Proving Information Inequalities and Identities with Symbolic Computation

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

***Abstract*—Proving linear inequalities and identities of Shannon's information measures, possibly with 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). In particular, an identity $f = 0$ is verified by solving two LPs, one for $f \geq 0$ and one for $f \leq 0$. In this paper, we develop a set of algorithms that can be implemented by symbolic computation. Based on these algorithms, procedures for verifying linear information inequalities and identities are devised. Compared with LP-based algorithms, our procedures can produce analytical proofs that are both human-verifiable and free of numerical errors. Our procedures are also more efficient computationally. For constrained inequalities, by taking advantage of the algebraic structure of the problem, the size of the LP that needs to be solved can be significantly reduced. For identities, instead of solving two LPs, the identity can be verified directly with very little computation.**

***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 [8] is a cloud-based 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.

Using the above LP-based approach, to prove an information identity $f = 0$, two LPs need to be solved, one for the inequality $f \geq 0$ and the other for 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 [10][12].

The LP-based approach is in general not computationally efficient because it does not take advantage of the special structure of the underlying LP. In this paper, we take a different approach. Instead of transforming the problem into a general LP to be solved numerically, we develop algorithms that can be implemented by symbolic computation, and based on these algorithms, procedures for proving information inequalities and identities are devised. Our specific contributions are:

1) Analytical proofs for information inequalities and identities that are free of numerical errors can be produced.
2) Compared with the LP-based approach, the computational efficiency of our procedure is in general much higher.
3) Information identities can be proved directly with very little computation instead of having to solve 2 LPs.

The rest of the paper is organized as follows. In Section II, we present the preliminaries for information inequalities. In Section III, we develop algorithms for simplifying a set of linear inequalities subject to linear inequality and equality constraints. In Section IV, the procedures for proving information inequalities and identities are presented. Two

This article was presented in part at the 2022 IEEE International Symposium on Information Theory. (Corresponding author: Raymond W. Yeung.)

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).

examples and an application are given in Sections V and VI respectively, to illustrate our procedures. Section VII concludes the paper.

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], [9, 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 equalities involving random variables X and Y, imply

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

again a basic equality involving X and Y. In order to eliminate such redundancies, the minimal subset of the basic inequalities was found in [1].

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 $\mathcal{N}_n$ 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 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

$$m = n + \sum_{r=0}^{n-2} \left(\begin{array}{c} n \\ r \end{array} \right) \left(\begin{array}{c} n-r \\ 2 \end{array} \right) = n + \left(\begin{array}{c} n \\ 2 \end{array} \right) 2^{n-2}.$$

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 the 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*

$$\left\{ \begin{array}{rcl} I(X;Z|Y) & = & 0 \\ I(X,Y;T|Z) & = & 0, \end{array} \right.$$

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 a set of linear constraint(s) on the 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., [11]. See [9, Ch. 15] for a discussion.

Shannon's information measures, with conditional mutual informations 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 $m \times (2^n - 1)$ 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 have zero row. 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. Here, $\mathbf{h} \in \mathbb{R}^{2^n-1}$ *is the variable vector.*

III. Linear inequalities and related algorithms

In this section, we will develop some algorithms for simplifying a linear inequality set constrained by a linear equality set. These algorithms will be used as building blocks for the procedures to be developed in Section IV 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 establish some related properties that are instrumental for developing the aforementioned algorithms. For some details, one can refer to [23], [26].

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 inequality sets have the form $S_f = \{f_i \geq 0, i \in \mathcal{N}_m\}$, with $f_i \neq 0$ and $f_i \in \mathbb{R}_h[\mathbf{x}]$, and all the equality sets have the form $E_{\tilde{f}} = \{\tilde{f}_i = 0, i \in \mathcal{N}_{\tilde{m}}\}$ with $\tilde{f}_i \neq 0$ 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 included 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:*

$$\mathbf{x} \textit{ satisfying } f_1 \geq 0, f_2 \geq 0, \ldots, f_m \geq 0 \textit{ implies } \mathbf{x} \textit{ satisfies } f \geq 0 \textit{ for all } \mathbf{x}.$$

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

Definition III.5. *Two inequalities* $f \geq 0$ *and* $g \geq 0$ *are trivially equivalent if* $f = c\,g$ *for some* $c \in \mathbb{R}_{>0}$*. Given two sets of inequalities* $S_f = \{f_i \geq 0, i \in \mathcal{N}_{m_1}\}$ *and* $S_g = \{g_i \geq 0, i \in \mathcal{N}_{m_2}\}$*, we say that* S_f *and* S_g *are trivially equivalent if*

1) *S_f and S_g have exactly the same number of inequalities;*
2) *for every $i \in \mathcal{N}_{m_1}$, $f_i \geq 0$ is trivially equivalent to $g_j \geq 0$ for some $j \in \mathcal{N}_{m_2}$;*
3) *for every $i \in \mathcal{N}_{m_2}$, $g_i \geq 0$ is trivially equivalent to $f_j \geq 0$ for some $j \in \mathcal{N}_{m_1}$.*

Furthermore, if S_f *and* S_g *are trivially equivalent, then we regard* S_f *and* S_g *as the same set of inequalities.*

Lemma III.1 (Farkas' Lemma[13], [14])**.** *Let* $\mathbf{A} \in \mathbb{R}^{m\times n}$ *and* $\mathbf{b} \in \mathbb{R}^n$*. Then exactly one the following two assertions is true:*

1. There exists an $\mathbf{x} \in \mathbb{R}^n$ *such that* $\mathbf{A}\mathbf{x} \geq 0$ *and* $\mathbf{b}^T\mathbf{x} < 0$.

2. There exists a $\mathbf{y} \in \mathbb{R}^m$ *such that* $\mathbf{A}^T\mathbf{y} = \mathbf{b}$ *and* $\mathbf{y} \geq 0$.

Lemma III.2 ([26])**.** *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* $h_1, \ldots, h_m$.

Note that this lemma generalizes Theorem 2 in [1].

Definition III.6. *Let* $f(\mathbf{x})$ *and* $g(\mathbf{x})$ *be two homogeneous linear polynomials. We say* $f(\mathbf{x}) \equiv g(\mathbf{x})$ *if and ony if* $f(\mathbf{x}) = g(\mathbf{x})$ *for any* $\mathbf{x} \in \mathbb{R}^n$.

Definition III.7. *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.3. *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* $i \in \mathcal{N}_m \backslash \{k\}$.

Proof. Assume (1) holds and let $\mathbf{x}$ be any solution of S_f. Then $f_k(\mathbf{x}) = \sum\limits_{i=1, i\neq k}^{m} p_i f_i(\mathbf{x}) \leq 0$ since $p_i \leq 0$ and $f_i(\mathbf{x}) \geq 0$, for $i \in \mathcal{N}_m \backslash \{k\}$. On the other hand, from $f_k(\mathbf{x}) \geq 0$, we obtain $f_k(\mathbf{x}) = 0$. Therefore, $f_k(\mathbf{x}) = 0$ for all solution $\mathbf{x}$ of S_f, i.e., $f_k = 0$ is an implied equality of S_f.

Now, assume that $f_k = 0$ is an implied equality of S_f, i.e., $f_k(\mathbf{x}) = 0$ for all solution $\mathbf{x}$ of S_f. This implies that if $\mathbf{x}$ is a solution of S_f, then $f_k(\mathbf{x}) \leq 0$. In other words, the inequality $f_k(\mathbf{x}) \leq 0$ is implied by the S_f. By Lemma III.2,

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.2023.3263178

there exist $q_i \geq 0$, $i \in \mathcal{N}_m$ such that

$$-f_k(\mathbf{x}) \equiv \sum_{i=1}^{m} q_i f_i(\mathbf{x}).$$

Then,

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

or

$$f_k(\mathbf{x}) \equiv \sum_{i=1, i\neq k}^{m} \left(-\frac{q_i}{1+q_k}\right) f_i(\mathbf{x}).$$

Upon letting $p_i = -\frac{q_i}{1+q_k}$, where $p_i \leq 0$ since $q_i \geq 0$, we obtain (1). This completes the proof. □

Let $E_{\tilde{f}}$ be the set of all implied equalities of S_f. Evidently, $\widetilde{\mathcal{R}}^{-1}(E_{\tilde{f}}) \subseteq \mathcal{R}^{-1}(S_f)$. Next, we give an example to show that if an equality set is imposed, a pure inequality set can become a non-pure inequality set.

Example III.1. *Let $S_f = \{f_1 \geq 0, f_2 \geq 0\}$, where $f_1 = x_1 + x_2$, $f_2 = x_1 - x_2$. Evidently, S_f is a pure inequality set. However, if we impose the constraint $x_1 = 0$, then S_f becomes $\{x_2 \geq 0, -x_2 \geq 0\}$, which is a non-pure inequality set.*

Proposition III.1. *A subset of a pure inequality set is a pure inequality set.*

Proof. The proposition follows immediately from Lemma III.3 and Definition III.7. □

Definition III.8. *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. If the solution sets of $S_{f'}$ and S_f are the same, then we say that S_f and $S_{f'}$ are equivalent.*

Proposition III.2. *If S_f and $S_{f'}$ are equivalent, then every inequality in S_f is implied by $S_{f'}$, and every inequality in $S_{f'}$ is implied by S_f.*

In the rest of the section, we will develop a few algorithms for simplifying a linear inequality set constrained by a linear equality set.

A. Dimension Reduction of a set of inequalities by an equality set

Let $S_f = \{f_i \geq 0, i \in \mathcal{N}_m\}$ be an inequality set and $E_{\tilde{f}} = \{\tilde{f}_i = 0, i \in \mathcal{N}_{\widetilde{m}}\}$ be an equality set. Recall that $P_f = \mathcal{R}^{-1}(S_f) = \{f_i, i \in \mathcal{N}_m\}$ and $P_{\tilde{f}} = \widetilde{\mathcal{R}}^{-1}(E_{\tilde{f}}) = \{\tilde{f}_i, i \in \mathcal{N}_{\widetilde{m}}\}$. The following proposition is well known (see for example [15, Chapter 1]).

Proposition III.3. *Under the variable order $x_1 \prec x_2 \prec \cdots \prec x_n$, the linear equation system $E_{\tilde{f}}$ can be reduced by Gauss-Jordan elimination to the unique form*

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

where $k_1 < k_2 < \cdots < k_{\widetilde{n}}$, x_{k_i} is the leading term of $x_{k_i} - U_i$, $\widetilde{n}$ is rank of the linear system $E_{\tilde{f}}$ 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 $k_{i+1} = n+1$ by convention. Furthermore, $\sum_{i\in\mathcal{N}_{\tilde{n}}} |U_i| = n - \tilde{n}$.[1]

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

Algorithm 1 Dimension Reduction

```
Input: $S_f$, $E_{\tilde{f}}$.
Output: The remainder set $R_f$.
Compute $\widetilde{E}$ with $E_{\tilde{f}}$ by Proposition III.3.
Substitute $x_{k_i}$ by $U_i$ in $P_f$ to obtain a set $R$.
Let $R_f = R \backslash \{0\}$.
return $\mathcal{R}(R_f)$.
```

We call the equality set $\widetilde{E}$ the *reduced row echelon form* of $E_{\tilde{f}}$. Likewise, we call the polynomial set $\widetilde{\mathcal{R}}^{-1}(\widetilde{E})$ the reduced row echelon form of $\widetilde{\mathcal{R}}^{-1}(E_{\tilde{f}})$. 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).

Example III.2. *Given a variable order $x_1 \prec x_2 \prec x_3$, let $S_f = \{f_1 \geq 0, f_2 \geq 0\}$ and $E_{\tilde{f}} = \{\tilde{f}_1 = 0, \tilde{f}_2 = 0, \tilde{f}_3 = 0\}$, where $f_1 = x_1 + x_2 - x_3$, $f_2 = x_2 + x_3$, $\tilde{f}_1 = x_1 + x_2 + x_3$, $\tilde{f}_2 = x_1 + x_2$, and $\tilde{f}_3 = x_3$. We write $P_f = \mathcal{R}^{-1}(S_f) = \{f_1, f_2\}$ and $P_{\tilde{f}} = \widetilde{\mathcal{R}}^{-1}(E_{\tilde{f}}) = \{\tilde{f}_1, \tilde{f}_2, \tilde{f}_3\}$.*

Firstly, we obtain that the rank of $E_{\tilde{f}}$ is $\tilde{n} = 2$. Then the reduced row echelon form of $E_{\tilde{f}}$ is given by $\widetilde{E} = \{x_{k_1} - U_1 = 0, x_{k_2} - U_2 = 0\}$, where $k_1 = 1,\ k_2 = 3$, $U_1 = -x_2$, $U_2 = 0$.

Using the equality constraints in $\tilde{E}$, we substitute $x_1 = -x_2$ and $x_3 = 0$ into $P_f = \{f_1, f_2\}$ to obtain $R = \{0, x_2\}$. Hence $R_f = R\backslash\{0\} = \{x_2\}$. In other words, the inequality set S_f is reduced to $\mathcal{R}(R_f) = \{x_2 \geq 0\}$ by the equality set $E_{\tilde{f}}$. Note that in $\mathcal{R}(R_f)$, only $n - \tilde{n} = 1$ variable, namely x_2, appears.

Remark III.1. *After the execution of Algorithm 1, the inequality set S_f constrained by the equality set $E_{\tilde{f}}$ is reduced to the inequality set $\mathcal{R}(R_f)$ constrained by the equality set $\tilde{E}$. Therefore, the solution set of 'S_f constrained by $E_{\tilde{f}}$' in $\mathbb{R}^n$ is the same as the solution set of '$\mathcal{R}(R_f)$ constrained by $\tilde{E}$' in $\mathbb{R}^n$.*

B. The implied equalities contained in a system of inequalities

In this subsection, we will show how to find all the implied equalities contained in a system of linear inequalities.

[1] We use $|U_i|$ to mean the number of variables of the polynomial U_i.

Firstly, we let $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$, $f \in \mathbb{R}_h[\mathbf{x}]$ and $g_i \in \mathbb{R}_h[\mathbf{x}]$, $i = 1, \dots, m$, and give an algorithm to check whether f is a conic combination of $\{g_i, i = 1, \dots, m\}$.

Algorithm 2 Conic Combination Discrimination Algorithm

```
Input: $f$, $g_i$, $i = 1, \dots, m$.
Output: The argument "$f$ is a conic combination of $g_i$, $i = 1, \dots, m$" is TRUE or FALSE.
Let $F = f - \sum_{i=1}^{m} p_i g_i$, where $P = \{p_i, i \in \mathcal{N}_m\}$ is a set of variables. Set $F \equiv \sum_{j=1}^{n} q_j x_j \equiv 0$. Then $Q = \{q_j = 0, j \in \mathcal{N}_n\}$ is a linear system in $P$.
if the linear system $Q$ has no solution then
    Declare the argument is 'FALSE' and terminate the algorithm.
else
    Solve the linear equations $\{q_j = 0, j \in \mathcal{N}_n\}$ by Gauss-Jordan elmination to obtain the solution set of $p_i$ in the form $\{p_i = P_i, i \in \mathcal{N}_m\}$, where $P_i$ is a linear function in $m - d$ variables of $P$ and $d$ is the rank of the linear system $Q$.
    if $P_i \in \mathbb{R}_{<0}$ (i.e. $P_i$ is a negative real number) for some $i \in \mathcal{N}_m$ then
        Declare the argument is 'FALSE' and terminate the algorithm.
    else
        Let $S_P$ be the set $\{P_i, i \in \mathcal{N}_m\}$, and let $\bar{S}_P = S_P \backslash \mathbb{R}$. Write $\bar{S}_P = \{\bar{P}_i,\ i \in \mathcal{N}_{m_1}\}$.
        if $\bar{S}_P$ is empty then
            Declare the argument is 'TRUE' and terminate the algorithm.
        else
            Solve Problem $\mathbf{P_3}$:
                $\min(0)$
                s.t. $\bar{P}_i \geq 0,\ i = 1, \dots, m_1.$
            if the above LP has a solution then
                Declare the argument is 'TRUE'.
            else
                Declare that the argument is 'FALSE'.
            end if
        end if
    end if
end if
return The argument "$f$ is a conic combination of $g_i$, $i = 1, \dots, m$" is 'TRUE' or 'FALSE'.
```

Next, let $S_f = \{f_i \geq 0, i \in \mathcal{N}_m\}$ be a given inequality set, where f_i is a linear function in $\mathbf{x}$. Based on Lemma III.3, we give the following algorithm, called the Implied Equalities Algorithm that can find all the implied equalities of S_f.

Algorithm 3 Implied Equalities Algorithm

```
Input: $S_f$.
Output: The implied equalities of $S_f$.
for $k$ from 1 to $m$ do
    Determine whether $-f_k$ is a conic combination of $\{f_i,\ i \in \mathcal{N}_m, i \neq k\}$ by Algorithm 2.
    if Algorithm 2 outputs 'TRUE' then
        Declare that the equality $f_k = 0$ is an implied equality of $S_f$.
    end if
end for
return All implied equalities $f_k = 0$ of $S_f$.
```

With Algorithm 3, we can obtain the set of implied equalities of S_f, denoted by $E_{\bar{f}}$.

Theorem III.1. *For any $g \in \mathbb{R}_h[\mathbf{x}]\backslash\{0\}$, the inequality set S_f implies $g(\mathbf{x}) = 0$ if and only if $E_{\bar{f}}$ is nonempty (i.e., S_f is not a pure inequality set) and $g(\mathbf{x})$ is a linear combination of the polynomials in $\widetilde{\mathcal{R}}^{-1}(E_{\bar{f}})$.*

Proof. "Only if" part:

S_f implies $g(\mathbf{x}) = 0$ means S_f implies $g(\mathbf{x}) \geq 0$ and $-g(\mathbf{x}) \geq 0$. By Lemma III.2, we have

$$g = p_1 f_1 + \cdots + p_m f_m, \tag{3}$$

$$-g = q_1 f_1 + \cdots + q_m f_m, \tag{4}$$

where $p_i \geq 0$ and $q_i \geq 0$ for $i \in \mathcal{N}_m$. Then we obtain

$$(p_1 + q_1) f_1 + \cdots + (p_m + q_m) f_m = 0. \tag{5}$$

Since $p_i + q_i \geq 0$ for $i \in \mathcal{N}_m$, we have the following two cases:

Case 1. $p_i + q_i = 0$. For this case, $p_i = q_i = 0$, which implies $g \equiv 0$, contradicting that $g \in \mathbb{R}_h(\mathbf{x})\backslash\{0\}$.

Case 2. $p_i + q_i > 0$. By (5), we see that $-f_i$ is a conic combination of f_j, $j \in \mathcal{N}_m\backslash\{i\}$, which implies $f_i \in \widetilde{\mathcal{R}}^{-1}(E_{\bar{f}})$.

Let $\alpha = \{i \in \mathcal{N}_m : p_i + q_i > 0\}$. Note that $\alpha \neq \emptyset$ because otherwise $g \equiv 0$. This also implies $\widetilde{\mathcal{R}}^{-1}(E_{\bar{f}})$ is nonempty. Then following (3), $g = \sum\limits_{i \in \alpha} p_i f_i$, where $f_i \in \widetilde{\mathcal{R}}^{-1}(E_{\bar{f}})$. Thus, g is a linear combination of the polynomials in $\widetilde{\mathcal{R}}^{-1}(E_{\bar{f}})$.

"If" part:

If $g(\mathbf{x})$ is a linear combination of $\widetilde{\mathcal{R}}^{-1}(E_{\bar{f}})$, then $E_{\bar{f}}$ implies $g(\mathbf{x}) = 0$. By Definition III.7, S_f implies $E_{\bar{f}}$. Hence, S_f implies $g(\mathbf{x}) = 0$.

□

The following example illustrates how we can apply Algorithm 3 and then Algorithm 1 to reduce a given inequality set.

Example III.3. *Fix the variable order $x_1 \prec x_2 \prec x_3$. Let $S_f = \{f_1 \geq 0, f_2 \geq 0, f_3 \geq 0, f_4 \geq 0, f_5 \geq 0\}$, where $f_1 =$*

x_1, $f_2 = x_2 - x_1$, $f_3 = -x_1$, $f_4 = -x_2$ and $f_5 = x_2 + x_3$. An application of Algorithm 3 to S_f yields the following:

We need to take turns to verify whether $f_i = 0$, $i \in \mathcal{N}_5$ is an implied equality. Next, we take f_1 as an example.

- *Firstly, we let $F = -f_1 - \sum_{i=1}^{4} p_i f_{i+1} = \sum_{j=1}^{3} q_j x_j$. Then we have $P = \{p_1, p_2, p_3, p_4\}$ and $Q = \{q_1 = 0, q_2 = 0, q_3 = 0\}$ with $q_1 = p_1 + p_2 - 1$, $q_2 = p_3 - p_4 - p_1$ and $q_3 = -p_4$.*
- *The rank of Q is $d = 3$. We then solve the linear equations Q by Gauss-Jordon elimination to obtain $\{p_i = P_i, i \in \mathcal{N}_4\}$, where $P_1 = p_3$, $P_2 = -p_3 + 1$, $P_3 = p_3$ and $P_4 = 0$, from which we can see that P_i's are linear functions of the variable p_3.*
- *$\bar{S}_P = \{\bar{P}_1, \bar{P}_2, \bar{P}_3\}$, where $\bar{P}_1 = p_3$, $\bar{P}_2 = -p_3 + 1$, $\bar{P}_3 = p_3$.*
- *Finally, we have the following linear programming problem:*
  $L_1: \quad \min(0) \quad \text{s.t.} \quad p_3 \geq 0, \ -p_3 + 1 \geq 0.$
- *We prove L_1 has a solution.*

We repeat the above steps for f_2, f_3, f_4 and f_5, respectively. Thus, we obtain the implied equality set, denoted by $E_{\bar{f}} = \{\bar{f}_1 = 0, \bar{f}_2 = 0, \bar{f}_3 = 0, \bar{f}_4 = 0\}$, where $\bar{f}_1 = x_1$, $\bar{f}_2 = x_2 - x_1$, $\bar{f}_3 = -x_1$ and $\bar{f}_4 = -x_2$.

Upon applying Algorithm 3, the inequality set S_f is reduced to the inequality set $S'_f = \{f_5 \geq 0\} = \{x_2 + x_3 \geq 0\}$ constrained by the equality set $E_{\bar{f}}$. Finally, apply Algorithm 1 with S'_f and $E_{\bar{f}}$ as inputs to obtain $R_f = \{x_3\}$. In other words, the inequality set S_f is reduced to $\{x_3 \geq 0\}$ constrained by the equality set $E_{\bar{f}}$ after the applications of Algorithm 3 and then Algorithm 1.

Remark III.2. *Note that finding all implied equalities given a system of inequalities is still not straightforward. This remains a bottleneck for solving large problems. Depending on the size of the problem, employing Algorithm 3 to find all implied equalities may not be practical. Hence, finding all or some implied equalities by studying the structure of the system of inequalities (e.g., as in [12], [17], [25]) remains an important approach that can potentially enable more efficient computation using the results of this paper.*

C. Minimal characterization set

In this subsection, we first define a minimal characterization set of an inequality set and prove its uniqueness. Then we present an algorithm to obtain this set.

Definition III.9. *Let $S_g = \{g_i \geq 0, i \in \mathcal{N}_m\}$ be an inequality set and $S_{g'} = \{g'_i \geq 0, i \in \mathcal{N}_{m'}\}$ be a subset of S_g. If*

1) S_g and $S_{g'}$ are equivalent, and

2) there is no redundant inequalities in $S_{g'}$,

we say that $S_{g'}$ is a minimal characterization set of S_g.

Proposition III.4. *Let $S_g = \{g_i \geq 0, i \in \mathcal{N}_m\}$ be an inequality set. If $S_{g'} = \{g'_i \geq 0, i \in \mathcal{N}_{m'}\}$ is a minimal characterization set of S_g, then $m' \leq m$ and $0 \notin \mathcal{R}^{-1}(S_{g'})$.*

Proof. Since $S_{g'} \subseteq S_g$ by Definition III.9, we have $m' \leq m$. In addition, if $0 \in \mathcal{R}^{-1}(S_{g'})$, then $0 \geq 0$ is a redundant inequality in $S_{g'}$, which contradicts that $S_{g'}$ is a minimal characterization set of S_g. Thus, $0 \notin \mathcal{R}^{-1}(S_{g'})$. □

The following corollary is immediate from Definition III.9 and Proposition III.1.

Corollary III.1. *A minimal characterization set of a pure inequality set is also a pure inequality set.*

Theorem III.2. *Let $h_1, \ldots, h_m \in \mathbb{R}_h[\mathbf{x}]$ and $S_h = \{h_i \geq 0, i \in \mathcal{N}_m\}$ be a pure inequality set. Then the minimal characterization set of S_h is unique.*

Proof. Consider two minimal characterization sets of a pure set of linear inequalities S_h, denoted by $\mathcal{S}_{h'} = \{h'_i \geq 0, i \in \mathcal{N}_{m_1}\}$ and $\mathcal{S}_{\bar{h}} = \{\bar{h}_i \geq 0, i \in \mathcal{N}_{m_2}\}$. By Definition III.9, $S_{h'}$ and $S_{\bar{h}}$ are equivalent, and by Corollary III.1, they are both pure inequality sets. We will prove by contradiction that $\mathcal{S}_{h'}$ and $\mathcal{S}_{\bar{h}}$ are trivially equivalent.

Assume that for some inequality $(h'_j \geq 0) \in S_{h'}$, we cannot find $(\bar{h}_i \geq 0) \in S_{\bar{h}}$ that is trivially equivalent to $h'_j \geq 0$. By Proposition III.2 and Lemma III.2, we have

$$h'_j \equiv \sum_{i=1}^{m_2} p_i \bar{h}_i,$$

with $p_i \geq 0$. Without loss of generality, assume that $p_i > 0$ for $i = 1, \ldots, \bar{m}_2$ and $p_i = 0$ for $i = \bar{m}_2 + 1, \ldots, m_2$, where $2 \leq \bar{m}_2 \leq m_2$. Again by Lemma III.2, for all $i \in \mathcal{N}_{m_2}$,

$$\bar{h}_i \equiv \sum_{k=1}^{m_1} q_{i,k} h'_k, \tag{6}$$

where $q_{i,k} \geq 0$. Then

$$h'_j \equiv \sum_{i=1}^{\bar{m}_2} p_i \bar{h}_i \equiv \sum_{i=1}^{\bar{m}_2} p_i \sum_{k=1}^{m_1} q_{i,k} h'_k. \tag{7}$$

Rewrite (7) as

$$\left(1 - \sum_{i=1}^{\bar{m}_2} p_i q_{i,j}\right) h'_j(\mathbf{x}) \equiv \sum_{i=1}^{\bar{m}_2} p_i \sum_{k \in \mathcal{N}_{m_1} \setminus \{j\}} q_{i,k} h'_k(\mathbf{x}). \tag{8}$$

By collecting the coefficients of $h'_k(\mathbf{x})$ on the RHS, we have

$$\left(1 - \sum_{i=1}^{\bar{m}_2} p_i q_{i,j}\right) h'_j(\mathbf{x}) \equiv \sum_{k \in \mathcal{N}_{m_1} \setminus \{j\}} a_k h'_k(\mathbf{x}). \tag{9}$$

where

$$a_k = \sum_{i=1}^{\bar{m}_2} p_i q_{i,k}. \tag{10}$$

Now in (6), for a fixed $i \in \mathcal{N}_{m_2}$, if $q_{i,k} = 0$ holds for all

$k = 1, \ldots, m_1$ such that $k \neq j$, then we have

$$\bar{h}_i \equiv \sum_{k=1}^{m_1} q_{i,k} h'_k \equiv q_{i,j} h'_j. \qquad (11)$$

If $q_{i,j} > 0$, then $\bar{h}_i$ and h'_j are trivially equivalent, contradicting our assumption that there exists no $\bar{h}_i \in \mathcal{S}_{\bar{h}}$ which is trivially equivalent to h'_j. On the other hand, if $q_{i,j} = 0$, then $\bar{h}_i \equiv 0$, which by Proposition III.4 contradicts the assumption that $\mathcal{S}_{\bar{h}}$ is a minimal characterization set of S_h. Thus we conclude that for every $i \in \mathcal{N}_{m_1}$, there exists at least one $k \in \mathcal{N}_{m_1} \backslash \{j\}$ such that $q_{i,k} > 0$. From this and (10), it is not difficult to see that on the RHS of (9), there exists at least one $k \in \mathcal{N}_{m_1} \backslash \{j\}$ such that $a_k > 0$.

Consider a solution $\mathbf{x}^*$ of $S_{h'}$ such that $h'_k(\mathbf{x}^*) > 0$ for all $k \in \mathcal{N}_{m_1}$. Such an $\mathbf{x}^*$ exists because $S_{h'}$ is a pure inequality set. Substituting $\mathbf{x} = \mathbf{x}^*$ in (9) to yield

$$\left(1 - \sum_{i=1}^{\bar{m}_2} p_i q_{i,j}\right) h'_j(\mathbf{x}^*) = \sum_{k \in \mathcal{N}_{m_1} \backslash \{j\}} a_k h'_k(\mathbf{x}^*). \qquad (12)$$

Since there exists at least one $k \in \mathcal{N}_{m_1} \backslash \{j\}$ such that $a_k > 0$, the RHS above is strictly positive, which implies that $1 - \sum\limits_{i=1}^{\bar{m}_2} p_i q_{i,j} > 0$. It then follows that h'_j can be written as a conic combination of h'_k, $k \in \mathcal{N}_{m_1} \backslash \{j\}$. In other words, $h'_j \geq 0$ is implied by $h'_k \geq 0$, $k \in \mathcal{N}_{m_1} \backslash \{j\}$. This contradicts that $\mathcal{S}_{h'}$ is a minimal characterization set of S_h.

Summarizing the above, we have proved that for every $(h'_j \geq 0) \in S_{h'}$, we can find an $(\bar{h}_i \geq 0) \in S_{\bar{h}}$ which is trivially equivalent to $h'_j \geq 0$. Moreover, $\bar{h}_i$ is unique, which can be seen as follows. If there exists another $(\bar{h}_{i'} \geq 0) \in S_{\bar{h}}$ which is trivially equivalent to $h'_j \geq 0$, then $\bar{h}_i \geq 0$ and $\bar{h}_{i'} \geq 0$ are also trivially equivalent to each other, contradicting that $S_{h'}$ is a minimal characterization set of S_h. In the same way, we can prove that for every $(\bar{h}_i \geq 0) \in S_{\bar{h}}$, we can find a unique $(h'_j \geq 0) \in S_{h'}$ which is trivially equivalent to $\bar{h}_i \geq 0$. Thus, $S_{h'}$ and $S_{\bar{h}}$ are trivially equivalent and have exactly the same number of inequalities, which means that the minimal characterization set of a pure inequality set S_h is unique. This completes the proof of the theorem. □

Theorem III.3. *Let $S_f = \{f_i \geq 0, i \in \mathcal{N}_{m_1}\}$ and $S_g = \{g_i, i \in \mathcal{N}_{m_2}\}$ be two pure inequality sets, and $S_{f'}$ and $S_{g'}$ be their minimal characterization sets respectively. If S_f and S_g are equivalent, then $S_{f'}$ and $S_{g'}$ are trivially equivalent.*

Proof. If the two pure inequality sets S_f and S_g are equivalent, then $S_{f'}$ and $S_{g'}$ are pure and equivalent. Thus the theorem follows immediately from the proof of Theorem III.2. □

Next, we give an example to show that the minimal characterization set of a non-pure inequality set may not be unique.

Example III.4. *Let $S_f = \{f_1 \geq 0, f_2 \geq 0, f_3 \geq 0, f_4 \geq 0,\}$ be an inequality set, where $f_1 = x_1 - x_2$, $f_2 = x_2$, $f_3 = -x_2$, $f_4 = x_1$. Evidently, S_f is a non-pure inequality set, and it can readily be seen that both $S_{f'} = \{f_1 \geq 0, f_2 \geq 0, f_3 \geq 0\}$ and $S_{f''} = \{f_2 \geq 0, f_3 \geq 0, f_4 \geq 0\}$ are minimal characterization sets of S_f. However, $S_{f'}$ and $S_{f''}$ are not trivially equivalent. Thus, the minimal characterization set of S_f isn't unique.*

Let $S_h = \{h_i \geq 0,\ i \in \mathcal{N}_m\}$ be an inequality set, where $h_i \in \mathbb{R}_h[\mathbf{x}]$. Based on Lemma III.2, the following algorithm, called Minimal Characterization Set Algorithm, can be used to obtain a minimal characterization set of S_h.

Algorithm 4 Minimal Characterization Set Algorithm

```
Input: $S_h$.
Output: A minimal characterization set of $S_h$.
1: Set $P_h := \mathcal{R}^{-1}(S_h)$.
2: for $k$ from 1 to $m$ do
3:     Determine whether $h_k$ is a conic combination of $P_h \backslash \{h_k\}$ by Algorithm 2.
4:     if Algorithm 2 outputs 'TRUE' then
5:         $P_h := P_h \backslash \{h_k\}$.
6:     end if
7: end for
8: return $\mathcal{R}(P_h)$.
```

Justification for Algorithm 4. Steps 2 to 5 remove the polynomial h_k from P_h if it can be expressed as a conic combination of $h_i, i \in \mathcal{M} \backslash \{k\}$. Iterating over all k from 1 to m, the output inequality set $\mathcal{R}(P_h)$ is equivalent to S_h and it is a pure inequality set. Hence, it is a minimal characterization set of S_h.

D. The reduced minimal characterization set

In this subsection, we first define the reduced minimal characterization set of a linear inequality set and prove its uniqueness. Then we present an algorithm to obtain this set.

Let $S_f = \{f_i \geq 0, i \in \mathcal{N}_m\}$ be a linear inequality set, and $E_{\bar{f}}$ be the set of implied equalities of S_f obtained by applying Algorithm 3. Then we obtain $\widetilde{E}$, the reduced row echelon form of $E_{\bar{f}}$, as in Proposition III.3. Let R_f be the remainder set obtained by reducing $\mathcal{R}^{-1}(S_f) \backslash \mathcal{R}^{-1}(E_{\bar{f}})$ by $\widetilde{\mathcal{R}}^{-1}(\widetilde{E})$ using Algorithm 1.

Theorem III.4. *The set $\mathcal{R}(R_f)$ is a pure inequality set.*

Proof. Let $\widetilde{E} = \{E_i = 0, i \in \mathcal{N}_{\tilde{n}}\}$, and assume there is an implied equality $(\bar{f} = 0) \in \mathcal{R}(R_f)$. In the process of obtaining $\bar{f}$, we substitute $x_{k_i} = U_i, i \in \widetilde{\mathcal{N}}_{\tilde{n}}$ into some polynomial $f \in \mathcal{R}^{-1}(S_f)$ (cf. equation (2)). Therefore, we can write

$$\bar{f} \equiv f - \sum_{i=1}^{\tilde{n}} c_i E_i, \qquad (13)$$

where c_i is the coefficient of x_{k_i} in f. Let $\mathbf{x}^*$ be a solution of S_f. From Remark III.1, we see that $\mathbf{x}^*$ is also a solution of $\mathcal{R}(R_f)$ constrained by $\widetilde{E}$, so that $E_i(\mathbf{x}^*) = 0$ for all $i \in \mathcal{N}_{\tilde{n}}$. From (13), we have

$$f(\mathbf{x}^*) = \bar{f}(\mathbf{x}^*) - \sum_{i=1}^{\tilde{n}} c_i E_i(\mathbf{x}^*).$$

Since $\bar{f} = 0$ is an implied equality of S_f, we have $\bar{f}(\mathbf{x}^*) = 0$. It follows from the above that $f(\mathbf{x}^*) = 0$. Since this holds for all solution $\mathbf{x}^*$ of S_f, we see that $f = 0$ is an implied equality of S_f, i.e., $(f = 0) \in E_{\bar{f}}$, which is a contradiction to $f \in \mathcal{R}^{-1}(S_f)\backslash\mathcal{R}^{-1}(E_{\bar{f}})$. The theorem is proved. □

Since $\mathcal{R}(R_f)$ is a pure inequality set, the minimal characterization set of $\mathcal{R}(R_f)$ is unique. We let $S_{r'}$ be the minimal characterization set of $\mathcal{R}(R_f)$.

Definition III.10. *The set $S_M = \widetilde{E} \cup S_{r'}$ is called the reduced minimal characterization set of S_f.*

Theorem III.5. *The reduced minimal characterization set of S_f is unique.*

Proof. Fix the variable order $x_1 \prec x_2 \prec \cdots \prec x_n$. By Proposition III.3, the reduced standard basis $\widetilde{\mathcal{R}}^{-1}(\widetilde{E})$ is unique, which yields that the remainder set R_f is unique. Since $\mathcal{R}(R_f)$ is a pure inequality set by Theorem III.2, the minimal characterization set of $\mathcal{R}(R_f)$ is unique. Hence, S_M is unique. □

In the following, we present an algorithm to find the reduced minimal characterization set of a linear inequality set.

Algorithm 5 Reduced Minimal Characterization Set Algorithm

```
Input: $S_f$.
Output: The reduced minimal characterization set of $S_f$.
1: Apply Algorithm 3 to find the implied equality set of $S_f$, denoted by $E_{\bar{f}}$.
2: Apply Algorithm 1 to reduce $\mathcal{R}^{-1}(S_f)\backslash\widetilde{\mathcal{R}}^{-1}(E_{\bar{f}})$ by $E_{\bar{f}}$ to obtain $R_f$ and $\widetilde{E}$, the reduced row echelon form of $E_{\bar{f}}$.
3: Apply Algorithm 4 to obtain the minimal characterization set of $\mathcal{R}(R_{\underline{f}})$, denoted by $S_{r'}$.
4: return $S_M = \widetilde{E} \cup S_{r'}$.
```

By Proposition III.3 and Theorems III.3 and III.5, we immediately obtain the following theorem.

Theorem III.6. *For two equivalent inequality sets, their reduced minimal characterization sets are same.*

Note that for a pure inequality set, the minimal characterization set is exactly the reduced minimal characterization set.

Remark III.3. *Since the basic inequalities contain no implied equality and hence form a pure inequality set, the elemental inequalities form the minimal characterization set of the basic inequalities. In fact, for a fixed number of random variables, Algorithm 5 can be used to compute the reduced minimal characterization set of the basic inequalities under the constraint of an equality set and possibly an inequality set (used for example, for including some non-Shannon-type inequalities).*

In this sense, the definition of "(reduced) minimal characterization set" (Definition III.9, Definition III.10) can be seen as a generalization of the "minimal characterization" discussed in [1], [12].

Remark III.4. *There are many works which have studied the reduction of polyhedral computation, such as the classical works [21], [22] and the standard methods [23], [24]. Both our approach and the known methods focus on the removal of redundant inequalities in a given set of inequalities (polyhedron), and aim to find a minimal non-redundant set. However, our method is different from the known methods.*

On the one hand, the known methods are developed for general LP problems, but we give algorithms specifically for homogeneous LP problems, so that the special algebraic structure (homogeneity and sparsity) of this type of information inequality problems can be better exploited. As a result, our method can in fact outperform the existing methods for such problems. See Section VI.

On the other hand, since there are typically very few implied equalities for non-homogeneous LP problems, the known methods pay more attention to the removal of redundant inequalities but ignore the implied equalities, so that in general only part of the redundancy in the orginal LP problem is removed. However, in homogeneous LP problems (in particular for proving information inequalities or identities), there are usually many implied equalities, and our methods can take full advantage of this property. First, we compute the implied equalities to reduce the number of variables in the LP. Then we remove the redundant inequalities to reduce the number of inequality constraints, so that we can remove all the redundancy in the original problem and finally obtain the minimal non-redundant set which is shown to be unique. See Algorithms 1 to 5.

Our methods can also be applied to non-homogeneous LP problems. However, since there are typically very few implied equalities, it may not be worth the extra computation to find them.

IV. Procedures for proving information inequalities and identities

In this section, we present two procedures for proving information inequalities and identities under the constraint of an inequality set and/or an equality set. They are designed in the spirit of Theorem II.2.

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.2023.3263178

A. Procedure I: Proving Information Inequalities

Input:
Objective information inequality: $\bar{F} \geq 0$.
Additional constraints: $\bar{C}_i = 0,\ i = 1, \dots, r_1$; $\bar{C}_j \geq 0,\ j = r_1 + 1, \dots, r_2$.
Element information inequalities: $\bar{C}_k \geq 0,\ k = r_2 + 1, \dots, r_3$.
// Here, $\bar{F}$, $\bar{C}_i$, $\bar{C}_j$, and $\bar{C}_k$ are linear combination of information measures.
Output: A proof of $\bar{F} \geq 0$ if feasible.

Step 1. Transform $\bar{F}$, $\bar{C}_i$, $\bar{C}_j$ and $\bar{C}_k$ to linear polynomials F, C_i, C_j and C_k in the joint entropies respectively.
// We need to solve
// **Problem $\mathbf{P_1}$**: Determine whether $F \geq 0$ is implied by

$$\begin{aligned} &C_i = 0,\ i = 1, \dots, r_1, \\ &C_j \geq 0,\ j = r_1 + 1, \dots, r_2, \\ &C_k \geq 0,\ k = r_2 + 1, \dots, r_3. \end{aligned}$$

Step 2. Apply Algorithm 1 to reduce $\{C_l, l \in \mathcal{N}_{r_3} \backslash \mathcal{N}_{r_1}\}$ by $\{C_l = 0, l \in \mathcal{N}_{r_1}\}$ to obtain the reduced row echelon form

of $\{C_l, l \in \mathcal{N}_{r_1}\}$, denoted by B, and the remainder set, denoted by $\mathbf{C}_1 = \{g_i, i \in \mathcal{N}_r\}$.

Step 3. Apply Algorithm 5 to obtain the reduced minimal characterization set of $\mathcal{R}(\mathbf{C}_1)$, denoted by

$S_M = \widetilde{E} \cup S_{r'}$. Write $S_{r'} = \{\mathbb{C}_j \geq 0, j \in \mathcal{N}_{t_2}\}$.

Step 4. Let $G = \widetilde{\mathcal{R}}^{-1}(\widetilde{E}) \cup B$ and compute the reduced row echelon form of G, denoted by $\mathcal{B} = \{\mathcal{C}_i, i \in \mathcal{N}_{t_1}\}$.

Step 5. Reduce F by $\widetilde{\mathcal{R}}(\mathcal{B})$ to obtain the remainder F_1.
// In the above, the inequality set $\mathcal{R}(\mathbf{C}_1)$ is generated by reducing $\{C_l \geq 0, l \in \mathcal{N}_{r_3} \backslash \mathcal{N}_{r_1}\}$ by B, and the inequality
// set $S_{r'}$ is generated by further reducing $\mathcal{R}(\mathbf{C}_1)$ by its own implied equalities, given by $\widetilde{E}$. On the other hand,
// the set $\mathcal{B}$ is generated by computing the reduced row echelon form of $\widetilde{\mathcal{R}}^{-1}(\widetilde{E}) \cup B$, and F_1 is generated by reducing F
// by $\widetilde{\mathcal{R}}(\mathcal{B})$. Therefore, only the free variables in the reduced row echelon form $\mathcal{B}$ are involved in F_1 and $S_{r'}$.
// The original **Problem $\mathbf{P_1}$** is now transformed into
// **Problem $\mathbf{P_2}$**: Determine whether $F_1 \geq 0$ is implied by the inequalities in $S_{r'}$, i.e.,

$$\mathbb{C}_i \geq 0,\ j = 1, \dots, t_2.$$

// Since the equality set $\widetilde{\mathcal{R}}(\mathcal{B})$ contains only constraints on the pivot variables in $\mathcal{B}$, it is ignored in formulation of
// **Problem $\mathbf{P_2}$**. The remaining steps follow Algorithm 4.

Step 6. Determine whether F_1 is a conic combination of $\{\mathbb{C}_i, i \in \mathcal{N}_{t_2}\}$ by Algorithm 2. If Algorithm 2 outputs 'TRUE', then the objective information inequality $\bar{F} \geq 0$ is proved. Otherwise, declare 'Not Provable'.

One can solve this problem by ITIP after obtaining **Problem $\mathbf{P_2}$**. In order to solve **Problem $\mathbf{P_2}$**, its dual problem will be considered. Now rewrite **Problem $\mathbf{P_2}$** in matrix form by letting $F_1 = \mathbf{c}^T\mathbf{x}$ and $(\mathbb{C}_1, \mathbb{C}_2, \cdots, \mathbb{C}_{t_2})^T = \mathbf{A}\mathbf{x}$. Then **Problem $\mathbf{P_2}$** can be rewritten as

$$\begin{aligned} &\text{minimize} && \mathbf{c}^T\mathbf{x} \\ &\text{subject to} && \mathbf{A}\mathbf{x} \geq \mathbf{0} \end{aligned} \tag{14}$$

such that the optimal value is nonnegative. The dual problem is
Problem $\mathbf{P_{2D}}$:

$$\begin{aligned} &\text{maximize} && 0 \\ &\text{subject to} && -\mathbf{A}^T\boldsymbol{\lambda} + \mathbf{c} = \mathbf{0} \\ &&& \boldsymbol{\lambda} \geq \mathbf{0}. \end{aligned} \tag{15}$$

Solving **Problem $\mathbf{P_{2D}}$** can be considerably easier than solving **Problem $\mathbf{P_2}$**, though the former may contain more variables. See the application example in Section VI.

Once the objective inequality has been processed by Procedure I (either the objective inequality is a Shannon-type inequality or it is not), the LP in **Problem $\mathbf{P_3}$** in Algorithm 2 is already solved. Let $N_v(\mathbf{P_1})$, $N_v(\mathbf{P_2})$, $N_v(\mathbf{P_{2D}})$ and $N_v(\mathbf{P_3})$ be the number of variables in **Problems $\mathbf{P_1}$, $\mathbf{P_2}$, $\mathbf{P_{2D}}$** and **$\mathbf{P_3}$** respectively. Let $N_c(\mathbf{P_1})$, $N_c(\mathbf{P_2})$, $N_c(\mathbf{P_{2D}})$ and $N_c(\mathbf{P_3})$ be the number of constraints in **Problems $\mathbf{P_1}$, $\mathbf{P_2}$, $\mathbf{P_{2D}}$** and **$\mathbf{P_3}$** respectively. It is clear that $N_v(\mathbf{P_1}) \geq N_v(\mathbf{P_2})$, $N_c(\mathbf{P_1}) \geq N_c(\mathbf{P_2})$, $N_v(\mathbf{P_{2D}}) \geq N_v(\mathbf{P_3})$ and $N_c(\mathbf{P_{2D}}) \geq N_c(\mathbf{P_3})$. The reduction of the number of variables and the number of constraints is in general significant. Since most of the computation in the procedure is attributed to solving the LP in **Problem $\mathbf{P_3}$**, compared with the approach using Theorem II.2 where a much larger LP needs to be solved, the efficiency can be significantly improved. Moreover, with our approach, an analytical proof can be generated automatically. Example V.1 will illustrate this point.

B. Procedure II: Proving Information Identities

Input:
Objective information identity: $\bar{F} = 0$.
Additional constraints: $\bar{C}_i = 0,\ i = 1, \dots, r_1$; $\bar{C}_j \geq 0,\ j = r_1 + 1, \dots, r_2$.
Element information inequalities: $\bar{C}_k \geq 0,\ k = r_2 + 1, \dots, r_3$.
// 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 feasible.

Step 1. Transform $\bar{F}$, $\bar{C}_i$, $\bar{C}_j$ and $\bar{C}_k$ to linear polynomials F, C_i, C_j and C_k in the joint entropies respectively.
// We need to solve
// **Problem $\mathbf{P_4}$**: Determine whether $F = 0$ is implied by

$$\begin{aligned} &C_i = 0,\ i = 1, \dots, r_1, \\ &C_j \geq 0,\ j = r_1 + 1, \dots, r_2, \\ &C_k \geq 0,\ k = r_2 + 1, \dots, r_3. \end{aligned}$$

Step 2. Apply Algorithm 1 to reduce $\{C_l, l \in \mathcal{N}_{r_3} \backslash \mathcal{N}_{r_1}\}$ by $\{C_l = 0, l \in \mathcal{N}_{r_1}\}$ to obtain the reduced row echelon form of $\{C_l, l \in \mathcal{N}_{r_1}\}$, denoted by B, and the remainder set, denoted by $\mathbf{C}_1 = \{g_i, i \in \mathcal{N}_r\}$.

Step 3. Apply Algorithm 5 to obtain the reduced minimal characterization set of $\mathcal{R}(\mathbf{C}_1)$, denoted by

$S_M = \widetilde{E} \cup S_{r'}$.

Step 4. Let $G = \widetilde{\mathcal{R}}^{-1}(\widetilde{E}) \cup B$ and compute the reduced row echelon form of G, denoted by $\mathcal{B} = \{\mathcal{C}_i, i \in \mathcal{N}_{t_1}\}$.

// The original problem $\mathbf{P_4}$ has been transformed into

// **Problem $\mathbf{P_5}$**: Determine whether $F = 0$ is implied by $\widetilde{\mathcal{R}}(\mathcal{B})$.

Step 5. Reduce F by $\widetilde{\mathcal{R}}(\mathcal{B})$ to obtain remainder F_1. If $F_1 \equiv 0$, then the objective identity $\bar{F} = 0$ is proved.

Otherwise, declare 'Not Provable'.

Justification. Steps 1 to 5 are exactly the same as Procedure I. We only need to justify Step 6. As explained in Procedure I, F_1 involves only the free variables in the reduced row echelon form $\mathcal{B}$. We now prove by contradiction that if $F_1 \not\equiv 0$, the free variables can be chosen such that F_1 is evaluated to a nonzero value. Assume $F_1 \not\equiv 0$ and that for any free variables satisfying $S_{r'}$, F_1 is evaluated to zero. Then $F_1 = 0$ is implied by $S_{r'}$. By Theorem III.1, $S_{r'}$ is not a pure inequality set, which is a contradiction because $S_{r'}$ is a pure inequality set by construction.

Remark IV.1. *An information identity $F = 0$ is equivalent to the two information inequalities $F \geq 0$ and $F \leq 0$. In the previous approach, in order to prove $F = 0$, $F \geq 0$ and $F \leq 0$ are proved separately by solving two LPs. In Procedure II, we transform the proof into a Gauss elimination problem, which greatly reduces the computational complexity.*

Remark IV.2. *Procedures I and II can be implemented on the computer by Maple for symbolic computation. Therefore, they can give explicit proofs of information inequalities and identities.*

V. Illustrative examples

In this section, we give two examples to illustrate Procedures I and II. The computation is performed by Maple. To simplify notations, we use $h_{1,2,3,4}$ to represent the joint entropy $H(X_1, X_2, X_3, X_4)$, so on and so forth.

A. Information Inequality under Equality Constraints

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

Proof. By symmetry of the problem, we only need to prove $H(X_1) \geq H(X_4)$. The proof is given according to Procedure I.

Input:

Objective information inequality:

$\bar{F} = H(X_1) - H(X_4) \geq 0$.

Equality Constraints: $\bar{C}_1 = I(X_1; X_4) = 0$,
$\bar{C}_2 = I(X_2; X_4) = 0$, $\bar{C}_3 = I(X_3; X_4) = 0$,
$\bar{C}_4 = H(X_4|X_1, X_2) = 0$, $\bar{C}_5 = H(X_4|X_1, X_3) = 0$,
$\bar{C}_6 = H(X_4|X_2, X_3) = 0$.

28 element information inequalities: $\bar{C}_k \geq 0,\ k \in \mathcal{N}_{34} \backslash \mathcal{N}_6$.

Step 1. We have $F = h_1 - h_4$, $C_1 = h_1 + h_4 - h_{1,4}$, $C_2 = h_2 + h_4 - h_{2,4}$, $C_3 = h_3 + h_4 - h_{3,4}$, $C_4 = h_{1,2,4} - h_{1,2}$, $C_5 = h_{1,3,4} - h_{1,3}$, $C_6 = h_{2,3,4} - h_{2,3}$, and 28 linear polynomials $C_k, k \in \mathcal{N}_{34} \backslash \mathcal{N}_6$ are obtained from the 28 element information inequalities.

Step 2. Fix the variable order $h_{1,2,3,4} \prec h_{2,3,4} \prec h_{1,3,4} \prec h_{1,2,4} \prec h_{1,2,3} \prec h_{3,4} \prec h_{2,4} \prec h_{2,3} \prec h_{1,4} \prec h_{1,3} \prec h_{1,2} \prec h_4 \prec h_3 \prec h_2 \prec h_1$. Compute the reduced row echelon form of $\{C_i, i \in \mathcal{N}_6\}$, $B = \{-h_{2,3} + h_{2,3,4}, -h_{1,3} + h_{1,3,4}, -h_{1,2} + h_{1,2,4}, -h_3 - h_4 + h_{3,4}, -h_2 - h_4 + h_{2,4}, -h_1 - h_4 + h_{1,4}\}$. Use Algorithm 1 to reduce $\{C_l, l \in \mathcal{N}_{34} \backslash \mathcal{N}_6\}$ by $\widetilde{\mathcal{R}}(B)$ to obtain the remainder set $\mathbf{C}_1 = \{g_i, i \in \mathcal{N}_{18}\}$.

Step 3. Use Algorithm 5 to obtain $S_M = \widetilde{E} \cup S_{r'}$ and $S_{r'} = \{\mathbb{C}_i, i = 1, \ldots, 10\}$, where

$$\begin{aligned}
&\mathbb{C}_1 = h_4,\\
&\mathbb{C}_2 = h_1 + h_2 - h_{1,2},\\
&\mathbb{C}_3 = h_1 + h_3 - h_{1,3},\\
&\mathbb{C}_4 = h_2 + h_3 - h_{2,3},\\
&\mathbb{C}_5 = -h_1 - h_4 + h_{1,2} + h_{1,3} - h_{1,2,3},\\
&\mathbb{C}_6 = -h_2 - h_4 + h_{1,2} + h_{2,3} - h_{1,2,3},\\
&\mathbb{C}_7 = -h_3 - h_4 + h_{1,3} + h_{2,3} - h_{1,2,3},\\
&\mathbb{C}_8 = -h_{1,2} + h_{1,2,3},\\
&\mathbb{C}_9 = -h_{1,3} + h_{1,2,3},\\
&\mathbb{C}_{10} = -h_{2,3} + h_{1,2,3}.
\end{aligned}$$

Step 4. Compute the reduced row echelon form $\mathcal{B} = \{-h_{1,2,3} + h_{1,2,3,4}, -h_{2,3} + h_{2,3,4}, -h_{1,3} + h_{1,3,4}, -h_{1,2} + h_{1,2,4}, -h_3 - h_4 + h_{3,4}, -h_2 - h_4 + h_{2,4}, -h_1 - h_4 + h_{1,4}\}$.

Step 5. Reduce F by $\widetilde{\mathcal{R}}(\mathcal{B})$ to obatain $F_1 = h_1 - h_4$.

Step 6. In Algorithm 2, we have $m = 10$, $n = 8$, $\bar{S}_P = \{\frac{1}{2}p_1, p_1, 1 + \frac{1}{2}p_1, 1 + \frac{1}{2}p_1 - p_3, -1 + \frac{1}{2}p_1 + p_3, 1 - p_3\}$ and $S_P = \bar{S}_P \cup \{1, 0\}$. Solve the LP in **Problem $\mathbf{P_3}$** to complete the proof. Alternatively, we can solve the inequality set $\mathcal{R}(\bar{S}_P)$ to obtain the solution $\{p_1 \geq 2 - 2p_3,\ p_3 \leq 1\}$. Substituting $p_1 = 0$ and $p_3 = 1$ to $\{p_i = P_i, i \in \mathcal{N}_{10}\}$ yields $\{p_1 = 0, p_2 = 0, p_3 = 1, p_4 = 0, p_5 = 0, p_6 = 0, p_7 = 1, p_8 = 0, p_9 = 0, p_{10} = 1\}$. Thus an explicit proof is given by $F_1 = \mathbb{C}_3 + \mathbb{C}_7 + \mathbb{C}_{10} \geq 0$. □

Table I shows the advantage of Procedure I for this example by comparing it with the Direct LP method induced by Theorem II.2 and with ITIP. Note that in both ITIP and Procedure I, the number of variables is first reduced by the equality constraints before solving the LP. However, in ITIP, the number of inequality constraints is not reduced.

B. Information Identity under Equality Constraints

Example V.2. $I(X_1; X_2|X_3) = 0$, $H(X_3) = I(X_2; X_3|X_1)$ $\Rightarrow H(X_1) = H(X_1|X_2, X_3)$.

TABLE I

	Number of variables	Number of equality constraints	Number of Inequality constraints
Direct LP method	15	6	28
ITIP	7	0	28
LP in **Problem $\mathbf{P_3}$**	2	0	6

Proof. The proof is given according to Procedure II.
Input:
Objective information identity:
$\bar{F} = H(X_1) - H(X_1|X_2, X_3) = 0$.
Equality Constraints: $\bar{C}_1 = I(X_1; X_2|X_3) = 0$,
$\bar{C}_2 = H(X_3) - I(X_2; X_3|X_1) = 0$.
9 element information inequalities: $\bar{C}_k \geq 0,\ k \in \mathcal{N}_{11} \backslash \mathcal{N}_2$.

Step 1. We have $F = h_1 + h_{2,3} - h_{1,2,3}$, $C_1 = h_{1,3} + h_{2,3} - h_{1,2,3} - h_3$, $C_2 = h_1 + h_3 + h_{1,2,3} - h_{1,2} - h_{1,3}$, $C_3 = h_{1,2,3} - h_{2,3}$, $C_4 = h_{1,2,3} - h_{1,3}$, $C_5 = h_{1,2,3} - h_{1,2}$, $C_6 = h_1 + h_2 - h_{1,2}$, $C_7 = h_{1,3} + h_{2,3} - h_{1,2,3} - h_3$, $C_8 = h_1 + h_3 - h_{1,3}$, $C_9 = h_{1,2} + h_{2,3} - h_{1,2,3} - h_2$, $C_{10} = h_2 + h_3 - h_{2,3}$ and $C_{11} = h_{1,2} + h_{1,3} - h_{1,2,3} - h_1$.

Step 2. Fix the variable order $h_{1,2,3} \prec h_{2,3} \prec h_{1,3} \prec h_{1,2} \prec h_3 \prec h_2 \prec h_1$. Compute the reduced row echelon form $B = \{h_1 + h_3 - h_{1,2} - h_{1,3} + h_{1,2,3}, h_1 - h_{1,2} + h_{2,3}\}$. Use Algorithm 1 to reduce $\{C_l, l \in \mathcal{N}_{11} \backslash \mathcal{N}_2\}$ by $\widetilde{\mathcal{R}}(B)$ to obtain the remainder set $\mathbf{C}_1 = \{g_i, i \in \mathcal{N}_8\}$, where $g_1 = -h_3 + h_{1,3}, g_2 = -h_1 - h_3 + h_{1,2}, g_3 = -h_1 - h_3 + h_{1,3}, g_4 = h_1 + h_2 - h_{1,2}, g_5 = h_1 + h_3 - h_{1,3}, g_6 = -h_2 + h_3 + h_{1,2} - h_{1,3}, g_7 = h_1 + h_2 + h_3 - h_{1,2}, g_8 = h_3$.

Step 3. Use Algorithm 5 to obtain $S_M = \widetilde{E} \cup S_{r'}$, where $\widetilde{E} = \{h_1 + h_3 - h_{1,3} = 0,\ h_1 + h_2 - h_{1,2} = 0\}$.

Step 4. Compute the Gauss-Jordon normal form $\mathcal{B} = \{-h_1 - h_2 + h_{1,2,3}, -h_2 + h_{2,3}, -h_1 - h_3 + h_{1,3}, -h_1 - h_2 + h_{1,2}\}$.

Step 5. Reduce F by $\mathcal{B}$ to obtain $F_1 \equiv 0$. Thus the information identity is proved.

□

VI. An Application

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 Tian [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 $(4, 3, 3)$.[2] It was the first time that a non-trivial information theory problem was solved using this LP framework.

In this section, we apply the results in the previous sections to Tian's problem and significantly reduce the required computation for solving the LP. We first give the abstract formulation of the problem.

Definition VI.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 the 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

$$4\alpha + 6\beta \geq 3B \tag{16}$$

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 $H(S_{i,j}) \leq \beta$, $S_{i,j} \in \mathcal{S}$.

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

Tian showed in Section III-B of [17] that it is not necessary to use all the 16 random variables for solving **Problem $\mathbf{P_6}$**, i.e., only a subset of the random variables in $\mathcal{W} \cup \mathcal{S}$ is needed. This idea for reducing the problem also is 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

$$4\alpha + 6\beta \geq 3B \tag{17}$$

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

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

Remark VI.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 give the proof of **Problem $\mathbf{P_7}$** according to Procedure I.

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

Step 1. The variable vector generated on $\mathcal{W}_1 \cup \mathcal{S}_1$ has $2^{12} - 1$ elements (joint entropies).

Step 2. According to conditions C1, C3, C4 and C6, write equality constraints in joint entropy forms: $C_i = 0,\ i \in \mathcal{N}_{22945}$. According to conditions C7, C8 and element information inequalities, write inequality constraints in joint entropy forms: $C_i,\ i \in \mathcal{N}_{90553} \backslash \mathcal{N}_{22945}$.

Step 3. Compute the reduced row echelon form of $\{C_i, i \in \mathcal{N}_{22945}\}$ and denote it by B. Then use Algorithm 1 to reduce $\{C_i,\ i \in \mathcal{N}_{90553} \backslash \mathcal{N}_{22945}\}$ by $\widetilde{\mathcal{R}}(B)$ to obtain the remainder set $\mathbf{C}_1 = \{g_i, i \in \mathcal{N}_{62981}\}$.

Step 4. Use Algorithm 4 to obtain $S_M = \widetilde{E} \cup S_{r'}$, and $S_{r'} = \{\mathbb{C}_i, i \in \mathcal{N}_{649}\}$. Here we list the formulas used below and omit the others.

$$\begin{aligned}
&\mathbb{C}_1 = 2h_5 - h_{9,12},\ \mathbb{C}_2 = 2h_{8,10,12} - h_{7,8,10,12} - h_{8,10},\\
&\mathbb{C}_3 = 2h_{2,3,8,9,10,11,12} - h_{2,3,5,6,7,8,10} - h_{3,4,7,9,10,11,12},\\
&\mathbb{C}_4 = 2h_{2,3,8,9,10,11,12} - h_{6,7,8,9,10,11,12} - h_{2,3,8,10,11},\\
&\mathbb{C}_5 = h_{9,12} + h_{8,10} - h_{8,10,11} - h_5,\\
&\mathbb{C}_6 = h_{3,8,9} + h_{8,9,10,12} - h_{3,5,7,9} - h_{8,11,12},\\
&\mathbb{C}_7 = h_{3,9,12} + h_{3,8,9} - h_{3,8,9,10} - h_{3,9},\\
&\mathbb{C}_8 = h_{8,11,12} + h_{8,10,11} - h_{8,9,10,12} - h_{11,12},\\
&\mathbb{C}_9 = h_{1,5,10,12} + h_{3,8,9,10} - h_{2,3,8,9,10,11,12} - h_{3,9,12},\\
&\mathbb{C}_{10} = h_{2,3,11,12} + h_{3,5,7,9,10,12} - h_{2,3,8,11,12} - h_{3,5,6,7,9,10},\\
&\mathbb{C}_{11} = h_{3,5,7,9} + h_{3,5,8,9} - h_{2,3,5,6,7,8,10} - h_{3,8,9},\\
&\mathbb{C}_{12} = h_{3,5,8,9} + h_{6,7,8,10,12} - h_{3,5,8,9,10} - h_{7,8,10,12},\\
&\mathbb{C}_{13} = h_{2,3,8,11,12} + h_{3,5,7,9,10,11,12} - h_{2,3,8,9,10,11,12}\\
&\qquad - h_{3,5,7,9,10,12},\\
&\mathbb{C}_{14} = h_{3,5,8,9,10} + h_{5,6,8,9,10,11} - h_{3,5,8,9,10,11} - h_{6,7,8,10,12},\\
&\mathbb{C}_{15} = h_{3,8,9,11,12} + h_{3,5,6,7,9,10} - h_{3,5,7,9,10,11} - h_{1,5,10,12},\\
&\mathbb{C}_{16} = h_{8,9,10,11,12} + h_{3,5,7,9,10,11} - h_{3,5,7,9,10,11,12}\\
&\qquad - h_{3,8,9,11,12},\\
&\mathbb{C}_{17} = h_{2,3,6,9,10,12} + h_{2,3,8,10,11} - h_{2,3,8,9,10,11,12} - h_{2,3,11,12},\\
&\mathbb{C}_{18} = h_{2,3,6,9,10,12} + h_{3,4,7,9,10,11,12} - h_{2,3,8,9,10,11,12}\\
&\qquad - h_{8,9,10,11,12},\\
&\mathbb{C}_{19} = h_{6,7,8,9,10,11,12} + h_{3,5,8,9,10,11} - h_{2,3,5,6,7,8,10}\\
&\qquad - h_{5,6,8,9,10,11},\\
&\mathbb{C}_{20} = 2h_5 - h_{11,12},\ \mathbb{C}_{21} = 2h_{3,8,9} - h_{2,3,6,9,10,12} - h_{3,9},\\
&\mathbb{C}_{22} = h_{3,9} + h_{8,11,12} - h_{3,8,9} - h_{8,12},\\
&\mathbb{C}_{23} = h_{11,12} + h_{8,12} - h_{8,11,12} - h_5,\\
&\mathbb{C}_{24} = h_{3,8,9} + h_{7,8,10,12} - h_{3,5,8,9} - h_{8,10,12},\\
&\mathbb{C}_{25} = \alpha - h_{3,9},\quad \mathbb{C}_{26} = \beta - h_5.
\end{aligned} \tag{18}$$

Step 5. Compute the Gauss-Jordan form $\mathcal{B} = \{T_i, i \in \mathcal{N}_{3997}\}$.

Step 6. Reduce F by $\widetilde{\mathcal{R}}(\mathcal{B})$ to obtain $F_1 = 4\alpha + 6\beta - 3h_{2,3,5,6,7,8,10}$.

In **Problem $\mathbf{P_2}$**, we have $t_2 = 649$ and 101 variables.

Step 7. In Algorithm 2, we have $m = 649$, $n = 101$. Let $\bar{S}_P$ be the polynomial set obtained in Algorithm 2. Then we have $S_P = \bar{S}_P \cup \{4, 6\}$. Solve the LP in **Problem $\mathbf{P_3}$** to complete the proof. An explicit proof is given by

$$F \geq F_1 = \sum_{i=1}^{26} p_i \mathbb{C}_i \geq 0, \tag{19}$$

where $p_j = 1, j \in \mathcal{N}_{19}$, and $p_{20} = 6, p_{21} = 2, p_{22} = 7, p_{23} = 7, p_{24} = 2, p_{25} = 4, p_{26} = 6$.

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 and Tian's method in [17].

We have obtained the LP in **Problem $\mathbf{P_2}$** by applying Procedure I with **Problem $\mathbf{P_7}$** as the input. Note that we do not have to simplify the LP by taking advantage of the symmetry of the problem and the problem-specific structures as in Tian's method; this is taken care of automatically by Procedure I. From the above table, we see that the LP in **Problem $\mathbf{P_2}$** is much simpler than the original LP, the LP solved in ITIP and the LP solved in Tian's method.

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 $\mathbf{Problem\ P_2}$	101	0	649
LP in $\mathbf{Problem\ P_{2D}}$	649	101	649
LP in $\mathbf{Problem\ P_3}$	548	0	647

To obtain an explicit proof, we need to solve the LP in $\mathbf{Problem\ P_3}$ which has 548 variables and 647 inequality constraints (these 647 inequality constraints contain 99 polynomial inequalities and the non-negativity of the 548 variables). If we only need to verify the inequality without yielding an explicit proof, then we only need to solve the LP in $\mathbf{Problem\ P_2}$.

It turns out that for this particular problem, solving the LP in $\mathbf{Problem\ P_3}$ is much simpler than solving the LP in $\mathbf{Problem\ P_2}$. Specifically, we used 20.0 seconds for solving the LP in $\mathbf{Problem\ P_2}$ but only 2.2 seconds for solving the LP in $\mathbf{Problem\ P_3}$ by MAPLE running on a desktop PC with an i7-6700 Core, 3.40GHz CPU and 16G memory. However, there is no guarantee that this is always the case.

VII. Discussion and Conclusion

In this paper, we develop a new method to prove linear information inequalities and identities. Instead of solving an LP, we transform the problem into a polynomial reduction problem. For the proof of information inequalities, compared with existing methods (ITIP and its variations), our method takes advantage of the algebraic structure of the problem and greatly reduces the computational complexity. For the proof of information identities, we give a simple direct proof method which is much more efficient than the existing methods.

Note that although our method provides an analytic proof for information inequalities/identities, the proof may not be intuitive because the final LP problems to be solved, though equivalent to the original LP problem, are in a different form. One way to obtain a more intuitive proof is to identify which of the original variables (i.e., the joint entropies) are eliminated and which of the original constraints are invoked in the process of obtaining the result. In principle this can be done, but the proof thus obtained would in general be much longer than the proof provided by our method. Nevertheless, this can be a direction for future research.

Our method can split a very large LP problem into several smaller ones that are easier to solve. With this approach, one may be able to solve certain large LP problems which are not solvable otherwise. However, a question worth asking is under what condition our methods are more efficient than solving the original LP problem directly, because it takes computation to split the large LP problem into smaller ones. While we do not have a definite answer, we remark that in general the more constraints there are in the problem, the more efficient our method is.

There are usually many constraints in information theory problems. If there are only very few constraints, it may be more efficient to solve the original LP problem directly. In the extreme case that there is no constraint and no additional inequalities in the problem, since the set of elemental inequalities is already minimal, applying our method would have no benefit at all.

In this paper, we only give the theoretical framework for the methods introduced. The actual implementation of these methods may involve the use of various existing computational software. For different parts of the methods, either symbolic computation or numerical computation can be employed. For symbolic computation, one can use software such as MATHEMATICA and MAPLE. For the part on numerical computation, one can consider CPLEX, GUROBI, etc. Symbolic computation has the advantage of being free of numerical errors and can provide an analytical proof, but it is computationally less efficient. We are in the process of developing a fully automated information inequality and identity prover, which will be reported in the near future.

Acknowledgment

The authors thank the reviewers for the comments which helped improve the presentation of results in the paper. Laigang Guo would like to thank Drs. Tao Guo and Qiaoqiao Zhou for the helpful discussions on Section V, and Prof. Chao Tian and Dr. Liang Chen for the insightful suggestions on Section VI. This work is partially supported by a research grant of the National Natural Science Foundation of China (11688101), and the Fundamental Research Funds for the Central Universities of China (2021NTST32).

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

[10] R. W. Yeung, A. Al-Bashabsheh, C. Chen, Q, Chen, and P. Moulin, “On information-theoretic characterizations of Markov random fields and subfields,” *IEEE Trans. Inf. Theory*, vol. 65, no. 3, pp. 1493-1511, 2018.

[11] 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.

[12] 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.

[13] J. Farkas, “Uber die Theorie der einfachen Ungleichungen,” *J. Reine Angew. Math.*, vol. 124, pp. 1-24, 1902.

[14] D. Achiya, “An elementary proof of Farkas’ lemma,” *SIAM Review*, vol. 39, no. 3, pp. 503-07, 1997.

[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] D. Bremner, K. Fukuda, A. Marzetta, “Primal-dual methods for vertex and facet enumeration (preliminary version)”, Proceedings of the thirteenth annual symposium on Computational geometry. pp: 49-56, 1997.

[22] K.L. Clarkson, “More output-sensitive geometric algorithms”, Proceedings 35th Annual Symposium on Foundations of Computer Science. IEEE, pp: 695-702, 1994.

[23] A. Schrijver, *Combinatorial optimization: polyhedra and efficiency*. Berlin: Springer, 2003.

[24] K. Fukuda (2022), Frequently Asked Questions in Polyhedral Computation. [Online]. Available:
https://people.inf.ethz.ch/fukudak//polyfaq/polyfaq.html

[25] S. Thakor, A. Grant and T. Chan, ”On Complexity Reduction of the LP Bound Computation and Related Problems,” 2011 International Symposium on Networking Coding, pp. 1-6, 2011.

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

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.