project of Jet Propulsion Laboratory, Pasadena, CA, USA, for salvaging the malfunctioning Galileo Spacecraft and NEC. His 25-bit synchronization marker was used onboard the Galileo Spacecraft for image synchronization. He is the author of the textbooks *A First Course in Information Theory* (Kluwer Academic/Plenum 2002) and its revision Information Theory and Network Coding (Springer 2008), which have been adopted by over 100 institutions around the world. This book has also been published in Chinese (Higher Education Press 2011, translation by Ning Cai et al.). He has coauthored with Shenghao Yang the monograph BATS Codes: Theory and Applications (Morgan & Claypool Publishers, 2017). In Spring 2014, he gave the first MOOC on information theory that reached over 25,000 students. His research interests include information theory and network coding.

Dr. Yeung is a fellow of the Hong Kong Academy of Engineering Sciences and the Hong Kong Institution of Engineers. He was a member of the Board of Governors of the IEEE Information Theory Society from 1999 to 2001. He has served on the committees of a number of information theory symposiums and workshops. He was a recipient of the Croucher Foundation Senior Research Fellowship from 2000 to 2001, the 2005 IEEE Information Theory Society Paper Award, the Friedrich Wilhelm Bessel Research Award of the Alexander von Humboldt Foundation in 2007, the 2016 IEEE Eric E. Sumner Award (for pioneering contributions to the field of network coding), the 2018 ACM SIGMOBILE Test-of-Time Paper Award, the 2021 IEEE Richard W. Hamming Medal (for fundamental contributions to information theory and pioneering network coding and its applications), and the 2022 Claude E. Shannon Award. In 2015, he was named (together with Zhen Zhang) an Outstanding Overseas Chinese Information Theorist by the China Information Theory Society. In 2018, with Shenghao Yang he co-founded n-hop technologies in Hong Kong that has successfully deployed BATS code in the Hong Kong Government's pilot smart lamppost system for wireless multi-hop transmission of sensor data. In 2019, his team won a Gold Medal with Congratulations of the Jury at the 47th International Exhibition of Inventions of Geneva for their invention "BATS: Enabling the Nervous System of Smart Cities." He was the General Chair of the First and the Fourth Workshops on Network, Coding, and Applications (NetCod 2005 and 2008), the Technical Co-Chair of the 2006 IEEE International Symposium on Information Theory and the 2006 IEEE Information Theory Workshop, Chengdu, China, and the General Co-Chair of the 2015 IEEE International Symposium on Information Theory. He currently serves as an Editor-at-Large for Communications in Information and Systems and an Editor of Foundation and Trends in Communications and Information Theory and Foundation and Trends in Networking. He was an Associate Editor for Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY from 2003 to 2005. From 2011 to 2012, he was a Distinguished Lecturer of the IEEE Information Theory Society.

Xiao-Shan Gao (Senior member, IEEE) received the PhD degree from the Chinese Academy of Sciences, Beijing, China, in 1988. Currently, he is a Professor with the Institute of Systems Science, Chinese Academy of Sciences, Beijing, China. His research interests include automated reasoning, symbolic computation, intelligent CAD and CAGD, and robotics. He is the Chief Editor of the Journal of System Science and Complexity, and the Editorial Board of the Journal of Symbolic Computation. He has published over 90 research papers, two monographs and edited four books or conference proceedings. He was the recipient of many awards, including the First Prize of Natural Science of the Chinese Academy of Sciences, the Second Prize of National Natural Science, and the ISSAC2011 Outstanding Paper Award issued by ACM SIGSAM.