

The Explicit Coding Rate Region of Symmetric Multilevel Diversity Coding

Tao Guo and Raymond W. Yeung

Abstract—It is well known that *superposition coding*, namely separately encoding the independent sources, is optimal for symmetric multilevel diversity coding (SMDC) (Yeung-Zhang 1999) for any $L \geq 2$, where L is the number of levels of the coding system. However, the characterization of the coding rate region therein involves uncountably many linear inequalities and the constant term (i.e., the lower bound) in each inequality is given in terms of the solution of a linear optimization problem. Thus this implicit characterization of the coding rate region does not enable the determination of the achievability of a given rate tuple. In principle, the achievability of a given rate tuple can be determined by direct computation, but the complexity is prohibitive even for $L = 5$. In this paper, for any fixed L , we obtain in closed form a finite set of linear inequalities for characterizing the coding rate region. We further show by the symmetry of the problem that only a much smaller subset of this finite set of inequalities needs to be verified in determining the achievability of a given rate tuple. Yet, the cardinality of this smaller set grows at least exponentially fast with L . We also present a subset entropy inequality, which together with our explicit characterization of the coding rate region, is sufficient for proving the optimality of superposition coding.

Index Terms—Symmetric multilevel diversity coding, superposition coding, network coding, closed-form, distributed data storage, robust network communication.

I. INTRODUCTION

Multilevel diversity coding was introduced by Yeung [2]. In a multilevel diversity coding system, an information source is encoded by a number of encoders. There are a set of decoders, which are partitioned into multiple *levels*. The reconstructions of the source by decoders within the same level are identical.

Here, we confine our discussion to multilevel diversity coding systems with symmetrical connectivity between the encoders and decoders, referred to as *symmetrical multilevel diversity coding* (SMDC) [3]–[5]. The SMDC system finds applications in distributed data storage [6], [7], secret sharing

[8]–[10], and robust network communication [11], [12]. It is a special case of multi-source network coding [13]–[15]. This problem can also be regarded as a lossless counterpart of the multiple descriptions problem [16]–[19]. The SMDC coding strategy in turn is used for approximating the multiple descriptions rate region in [20], [21].

In the SMDC problem, there are L ($L \geq 2$) independent discrete memoryless sources $\{X_l(t) : t = 1, 2, \dots\}$, $l = 1, 2, \dots, L$, where for each l , $X_l(t)$ are independent and identically distributed copies of a generic random variable X_l . The importance of the sources is in the order $X_1(t) > X_2(t) > \dots > X_L(t)$. The sources are encoded by L encoders. There are totally $2^L - 1$ decoders, each of which has access to a distinct subset of the encoders. A decoder which can access any α encoders, called a Level α decoder, is required to reconstruct the first α sources. Such a system is called a symmetric L -level diversity coding system. The system is symmetric in the sense that the problem is unchanged by permuting the L encoders, which is evident from the reconstruction requirements of the decoders.

The SMDC problem was treated for $L = 3$ in [4] and in full generality by Yeung and Zhang [5], where a coding method called *superposition coding* (to be formally defined in Section II.B) was proved to be optimal. In this method, the independent sources $\{X_l(t)\}$, $l = 1, 2, \dots, L$ are encoded separately. Albanese *et al.* studied the *priority encoding transmission* (PET) problem in [22], which is almost the same as SMDC. In [22], they proposed a coding scheme using the same idea as superposition coding and further obtained a sum-rate lower bound which is also given in [4]. The problem has subsequently been generalized in different directions. The *secure* communication setting was considered by Balasubramanian *et al.* [23] and Jiang *et al.* [24]. In [24], they also extended the original SMDC setting by introducing an *all-access* encoder which is accessible by all the decoders. In both of the above settings, superposition coding is shown to be optimal. Xiao *et al.* [25] studied the problem of *distributed* multilevel diversity coding where each source is decomposed into L components, each of which is accessed by one distinct encoder. Tian and Liu [26] considered the problem with *regeneration*, where the storage versus repair-bandwidth tradeoff was investigated. Mohajer *et al.* [27] considered the *asymmetric* multilevel diversity coding problem and proved that superposition coding is in general suboptimal. Li *et al.* [28] studied the multilevel diversity coding problem with at most 3 sources and 4 encoders in a systematic way and obtained the exact rate region of each of the over 7,000 instances with the aid of computation. In the current paper, we focus on some

This paper was presented in part at 2018 IEEE International Symposium on Information Theory (ISIT) [1].

Tao Guo was with the Institute of Network Coding and the Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong SAR, China. He is now with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA. (e-mail: guotao@tamu.edu)

Raymond Yeung is with the Institute of Network Coding and the Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong SAR, China. He also holds an adjunct position at the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China. (e-mail: whyeung@ie.cuhk.edu.hk)

Their work was funded in part by the University Grant Committee of the Hong Kong SAR, China (Project No. AoE/E-02/08), the National Natural Science Foundation of China (Grant No. 61471215), and the Science, Technology and Innovation Commission of Shenzhen Municipality (Project No. JSGG20160301170514984).

fundamental issues pertaining to the original SMDC problem discussed in [4], [5].

It was proved in [4] that superposition coding is optimal for $L = 3$, and the corresponding coding rate region, referred to as the superposition coding rate region, can be explicitly characterized by 10 linear inequalities in the coding rates of the 3 encoders. Thus, the achievability of any given rate triple can be determined by verifying these 10 inequalities.

However, the optimality proof in [4] is not readily generalizable to a general L . Here is an outline of the proof in [4]. The superposition coding rate region is first characterized by the aforementioned 10 inequalities. This involves the determination of the extreme points of the region. Then the necessity of these 10 inequalities are established by means of conventional techniques for proving converse coding theorems. The difficulty for generalizing the proof to a general L is two-fold:

- 1) It is observed through computation that both the number of linear inequalities needed for characterizing the superposition coding rate region and the number of extreme points of this region grow with L . As such, it is impossible to determine all of them for a general L .
- 2) For a fixed L , once the superposition coding rate region is characterized by a finite set of linear inequalities, their necessity needs to be proved. With conventional techniques, this needs to be done for each inequality in a way that depends on the coefficients of coding rates. It is observed through computation that the number of these inequalities grows with L . Therefore, for a general L , it is not possible to prove the necessity of all of these inequalities.

For a fixed L , the extreme points of the superposition coding rate region and the set of linear inequalities characterizing the region can in principle be found by computation. However, the complexity grows very quickly with L and becomes prohibitive even for $L = 5$. On a notebook computer, by using the Fourier-Motzkin elimination algorithm [29], we were able to compute all the linear inequalities needed for characterizing the superposition coding rate region for $L = 4$ in less than 2 minutes. However, the computation involved for $L = 5$ is already unmanageable.

In [5], the optimality of superposition coding was established for a general L by means of a highly sophisticated method that does not involve any explicit characterization of the coding rate region. Instead of a fixed L , the problem is tackled for a general L . As L is not fixed, the number of linear inequalities needed for the characterization of the superposition coding rate region may be unbounded. To get around the problem, the coding rate region is characterized by an uncountable collection of linear inequalities, where for each inequality, the coefficients associated with the rates are arbitrary nonnegative real numbers with at least one of them being nonzero. The constant terms (i.e., the lower bounds) in these inequalities are given implicitly in terms of the solution of a *common* linear optimization problem with the coefficients associated with the rates as parameters. In other words, although the coding rate region is characterized by uncountably many linear inequalities, they have a common

form and the necessity of these inequalities can be established in a unified manner.

Although the optimality of superposition coding for a general L has been established in [5], this result does not yield an explicit characterization of the coding rate region for any fixed L . In particular, it does not enable the determination of the achievability of a given rate tuple, even for a fixed L , for the following two reasons. First, the characterization of the coding rate region in [5] involves uncountably many inequalities. Second, each inequality in the characterization is implicit, and can be made explicit only by solving a linear optimization problem.

In the present paper, we develop fundamental results pertaining to SMDC. Our main contributions are summarized as follows:

- 1) We obtain an explicit characterization of the coding rate region for a general L . This is done by first solving in closed form the linear optimization problem in [5] that gives an implicit characterization of the coding rate region. Then among all the uncountably many inequalities involved in characterizing the coding rate region, we identify a finite subset that is sufficient for characterizing the coding rate region. It is further proved that there is no redundancy in this finite set of inequalities. Thus for a fixed L , the achievability of any given rate tuple can be determined.
- 2) By taking advantage of the symmetry of the problem, we show that in determining the achievability of a given rate tuple, it suffices to verify a much smaller subset of the set of inequalities identified in 1). Yet, the cardinality of this smaller set of inequalities grows at least exponentially fast with L . This reveals the inherent complexity of the problem.
- 3) A subset entropy inequality, which plays a key role in the converse proof in [5], requires a painstaking and extremely technical proof. We present a weaker version of this subset entropy inequality whose proof is considerably simpler. With our explicit characterization of the coding rate region, this weaker version of the subset entropy inequality is sufficient for proving the optimality of superposition coding.

The rest of the paper is organized as follows. We first formulate the problem and state some existing results in Section II. In Section III, we present a closed-form solution of the linear optimization problem in [5] and establish some basic properties of the solution. In Section IV, we identify a finite set of inequalities that characterizes the superposition coding rate region and show that this set contains no redundancy. In Section V, we further identify a subset of inequalities we need to verify in determining the achievability of a given rate tuple. We also provide a lower bound and an upper bound on the cardinality of this set. In Section VI, we present a weaker version of the subset entropy inequality in [5]. We conclude the paper in Section VII. Some essential proofs can be found in the appendices.

II. PROBLEM FORMULATION AND EXISTING RESULTS

A. Problem Formulation

An L -level SMDC system, $L \geq 2$, is depicted in Fig. 1. The problem is defined as follows. Let

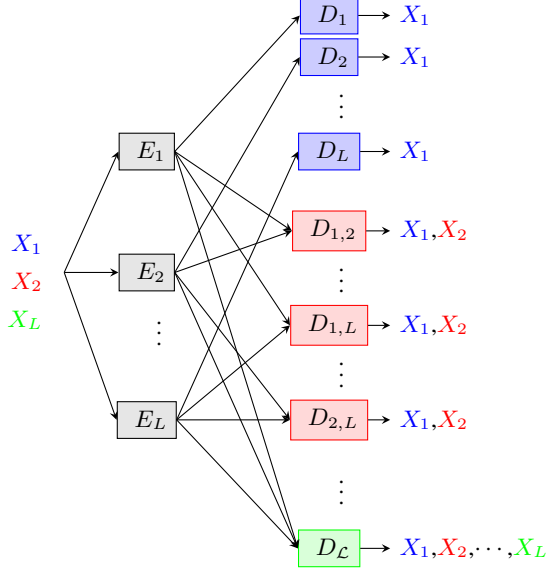


Fig. 1: The symmetrical multilevel diversity coding system.

$\mathcal{L} = \{1, 2, \dots, L\}$. Let t be the time index and $\{(X_1(t), X_2(t), \dots, X_L(t)) : t = 1, 2, \dots\}$ be a collection of L independent discrete memoryless information sources with an L -tuple of generic random variables (X_1, X_2, \dots, X_L) taking values in $\mathcal{X}_1 \times \mathcal{X}_2 \times \dots \times \mathcal{X}_L$, where $\mathcal{X}_i, i \in \mathcal{L}$ are finite alphabets. There are L encoders, indexed by \mathcal{L} , each of which can access all the L information sources. There are also $2^L - 1$ decoders. For each $\mathcal{U} \subseteq \mathcal{L}$ such that $\mathcal{U} \neq \emptyset$, Decoder- \mathcal{U} can access the subset of encoders indexed by \mathcal{U} . Without loss of generality, assume the elements in \mathcal{U} are in an ascending order. For $1 \leq \alpha \leq L$ and \mathcal{U} such that $|\mathcal{U}| = \alpha$, Decoder- \mathcal{U} can reconstruct the first α sources $\{X_1(t), X_2(t), \dots, X_\alpha(t)\}$ perfectly asymptotically, which will be defined later.

An $(n, M_1, M_2, \dots, M_L)$ code is defined by the encoding functions

$$E_l : \prod_{i=1}^L \mathcal{X}_i^n \rightarrow \{1, 2, \dots, M_l\}, \text{ for } l \in \mathcal{L}$$

and decoding functions

$$D_{\mathcal{U}} : \prod_{l \in \mathcal{U}} \{1, 2, \dots, M_l\} \rightarrow \prod_{i=1}^{|\mathcal{U}|} \mathcal{X}_i^n, \text{ for } \mathcal{U} \subseteq \mathcal{L} \text{ and } \mathcal{U} \neq \emptyset.$$

For $1 \leq \alpha \leq L$, let $\mathbf{X}_\alpha = (X_\alpha(1), X_\alpha(2), \dots, X_\alpha(n))$. Let $W_l = E_l(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_L)$ be the output of Encoder- l and $W_{\mathcal{U}} = (W_i : i \in \mathcal{U})$ for $\mathcal{U} \subseteq \mathcal{L}$.¹ A nonnegative rate tuple (R_1, R_2, \dots, R_L) is *achievable* if for any $\epsilon > 0$, there exists

¹Here $E_l(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_L)$ is a function of random vectors and hence W_l is a random variable. The reader should not confuse E_l with the expectation of a random variable.

for sufficiently large n an $(n, M_1, M_2, \dots, M_L)$ code such that

$$\frac{1}{n} \log M_l \leq R_l + \epsilon, \forall l \in \mathcal{L},$$

and

$$\Pr\{D_{\mathcal{U}}(W_{\mathcal{U}}) \neq (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_\alpha)\} \leq \epsilon,$$

for all $\alpha = 1, 2, \dots, L$ and $\mathcal{U} \subseteq \mathcal{L}$ such that $|\mathcal{U}| = \alpha$. The achievable rate region \mathcal{R} is defined as the collection of all achievable rate tuples.

B. Existing Results

We adopt the terminologies and notations in [5]. Let \mathcal{R}_{sup} be the rate region induced by superposition coding. Then \mathcal{R}_{sup} is the set of nonnegative rate tuples $\mathbf{R} = (R_1, R_2, \dots, R_L)$ such that

$$R_l = \sum_{\alpha=1}^L r_l^\alpha, \text{ for } l \in \mathcal{L} \quad (1)$$

for some $r_l^\alpha \geq 0$, $1 \leq \alpha \leq L$, satisfying

$$\sum_{l \in \mathcal{U}} r_l^\alpha \geq H(X_\alpha), \text{ for all } \mathcal{U} \subseteq \mathcal{L} \text{ and } |\mathcal{U}| = \alpha. \quad (2)$$

For an elaborative discussion on superposition coding for the 3-level SMDC system, we refer the reader to [4].

For a fixed L , based on (1) and (2), one can apply the Fourier-Motzkin algorithm to eliminate r_l^α for $l, \alpha \in \mathcal{L}$. The output is a set of linear inequalities involving $R_l, l \in \mathcal{L}$ that gives an explicit characterization of \mathcal{R}_{sup} . However, as mentioned in Section I, the computation involved for $L \geq 5$ is unmanageable.

Let $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_L)$ and

$$\mathbb{R}_+^L = \{\boldsymbol{\lambda} : \boldsymbol{\lambda} \neq \mathbf{0} \text{ and } \lambda_i \in \mathbb{R}, \lambda_i \geq 0 \text{ for } i \in \mathcal{L}\}. \quad (3)$$

Let $\Omega_L^\alpha = \{\mathbf{v} \in \{0, 1\}^L : |\mathbf{v}| = \alpha\}$, where $|\mathbf{v}|$ is the Hamming weight of a vector $\mathbf{v} = (v_1, v_2, \dots, v_L)$. Note that there is a one-to-one correspondence between a vector $\mathbf{v} \in \{0, 1\}^L$ and Decoder- \mathcal{U} , where $\mathcal{U} = \{i : v_i = 1\}$. For any $\mathbf{v} \in \Omega_L^\alpha$, let $c_\alpha(\mathbf{v})$ be any nonnegative real number. For any $\boldsymbol{\lambda} \in \mathbb{R}_+^L$ and $\alpha \in \mathcal{L}$, let $f_\alpha(\boldsymbol{\lambda})$ be the optimal solution to the following optimization problem:

$$f_\alpha(\boldsymbol{\lambda}) \triangleq \max_{\mathbf{v} \in \Omega_L^\alpha} \sum c_\alpha(\mathbf{v}) \quad (4)$$

$$\text{s.t. } \sum_{\mathbf{v} \in \Omega_L^\alpha} c_\alpha(\mathbf{v}) \mathbf{v} \leq \boldsymbol{\lambda} \quad (5)$$

$$c_\alpha(\mathbf{v}) \geq 0, \forall \mathbf{v} \in \Omega_L^\alpha. \quad (6)$$

Note that the functions $f_\alpha(\cdot)$ and $c_\alpha(\cdot)$ above depend on L , but for simplicity we omit this dependency in the notations. Thus, if the length of $\boldsymbol{\lambda}$ is given, then $f_\alpha(\boldsymbol{\lambda})$ can be defined accordingly. A set $\{c_\alpha(\mathbf{v}) : \mathbf{v} \in \Omega_L^\alpha\}$ is called an α -resolution for $\boldsymbol{\lambda}$ if (5) and (6) are satisfied and it will be abbreviated as $\{c_\alpha(\mathbf{v})\}$ if there is no ambiguity. Furthermore, an α -resolution is called *optimal* if it achieves the optimal value $f_\alpha(\boldsymbol{\lambda})$.

Remark 1. Here is an intuitive explanation of $f_\alpha(\boldsymbol{\lambda})$: Consider a set of items from L different types indexed by \mathcal{L} , where the number of items of type i ($i \in \mathcal{L}$) is λ_i . An α -type group is

defined as a group of α items of different types. The goal is to cluster the items into α -type groups so that the total number of such groups is maximized. This maximum is defined as $f_\alpha(\lambda)$.

Let \mathcal{R}_h be the collection of nonnegative rate tuples \mathbf{R} such that

$$\sum_{l=1}^L \lambda_l R_l \geq \sum_{\alpha=1}^L f_\alpha(\lambda) H(X_\alpha), \text{ for all } \lambda \in \mathbb{R}_+^L. \quad (7)$$

It was proved in [5] that the superposition region \mathcal{R}_{sup} can be alternatively characterized by \mathcal{R}_h . This means that in addition to being the optimal value of the optimization problem in (4), for every fixed $\lambda \in \mathbb{R}_+^L$, $f_\alpha(\lambda)$ also gives a tightest possible linear outer bound on \mathcal{R}_{sup} via (7). It was further proved in [5] that \mathcal{R}_h is an outer bound on \mathcal{R} . Then

$$\mathcal{R}_{\text{sup}} \subseteq \mathcal{R} \subseteq \mathcal{R}_h$$

which implies

$$\mathcal{R} = \mathcal{R}_h = \mathcal{R}_{\text{sup}}, \quad (8)$$

i.e., superposition coding is optimal.

The following lemma is a direct consequence of Lemma 4 and 7 in [5]. It will be used in the proof of our main result in the next section.

Lemma 1. Assume $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L$. For $\alpha \geq 2$, if $\lambda_1 \leq \frac{\lambda_2 + \lambda_3 + \dots + \lambda_L}{\alpha - 1}$, then $f_\alpha(\lambda) = \frac{1}{\alpha} \sum_{i=1}^L \lambda_i$.

III. OPTIMAL α -RESOLUTION

For any $\lambda \in \mathbb{R}_+^L$ and any permutation ω on $\{1, 2, \dots, L\}$, with an abuse of notation, we denote $(\lambda_{\omega(1)}, \lambda_{\omega(2)}, \dots, \lambda_{\omega(L)})$ by $\omega(\lambda)$. For each $\alpha \in \mathcal{L}$, due to the symmetry of the system, it is intuitive that the values of $f_\alpha(\omega(\lambda))$ are the same for all ω . This important property of $f_\alpha(\lambda)$ is formally proved in the following lemma.

Lemma 2. $f_\alpha(\omega(\lambda)) = f_\alpha(\lambda)$ for any $\alpha \in \mathcal{L}$.

Proof. For any $\alpha \in \mathcal{L}$, let $\{c_\alpha(v) : v \in \Omega_L^\alpha\}$ be an optimal α -resolution for λ . Then we have

$$\begin{aligned} \sum_{v \in \Omega_L^\alpha} c_\alpha(v) v &\leq \lambda, \\ c_\alpha(v) &\geq 0, \forall v \in \Omega_L^\alpha, \end{aligned} \quad (9)$$

and

$$f_\alpha(\lambda) = \sum_{v \in \Omega_L^\alpha} c_\alpha(v).$$

Let $\sum_{v \in \Omega_L^\alpha} c_\alpha(v) v = \tilde{\lambda}$. Then by (9), we have $\tilde{\lambda} \leq \lambda$. For any permutation ω on $\{1, 2, \dots, L\}$, we can check that

$$\sum_{v \in \Omega_L^\alpha} c_\alpha(v) \omega(v) = \omega \left(\sum_{v \in \Omega_L^\alpha} c_\alpha(v) v \right) = \omega(\tilde{\lambda}) \leq \omega(\lambda). \quad (10)$$

For any $v \in \Omega_L^\alpha$, let

$$c'_\alpha(\omega(v)) = c_\alpha(v).$$

It is immediate that for all $v \in \Omega_L^\alpha$,

$$c'_\alpha(\omega(v)) \geq 0. \quad (11)$$

Since ω is a one-to-one mapping from Ω_L^α to Ω_L^α , we have $v \in \Omega_L^\alpha$ if and only if $\omega(v) \in \Omega_L^\alpha$ for any ω . Thus,

$$\begin{aligned} \sum_{\omega(v) \in \Omega_L^\alpha} c'_\alpha(\omega(v)) \omega(v) &= \sum_{v \in \Omega_L^\alpha} c'_\alpha(\omega(v)) \omega(v) \\ &= \sum_{v \in \Omega_L^\alpha} c_\alpha(v) \omega(v) \\ &\leq \omega(\lambda), \end{aligned} \quad (12)$$

where the inequality follows from (10). By (11) and (12), we see that $\{c'_\alpha(\omega(v)) : v \in \Omega_L^\alpha\}$ is an α -resolution for $\omega(\lambda)$. In light of the definition of $f_\alpha(\lambda)$ in (4)-(6), we have

$$\begin{aligned} f_\alpha(\omega(\lambda)) &\geq \sum_{\omega(v) \in \Omega_L^\alpha} c'_\alpha(\omega(v)) \\ &= \sum_{v \in \Omega_L^\alpha} c'_\alpha(\omega(v)) \\ &= \sum_{v \in \Omega_L^\alpha} c_\alpha(v) \\ &= f_\alpha(\lambda), \end{aligned}$$

and so

$$f_\alpha(\omega(\lambda)) \geq f_\alpha(\lambda). \quad (13)$$

Let ω^{-1} be the inverse permutation of ω . By the same argument, we can obtain

$$f_\alpha(\omega^{-1}(\omega(\lambda))) \geq f_\alpha(\omega(\lambda)). \quad (14)$$

Since $\omega^{-1}(\omega(\lambda)) = \lambda$, we see from (13) and (14) that

$$f_\alpha(\lambda) = f_\alpha(\omega(\lambda)) \quad \text{for all } \alpha \in \mathcal{L}.$$

The lemma is proved. \square

If a vector λ satisfies

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L, \quad (15)$$

we call λ an *ordered vector*. Throughout this section, except for Lemma 8, in light of Lemma 2, we assume without loss of generality that λ is an ordered vector. For any $\alpha \in \mathcal{L}$, it is easy to see that

$$f_\alpha(\mu\lambda) = \mu f_\alpha(\lambda) \quad (16)$$

for all $\mu \in \mathbb{R}$ such that $\mu > 0$. In view of (7) and (16), we will consider only λ 's whose minimum nonzero element is equal to 1. Then there exists a $\zeta \in \mathcal{L}$ such that

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_\zeta = 1$$

and $\lambda_i = 0$ for all $i = \zeta + 1, \zeta + 2, \dots, L$.

Fix λ , it is easy to see that

$$f_1(\lambda) = \sum_{i=1}^L \lambda_i, \quad (17)$$

and

$$f_\zeta(\lambda) = 1,$$

and for $\alpha \geq \zeta + 1$,

$$f_\alpha(\lambda) = 0.$$

For other cases, determining the value of $f_\alpha(\lambda)$ is highly nontrivial.

For $\alpha \in \mathcal{L}$ and $\beta = 0, 1, \dots, \alpha - 1$, let

$$g_{\alpha, \lambda}(\beta) = \frac{1}{\alpha - \beta} \sum_{i=\beta+1}^L \lambda_i. \quad (18)$$

Let β_α^* be a value of β (not necessarily unique) that achieves the minimum $\min_{\beta \in \{0, 1, \dots, \alpha-1\}} g_{\alpha, \lambda}(\beta)$, i.e.,

$$g_{\alpha, \lambda}(\beta_\alpha^*) = \min_{\beta \in \{0, 1, \dots, \alpha-1\}} g_{\alpha, \lambda}(\beta). \quad (19)$$

The following theorem, a main result of the current paper, gives a closed-form solution for $f_\alpha(\lambda)$.

Theorem 1. For any $\alpha \in \mathcal{L}$, $f_\alpha(\lambda) = g_{\alpha, \lambda}(\beta_\alpha^*)$.

Proof. Fix an $\alpha \in \mathcal{L}$, and denote β_α^* by β^* for simplicity. We prove the theorem by proving i) $f_\alpha(\lambda) \leq g_{\alpha, \lambda}(\beta^*)$; ii) there exists a solution for the optimization problem (4) that achieves $g_{\alpha, \lambda}(\beta^*)$, so that $f_\alpha(\lambda) \geq g_{\alpha, \lambda}(\beta^*)$.

i) $f_\alpha(\lambda) \leq g_{\alpha, \lambda}(\beta^*)$.

For $0 \leq \beta \leq \alpha - 1$, let e_β be an L -vector with the first β components being 0 and the last $L - \beta$ components being 1. For any $v \in \Omega_L^\alpha$, since $\sum_{i=1}^\beta v_i \leq \beta$, we have

$$v \cdot e_\beta = \sum_{i=\beta+1}^L v_i \geq \alpha - \beta.$$

Then for any solutions $\{c_\alpha(v)\}$ to the optimization problem in (4), we have

$$\begin{aligned} \sum_{i=\beta+1}^L \lambda_i &= \lambda \cdot e_\beta \\ &\geq \left(\sum_{v \in \Omega_L^\alpha} c_\alpha(v) v \right) \cdot e_\beta \\ &= \sum_{v \in \Omega_L^\alpha} c_\alpha(v) (v \cdot e_\beta) \\ &\geq \sum_{v \in \Omega_L^\alpha} c_\alpha(v) (\alpha - \beta) \\ &= (\alpha - \beta) \sum_{v \in \Omega_L^\alpha} c_\alpha(v). \end{aligned}$$

This implies that for all $0 \leq \beta \leq \alpha - 1$,

$$f_\alpha(\lambda) \leq \frac{1}{\alpha - \beta} \sum_{i=\beta+1}^L \lambda_i = g_{\alpha, \lambda}(\beta).$$

Thus, we have

$$f_\alpha(\lambda) \leq g_{\alpha, \lambda}(\beta^*).$$

ii) $f_\alpha(\lambda) \geq g_{\alpha, \lambda}(\beta^*)$.

We now show that there exists a solution that achieves $g_{\alpha, \lambda}(\beta^*)$. For any $\alpha \in \mathcal{L}$ and $\beta^* \in \{0, 1, \dots, \alpha - 2\}$, by (19), we have

$$\frac{1}{\alpha - \beta^*} \sum_{i=\beta^*+1}^L \lambda_i \leq \frac{1}{\alpha - (\beta^* + 1)} \sum_{i=\beta^*+2}^L \lambda_i,$$

which is equivalent to

$$\lambda_{\beta^*+1} \leq \frac{1}{(\alpha - \beta^*) - 1} \sum_{i=\beta^*+2}^L \lambda_i. \quad (20)$$

Denote the $(L - \beta^*)$ -vector $(\lambda_{\beta^*+1}, \lambda_{\beta^*+2}, \dots, \lambda_L)$ by λ' . In view of (20), by Lemma 1, (18), and (19), we have

$$f_{\alpha-\beta^*}(\lambda') = \frac{1}{\alpha - \beta^*} \sum_{i=\beta^*+1}^L \lambda_i = g_{\alpha, \lambda}(\beta^*). \quad (21)$$

In view of (17) and (18) with $\beta = \beta^*$, it is easy to check that (21) is also satisfied for $\beta^* = \alpha - 1$. Without loss of generality, let $\{c_{\alpha-\beta^*}(\mathbf{u}) : \mathbf{u} \in \Omega_{L-\beta^*}^{\alpha-\beta^*}\}$ be an optimal $(\alpha - \beta^*)$ -resolution for λ' . Then it follows from (21) that

$$\sum_{\mathbf{u} \in \Omega_{L-\beta^*}^{\alpha-\beta^*}} c_{\alpha-\beta^*}(\mathbf{u}) = f_{\alpha-\beta^*}(\lambda') = g_{\alpha, \lambda}(\beta^*).$$

For any $v \in \Omega_L^\alpha$, let

$$c_\alpha(v) = \begin{cases} c_{\alpha-\beta^*}(\mathbf{u}), & \text{if } v = (11 \cdots 1\mathbf{u}) \\ & \text{for some } \mathbf{u} \in \Omega_{L-\beta^*}^{\alpha-\beta^*} \\ 0, & \text{otherwise.} \end{cases} \quad (22)$$

Then we have

$$\sum_{v \in \Omega_L^\alpha} c_\alpha(v) = \sum_{\mathbf{u} \in \Omega_{L-\beta^*}^{\alpha-\beta^*}} c_{\alpha-\beta^*}(\mathbf{u}) = g_{\alpha, \lambda}(\beta^*). \quad (23)$$

Again, by (19), we have

$$\frac{1}{\alpha - (\beta^* - 1)} \sum_{i=\beta^*}^L \lambda_i \geq \frac{1}{\alpha - \beta^*} \sum_{i=\beta^*+1}^L \lambda_i.$$

Then

$$\lambda_{\beta^*} \geq \frac{1}{\alpha - \beta^*} \sum_{i=\beta^*+1}^L \lambda_i = g_{\alpha, \lambda}(\beta^*),$$

where the equality above follows from (18). Thus,

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{\beta^*} \geq g_{\alpha, \lambda}(\beta^*).$$

For $i = 1, 2, \dots, \beta^*$, since $c_\alpha(v) = 0$ if $v_i = 0$, we have

$$\sum_{v \in \Omega_L^\alpha: v_i=1} c_\alpha(v) = \sum_{v \in \Omega_L^\alpha} c_\alpha(v) = g_{\alpha, \lambda}(\beta^*) \leq \lambda_i, \quad (24)$$

where the second equality follows from (23). For $i = \beta^* + 1, \beta^* + 2, \dots, L$,

$$\begin{aligned} \sum_{v \in \Omega_L^\alpha: v_i=1} c_\alpha(v) &= \sum_{\substack{v \in \Omega_L^\alpha: v_i=1, \\ (v_1, \dots, v_{\beta^*})=1}} c_\alpha(v) + \sum_{\substack{v \in \Omega_L^\alpha: v_i=1, \\ (v_1, \dots, v_{\beta^*}) \neq 1}} c_\alpha(v) \\ &= \sum_{\substack{\mathbf{u} \in \Omega_{L-\beta^*}^{\alpha-\beta^*}: \\ u_{i-\beta^*}=1}} c_{\alpha-\beta^*}(\mathbf{u}) + 0 \\ &\leq \lambda_i, \end{aligned} \quad (25)$$

since $\{c_{\alpha-\beta^*}(\mathbf{u}) : \mathbf{u} \in \Omega_{L-\beta^*}^{\alpha-\beta^*}\}$ is an optimal $(\alpha - \beta^*)$ -resolution for λ' . From (23), (24), and (25), we can see

that $\{c_\alpha(\mathbf{v}) : \mathbf{v} \in \Omega_L^\alpha\}$ defined by (22) is an α -resolution for λ that achieves $g_{\alpha,\lambda}(\beta^*)$. Thus, we have

$$f_\alpha(\lambda) \geq g_{\alpha,\lambda}(\beta^*).$$

□ and

The following lemma provides an important insight into the minimum in (19).

Lemma 3. For any $\alpha \in \{2, 3, \dots, L\}$ and $0 \leq \beta \leq \alpha - 2$,

(i) if $g_{\alpha,\lambda}(\beta) \geq g_{\alpha,\lambda}(\beta + 1)$, then

$$g_{\alpha,\lambda}(0) \geq g_{\alpha,\lambda}(1) \geq \dots \geq g_{\alpha,\lambda}(\beta + 1);$$

(ii) if $g_{\alpha,\lambda}(\beta) \leq g_{\alpha,\lambda}(\beta + 1)$, then

$$g_{\alpha,\lambda}(\beta) \leq g_{\alpha,\lambda}(\beta + 1) \leq \dots \leq g_{\alpha,\lambda}(\alpha - 1).$$

Remark 2. In Lemma 3, if all the non-strict inequalities are replaced by strict inequalities, the lemma remains valid. This alternative version of Lemma 3 can be proved by modifying the proof below accordingly.

Remark 3. Lemma 3 reveals the *pseudo-convexity* [30] of the function $g_{\alpha,\lambda}(\beta)$.

Proof of Lemma 3. In the following, we only prove (ii). The proof for (i) can be obtained similarly.

For $\alpha = 2$, the lemma is immediate. For $3 \leq \alpha \leq L$ and $\beta = \alpha - 2$, (ii) is immediate. For $0 \leq \beta \leq \alpha - 3$, from the definition of $g_{\alpha,\lambda}(\cdot)$ in (18), the condition $g_{\alpha,\lambda}(\beta) \leq g_{\alpha,\lambda}(\beta + 1)$ is equivalent to

$$\frac{1}{\alpha - \beta} \sum_{i=\beta+1}^L \lambda_i \leq \frac{1}{\alpha - (\beta + 1)} \sum_{i=\beta+2}^L \lambda_i,$$

or

$$\lambda_{\beta+1} \leq \frac{1}{\alpha - (\beta + 1)} \sum_{i=\beta+2}^L \lambda_i.$$

Then by the assumption in (15), we have

$$\lambda_{\beta+2} \leq \frac{1}{\alpha - (\beta + 1)} \sum_{i=\beta+2}^L \lambda_i,$$

or

$$\lambda_{\beta+2} \leq \frac{1}{\alpha - (\beta + 2)} \sum_{i=\beta+3}^L \lambda_i,$$

which is also equivalent to

$$\frac{1}{\alpha - (\beta + 1)} \sum_{i=\beta+2}^L \lambda_i \leq \frac{1}{\alpha - (\beta + 2)} \sum_{i=\beta+3}^L \lambda_i.$$

From (18), we have

$$g_{\alpha,\lambda}(\beta + 1) \leq g_{\alpha,\lambda}(\beta + 2).$$

Then we see inductively that for all $\beta + 1 \leq \beta' \leq \alpha - 2$,

$$g_{\alpha,\lambda}(\beta') \leq g_{\alpha,\lambda}(\beta' + 1).$$

The lemma is proved. □

For any $\alpha \in \{2, 3, \dots, L\}$ and any $\beta \in \{0, 1, \dots, \alpha - 1\}$, we can readily see from Lemma 3 that $\beta_\alpha^* = \beta$ if and only if

$$g_{\alpha,\lambda}(0) \geq g_{\alpha,\lambda}(1) \geq \dots \geq g_{\alpha,\lambda}(\beta)$$

$$g_{\alpha,\lambda}(\beta) \leq g_{\alpha,\lambda}(\beta + 1) \leq \dots \leq g_{\alpha,\lambda}(\alpha - 1).$$

This provides a method to find the optimal value β_α^* conveniently. We only need to compare $g_{\alpha,\lambda}(\beta)$ and $g_{\alpha,\lambda}(\beta + 1)$ for $\beta = 0, 1, \dots, \alpha - 2$ successively and stop at the first β such that $g_{\alpha,\lambda}(\beta) \leq g_{\alpha,\lambda}(\beta + 1)$. Then this β gives a value of β_α^* that achieves the minimum in (19).

Lemma 4. $0 = \beta_1^* \leq \beta_2^* \leq \dots \leq \beta_L^*$.

Proof. It is easy to see from (17) that $\beta_1^* = 0$. This implies that $\beta_1^* \leq \beta_2^*$. Now, we prove the lemma by showing that $\beta_{\alpha-1}^* \leq \beta_\alpha^*$ for any $3 \leq \alpha \leq L$. If $\beta_\alpha^* \in \{\alpha - 2, \alpha - 1\}$, since for a fixed $\alpha \in \mathcal{L}$ we have $0 \leq \beta \leq \alpha - 1$, it is obvious that

$$\beta_{\alpha-1}^* \leq \alpha - 2 \leq \beta_\alpha^*.$$

Otherwise, $\beta_\alpha^* \in \{0, 1, \dots, \alpha - 3\}$. Since β_α^* achieves the minimum in (19), we have

$$\frac{1}{\alpha - \beta_\alpha^*} \sum_{i=\beta_\alpha^*+1}^L \lambda_i \leq \frac{1}{\alpha - (\beta_\alpha^* + 1)} \sum_{i=\beta_\alpha^*+2}^L \lambda_i,$$

which is equivalent to

$$\lambda_{\beta_\alpha^*+1} \leq \frac{1}{\alpha - (\beta_\alpha^* + 1)} \sum_{i=\beta_\alpha^*+2}^L \lambda_i.$$

This implies that

$$\lambda_{\beta_\alpha^*+1} \leq \frac{1}{\alpha - (\beta_\alpha^* + 2)} \sum_{i=\beta_\alpha^*+2}^L \lambda_i,$$

which is equivalent to

$$\frac{1}{\alpha - (\beta_\alpha^* + 1)} \sum_{i=\beta_\alpha^*+1}^L \lambda_i \leq \frac{1}{\alpha - (\beta_\alpha^* + 2)} \sum_{i=\beta_\alpha^*+2}^L \lambda_i.$$

Thus, we have

$$\frac{1}{(\alpha - 1) - \beta_\alpha^*} \sum_{i=\beta_\alpha^*+1}^L \lambda_i \leq \frac{1}{(\alpha - 1) - (\beta_\alpha^* + 1)} \sum_{i=\beta_\alpha^*+2}^L \lambda_i,$$

which by (18) implies that

$$g_{\alpha-1,\lambda}(\beta_\alpha^*) \leq g_{\alpha-1,\lambda}(\beta_\alpha^* + 1).$$

By the discussion following Lemma 3, we conclude that

$$\beta_{\alpha-1}^* \leq \beta_\alpha^*.$$

□

The following lemma will be used for proving Lemma 6.

Lemma 5. Let $\lambda_1 = (\lambda_{1,1}, \lambda_{1,2}, \dots, \lambda_{1,L})$ and $\lambda_2 = (\lambda_{2,1}, \lambda_{2,2}, \dots, \lambda_{2,L})$ be two vectors such that $\lambda_{1,1} > \lambda_{2,1}$ and $\lambda_{1,i} = \lambda_{2,i}$ for all $2 \leq i \leq L$. For any $\alpha_0 \in \mathcal{L}$, if $f_{\alpha_0}(\lambda_1) = f_{\alpha_0}(\lambda_2)$, then $f_\alpha(\lambda_1) = f_\alpha(\lambda_2)$ for all $\alpha \geq \alpha_0$. □

Proof. For $\alpha \in \mathcal{L}$, let β_α^1 and β_α^2 be the values (not necessarily unique) that achieve $f_\alpha(\lambda_1)$ and $f_\alpha(\lambda_2)$, respectively. We first prove the claim that among all the possible values of $\beta_{\alpha_0}^1$ and $\beta_{\alpha_0}^2$, there exists a pair of $(\beta_{\alpha_0}^1, \beta_{\alpha_0}^2)$ such that $\beta_{\alpha_0}^1 \geq 1$ and $\beta_{\alpha_0}^2 \geq 1$. Consider the following four cases for all the possible values of $(\beta_{\alpha_0}^1, \beta_{\alpha_0}^2)$:

- i) $\beta_{\alpha_0}^1 = \beta_{\alpha_0}^2 = 0$;
- ii) $\beta_{\alpha_0}^1 \geq 1, \beta_{\alpha_0}^2 = 0$;
- iii) $\beta_{\alpha_0}^1 = 0, \beta_{\alpha_0}^2 \geq 1$;
- iv) $\beta_{\alpha_0}^1 \geq 1, \beta_{\alpha_0}^2 \geq 1$.

Since $f_{\alpha_0}(\lambda_1) = f_{\alpha_0}(\lambda_2)$, it is easy to see that i) and ii) are impossible. If iii) is true, we have

$$\frac{1}{\alpha_0} \sum_{i=1}^L \lambda_{1,i} = \frac{1}{\alpha_0 - \beta_{\alpha_0}^2} \sum_{i=\beta_{\alpha_0}^2+1}^L \lambda_{2,i} = \frac{1}{\alpha_0 - \beta_{\alpha_0}^2} \sum_{i=\beta_{\alpha_0}^2+1}^L \lambda_{1,i},$$

where the second equality follows from $\beta_{\alpha_0}^2 \geq 1$. This implies that

$$f_{\alpha_0}(\lambda_1) = \frac{1}{\alpha_0 - \beta_{\alpha_0}^2} \sum_{i=\beta_{\alpha_0}^2+1}^L \lambda_{1,i},$$

i.e. $(\beta_{\alpha_0}^2, \beta_{\alpha_0}^2)$ is a possible pair. This proves the claim. For all $\alpha \geq \alpha_0$, by Lemma 4, we have $\beta_\alpha^1 \geq 1$ and $\beta_\alpha^2 \geq 1$. Then by Theorem 1, we have

$$f_\alpha(\lambda_1) = f_\alpha(\lambda_2) = \min_{\beta \in \{1, 2, \dots, \alpha-1\}} g_{\alpha, \lambda_1}(\beta).$$

The lemma is proved. \square

Let $\lambda^{[1]}$ be the L -vector with the first component being 1 and the rest being 0, i.e.,

$$\lambda^{[1]} = (1, 0, 0, \dots, 0). \quad (26)$$

Lemma 6. If $\lambda_1 > \sum_{i=2}^L \lambda_i$, let $\lambda' = (\sum_{i=2}^L \lambda_i, \lambda_2, \lambda_3, \dots, \lambda_L)$. Then for all $\alpha \in \mathcal{L}$,

$$f_\alpha(\lambda) = \left(\lambda_1 - \sum_{i=2}^L \lambda_i \right) f_\alpha(\lambda^{[1]}) + f_\alpha(\lambda'). \quad (27)$$

Proof. By Theorem 1, we have

$$f_2(\lambda') = \sum_{i=2}^L \lambda_i.$$

The condition $\lambda_1 > \sum_{i=2}^L \lambda_i$ implies that

$$g_{2, \lambda}(0) > g_{2, \lambda}(1).$$

Thus by Theorem 1, we have

$$f_2(\lambda) = g_{2, \lambda}(1) = \sum_{i=2}^L \lambda_i.$$

Then

$$f_2(\lambda) = f_2(\lambda'),$$

and by Lemma 5, we have

$$f_\alpha(\lambda) = f_\alpha(\lambda'), \text{ for all } 2 \leq \alpha \leq L. \quad (28)$$

For $2 \leq \alpha \leq L$, since $f_\alpha(\lambda^{[1]}) = 0$, the equation (27) is satisfied by virtue of (28). For $\alpha = 1$, we can check that

$$\begin{aligned} f_1(\lambda) &= \sum_{i=1}^L \lambda_i \\ &= \left(\lambda_1 - \sum_{i=2}^L \lambda_i \right) \cdot 1 + 2 \sum_{i=2}^L \lambda_i \\ &= \left(\lambda_1 - \sum_{i=2}^L \lambda_i \right) f_1(\lambda^{[1]}) + f_1(\lambda'), \end{aligned}$$

so that (27) is also satisfied. This proves the lemma. \square

Lemma 7. For any $\eta \in \{1, 2, \dots, L-1\}$,

- (i) if $\lambda_1 \leq \frac{1}{\eta} \sum_{i=2}^L \lambda_i$, then $f_\alpha(\lambda) = g_{\alpha, \lambda}(0)$, for $\alpha = 1, 2, \dots, \eta+1$;
- (ii) if $\lambda_1 \geq \frac{1}{\eta} \sum_{i=2}^L \lambda_i$, then $f_\alpha(\lambda) = f_{\alpha-1}(\lambda_2, \lambda_3, \dots, \lambda_L)$, for $\alpha = \eta+1, \eta+2, \dots, L$.

Remark 4. If $\lambda_1 = \frac{1}{\eta} \sum_{i=2}^L \lambda_i$, we have from the lemma that

$$f_{\eta+1}(\lambda) = f_\eta(\lambda_2, \lambda_3, \dots, \lambda_L) = \frac{1}{\eta+1} \sum_{i=1}^L \lambda_i.$$

In this case,

$$f_\alpha(\lambda) = \begin{cases} \frac{1}{\alpha} \sum_{i=1}^L \lambda_i, & \text{for } \alpha \leq \eta+1 \\ f_{\alpha-1}(\lambda_2, \lambda_3, \dots, \lambda_L), & \text{for } \alpha \geq \eta+1. \end{cases}$$

Proof. We first prove (i). For $\alpha \leq \eta+1$, it is easy to check that

$$\frac{1}{\alpha} \left(1 + \frac{1}{\eta} \right) \leq \frac{1}{\alpha-1}. \quad (29)$$

Thus,

$$\begin{aligned} \frac{1}{\alpha} \sum_{i=1}^L \lambda_i &= \frac{1}{\alpha} \lambda_1 + \frac{1}{\alpha} \sum_{i=2}^L \lambda_i \\ &\leq \frac{1}{\alpha} \left(\frac{1}{\eta} \sum_{i=2}^L \lambda_i \right) + \frac{1}{\alpha} \sum_{i=2}^L \lambda_i \end{aligned} \quad (30)$$

$$\begin{aligned} &= \frac{1}{\alpha} \left(1 + \frac{1}{\eta} \right) \sum_{i=2}^L \lambda_i \\ &\leq \frac{1}{\alpha-1} \sum_{i=2}^L \lambda_i, \end{aligned} \quad (31)$$

where (30) follows from the assumption that $\lambda_1 \leq \frac{1}{\eta} \sum_{i=2}^L \lambda_i$ and (31) follows from (29). Then by the discussion following Lemma 3, we have

$$f_\alpha(\lambda) = \frac{1}{\alpha} \sum_{i=1}^L \lambda_i = g_{\alpha, \lambda}(0).$$

Next, we prove (ii). For $\alpha \geq \eta+1$, it is easy to check that

$$\frac{1}{\alpha} \left(1 + \frac{1}{\eta} \right) \geq \frac{1}{\alpha-1}. \quad (32)$$

Similar to the derivation of (31), with the assumption that $\lambda_1 \geq \frac{1}{\eta} \sum_{i=2}^L \lambda_i$, (32) implies that

$$\frac{1}{\alpha} \sum_{i=1}^L \lambda_i \geq \frac{1}{\alpha-1} \sum_{i=2}^L \lambda_i. \quad (33)$$

Thus, we have

$$\begin{aligned} f_\alpha(\boldsymbol{\lambda}) &= \min_{\beta \in \{0,1,\dots,\alpha-1\}} \left\{ \frac{1}{\alpha-\beta} \sum_{i=\beta+1}^L \lambda_i \right\} \\ &= \min_{\beta \in \{1,2,\dots,\alpha-1\}} \left\{ \frac{1}{\alpha-\beta} \sum_{i=\beta+1}^L \lambda_i \right\} \\ &= \min_{\beta \in \{0,1,\dots,\alpha-2\}} \left\{ \frac{1}{(\alpha-1)-\beta} \sum_{i=\beta+2}^L \lambda_i \right\} \\ &= f_{\alpha-1}(\lambda_2, \lambda_3, \dots, \lambda_L), \end{aligned} \quad (34)$$

where (34) follows from (33). This proves the lemma. \square

The following lemma implies that $f_\alpha(\boldsymbol{\lambda})$ is a concave function of $\boldsymbol{\lambda} \in \mathbb{R}_+^L$ for all $\alpha \in \mathcal{L}$. Note that the vectors in this lemma are not necessarily ordered.

Lemma 8. For any $\alpha \in \mathcal{L}$,

$$f_\alpha(\mu_1 \boldsymbol{\lambda}_1 + \mu_2 \boldsymbol{\lambda}_2) \geq \mu_1 f_\alpha(\boldsymbol{\lambda}_1) + \mu_2 f_\alpha(\boldsymbol{\lambda}_2)$$

for any $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \in \mathbb{R}_+^L$ and $\mu_1, \mu_2 \geq 0$.

Proof. Let $\boldsymbol{\lambda}_1 = (\lambda_{1,1}, \lambda_{1,2}, \dots, \lambda_{1,L})$ and $\boldsymbol{\lambda}_2 = (\lambda_{2,1}, \lambda_{2,2}, \dots, \lambda_{2,L})$. Let $\pi_1(\cdot), \pi_2(\cdot)$ be two permutations of $\{1, 2, \dots, L\}$ such that

$$\lambda_{1,\pi_1(1)} \geq \lambda_{1,\pi_1(2)} \geq \dots \geq \lambda_{1,\pi_1(L)}$$

and

$$\lambda_{2,\pi_2(1)} \geq \lambda_{2,\pi_2(2)} \geq \dots \geq \lambda_{2,\pi_2(L)}.$$

Denote the ordered vectors by $\pi_1(\boldsymbol{\lambda}_1)$ and $\pi_2(\boldsymbol{\lambda}_2)$, respectively. For any $\beta = 0, 1, \dots, \alpha-1$, it is easy to see that

$$\frac{1}{\alpha-\beta} \sum_{i=\beta+1}^L \lambda_{1,i} \geq \frac{1}{\alpha-\beta} \sum_{i=\beta+1}^L \lambda_{1,\pi_1(i)}$$

and

$$\frac{1}{\alpha-\beta} \sum_{i=\beta+1}^L \lambda_{2,i} \geq \frac{1}{\alpha-\beta} \sum_{i=\beta+1}^L \lambda_{2,\pi_2(i)}.$$

Thus, we have

$$\begin{aligned} &\frac{1}{\alpha-\beta} \sum_{i=\beta+1}^L (\mu_1 \lambda_{1,i} + \mu_2 \lambda_{2,i}) \\ &\geq \frac{1}{\alpha-\beta} \sum_{i=\beta+1}^L (\mu_1 \lambda_{1,\pi_1(i)} + \mu_2 \lambda_{2,\pi_2(i)}). \end{aligned}$$

This implies that

$$f_\alpha(\mu_1 \boldsymbol{\lambda}_1 + \mu_2 \boldsymbol{\lambda}_2) \geq f_\alpha(\mu_1 \pi_1(\boldsymbol{\lambda}_1) + \mu_2 \pi_2(\boldsymbol{\lambda}_2)).$$

For any $\alpha \in \mathcal{L}$, it is easy to check that

$$f_\alpha(\pi_1(\boldsymbol{\lambda}_1)) = f_\alpha(\boldsymbol{\lambda}_1)$$

and

$$f_\alpha(\pi_2(\boldsymbol{\lambda}_2)) = f_\alpha(\boldsymbol{\lambda}_2).$$

Therefore, if the lemma holds for any ordered vectors $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$, then the lemma holds for any vectors $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ (not necessarily ordered), because

$$\begin{aligned} f_\alpha(\mu_1 \boldsymbol{\lambda}_1 + \mu_2 \boldsymbol{\lambda}_2) &\geq f_\alpha(\mu_1 \pi_1(\boldsymbol{\lambda}_1) + \mu_2 \pi_2(\boldsymbol{\lambda}_2)) \\ &\geq \mu_1 f_\alpha(\pi_1(\boldsymbol{\lambda}_1)) + \mu_2 f_\alpha(\pi_2(\boldsymbol{\lambda}_2)) \\ &= \mu_1 f_\alpha(\boldsymbol{\lambda}_1) + \mu_2 f_\alpha(\boldsymbol{\lambda}_2). \end{aligned}$$

Thus without loss of generality, we assume that $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ are ordered. Then for any $\beta = 0, 1, \dots, \alpha-1$, we have from Theorem 1 that

$$\frac{1}{\alpha-\beta} \sum_{i=\beta+1}^L \lambda_{1,i} \geq f_\alpha(\boldsymbol{\lambda}_1)$$

and

$$\frac{1}{\alpha-\beta} \sum_{i=\beta+1}^L \lambda_{2,i} \geq f_\alpha(\boldsymbol{\lambda}_2),$$

which implies

$$\frac{1}{\alpha-\beta} \sum_{i=\beta+1}^L (\mu_1 \lambda_{1,i} + \mu_2 \lambda_{2,i}) \geq \mu_1 f_\alpha(\boldsymbol{\lambda}_1) + \mu_2 f_\alpha(\boldsymbol{\lambda}_2).$$

By taking the minimum over all $\beta = 0, 1, \dots, \alpha-1$, we obtain

$$\begin{aligned} &\min_{\beta \in \{0,1,\dots,\alpha-1\}} \left\{ \frac{1}{\alpha-\beta} \sum_{i=\beta+1}^L (\mu_1 \lambda_{1,i} + \mu_2 \lambda_{2,i}) \right\} \\ &\geq \mu_1 f_\alpha(\boldsymbol{\lambda}_1) + \mu_2 f_\alpha(\boldsymbol{\lambda}_2), \end{aligned}$$

which by Theorem 1 is equivalent to

$$f_\alpha(\mu_1 \boldsymbol{\lambda}_1 + \mu_2 \boldsymbol{\lambda}_2) \geq \mu_1 f_\alpha(\boldsymbol{\lambda}_1) + \mu_2 f_\alpha(\boldsymbol{\lambda}_2).$$

This proves the lemma. \square

IV. THE MINIMUM SUFFICIENT SET OF INEQUALITIES

Even though the superposition region \mathcal{R}_{sup} (cf. (7) and (8)) can be explicitly characterized by solving $f_\alpha(\boldsymbol{\lambda})$ in Theorem 1, an uncountable number of inequalities are involved. For a fixed L , among all these inequalities, only a finite number of them are needed because \mathcal{R}_{sup} is a polytope. In this section, we provide a method to determine this minimum sufficient set of inequalities.

For any $\boldsymbol{\lambda} \in \mathbb{R}_+^L$, let $\pi(\cdot)$ be a permutation of $\{1, 2, \dots, L\}$ such that

$$\lambda_{\pi(1)} \geq \lambda_{\pi(2)} \geq \dots \geq \lambda_{\pi(L)}. \quad (35)$$

Recall that we consider only $\boldsymbol{\lambda}$'s whose minimum nonzero element is equal to 1. Let $\zeta \in \mathcal{L}$ be such that

$$\lambda_{\pi(\zeta)} = 1 \quad (36)$$

and for $j = \zeta + 1, \zeta + 2, \dots, L$,

$$\lambda_{\pi(j)} = 0. \quad (37)$$

Toward listing all the inequalities defining \mathcal{R}_{sup} , we consider a certain finite subset of \mathbb{R}_+^L defined as follows. Let \mathcal{G}_L be the collection of all $\lambda \in \mathbb{R}_+^L$ such that for $j = \zeta - 1, \zeta - 2, \dots, 1$,

$$\lambda_{\pi(j)} \in \left\{ \sum_{i=j+1}^L \lambda_{\pi(i)}, \frac{1}{2} \sum_{i=j+1}^L \lambda_{\pi(i)}, \dots, \frac{1}{\theta_{j+1} + 1} \sum_{i=j+1}^L \lambda_{\pi(i)} \right\} \quad (38)$$

where $\theta_\zeta = 0$ and for $j = \zeta - 2, \dots, 1$, θ_{j+1} is the integer such that

$$\lambda_{\pi(j+1)} = \frac{1}{\theta_{j+1}} \sum_{i=j+2}^L \lambda_{\pi(i)}. \quad (39)$$

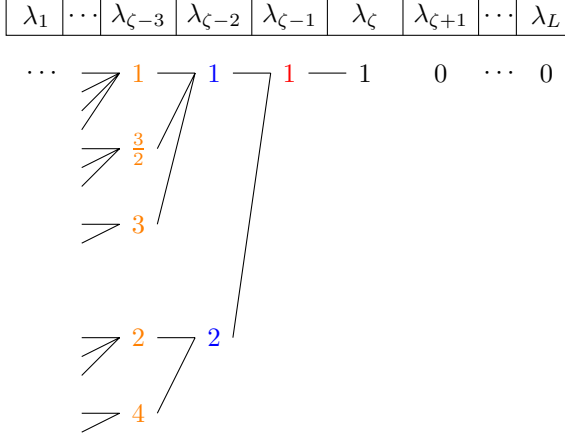


Fig. 2: Recursive generation of vectors in \mathcal{G}_L^0

Here, (36)-(39) not only defines \mathcal{G}_L but in fact provides a method to exhaust all $\lambda \in \mathcal{G}_L$. For $\zeta = 1$, the only possible λ are $\lambda^{[1]} = (1, 0, 0, \dots)$ and its permutations. For $\zeta \geq 2$, starting with $\lambda_{\pi(\zeta)} = 1$, the values of $\lambda_{\pi(\zeta-1)}, \lambda_{\pi(\zeta-2)}, \dots, \lambda_{\pi(1)}$ can be chosen recursively according to (38). It is easy to check that

$$\theta_j \in \{1, 2, \dots, \theta_{j+1} + 1\} \quad (40)$$

and

$$1 \leq \theta_j \leq \zeta - j \quad (41)$$

for $1 \leq j \leq \zeta - 1$. Furthermore, for the last element of the set in (38) which is the smallest in the set, we have

$$\begin{aligned} & \frac{1}{\theta_{j+1} + 1} \sum_{i=j+1}^L \lambda_{\pi(i)} \\ &= \frac{1}{\theta_{j+1} + 1} \left(\frac{1}{\theta_{j+1}} \sum_{i=j+2}^L \lambda_{\pi(i)} + \sum_{i=j+2}^L \lambda_{\pi(i)} \right) \\ &= \frac{1}{\theta_{j+1} + 1} \frac{\theta_{j+1} + 1}{\theta_{j+1}} \sum_{i=j+2}^L \lambda_{\pi(i)} \\ &= \frac{1}{\theta_{j+1}} \sum_{i=j+2}^L \lambda_{\pi(i)} \\ &= \lambda_{\pi(j+1)}, \end{aligned}$$

so that $\lambda_{\pi(j)} \geq \lambda_{\pi(j+1)}$ as required by (35). Also, we see from (38) that

- 1) for $\zeta \geq 2$, $\lambda_{\pi(\zeta-1)} = \lambda_{\pi(\zeta)} = 1$;
- 2) for $\zeta \geq 3$, $\lambda_{\pi(j+1)}$ is always a possible choice for $\lambda_{\pi(j)}$ for $1 \leq j \leq \zeta - 2$.

Denote the cardinality of \mathcal{G}_L by S_L . Let $\mathcal{G}_L^0 = \{\lambda \in \mathcal{G}_L : \lambda \text{ is ordered}\}$, and denote its cardinality by $|\mathcal{G}_L^0| = S_L^0$. (42)

The vectors in \mathcal{G}_L^0 are generated recursively as illustrated in Fig. 2. For the ease of notation, we let

$$\mathcal{G}_L^0 = \{\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(S_L^0)}\} \quad (43)$$

with $\lambda^{(1)} = \lambda^{[1]}$ (cf. (26)) and

$$\mathcal{G}_L = \mathcal{G}_L^0 \cup \{\lambda^{(S_L^0+1)}, \lambda^{(S_L^0+2)}, \dots, \lambda^{(S_L)}\}.$$

In other words, the set \mathcal{G}_L is the collection of all possible permutations of the vectors in \mathcal{G}_L^0 .

For $i = 1, 2, \dots, S_L$, let $\pi_i(\cdot)$ be a permutation of $\{1, 2, \dots, L\}$ such that

$$\lambda_{\pi_i(1)}^{(i)} \geq \lambda_{\pi_i(2)}^{(i)} \geq \dots \geq \lambda_{\pi_i(L)}^{(i)}. \quad (44)$$

For any $\lambda \in \mathbb{R}_+^L$, let $\mathbf{f}(\lambda) = (f_1(\lambda), f_2(\lambda), \dots, f_L(\lambda))$. The following technical lemma will be instrumental for the proof of our main theorem.

Lemma 9. Consider any ordered vector $\lambda \in \mathbb{R}_+^L$ such that $\lambda \neq \lambda^{[1]}$. Assume there exists $c_i \geq 0$, $i = 1, 2, \dots, S_L$ such that

$$(\lambda, \mathbf{f}(\lambda)) = \sum_{i=1}^{S_L} c_i \cdot (\lambda^{(i)}, \mathbf{f}(\lambda^{(i)})). \quad (45)$$

Let $\mathcal{I} = \{i \in \{1, 2, \dots, S_L\} : c_i \neq 0\}$. For any $\eta \in \{1, 2, \dots, \zeta - 1\}$,

- (i) if $\lambda_1 \leq \frac{1}{\eta} \sum_{j=2}^L \lambda_j$, then $\lambda_{\pi_i(1)}^{(i)} \leq \frac{1}{\eta} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)}$ for all $i \in \mathcal{I}$;
- (ii) if $\lambda_1 \geq \frac{1}{\eta} \sum_{j=2}^L \lambda_j$, then $\lambda_{\pi_i(1)}^{(i)} \geq \frac{1}{\eta} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)}$ for all $i \in \mathcal{I}$.

Remark 5. In the above, since λ is ordered, we have

$$\lambda_1 \geq \frac{1}{\zeta - 1} \sum_{j=2}^L \lambda_j. \quad (46)$$

Therefore, when $\eta = \zeta - 1$, the condition in (i) can only be satisfied with an equality, i.e., $\lambda_1 = \frac{1}{\zeta - 1} \sum_{j=2}^L \lambda_j$.

Proof. See Appendix A. □

The assumption that $\lambda^{(i)} \in \mathcal{G}_L$ for $1 \leq i \leq S_L$ is not invoked in the proof of Lemma 9. By taking this assumption into account, Lemma 9 can be further strengthened with the following setup. For any ordered vector $\lambda \in \mathbb{R}_+^L$ not equal to $\lambda^{[1]}$, by the constraint in (46), there exists a unique $\eta \in \{1, 2, \dots, \zeta - 1\}$ such that

$$\frac{1}{\eta} \sum_{j=2}^L \lambda_j \leq \lambda_1 < \frac{1}{\eta - 1} \sum_{j=2}^L \lambda_j. \quad (47)$$

In the sequel, we adopt the convention that

$$\frac{1}{0} \cdot c = \begin{cases} \infty, & \text{if } c \neq 0 \\ 1, & \text{if } c = 0. \end{cases}$$

Then the upper bound in (47) is ∞ when $\eta = 1$.

Lemma 10. For any ordered vector $\lambda \in \mathbb{R}_+^L$ such that $\lambda \neq \lambda^{[1]}$, assume there exists $c_i \geq 0$, $i = 1, 2, \dots, S_L$ such that

$$(\lambda, f(\lambda)) = \sum_{i=1}^{S_L} c_i \cdot (\lambda^{(i)}, f(\lambda^{(i)})). \quad (48)$$

Then for all $i \in \mathcal{I}$,

$$\lambda_{\pi_i(1)}^{(i)} \in \left\{ \frac{1}{\eta} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)}, \frac{1}{\eta-1} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)} \right\}, \quad (49)$$

where η depends on λ and is defined in (47). In particular, if the lower bound in (47) is tight, then $\lambda_{\pi_i(1)}^{(i)} = \frac{1}{\eta} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)}$ for all $i \in \mathcal{I}$.

Proof. The lemma can be easily obtained from Lemma 9. See details in Appendix B. \square

Remark 6. For all $1 \leq i \leq S_L$, $\lambda_{\pi_i(1)}^{(i)}$ can in general take one of the $\theta_2 + 1$ values prescribed in (38). However, under the constraint (48), the above lemma asserts that for all $i \in \mathcal{I}$, $\lambda_{\pi_i(1)}^{(i)}$ can only take one of the two values prescribed in (49).

Let \mathcal{R}^* be the collection of nonnegative rate tuples \mathbf{R} such that

$$\sum_{l=1}^L \lambda_l R_l \geq \sum_{\alpha=1}^L f_\alpha(\lambda) H(X_\alpha), \text{ for all } \lambda \in \mathcal{G}_L. \quad (50)$$

The next theorem shows that \mathcal{R}^* provides an equivalent characterization of \mathcal{R}_{sup} . Note that \mathcal{R}^* is the intersection of only a finite set of halfspaces, because the cardinality of \mathcal{G}_L is finite in view of its definition in (36)-(38). Thus, \mathcal{R}^* is more explicit than \mathcal{R}_h . For $L = 1, 2, \dots, 5$, all the rate constraints of \mathcal{R}^* with ordered coefficient vectors are listed in Appendix I.

Theorem 2. $\mathcal{R}_{\text{sup}} = \mathcal{R}^*$.

Proof. We prove the theorem by showing that $\mathcal{R}_h = \mathcal{R}^*$. Since $\mathcal{G}_L \subseteq \mathbb{R}_+^L$, we have $\mathcal{R}_h \subseteq \mathcal{R}^*$. To show $\mathcal{R}^* \subseteq \mathcal{R}_h$, we consider the following. Define three sets of $(2L)$ -vectors by

$$\mathcal{F}_L^1 = \{(\lambda, f(\lambda)) : \lambda \in \mathbb{R}_+^L\},$$

$$\mathcal{F}_L^2 = \{(\lambda, f(\lambda)) : \lambda \in \mathcal{G}_L\},$$

and

$$\mathcal{F}_L^3 = \{(\lambda, f(\lambda)) : \lambda \in \mathcal{G}_L^0\}.$$

Note that none of \mathcal{F}_L^2 and \mathcal{F}_L^3 is a vector space since they are not closed under vector addition. We prove $\mathcal{R}^* \subseteq \mathcal{R}_h$ by proving the claim that any $(\lambda, f(\lambda)) \in \mathcal{F}_L^1$ is a conic combination of the vectors in \mathcal{F}_L^2 .

Without loss of generality, we consider only λ such that $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L$, and show that $(\lambda, f(\lambda))$ for any such λ is a conic combination of the vectors in \mathcal{F}_L^2 . We prove the claim by induction on L for $L \geq 1$. Since we consider only

λ 's whose minimum nonzero element is equal to 1, it is easy to see that $\mathcal{F}_L^1 = \mathcal{F}_L^3 = \{(1, 1)\}$ and thus the claim is true for $L = 1$.

Assume the claim is true for $L = N$. We will show that the claim is true for $L = N + 1$. This can readily be verified for $\lambda \in \mathbb{R}_+^{N+1}$ such that $\zeta = 1$. Thus, we consider only $\lambda \in \mathbb{R}_+^{N+1}$ such that $\zeta \geq 2$. For any ordered vector $\lambda_N = (\lambda_2, \lambda_3, \dots, \lambda_{N+1}) \in \mathbb{R}_+^N$, let $\lambda_{N+1} = (\lambda_1, \lambda_2, \dots, \lambda_{N+1})$ where $\lambda_1 \geq \lambda_2$. By the induction hypothesis, there exist $c_i \geq 0$, $i = 1, 2, \dots, S_N^0$ such that

$$(\lambda_N, f(\lambda_N)) = \sum_{i=1}^{S_N^0} c_i (\lambda_N^{(i)}, f(\lambda_N^{(i)})), \quad (51)$$

where $\lambda_N^{(i)}$, $i = 1, 2, \dots, S_N^0$ are distinct elements of \mathcal{G}_N^0 . Let $\lambda_N = (\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_{N+1}^{(i)})$. Recall that $\mathcal{I} = \{i \in \{1, 2, \dots, S_N\} : c_i \neq 0\}$ in Lemma 9. For simplicity, let $c_i = 0$ for all $i \in \{S_N^0 + 1, S_N^0 + 2, \dots, S_N\}$. For any $i \in \mathcal{I}$, by Lemma 10, we have

$$\lambda_2^{(i)} \in \left\{ \frac{1}{\eta'} \sum_{j=3}^{N+1} \lambda_j^{(i)}, \frac{1}{\eta' - 1} \sum_{j=3}^{N+1} \lambda_j^{(i)} \right\}, \quad (52)$$

where $\eta' \in \{1, 2, \dots, N - 1\}$ is unique and determined by

$$\frac{1}{\eta'} \sum_{j=3}^{N+1} \lambda_j \leq \lambda_2 < \frac{1}{\eta' - 1} \sum_{j=3}^{N+1} \lambda_j. \quad (53)$$

Since the second inequality in (53) is equivalent to $\lambda_2 < \frac{1}{\eta'} \sum_{j=2}^{N+1} \lambda_j$, we consider the following three cases for λ_1 :

- a) $\frac{1}{\eta'} \sum_{j=2}^{N+1} \lambda_j < \lambda_1 \leq \sum_{j=2}^{N+1} \lambda_j$;
- b) $\lambda_2 \leq \lambda_1 \leq \frac{1}{\eta'} \sum_{j=2}^{N+1} \lambda_j$;
- c) $\lambda_1 > \sum_{j=2}^{N+1} \lambda_j$.

Case a): If $\frac{1}{\eta'} \sum_{j=2}^{N+1} \lambda_j < \lambda_1 \leq \sum_{j=2}^{N+1} \lambda_j$, there exists a unique $\varphi \in \{1, 2, \dots, \eta' - 1\}$ such that

$$\frac{1}{\varphi + 1} \sum_{j=2}^{N+1} \lambda_j < \lambda_1 \leq \frac{1}{\varphi} \sum_{j=2}^{N+1} \lambda_j. \quad (54)$$

Then by Lemma 7, we have

$$f_\alpha(\lambda_{N+1}) = \begin{cases} \frac{1}{\alpha} \sum_{j=1}^{N+1} \lambda_j, & \text{for } 1 \leq \alpha \leq \varphi + 1 \\ f_{\alpha-1}(\lambda_N), & \text{for } \varphi + 2 \leq \alpha \leq N + 1. \end{cases} \quad (55)$$

For all $i \in \mathcal{I}$, let

$$\lambda_1^{(1,i)} = \frac{1}{\varphi + 1} \sum_{j=2}^{N+1} \lambda_j^{(i)} \quad (56)$$

and

$$\lambda_1^{(2,i)} = \frac{1}{\varphi} \sum_{j=2}^{N+1} \lambda_j^{(i)}. \quad (57)$$

For $j \in \{2, 3, \dots, N + 1\}$, for notational simplicity, let

$$\lambda_j^{(1,i)} = \lambda_j^{(2,i)} = \lambda_j^{(i)}. \quad (58)$$

Let

$$\lambda_{N+1}^{(1,i)} = (\lambda_1^{(1,i)}, \lambda_2^{(1,i)}, \dots, \lambda_{N+1}^{(1,i)})$$

and

$$\lambda_{N+1}^{(2,i)} = (\lambda_1^{(2,i)}, \lambda_2^{(2,i)}, \dots, \lambda_{N+1}^{(2,i)}).$$

From (52), (38), and the range of φ , we can check that $\lambda_{N+1}^{(1,i)}, \lambda_{N+1}^{(2,i)} \in \mathcal{G}_{N+1}^0$. By Remark 4 following Lemma 7, we have

$$f_\alpha(\lambda_{N+1}^{(1,i)}) = \begin{cases} \frac{1}{\alpha} \sum_{j=1}^{N+1} \lambda_j^{(1,i)}, & \text{for } 1 \leq \alpha \leq \varphi + 2 \\ f_{\alpha-1}(\lambda_N^{(i)}), & \text{for } \varphi + 2 \leq \alpha \leq N + 1, \end{cases} \quad (59)$$

and

$$f_\alpha(\lambda_{N+1}^{(2,i)}) = \begin{cases} \frac{1}{\alpha} \sum_{j=1}^{N+1} \lambda_j^{(2,i)}, & \text{for } 1 \leq \alpha \leq \varphi + 1 \\ f_{\alpha-1}(\lambda_N^{(i)}), & \text{for } \varphi + 1 \leq \alpha \leq N + 1. \end{cases} \quad (60)$$

Consider the conic combination of $\lambda_1^{(1,i)}$ for $i \in \mathcal{I}$,

$$\sum_{i \in \mathcal{I}} c_i \lambda_1^{(1,i)} = \frac{1}{\varphi + 1} \sum_{i \in \mathcal{I}} c_i \sum_{j=2}^{N+1} \lambda_j^{(i)} \quad (61)$$

$$= \frac{1}{\varphi + 1} \sum_{j=2}^{N+1} \left(\sum_{i \in \mathcal{I}} c_i \lambda_j^{(i)} \right)$$

$$= \frac{1}{\varphi + 1} \sum_{j=2}^{N+1} \lambda_j \quad (62)$$

$$< \lambda_1, \quad (63)$$

where (61) follows from (56), (62) follows from (51), and (63) follows from (54). Similarly, from (57), (51), and (54), we have

$$\sum_{i \in \mathcal{I}} c_i \lambda_1^{(2,i)} \geq \lambda_1.$$

Let $u_1 = \sum_{i \in \mathcal{I}} c_i \lambda_1^{(1,i)}$ and $u_2 = \sum_{i \in \mathcal{I}} c_i \lambda_1^{(2,i)}$. Then we have $u_1 < \lambda_1 \leq u_2$. For all $i \in \mathcal{I}$, let

$$\lambda_1^{(i)} = \frac{u_2 - \lambda_1}{u_2 - u_1} \lambda_1^{(1,i)} + \frac{\lambda_1 - u_1}{u_2 - u_1} \lambda_1^{(2,i)}. \quad (64)$$

It is easy to check that

$$\sum_{i \in \mathcal{I}} c_i \lambda_1^{(i)} = \lambda_1. \quad (65)$$

Let $c_i^{(1)} = \frac{u_2 - \lambda_1}{u_2 - u_1} \cdot c_i$ and $c_i^{(2)} = \frac{\lambda_1 - u_1}{u_2 - u_1} \cdot c_i$. It is readily seen that $c_i^{(1)}$ and $c_i^{(2)}$ are nonnegative, and we can check that

$$\begin{cases} c_i^{(1)} + c_i^{(2)} = c_i \\ c_i^{(1)} \lambda_1^{(1,i)} + c_i^{(2)} \lambda_1^{(2,i)} = c_i \lambda_1^{(i)}. \end{cases} \quad (66)$$

Then we have from (55), (65), and (66) that

$$\lambda_{N+1} = \sum_{i \in \mathcal{I}} \left(c_i^{(1)} \lambda_{N+1}^{(1,i)} + c_i^{(2)} \lambda_{N+1}^{(2,i)} \right). \quad (67)$$

Following (55), we have for $1 \leq \alpha \leq \varphi + 1$ that

$$\begin{aligned} f_\alpha(\lambda_{N+1}) &= \frac{1}{\alpha} \sum_{j=1}^{N+1} \lambda_j \\ &= \frac{1}{\alpha} \sum_{j=1}^{N+1} \sum_{i \in \mathcal{I}} c_i \lambda_j^{(i)} \end{aligned} \quad (68)$$

$$= \frac{1}{\alpha} \sum_{j=1}^{N+1} \left[\sum_{i \in \mathcal{I}} \left(c_i^{(1)} \lambda_j^{(1,i)} + c_i^{(2)} \lambda_j^{(2,i)} \right) \right] \quad (69)$$

$$\begin{aligned} &= \sum_{i \in \mathcal{I}} \left[c_i^{(1)} \left(\frac{1}{\alpha} \sum_{j=1}^{N+1} \lambda_j^{(1,i)} \right) + c_i^{(2)} \left(\frac{1}{\alpha} \sum_{j=1}^{N+1} \lambda_j^{(2,i)} \right) \right] \\ &= \sum_{i \in \mathcal{I}} \left[c_i^{(1)} f_\alpha(\lambda_{N+1}^{(1,i)}) + c_i^{(2)} f_\alpha(\lambda_{N+1}^{(2,i)}) \right] \end{aligned} \quad (70)$$

where (68) follows from (51), (69) follows from (58) and (66), and (70) follows from (59) and (60). Similarly, for $\varphi + 2 \leq \alpha \leq N + 1$, following (55), we have

$$\begin{aligned} f_\alpha(\lambda_{N+1}) &= f_{\alpha-1}(\lambda_N) \\ &= \sum_{i \in \mathcal{I}} c_i f_{\alpha-1}(\lambda_N^{(i)}) \end{aligned} \quad (71)$$

$$= \sum_{i \in \mathcal{I}} \left(c_i^{(1)} + c_i^{(2)} \right) f_{\alpha-1}(\lambda_N^{(i)}) \quad (72)$$

$$= \sum_{i \in \mathcal{I}} \left[c_i^{(1)} f_\alpha(\lambda_{N+1}^{(1,i)}) + c_i^{(2)} f_\alpha(\lambda_{N+1}^{(2,i)}) \right], \quad (73)$$

where (71) follows from (51), (72) follows from (66), and (73) follows from (59) and (60). In other words, (70) or (73) holds for all $1 \leq \alpha \leq N + 1$. Summarizing the above, we have

$$\begin{aligned} &(\lambda_{N+1}, f(\lambda_{N+1})) \\ &= \sum_{i \in \mathcal{I}} \left[c_i^{(1)} \left(\lambda_{N+1}^{(1,i)}, f_\alpha(\lambda_{N+1}^{(1,i)}) \right) + c_i^{(2)} \left(\lambda_{N+1}^{(2,i)}, f_\alpha(\lambda_{N+1}^{(2,i)}) \right) \right], \end{aligned} \quad (74)$$

and thus $(\lambda_{N+1}, f(\lambda_{N+1}))$ is a conic combination of vectors in \mathcal{F}_{N+1}^3 .

Case b): If $\lambda_2 \leq \lambda_1 \leq \frac{1}{\eta'} \sum_{j=2}^{N+1} \lambda_j$, since the condition $\frac{1}{\eta'} \sum_{j=3}^{N+1} \lambda_j \leq \lambda_2$ in (53) is equivalent to

$$\lambda_2 \geq \frac{1}{\eta' + 1} \sum_{j=2}^{N+1} \lambda_j,$$

we have

$$\frac{1}{\eta' + 1} \sum_{j=2}^{N+1} \lambda_j \leq \lambda_1 \leq \frac{1}{\eta'} \sum_{j=2}^{N+1} \lambda_j.$$

By Lemma 7, we obtain

$$f_\alpha(\lambda_{N+1}) = \begin{cases} \frac{1}{\alpha} \sum_{j=1}^{N+1} \lambda_j, & \text{for } 1 \leq \alpha \leq \eta' + 1 \\ f_{\alpha-1}(\lambda_N), & \text{for } \eta' + 2 \leq \alpha \leq N + 1. \end{cases}$$

In light of (52), let $\mathcal{I}_1 = \left\{ i \in \mathcal{I} : \lambda_2^{(i)} = \frac{1}{\eta'-1} \sum_{j=3}^{N+1} \lambda_j^{(i)} \right\}$ and $\mathcal{I}_2 = \left\{ i \in \mathcal{I} : \lambda_2^{(i)} = \frac{1}{\eta'} \sum_{j=3}^{N+1} \lambda_j^{(i)} \right\}$, where $\mathcal{I}_1 \cup \mathcal{I}_2 = \mathcal{I}$. For $i \in \mathcal{I}_1$, let

$$\lambda_1^{(1,i)} = \lambda_1^{(2,i)} = \frac{1}{\eta'} \sum_{j=2}^{N+1} \lambda_j^{(i)}. \quad (75)$$

For $i \in \mathcal{I}_2$, let

$$\lambda_1^{(1,i)} = \frac{1}{\eta' + 1} \sum_{j=2}^{N+1} \lambda_j^{(i)} \quad (76)$$

and

$$\lambda_1^{(2,i)} = \frac{1}{\eta'} \sum_{j=2}^{N+1} \lambda_j^{(i)}.$$

Again, from (52) and (38), we can check that $\lambda_{N+1}^{(1,i)}, \lambda_{N+1}^{(2,i)} \in \mathcal{G}_{N+1}^0$ for all $i \in \mathcal{I}$. By Remark 4 following Lemma 7, we have for $i \in \mathcal{I}_1$ that

$$\begin{aligned} f_\alpha(\lambda_{N+1}^{(1,i)}) &= f_\alpha(\lambda_{N+1}^{(2,i)}) \\ &= \begin{cases} \frac{1}{\alpha} \sum_{j=1}^{N+1} \lambda_j^{(1,i)}, & \text{for } 1 \leq \alpha \leq \eta' + 1 \\ f_{\alpha-1}(\lambda_N^{(i)}), & \text{for } \eta' + 1 \leq \alpha \leq N + 1, \end{cases} \end{aligned}$$

and for $i \in \mathcal{I}_2$,

$$f_\alpha(\lambda_{N+1}^{(1,i)}) = \begin{cases} \frac{1}{\alpha} \sum_{j=1}^{N+1} \lambda_j^{(2,i)}, & \text{for } 1 \leq \alpha \leq \eta' + 2 \\ f_{\alpha-1}(\lambda_N^{(i)}), & \text{for } \eta' + 2 \leq \alpha \leq N + 1, \end{cases}$$

and

$$f_\alpha(\lambda_{N+1}^{(2,i)}) = \begin{cases} \frac{1}{\alpha} \sum_{j=1}^{N+1} \lambda_j^{(2,i)}, & \text{for } 1 \leq \alpha \leq \eta' + 1 \\ f_{\alpha-1}(\lambda_N^{(i)}), & \text{for } \eta' + 1 \leq \alpha \leq N + 1. \end{cases}$$

Following from (75) and (76), we have

$$\begin{aligned} \sum_{i \in \mathcal{I}} c_i \lambda_1^{(1,i)} &= \sum_{i \in \mathcal{I}_1} c_i \frac{1}{\eta'} \sum_{j=2}^{N+1} \lambda_j^{(i)} + \sum_{i \in \mathcal{I}_2} c_i \frac{1}{\eta' + 1} \sum_{j=2}^{N+1} \lambda_j^{(i)} \\ &= \sum_{i \in \mathcal{I}_1} c_i \frac{1}{\eta'} \left(1 + \frac{1}{\eta' - 1} \right) \sum_{j=3}^{N+1} \lambda_j^{(i)} \\ &\quad + \sum_{i \in \mathcal{I}_2} c_i \frac{1}{\eta' + 1} \left(1 + \frac{1}{\eta'} \right) \sum_{j=3}^{N+1} \lambda_j^{(i)} \quad (77) \end{aligned}$$

$$\begin{aligned} &= \sum_{i \in \mathcal{I}_1} c_i \frac{1}{\eta' - 1} \sum_{j=3}^{N+1} \lambda_j^{(i)} + \sum_{i \in \mathcal{I}_2} c_i \frac{1}{\eta'} \sum_{j=3}^{N+1} \lambda_j^{(i)} \\ &= \sum_{i \in \mathcal{I}_1} c_i \lambda_2^{(i)} + \sum_{i \in \mathcal{I}_2} c_i \lambda_2^{(i)} \quad (78) \\ &= \lambda_2 \\ &\leq \lambda_1, \end{aligned}$$

where (77) and (78) follow from the definition of \mathcal{I}_1 and \mathcal{I}_2 . Similar to (61)-(63), we have

$$\sum_{i \in \mathcal{I}} c_i \lambda_1^{(2,i)} \geq \lambda_1.$$

For $i \in \mathcal{I}$, similar to (64)-(66), let $c_i^{(1)} = \frac{u_2 - \lambda_1}{u_2 - u_1} \cdot c_i$, $c_i^{(2)} = \frac{\lambda_1 - u_1}{u_2 - u_1} \cdot c_i$, and

$$\lambda_1^{(i)} = \frac{u_2 - \lambda_1}{u_2 - u_1} \lambda_1^{(1,i)} + \frac{\lambda_1 - u_1}{u_2 - u_1} \lambda_1^{(2,i)}.$$

We can check that

$$\sum_{i \in \mathcal{I}} c_i \lambda_1^{(i)} = \lambda_1$$

and for all $i \in \mathcal{I}$,

$$\begin{cases} c_i^{(1)} + c_i^{(2)} = c_i \\ c_i^{(1)} \lambda_1^{(1,i)} + c_i^{(2)} \lambda_1^{(2,i)} = c_i \lambda_1^{(i)}. \end{cases}$$

Then similar to (67)-(74), we have

$$\begin{aligned} &(\lambda_{N+1}, \mathbf{f}(\lambda_{N+1})) \\ &= \sum_{i \in \mathcal{I}} \left[c_i^{(1)} \left(\lambda_{N+1}^{(1,i)}, f_\alpha(\lambda_{N+1}^{(1,i)}) \right) + c_i^{(2)} \left(\lambda_{N+1}^{(2,i)}, f_\alpha(\lambda_{N+1}^{(2,i)}) \right) \right]. \end{aligned}$$

Case c): If $\lambda_1 > \sum_{j=2}^{N+1} \lambda_j$, let $\lambda'_{N+1} = \left(\sum_{j=2}^{N+1} \lambda_j, \lambda_2, \dots, \lambda_{N+1} \right)$ and $\lambda_{N+1}^{(1)}$ be the $(N+1)$ -vector with the first component being 1 and the rest being 0. From Lemma 6, we have

$$\begin{aligned} &(\lambda_{N+1}, \mathbf{f}(\lambda_{N+1})) \\ &= \left(\lambda_1 - \sum_{j=2}^{N+1} \lambda_j \right) \left(\lambda_{N+1}^{(1)}, \mathbf{f}(\lambda_{N+1}^{(1)}) \right) + (\lambda'_{N+1}, \mathbf{f}(\lambda'_{N+1})). \end{aligned}$$

It is easy to see that

$$\left(\lambda_{N+1}^{(1)}, \mathbf{f}(\lambda_{N+1}^{(1)}) \right) \in \mathcal{F}_{N+1}^3.$$

Note that λ'_{N+1} satisfies the condition for Case a) provided that $\eta' \neq 1$. Otherwise, it satisfies the condition for Case b). Thus we see that $(\lambda'_{N+1}, \mathbf{f}(\lambda'_{N+1}))$ is always a conic combination of the vectors in \mathcal{F}_{N+1}^3 . This implies that $(\lambda_{N+1}, \mathbf{f}(\lambda_{N+1}))$ is a conic combination of the vectors in \mathcal{F}_{N+1}^3 , as is to be proved. \square

For any $\lambda \in \mathbb{R}_+^L$, let $\lambda_{L-1} = (\lambda_2, \lambda_3, \dots, \lambda_L)$ and $\mathbf{f}(\lambda_{L-1}) = (f_1(\lambda_{L-1}), f_2(\lambda_{L-1}), \dots, f_{L-1}(\lambda_{L-1}))$. The following lemma provides a method for finding a set of conic combination coefficients for $(\lambda_{L-1}, \mathbf{f}(\lambda_{L-1}))$ from the conic combination for $(\lambda, \mathbf{f}(\lambda))$.

Lemma 11. Consider any ordered vector $\lambda \in \mathbb{R}_+^L$ such that $\lambda \neq \lambda^{[1]}$. Assume there exists $c_i \geq 0$, $i = 1, 2, \dots, S_L$ such that

$$(\lambda, \mathbf{f}(\lambda)) = \sum_{i=1}^{S_L} c_i \cdot (\lambda^{(i)}, \mathbf{f}(\lambda^{(i)})), \quad (79)$$

Then we have

$$\begin{aligned} &(\lambda_{L-1}, \mathbf{f}(\lambda_{L-1})) \\ &= \sum_{i=1}^{S_L} c_i \cdot ((\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}), \mathbf{f}(\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)})). \end{aligned}$$

Proof. See Appendix C. \square

Lemma 12. For any $\lambda^{(i)} \in \mathcal{G}_L$, if $(\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}) \in \mathcal{G}_{L-1}$, then $\lambda_1^{(i)} = 0$ or $\lambda_{\pi_i(1)}^{(i)}$.

Proof. See Appendix D. \square

Lemma 13. For any $i_0 \in \{1, 2, \dots, S_L\}$, there does not exist $(c_1, c_2, \dots, c_{S_L}) \in \mathbb{R}_+^{S_L}$ such that $c_{i_0} = 0$ and

$$(\lambda^{(i_0)}, f(\lambda^{(i_0)})) = \sum_{i=1}^{S_L} c_i \cdot (\lambda^{(i)}, f(\lambda^{(i)})).$$

Proof. See Appendix E. \square

Theorem 2 gives a rate region \mathcal{R}^* that simplifies the characterization of the superposition region. The following theorem shows that there is no redundancy in the specification of \mathcal{R}^* .

Theorem 3. For the inequalities specifying \mathcal{R}^* in (50), none of them is implied by the others.

Proof. For any $i_0 \in \{1, 2, \dots, S_L\}$, consider the following linear program:

$$\begin{aligned} m_p = \min \quad & \sum_{l=1}^L \lambda_l^{(i_0)} R_l \\ \text{s.t.} \quad & \sum_{l=1}^L \lambda_l^{(i)} R_l \geq \sum_{\alpha=1}^L f_\alpha(\lambda^{(i)}) H(X_\alpha), \\ & \forall 1 \leq i \leq S_L, i \neq i_0. \end{aligned}$$

To prove Theorem 3, it suffices to show the following: for any $i_0 \in \{1, 2, \dots, S_L\}$,

$$\sum_{\alpha=1}^L f(\lambda^{(i_0)}) H(X_\alpha) > m_p.$$

By strong duality, $m_p = m_d$, where m_d is the optimal value of the dual problem

$$\begin{aligned} m_d = \max \quad & \sum_{1 \leq i \leq S_L, i \neq i_0} c_i \left(\sum_{\alpha=1}^L f_\alpha(\lambda^{(i)}) H(X_\alpha) \right) \\ \text{s.t.} \quad & \sum_{1 \leq i \leq S_L, i \neq i_0} c_i \lambda_l^{(i)} = \lambda_l^{(i_0)}, \forall l \in \mathcal{L} \\ & c_i \geq 0, 1 \leq i \leq S_L, i \neq i_0. \end{aligned} \quad (80)$$

Then it suffices to show that for any $i_0 \in \{1, 2, \dots, S_L\}$,

$$\sum_{\alpha=1}^L f(\lambda^{(i_0)}) H(X_\alpha) > m_d, \quad (81)$$

for all possible values of $H(X_\alpha)$, $\alpha \in \mathcal{L}$. For notational simplicity, let $c_{i_0} = 0$. By Lemma 8, (80) implies that

$$f_\alpha(\lambda^{(i_0)}) \geq \sum_{i=1}^{S_L} c_i f_\alpha(\lambda^{(i)}), \text{ for all } \alpha \in \mathcal{L}. \quad (82)$$

Upon multiplying by $H(X_\alpha)$ and summing over all $\alpha \in \mathcal{L}$, we obtain

$$\sum_{\alpha=1}^L f_\alpha(\lambda^{(i_0)}) H(X_\alpha) \geq \sum_{\alpha=1}^L \left(\sum_{i=1}^{S_L} c_i f_\alpha(\lambda^{(i)}) \right) H(X_\alpha),$$

which is equivalent to (81) except that the inequality above is nonstrict. Thus to prove (81), we only need to show that there exists at least one $\alpha \in \mathcal{L}$ such that

$$f_\alpha(\lambda^{(i_0)}) > \sum_{i=1}^{S_L} c_i f_\alpha(\lambda^{(i)}).$$

Assume the contrary is true, i.e. equality holds in (82) for all $\alpha \in \mathcal{L}$. Then this implies

$$(\lambda^{(i_0)}, f(\lambda^{(i_0)})) = \sum_{i=1}^{S_L} c_i \cdot (\lambda^{(i)}, f(\lambda^{(i)})), \quad (83)$$

which is a contradiction to Lemma 13. This completes the proof of the theorem. \square

V. CHECKING THE ACHIEVABILITY OF A RATE TUPLE

A. Checking Achievability

Given the superposition coding rate region \mathcal{R}_{sup} characterized by the constraints in (1) and (2), it is readily seen that a rate tuple is achievable if and only if there exist nonnegative variables r_i^α ($i, \alpha \in \mathcal{L}$) satisfying (1) and (2). Thus, we can check the achievability of a given rate tuple \mathbf{R} by determining whether there exists a set of *feasible solutions* r_i^α ($i, \alpha \in \mathcal{L}$) for the optimization problem:

$$\begin{aligned} \min \quad & 0 \\ \text{s.t.} \quad & \sum_{\alpha=1}^L r_i^\alpha = R_i, \quad \forall i \in \mathcal{L} \end{aligned} \quad (84)$$

$$\begin{aligned} & \sum_{i \in \mathcal{U}} r_i^\alpha \geq H(X_\alpha), \quad \forall \mathcal{U} \subseteq \mathcal{L}, |\mathcal{U}| = \alpha \\ & r_i^\alpha \geq 0, \quad 1 \leq i, \alpha \leq L. \end{aligned} \quad (85)$$

This can be easily achieved through the MATLAB linear programming function:

$$\mathbf{x} = \text{linprog}(\mathbf{f}, \mathbf{A}, \mathbf{b}, \mathbf{Aeq}, \mathbf{beq}).$$

We have run numerical tests of the “linprog” function on a notebook computer to determine the achievability of a given rate tuple for $L \leq 20$. For $L = 21$, the program runs out of memory, since the size of the constraint matrices “A” and “Aeq” become prohibitively large.

It is also natural to check the achievability of a given rate tuple by verifying the inequalities specifying \mathcal{R}_{sup} in Theorem 2. Even though by Theorem 3 there is no redundancy in the set of inequalities in (50) that specifies \mathcal{R}^* , by taking advantage of the symmetry of the problem, we in fact do not need to check all these inequalities. The next lemma identifies the subset of these inequalities we need to check. As we will see, the number of such inequalities is significantly smaller than the total number of inequalities specifying \mathcal{R}^* .

For any $\mathbf{R} \in \mathbb{R}_+^L$ and any permutation ω on $\{1, 2, \dots, L\}$, similar to the definition of $\omega(\lambda)$ at the beginning of Section III, let

$$\omega(\mathbf{R}) = (R_{\omega(1)}, R_{\omega(2)}, \dots, R_{\omega(L)}).$$

Due to the symmetry of the problem, for any $\lambda \in \mathcal{G}_L$, the inequality

$$\sum_{l=1}^L \lambda_l R_l \geq \sum_{\alpha=1}^L f_\alpha(\lambda) H(X_\alpha)$$

implies the inequality

$$\sum_{l=1}^L \lambda_{\omega(l)} R_{\omega(l)} \geq \sum_{\alpha=1}^L f_{\alpha}(\omega(\lambda)) H(X_{\alpha}),$$

and vice versa. Thus, \mathbf{R} is achievable if and only if $\omega(\mathbf{R})$ is achievable for all ω . As such, we only need to consider rate tuples $\mathbf{R} \in \mathbb{R}_+^L$ such that

$$R_1 \leq R_2 \leq \dots \leq R_L. \quad (86)$$

Lemma 14. *For any nonnegative rate tuple \mathbf{R} such that (86) is satisfied and any $\lambda \in \mathbb{R}_+^L$, we have*

$$\sum_{i=1}^L \lambda_i R_i \geq \sum_{i=1}^L \lambda_{\pi(i)} R_i.$$

Remark 7. The inequality in Lemma 14 is sometimes called the *rearrangement inequality* [31, Chapter 5].

Proof. See Appendix H. \square

From Lemma 2, we can see that RHS of the inequality in (50) does not change with λ replaced by $\pi(\lambda)$. Thus, in order to check the achievability of a given rate tuple (assume satisfying (86)), by Lemma 14, we only need to check those inequalities for which the coefficients are in descending order, i.e.

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L.$$

All the other inequalities are redundant for this rate tuple. Then, the number of inequalities we need to check is only S_L^0 (cf. (42)), which is bounded in the following theorem.

Theorem 4. $2^{L-1} \leq S_L^0 \leq L!$.

Remark 8. We will see from the proof that both the above inequalities become strict for $L \geq 3$.

Remark 9. On a notebook computer, it took about 8 days to list all the S_L^0 inequalities for all $L \leq 15$. For $L = 16$, the computation involved appears to be prohibitive.

Proof. We can see from Appendix I that $S_1^0 = 1$ and $S_2^0 = 2$. It is easy to check that the theorem is true for $L = 1$ and $L = 2$.

For $L \geq 2$, let $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_L)$ and $\mathcal{G}_L^* = \{\lambda \in \mathcal{G}_L^0 : \lambda_L = 0\}$. For any $\lambda \in \mathcal{G}_L^*$, it is easy to check from (38) that $(\lambda_1, \lambda_2, \dots, \lambda_{L-1}) \in \mathcal{G}_{L-1}^0$. On the other hand, for any $(\lambda_1, \lambda_2, \dots, \lambda_{L-1}) \in \mathcal{G}_{L-1}^0$, we have $(\lambda_1, \lambda_2, \dots, \lambda_{L-1}, 0) \in \mathcal{G}_L^*$. Thus, there is a one-to-one correspondence between \mathcal{G}_L^* and \mathcal{G}_{L-1}^0 , which implies that

$$|\mathcal{G}_L^*| = S_{L-1}^0. \quad (87)$$

For $k \geq 2$, let $D_k = |\mathcal{G}_k^0 \setminus \mathcal{G}_k^*|$. By (87) and the fact that $\mathcal{G}_k^* \subseteq \mathcal{G}_k^0$, we have

$$D_k = S_k^0 - S_{k-1}^0,$$

which implies that

$$S_L^0 = S_1^0 + \sum_{k=2}^L D_k.$$

Now, we only need to calculate D_k for $k \geq 2$. For any $(\lambda_{L-k+1}, \lambda_{L-k+2}, \dots, \lambda_L) \in \mathcal{G}_k^0 \setminus \mathcal{G}_k^*$, where $\lambda_L = 1$, we can see from (38) that $(\lambda_{L-k+2}, \lambda_{L-k+3}, \dots, \lambda_L) \in \mathcal{G}_{k-1}^0 \setminus \mathcal{G}_{k-1}^*$ by construction. Thus, all $(\lambda_{L-k+1}, \lambda_{L-k+2}, \dots, \lambda_L) \in \mathcal{G}_k^0 \setminus \mathcal{G}_k^*$ can be generated from $(\lambda_{L-k+2}, \lambda_{L-k+3}, \dots, \lambda_L) \in \mathcal{G}_{k-1}^0 \setminus \mathcal{G}_{k-1}^*$ with a proper choice of λ_{L-k+1} . Since $\lambda_L = 1$, we have $\zeta = L$. Recall from (39) that $\theta_L = 0$ and for $j = L-1, L-2, \dots, L-k+1$, θ_j is the integer such that

$$\lambda_j = \frac{1}{\theta_j} \sum_{i=j+1}^L \lambda_i. \quad (88)$$

According to (88), the k -vector $(\lambda_{L-k+1}, \lambda_{L-k+2}, \dots, \lambda_L) \in \mathcal{G}_k^0 \setminus \mathcal{G}_k^*$ is uniquely determined by the tuple $(\theta_{L-k+1}, \dots, \theta_{L-1}, \theta_L)$. Thus D_k is equal to the cardinality of the set

$$\Theta_k = \{(\theta_{L-k+1}, \dots, \theta_{L-1}, \theta_L) : 1 \leq \theta_j \leq \theta_{j+1} + 1 \text{ for } j = L-1, L-2, \dots, L-k+1\}.$$

By straightforward counting, we can obtain

$$D_k = |\Theta_k| = \sum_{\theta_L=0}^0 \sum_{\theta_{L-1}=1}^{\theta_L+1} \sum_{\theta_{L-2}=1}^{\theta_{L-1}+1} \dots \sum_{\theta_{L-k+1}=1}^{\theta_{L-k+2}+1} 1. \quad (89)$$

Now we bound D_k according to (89). Observe that $\theta_L = 0$ and $\theta_{L-1} = 1$ always hold. Then for $k \geq 3$, (89) can be rewritten as

$$D_k = \sum_{\theta_{L-2}=1}^2 \sum_{\theta_{L-3}=1}^{\theta_{L-2}+1} \dots \sum_{\theta_{L-k+1}=1}^{\theta_{L-k+2}+1} 1.$$

Let

$$D_k^{(1)} = \sum_{\theta_{L-2}=1}^2 \sum_{\theta_{L-3}=1}^2 \dots \sum_{\theta_{L-k+1}=1}^2 1$$

and

$$D_k^{(2)} = \sum_{\theta_{L-2}=1}^2 \sum_{\theta_{L-3}=1}^3 \dots \sum_{\theta_{L-k+1}=1}^{k-1} 1.$$

From (41), it is easy to check that

$$D_k^{(1)} \leq D_k \leq D_k^{(2)}, \quad (90)$$

and we have

$$D_k^{(1)} = 2^{k-2}$$

and

$$D_k^{(2)} = (k-1)!.$$

Thus, for $L \geq 3$, we have

$$\sum_{k=3}^L D_k^{(1)} = \sum_{k=3}^L 2^{k-2} = 2^{L-1} - 2$$

and

$$\begin{aligned} \sum_{k=3}^L D_k^{(2)} &= \sum_{k=3}^L (k-1)! \\ &\leq (L-1)! \times (L-2) \\ &\leq L! - 2. \end{aligned}$$

Then, by (V-A), (90), and the fact that $S_1^0 = 1$ and $D_2 = 1$, we have for $L \geq 3$ that

$$2^{L-1} \leq S_L^0 \leq L!.$$

This proves the theorem. \square

B. Comparison of complexity

In this section, we compare the complexity of checking the achievability of a given rate tuple through different methods described in the last section. In Figure 3, we compare the program running time of checking the LP feasibility using the MATLAB “linprog” function and that of checking the achievability of a rate tuple through the inequalities with ordered coefficients. We can see from the figure that checking

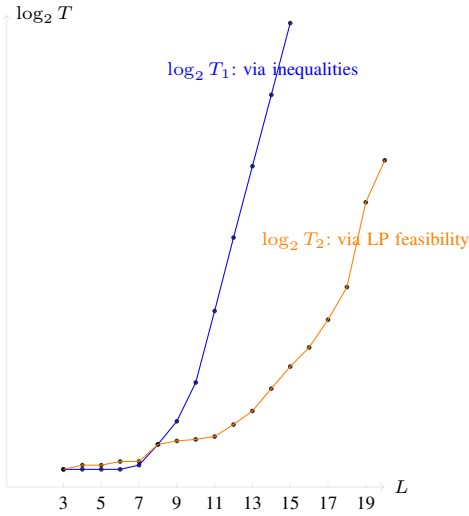


Fig. 3: logarithm of running time

the LP feasibility uses much less time than checking the achievability of a rate tuple through the inequalities with ordered coefficients. We also observe that the running time of checking the inequalities with ordered coefficients grows exponentially with L for $L \geq 10$, even though it was shown in Theorem 4 that the number of such inequalities may grow at a rate higher than exponential in L .

The time of listing the inequalities with ordered coefficients and the time of constructing the parameters of “linprog” are involved in the comparison in Figure 3. If we want to check the achievability of a large number of rate tuples, the more efficient way is to save the inequalities with ordered coefficients and the parameters of “linprog” in advance. Then we can use the “load” function in MATLAB to invoke these data when checking the achievability of rate tuples. In this case, we start counting the program running time right after the “load” function, and we call this the *pure running time*. In Figure 4, we compare the pure running time of checking the LP feasibility using “linprog” function and that of checking the achievability of a rate tuple through the inequalities with ordered coefficients. We see from the figure that checking the achievability of a rate tuple through inequalities in turn uses

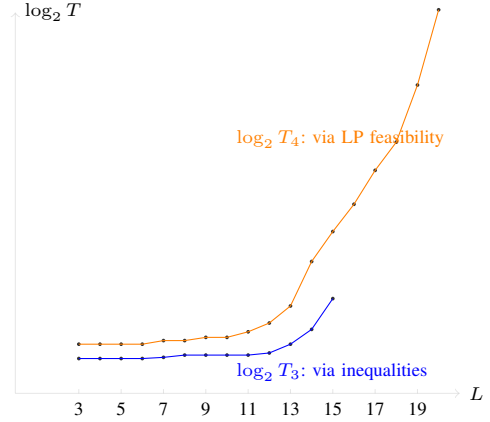


Fig. 4: logarithm of pure running time

much less time than checking the LP feasibility. This is not surprising because the time for checking the achievability through inequalities is mainly spent on listing these inequalities.

The bottleneck of checking the LP feasibility lies in that the parameters of the “linprog” function (i.e., the coefficients of the LP conditions (84)-(85)) use a mass of memory, which exceeds the capacity of the hard disk on the notebook computer for $L \geq 21$. The bottleneck of checking the achievability through the inequalities with ordered coefficients lies in that the complexity of listing these inequalities grows exponentially with L , which becomes unmanageable for $L \geq 15$.

VI. SUBSET ENTROPY INEQUALITY

In [5], the proof of the optimality of superposition coding was established through a subset entropy inequality, namely Theorem 3 therein. As we will point out, this subset entropy inequality is in fact a generalization of Han’s inequality [32]. The proof of Theorem 3 in [5], however, is painstaking. In this section, we present a weaker version of this theorem, namely Theorem 5 below, whose proof is considerably simpler. With our explicit characterization of \mathcal{R}_{sup} in Theorem 2, Theorem 5 is sufficient for proving the optimality of superposition coding.

Theorem 5 (Subset entropy inequality). *Let $L \geq 2$ and for any $\mathbf{u} \in \{0, 1\}^L$, let $H_{\mathbf{u}} = H(W_i : u_i = 1)$. For any $\boldsymbol{\lambda} \in \mathcal{G}_L$, there exists $\{c_{\alpha}(\mathbf{u})\}$, $\alpha \in \mathcal{L}$, where $\{c_{\alpha}(\mathbf{u})\}$ is an optimal α -resolution for $\boldsymbol{\lambda}$, such that*

$$\sum_{\mathbf{u} \in \Omega_L^{\alpha-1}} c_{\alpha-1}(\mathbf{u}) H_{\mathbf{u}} \geq \sum_{\mathbf{u} \in \Omega_L^{\alpha}} c_{\alpha}(\mathbf{u}) H_{\mathbf{u}} \quad (91)$$

for all $\alpha = 2, 3, \dots, L$.

Remark 10. Theorem 3 in [5] is the same as Theorem 5 above except that the former is for all $\boldsymbol{\lambda} \in \mathbb{R}_+^L$. By the explicit characterization of \mathcal{R}_{sup} in Theorem 2, namely \mathcal{R}^* , Theorem 5 is sufficient for proving the tightness of \mathcal{R}^* .

Remark 11. For $\alpha \in \mathcal{L}$ and $\mathbf{u} \in \Omega_L^{\alpha}$, let

$$\tilde{c}_{\alpha}(\mathbf{u}) = \frac{1}{\binom{L-1}{\alpha-1}}.$$

It is not difficult to see that for all $\alpha \in \mathcal{L}$, $\{\tilde{c}_\alpha(\mathbf{u}) : \mathbf{u} \in \Omega_L^\alpha\}$ is the unique optimal α -resolution for $\boldsymbol{\lambda} = \mathbf{1}$. Then (91) in Theorem 5 becomes

$$\frac{1}{\binom{L-1}{\alpha-2}} \sum_{\mathbf{u} \in \Omega_L^{\alpha-1}} H_{\mathbf{u}} \geq \frac{1}{\binom{L-1}{\alpha-1}} \sum_{\mathbf{u} \in \Omega_L^\alpha} H_{\mathbf{u}},$$

which is Han's inequality [32]. It was proved in [24] that both Han's inequality and the subset entropy inequality in [5] can be established from the subset entropy inequality of Madiman and Tetali [33].

Proof of Theorem 5. By symmetry, we only have to prove the theorem for $\boldsymbol{\lambda} \in \mathcal{G}_L^0$. We will prove the theorem by induction on L . It is easy to check that the theorem is true for $L = 2$. Assume the theorem is true for $L = N - 1$, we will show that the theorem is also true for $L = N$. This can be readily verified for $\boldsymbol{\lambda} \in \mathcal{G}_N^0$ such that $\zeta = 1$. Thus, we only need to consider $\boldsymbol{\lambda} \in \mathcal{G}_N^0$ such that $\zeta \geq 2$.

For any $\boldsymbol{\lambda}_N = (\lambda_1, \lambda_2, \dots, \lambda_N) \in \mathcal{G}_N^0$, by the construction in (38), we have $\boldsymbol{\lambda}_{N-1} = (\lambda_2, \lambda_3, \dots, \lambda_N) \in \mathcal{G}_{N-1}^0$. For $\alpha \in \{1, 2, \dots, N-1\}$, by the induction hypothesis, let $\{\tilde{c}_\alpha(\mathbf{u}) : \mathbf{u} \in \Omega_{N-1}^\alpha\}$ be an optimal α -resolution for $\boldsymbol{\lambda}_{N-1}$ such that (91) is satisfied for all $\alpha = 2, 3, \dots, N-1$. Now we need to design a proper optimal α -resolution $\{c_\alpha(\mathbf{w}) : \mathbf{w} \in \Omega_N^\alpha\}$ for $\boldsymbol{\lambda}_N$ that satisfies (91) for all $\alpha = 2, 3, \dots, N$.

From (38), there exists a $\theta \in \{1, 2, \dots, N-1\}$ such that

$$\lambda_1 = \frac{1}{\theta} \sum_{i=2}^N \lambda_i. \quad (92)$$

For any $\mathbf{u} \in \{0, 1\}^{N-1}$ and $\mathbf{w} \in \{0, 1\}^N$, let $\mathbf{u} = (u_2, u_3, \dots, u_N)$ and $\mathbf{w} = (w_1, w_2, \dots, w_N)$. For $\alpha = 1, 2, \dots, N$, we now construct an α -resolution for $\boldsymbol{\lambda}_N$ in (i) and (ii) in the following.

(i) Design $\{c_\alpha(\mathbf{w})\}$ for $\alpha = \theta + 1, \theta + 2, \dots, N$.

For $\alpha \geq \theta + 1$ and $\mathbf{w} \in \Omega_N^\alpha$, let

$$c_\alpha(\mathbf{w}) = \begin{cases} \tilde{c}_{\alpha-1}(\mathbf{u}), & \text{if } \mathbf{w} = (1, \mathbf{u}), \mathbf{u} \in \Omega_{N-1}^{\alpha-1} \\ 0, & \text{otherwise.} \end{cases}$$

From (92), we have

$$\lambda_1 \geq \frac{1}{\alpha-1} \sum_{i=2}^N \lambda_i \quad \text{for all } \alpha = \theta + 1, \theta + 2, \dots, N. \quad (93)$$

Lemma 9 in [5] states that $\{c_\alpha(\mathbf{w})\}$ is an optimal α -resolution for $\boldsymbol{\lambda}_N$ if

$$\lambda_1 > \frac{1}{\alpha-1} \sum_{i=2}^N \lambda_i \quad \text{for all } \alpha = \theta + 1, \theta + 2, \dots, N. \quad (94)$$

We observe that the lemma can be strengthened by replacing the strict inequality in (94) by a non-strict inequality (i.e., the condition in (93)) with essentially no change in the proof. Thus, by invoking this strengthened version of the lemma, we conclude that $\{c_\alpha(\mathbf{w})\}$ is an optimal α -resolution for $\boldsymbol{\lambda}_N$.

For all $\alpha = \theta + 2, \theta + 3, \dots, N$, following the steps leading to (48) in [5], we can check that

$$\sum_{\mathbf{w} \in \Omega_N^{\alpha-1}} c_{\alpha-1}(\mathbf{w}) H_{\mathbf{w}} \geq \sum_{\mathbf{w} \in \Omega_N^\alpha} c_\alpha(\mathbf{w}) H_{\mathbf{w}}.$$

(ii) Design $\{c_\alpha(\mathbf{w})\}$ for $\alpha = 1, 2, \dots, \theta$.

For $\alpha = \theta + 1, \theta, \dots, 2$ and any optimal α -resolution $\{c_\alpha(\mathbf{w}) : \mathbf{w} \in \Omega_N^\alpha\}$ for $\boldsymbol{\lambda}_N$, we claim that there exists an optimal $(\alpha - 1)$ -resolution $\{c_{\alpha-1}(\mathbf{w}) : \mathbf{w} \in \Omega_N^{\alpha-1}\}$ for $\boldsymbol{\lambda}_N$ such that (91) is satisfied. Since $\lambda_1 \leq \frac{1}{\alpha-1} \sum_{i=2}^N \lambda_i$, this is exactly the first case of the proof of Proposition 1 in [5], which is relatively straightforward.

In (i) and (ii) above, we have constructed an optimal α -resolution $\{c_\alpha(\mathbf{w})\}$ for any $\boldsymbol{\lambda}_N \in \mathcal{G}_N^0$ that satisfies (91) for all $\alpha = 2, 3, \dots, N$. This proves the theorem. \square

VII. CONCLUSION AND REMARKS

In this paper, we studied the SMDC problem for which superposition coding was proved to be optimal in [4], [5]. We enhanced their results by obtaining in closed form the minimum set of inequalities that is needed for characterizing \mathcal{R}_{sup} , the superposition coding rate region. We further show by the symmetry of the problem that only a much smaller subset of these inequalities needs to be verified in determining the achievability of a given rate tuple. Yet, the cardinality of this smaller set grows at least exponentially fast with L , the number of levels of the coding system, thus revealing the inherent complexity of the problem. A subset entropy inequality, which plays a key role in the converse proof in [5], requires a painstaking and extremely technical proof. We present a weaker version of this subset entropy inequality whose proof is considerably simpler. With our explicit characterization of the coding rate region, this weaker version of the subset entropy inequality is sufficient for proving the optimality of superposition coding. Some of our results may be extensible to the more general settings in [23]–[27].

While the coding rate region needs to be characterized by a set of inequalities whose size grows at least exponentially with L , if these inequalities are used directly for checking whether a certain rate tuple is within the coding rate region, then inevitably it requires at least an exponential amount of time. However, given that these inequalities are not arbitrary but instead highly structured, it may still be possible to devise a polynomial-time algorithm to preform the checking. If such an algorithm indeed exists, then the results in this paper can well be an important handle for finding it. This is an interesting problem for future research.

VIII. ACKNOWLEDGEMENTS

The authors wish to thank Professor Chandra Nair and the Associate Editor for the suggestions on the linear programming method for checking rate tuple achievability. The authors also wish to thank Professor Chao Tian for discussions on the duality of linear programs in the proof of Theorem 3. The authors also wish to thank the reviewers for their helpful and detailed comments.

APPENDIX A
PROOF OF LEMMA 9

We first prove (i). By Lemma 7 (i), the condition $\lambda_1 \leq \frac{1}{\eta} \sum_{j=2}^L \lambda_j$ implies that

$$f_{\eta+1}(\lambda) = \frac{1}{\eta+1} \sum_{j=1}^L \lambda_j. \quad (95)$$

In the following, we prove the claim by contradiction. Assume there exists a nonempty subset $\mathcal{I}_1 \subseteq \mathcal{I}$ such that $i \in \mathcal{I}_1$ if and only if

$$\lambda_{\pi_i(1)}^{(i)} > \frac{1}{\eta} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)}, \quad (96)$$

which is equivalent to

$$\sum_{j=1}^L \lambda_{\pi_i(j)}^{(i)} > \left(1 + \frac{1}{\eta}\right) \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)},$$

or

$$\frac{1}{\eta+1} \sum_{j=1}^L \lambda_{\pi_i(j)}^{(i)} > \frac{1}{\eta} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)}. \quad (97)$$

For all $i \in \mathcal{I}$, by Lemma 7 (ii), the condition in (96) implies that

$$f_{\eta+1}(\lambda^{(i)}) = f_{\eta}(\lambda_{\pi_i(2)}^{(i)}, \lambda_{\pi_i(3)}^{(i)}, \dots, \lambda_{\pi_i(L)}^{(i)}). \quad (98)$$

By Theorem 1, we have

$$f_{\eta}(\lambda_{\pi_i(2)}^{(i)}, \lambda_{\pi_i(3)}^{(i)}, \dots, \lambda_{\pi_i(L)}^{(i)}) \leq \frac{1}{\eta} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)}. \quad (99)$$

Then by (98), (99), and (97) we obtain

$$f_{\eta+1}(\lambda^{(i)}) < \frac{1}{\eta+1} \sum_{j=1}^L \lambda_{\pi_i(j)}^{(i)} = \frac{1}{\eta+1} \sum_{j=1}^L \lambda_j^{(i)}. \quad (100)$$

For $i \in \mathcal{I} \setminus \mathcal{I}_1$, we have

$$\lambda_{\pi_i(1)}^{(i)} \leq \frac{1}{\eta} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)},$$

which by Lemma 7 (i) implies that

$$f_{\eta+1}(\lambda^{(i)}) = \frac{1}{\eta+1} \sum_{j=1}^L \lambda_j^{(i)}. \quad (101)$$

Thus, we have from (45) that

$$\begin{aligned} f_{\eta+1}(\lambda) &= \sum_{i=1}^{S_L} c_i f_{\eta+1}(\lambda^{(i)}) \\ &= \sum_{i \in \mathcal{I}_1} c_i f_{\eta+1}(\lambda^{(i)}) + \sum_{i \in \mathcal{I} \setminus \mathcal{I}_1} c_i f_{\eta+1}(\lambda^{(i)}) \\ &< \sum_{i=1}^{S_L} c_i \cdot \left(\frac{1}{\eta+1} \sum_{j=1}^L \lambda_j^{(i)} \right) \\ &= \frac{1}{\eta+1} \sum_{j=1}^L \left(\sum_{i=1}^{S_L} c_i \lambda_j^{(i)} \right) \\ &= \frac{1}{\eta+1} \sum_{j=1}^L \lambda_j, \end{aligned}$$

where the inequality follows from (100) and (101). This is a contradiction to (95). Thus, the assumption in (96) is false and we have for all $i \in \mathcal{I}$ that

$$\lambda_{\pi_i(1)}^{(i)} \leq \frac{1}{\eta} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)}.$$

Next, we prove (ii) by contradiction. Assume there exists a nonempty subset $\mathcal{I}_2 \subseteq \mathcal{I}$ such that $i \in \mathcal{I}_2$ if and only if

$$\lambda_{\pi_i(1)}^{(i)} < \frac{1}{\eta} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)}, \quad (102)$$

which is equivalent to

$$\frac{1}{\eta+1} \sum_{j=1}^L \lambda_{\pi_i(j)}^{(i)} < \frac{1}{\eta} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)},$$

or

$$g_{\eta+1, \lambda^{(i)}}(0) < g_{\eta+1, \lambda^{(i)}}(1). \quad (103)$$

For any $i \in \mathcal{I}_2$, by Lemma 7 (i), (102) implies that

$$f_{\eta+1}(\lambda^{(i)}) = g_{\eta+1, \lambda^{(i)}}(0).$$

For any $t \in \{1, 2, \dots, \eta\}$, in light of (103), by applying the alternative version of Lemma 3 (ii) (see the remark below Lemma 3), we obtain

$$f_{\eta+1}(\lambda^{(i)}) = g_{\eta+1, \lambda^{(i)}}(0) < g_{\eta+1, \lambda^{(i)}}(1) < \dots < g_{\eta+1, \lambda^{(i)}}(t).$$

Then it follows from the definition of $\pi_i(\cdot)$ in (44) that

$$\begin{aligned} f_{\eta+1}(\lambda^{(i)}) &< g_{\eta+1, \lambda^{(i)}}(t) \\ &= \frac{1}{\eta+1-t} \sum_{j=t+1}^L \lambda_{\pi_i(j)}^{(i)} \\ &\leq \frac{1}{\eta+1-t} \sum_{j=t+1}^L \lambda_j^{(i)}, \end{aligned} \quad (104)$$

and so

$$f_{\eta+1}(\lambda^{(i)}) < \frac{1}{\eta+1-t} \sum_{j=t+1}^L \lambda_j^{(i)}. \quad (105)$$

For all $i \in \mathcal{I} \setminus \mathcal{I}_2$ and any $t \in \{1, 2, \dots, \eta\}$, by Theorem 1, similar to (104), we have

$$\begin{aligned} f_{\eta+1}(\lambda^{(i)}) &\leq g_{\eta+1, \lambda^{(i)}}(t) = \frac{1}{\eta+1-t} \sum_{j=t+1}^L \lambda_{\pi_i(j)}^{(i)} \\ &\leq \frac{1}{\eta+1-t} \sum_{j=t+1}^L \lambda_j^{(i)}, \end{aligned}$$

and so

$$f_{\eta+1}(\lambda^{(i)}) \leq \frac{1}{\eta+1-t} \sum_{j=t+1}^L \lambda_j^{(i)}. \quad (106)$$

Thus, by (45), we have for any $t \in \{1, 2, \dots, \eta\}$ that

$$\begin{aligned}
 f_{\eta+1}(\lambda) &= \sum_{i=1}^{S_L} c_i f_{\eta+1}(\lambda^{(i)}) \\
 &< \sum_{i=1}^{S_L} c_i \cdot \left(\frac{1}{\eta+1-t} \sum_{j=t+1}^L \lambda_j^{(i)} \right) \\
 &= \frac{1}{\eta+1-t} \sum_{j=t+1}^L \lambda_j \\
 &= g_{\eta+1, \lambda}(t),
 \end{aligned} \tag{107}$$

where the inequality follows from (105) and (106).

The condition $\lambda_1 \geq \frac{1}{\eta} \sum_{j=2}^L \lambda_j$ is equivalent to

$$g_{\eta+1, \lambda}(0) \geq g_{\eta+1, \lambda}(1).$$

Then by Theorem 1, we have

$$f_{\eta+1}(\lambda) = \min_{\beta \in \{0, 1, \dots, \eta\}} g_{\eta+1, \lambda}(\beta) = \min_{\beta \in \{1, 2, \dots, \eta\}} g_{\eta+1, \lambda}(\beta).$$

Thus, there must exist a $t \in \{1, 2, \dots, \eta\}$ such that

$$f_{\eta+1}(\lambda) = g_{\eta+1, \lambda}(t),$$

which is a contradiction to (107). Thus the assumption in (102) is false and we have for all $i \in \mathcal{I}$ that

$$\lambda_{\pi_i(1)}^{(i)} \geq \frac{1}{\eta} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)}.$$

APPENDIX B PROOF OF LEMMA 10

For all $i \in \mathcal{I}$, since $\lambda^{(i)} \in \mathcal{G}_L$ in light of (38), we only need to prove $\lambda_{\pi_i(1)}^{(i)} \leq \frac{1}{\eta-1} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)}$ and $\lambda_{\pi_i(1)}^{(i)} \geq \frac{1}{\eta} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)}$.

We first prove the upper bound on $\lambda_{\pi_i(1)}^{(i)}$. For $\eta = 1$ and $i = 1$, we have

$$\lambda_{\pi_i(1)}^{(i)} = 1 = \frac{1}{\eta-1} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)}.$$

For $\eta = 1$ and $i \in \mathcal{I} \setminus \{1\}$, it is obvious that

$$\lambda_{\pi_i(1)}^{(i)} < \frac{1}{\eta-1} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)} = \infty.$$

For $\eta \in \{2, 3, \dots, \zeta-1\}$, the upper bound in (47) can be rewritten as

$$\lambda_1 < \frac{1}{\eta'} \sum_{j=2}^L \lambda_j,$$

where $\eta' = \eta - 1$ and $\eta' \in \{1, 2, \dots, \zeta-2\}$. By Lemma 9 (i), this implies that

$$\lambda_{\pi_i(1)}^{(i)} \leq \frac{1}{\eta'} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)} = \frac{1}{\eta-1} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)}.$$

Thus, the upper bound on $\lambda_{\pi_i(1)}^{(i)}$ is proved.

Now we prove the lower bound on $\lambda_{\pi_i(1)}^{(i)}$. For $\eta \in \{1, 2, \dots, \zeta-1\}$, the lower bound in (47) is

$$\lambda_1 \geq \frac{1}{\eta} \sum_{j=2}^L \lambda_j,$$

and so by Lemma 9 (ii), we have

$$\lambda_{\pi_i(1)}^{(i)} \geq \frac{1}{\eta} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)}.$$

If the lower bound in (47) is tight, it follows immediately from Lemma 9 that for any $\eta \in \{1, 2, \dots, \zeta-1\}$,

$$\lambda_{\pi_i(1)}^{(i)} = \frac{1}{\eta} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)}.$$

This proves the lemma.

APPENDIX C PROOF OF LEMMA 11

We only need to show that for $\alpha = 1, 2, \dots, L-1$,

$$f_\alpha(\lambda_{L-1}) = \sum_{i=1}^{S_L} c_i f_\alpha(\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}).$$

By (47) and Lemma 10, we have for $i \in \mathcal{I}$ that

$$\lambda_{\pi_i(1)}^{(i)} \in \left\{ \frac{1}{\eta} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)}, \frac{1}{\eta-1} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)} \right\}. \tag{108}$$

Consider the following two cases:

- i) $\alpha = 1, 2, \dots, \eta-1$;
- ii) $\alpha = \eta, \eta+1, \dots, L-1$.

Case i): For $\alpha = 1, 2, \dots, \eta-1$, if $\eta = 2$, then α can only be 1 and it is easy to see that

$$\begin{aligned}
 f_1(\lambda_{L-1}) &= \sum_{j=2}^L \lambda_j = \sum_{j=2}^L \sum_{i=1}^{S_L} c_i \lambda_j^{(i)} = \sum_{i=1}^{S_L} c_i \sum_{j=2}^L \lambda_j^{(i)} \\
 &= \sum_{i=1}^{S_L} c_i f_1(\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}).
 \end{aligned}$$

If $\eta > 2$, consider the following. The second inequality in (47) is equivalent to

$$\frac{1}{\eta} \sum_{j=1}^L \lambda_j < \frac{1}{\eta-1} \sum_{j=2}^L \lambda_j \tag{109}$$

or

$$g_{\eta, \lambda}(0) < g_{\eta, \lambda}(1).$$

By applying the alternative version of Lemma 3 (ii) (see the remark below Lemma 3), we obtain

$$g_{\eta, \lambda}(1) < g_{\eta, \lambda}(2),$$

which is equivalent to

$$\frac{1}{\eta-1} \sum_{j=2}^L \lambda_j < \frac{1}{\eta-2} \sum_{j=3}^L \lambda_j$$

or

$$\lambda_2 < \frac{1}{\eta-2} \sum_{j=3}^L \lambda_j. \quad (110)$$

Then by Lemma 7 (i), we have

$$f_\alpha(\lambda_{L-1}) = g_{\alpha, \lambda_{L-1}}(0) = \frac{1}{\alpha} \sum_{j=2}^L \lambda_j. \quad (111)$$

Since (108) implies

$$\lambda_{\pi_i(1)}^{(i)} \leq \frac{1}{\eta-1} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)}, \quad (112)$$

similar to (109)-(110) (with all $<$'s replaced by \leq 's), we have

$$\lambda_{\pi_i(2)}^{(i)} \leq \frac{1}{\eta-2} \sum_{j=3}^L \lambda_{\pi_i(j)}^{(i)}. \quad (113)$$

Thus, following (112) and (113), we have

$$\begin{aligned} \lambda_{\pi_i(1)}^{(i)} &\leq \frac{1}{\eta-1} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)} \\ &\leq \frac{1}{\eta-1} \left(\frac{1}{\eta-2} + 1 \right) \sum_{j=3}^L \lambda_{\pi_i(j)}^{(i)} \\ &= \frac{1}{\eta-2} \sum_{j=3}^L \lambda_{\pi_i(j)}^{(i)} \\ &\leq \frac{1}{\eta-2} \sum_{j \in \{2,3,\dots,L\} \setminus \{j_0\}} \lambda_{\pi_i(j)}^{(i)} \end{aligned} \quad (114)$$

for any $j_0 \in \{2,3,\dots,L\}$. Let $\pi'_i(\cdot)$ be a permutation of $\{2,3,\dots,L\}$ defined as follows:

- a) if $\pi_i(1) = 1$, then $\pi'_i(j) = \pi_i(j)$ for all $j \in \{2,3,\dots,L\}$;
- b) if $\pi_i(j_0) = 1$ for some $j_0 \in \{2,3,\dots,L\}$, then

$$\pi'_i(j) = \begin{cases} \pi_i(j-1), & \text{for } j = 2, 3, \dots, j_0 \\ \pi_i(j), & \text{for } j = j_0 + 1, \dots, L. \end{cases} \quad (115)$$

It is easy to check that

$$\lambda_{\pi'_i(2)}^{(i)} \geq \lambda_{\pi'_i(3)}^{(i)} \geq \dots \geq \lambda_{\pi'_i(L)}^{(i)}.$$

If a) holds, then by (113), we have

$$\lambda_{\pi'_i(2)}^{(i)} = \lambda_{\pi_i(2)}^{(i)} \leq \frac{1}{\eta-2} \sum_{j=3}^L \lambda_{\pi_i(j)}^{(i)} = \frac{1}{\eta-2} \sum_{j=3}^L \lambda_{\pi'_i(j)}^{(i)}.$$

If b) holds, then by (114), we have

$$\begin{aligned} \lambda_{\pi'_i(2)}^{(i)} &= \lambda_{\pi_i(1)}^{(i)} \\ &\leq \frac{1}{\eta-2} \sum_{j \in \{2,3,\dots,L\} \setminus \{j_0\}} \lambda_{\pi_i(j)}^{(i)} \\ &= \frac{1}{\eta-2} \sum_{j=3}^L \lambda_{\pi'_i(j)}^{(i)}. \end{aligned}$$

Summarizing the two cases, we see that

$$\lambda_{\pi'_i(2)}^{(i)} \leq \frac{1}{\eta-2} \sum_{j=3}^L \lambda_{\pi'_i(j)}^{(i)}$$

always holds. By Lemma 7 (i), this implies that

$$f_\alpha(\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}) = \frac{1}{\alpha} \sum_{j=2}^L \lambda_{\pi'_i(j)}^{(i)} = \frac{1}{\alpha} \sum_{j=2}^L \lambda_j^{(i)}. \quad (116)$$

Following (111), we have

$$\begin{aligned} f_\alpha(\lambda_{L-1}) &= \frac{1}{\alpha} \sum_{j=2}^L \lambda_j \\ &= \frac{1}{\alpha} \sum_{j=2}^L \left(\sum_{i=1}^{S_L} c_i \lambda_j^{(i)} \right) \\ &= \sum_{i=1}^{S_L} c_i \left(\frac{1}{\alpha} \sum_{j=2}^L \lambda_j^{(i)} \right) \\ &= \sum_{i=1}^{S_L} c_i f_\alpha(\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}), \end{aligned} \quad (117)$$

where (117) follows from (79) and (118) follows from (116).

Case ii): For $\alpha = \eta, \eta+1, \dots, L-1$, by Lemma 7 (ii), the first inequality in (47) implies that

$$f_\alpha(\lambda_{L-1}) = f_{\alpha+1}(\lambda). \quad (119)$$

For any $i \in \{1, 2, \dots, S_L\}$, since (108) implies $\lambda_{\pi_i(1)}^{(i)} \geq \frac{1}{\eta} \sum_{j=2}^L \lambda_{\pi_i(j)}^{(i)}$, we have by Lemma 7 (ii) that

$$f_{\alpha+1}(\lambda^{(i)}) = f_\alpha(\lambda_{\pi_i(2)}^{(i)}, \lambda_{\pi_i(3)}^{(i)}, \dots, \lambda_{\pi_i(L)}^{(i)}). \quad (120)$$

From the definition of $\pi'_i(\cdot)$, it is readily seen that

$$\lambda_{\pi_i(j)}^{(i)} \leq \lambda_{\pi'_i(j)}^{(i)}, \text{ for all } j = 2, 3, \dots, L.$$

Thus, we have for any $\beta = 1, 2, \dots, \alpha-1$ that

$$\frac{1}{\alpha-\beta} \sum_{j=\beta+1}^L \lambda_{\pi_i(j)}^{(i)} \leq \frac{1}{\alpha-\beta} \sum_{j=\beta+1}^L \lambda_{\pi'_i(j)}^{(i)}.$$

By Theorem 1, this implies that

$$f_\alpha(\lambda_{\pi_i(2)}^{(i)}, \lambda_{\pi_i(3)}^{(i)}, \dots, \lambda_{\pi_i(L)}^{(i)}) \leq f_\alpha(\lambda_{\pi'_i(2)}^{(i)}, \lambda_{\pi'_i(3)}^{(i)}, \dots, \lambda_{\pi'_i(L)}^{(i)}),$$

and thus by (120), we have

$$f_{\alpha+1}(\lambda^{(i)}) \leq f_\alpha(\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}). \quad (121)$$

Following (119), we have

$$f_\alpha(\lambda_{L-1}) = f_{\alpha+1}(\lambda) \quad (122)$$

$$= \sum_{i=1}^{S_L} c_i f_{\alpha+1}(\lambda^{(i)}) \quad (123)$$

$$\leq \sum_{i=1}^{S_L} c_i f_\alpha(\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}) \quad (124)$$

$$\leq f_\alpha \left(\sum_{i=1}^{S_L} c_i \cdot (\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}) \right) \quad (125)$$

$$= f_\alpha \left(\sum_{i=1}^{S_L} c_i \lambda_2^{(i)}, \sum_{i=1}^{S_L} c_i \lambda_3^{(i)}, \dots, \sum_{i=1}^{S_L} c_i \lambda_L^{(i)} \right) \quad (126)$$

$$= f_\alpha(\lambda_2, \lambda_3, \dots, \lambda_L) \quad (127)$$

where both (123) and (126) follow from (79), (124) follows from (121), and (125) follows from Lemma 8. Upon observing that the LHS of (122) is the same as the RHS of (127), we conclude that the inequalities in both (124) and (125) are tight, and hence

$$f_\alpha(\lambda_{L-1}) = \sum_{i=1}^{S_L} c_i f_\alpha(\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}).$$

The lemma is proved.

APPENDIX D PROOF OF LEMMA 12

Fix $i \in \{1, 2, \dots, S_L\}$ and assume that $(\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}) \in \mathcal{G}_{L-1}$. Let γ_j , $j = 1, 2, \dots, \zeta - 1$ be the integer such that

$$\lambda_{\pi_i(j)}^{(i)} = \frac{1}{\gamma_j} \sum_{k=j+1}^L \lambda_{\pi_i(k)}^{(i)},$$

and let $\gamma_\zeta = 0$. Note that the role of γ_j for $\lambda^{(i)}$ is the same as the role of θ_j for λ (cf. (39)). Also note that ζ and γ_j depend on i , but since we fix i , this dependence is omitted to simplify notation.

Let $j_0 \in \{1, 2, \dots, L\}$ be such that

$$\lambda_{\pi_i(j_0)}^{(i)} = \lambda_1^{(i)}. \quad (128)$$

If $j_0 \geq \zeta + 1$, $\lambda_1^{(i)} = 0$ and thus the lemma is proved. If $j_0 = 1$, the lemma is immediate from (128). If $2 \leq j_0 \leq \zeta$, we claim that $\gamma_j = \gamma_{j_0} + (j_0 - j)$ and $\lambda_{\pi_i(j)}^{(i)} = \lambda_{\pi_i(j_0)}^{(i)}$ for all $j = j_0, j_0 - 1, \dots, 1$. Then the lemma follows from the claim for $j = 1$. In the following, we prove the claim by induction on j for $j \leq j_0$. The claim is immediate for $j = j_0$. Assume the claim is true for $j = j_0, j_0 - 1, \dots, N$ for some $N \in \{2, 3, \dots, j_0\}$, and we will show that the claim is also true for $j = N - 1$. By the induction hypothesis, we have

$$\gamma_N = \gamma_{j_0} + (j_0 - N),$$

and for all $j = j_0, j_0 - 1, \dots, N$,

$$\lambda_{\pi_i(j)}^{(i)} = \lambda_{\pi_i(j_0)}^{(i)} = \frac{1}{\gamma_{j_0}} \sum_{k=j_0+1}^L \lambda_{\pi_i(k)}^{(i)}. \quad (129)$$

By (38) and (40), there exists

$$\gamma_{N-1} \in \{1, 2, \dots, \gamma_{j_0} + (j_0 - N + 1)\} \quad (130)$$

such that

$$\lambda_{\pi_i(N-1)}^{(i)} = \frac{1}{\gamma_{N-1}} \sum_{k=N}^L \lambda_{\pi_i(k)}^{(i)}. \quad (131)$$

Thus, we have

$$\begin{aligned} \lambda_{\pi_i(N-1)}^{(i)} &= \frac{1}{\gamma_{N-1}} \sum_{k=N}^L \lambda_{\pi_i(k)}^{(i)} \\ &= \frac{1}{\gamma_{N-1}} \left(1 + \sum_{k=N}^{j_0} \frac{1}{\gamma_{j_0}} \right) \sum_{k=j_0+1}^L \lambda_{\pi_i(k)}^{(i)} \quad (132) \\ &= \frac{\gamma_{j_0} + (j_0 - N + 1)}{\gamma_{j_0} \gamma_{N-1}} \sum_{k=j_0+1}^L \lambda_{\pi_i(k)}^{(i)}, \quad (133) \end{aligned}$$

where (132) follows from (129). In light of (128) and $j_0 \geq 2$, recall the definition of $\pi'_i(\cdot)$ in (115). With the assumption that $(\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}) \in \mathcal{G}_{L-1}$, by (38), there exists an integer γ'_{N-1} such that

$$\lambda_{\pi'_i(N)}^{(i)} = \frac{1}{\gamma'_{N-1}} \sum_{k=N+1}^L \lambda_{\pi'_i(k)}^{(i)}$$

or

$$\lambda_{\pi_i(N-1)}^{(i)} = \frac{1}{\gamma'_{N-1}} \left(\sum_{k=N}^{j_0-1} \lambda_{\pi_i(k)}^{(i)} + \sum_{k=j_0+1}^L \lambda_{\pi_i(k)}^{(i)} \right). \quad (134)$$

Comparing the RHS of (131) and (134), since $\lambda_{\pi_i(j_0)}^{(i)} \geq 1$, we have

$$\sum_{k=N}^L \lambda_{\pi_i(k)}^{(i)} > \sum_{k=N}^{j_0-1} \lambda_{\pi_i(k)}^{(i)} + \sum_{k=j_0+1}^L \lambda_{\pi_i(k)}^{(i)}.$$

Since the LHS of (131) and (134) are the same, we see that $\gamma'_{N-1} < \gamma_{N-1}$, which implies that

$$\gamma'_{N-1} \leq \gamma_{j_0} + (j_0 - N).$$

Following (134), we have

$$\begin{aligned} \lambda_{\pi_i(N-1)}^{(i)} &= \frac{1}{\gamma'_{N-1}} \left(\sum_{k=N}^{j_0-1} \lambda_{\pi_i(k)}^{(i)} + \sum_{k=j_0+1}^L \lambda_{\pi_i(k)}^{(i)} \right) \\ &= \frac{1}{\gamma'_{N-1}} \left(1 + \sum_{k=N}^{j_0-1} \frac{1}{\gamma_{j_0}} \right) \sum_{k=j_0+1}^L \lambda_{\pi_i(k)}^{(i)} \quad (135) \\ &= \frac{\gamma_{j_0} + (j_0 - N)}{\gamma_{j_0} \gamma'_{N-1}} \sum_{k=j_0+1}^L \lambda_{\pi_i(k)}^{(i)}, \quad (136) \end{aligned}$$

where (135) follows from (129). Comparing (133) and (136), it is easy to see that

$$\gamma'_{N-1} = \frac{\gamma_{N-1}[\gamma_{j_0} + (j_0 - N)]}{\gamma_{j_0} + (j_0 - N + 1)}.$$

Since $\gamma_{j_0} + (j_0 - N)$ and $\gamma_{j_0} + (j_0 - N + 1)$ are coprime and $\gamma_{N-1} \leq \gamma_{j_0} + (j_0 - N + 1)$ by (130), we have

$$\gamma_{N-1} = \gamma_{j_0} + (j_0 - N + 1). \quad (137)$$

Substituting (137) into (133) and invoking (129), we have $\lambda_{\pi_i(N-1)}^{(i)} = \lambda_{\pi_i(j_0)}^{(i)}$. This implies that the claim is true for $j = N - 1$. The lemma is proved.

APPENDIX E PROOF OF LEMMA 13

Since there is only one vector in \mathcal{G}_1 , we only need to consider $L \geq 2$. If $\zeta = 1$ for $\lambda^{(i_0)}$, it is obvious that $\lambda^{(i_0)}$ cannot be a conic combination of the other vectors in \mathcal{G}_L . Thus, we consider only $\lambda^{(i_0)} \in \mathcal{G}_L$ such that $\zeta \geq 2$. We prove the lemma by induction on L for $L \geq 2$. We first check that the claim is true for $L = 2$. It is easy to see from (38) that $\mathcal{G}_2 = \{(1, 0), (0, 1), (1, 1)\}$. Then $\mathbf{f}((1, 0)) = \mathbf{f}((0, 1)) = (1, 0)$ and $\mathbf{f}((1, 1)) = (2, 1)$. Since

$$(1, 1) = (1, 0) + (0, 1)$$

whereas

$$f_2((1, 1)) > f_2((1, 0)) + f_2((0, 1)),$$

we see that $((1, 1), \mathbf{f}((1, 1)))$ cannot be a conic combination of $((1, 0), \mathbf{f}((1, 0)))$ and $((0, 1), \mathbf{f}((0, 1)))$. Thus, the lemma is true for $L = 2$. For any $L \geq 3$, the lemma will be proved by contradiction via the following proposition, whose proof is given in Appendix F.

Proposition 1. *For any $L \geq 3$, if Lemma 13 is false, then the lemma is false for $L - 1$.*

By backward induction, if Lemma 13 is false for any $L \geq 3$, then the lemma is false for $L = 2$. This is a contradiction because we already have shown that the lemma is true for $L = 2$. This proves the lemma for all $L \geq 2$.

APPENDIX F PROOF OF PROPOSITION 1

Assume Lemma 13 is false for some $L \geq 3$, i.e., for some $i_0 \in \{1, 2, \dots, S_L\}$, there exists $(c_1, c_2, \dots, c_{S_L}) \in \mathbb{R}_+^{S_L}$ such that $c_{i_0} = 0$ and

$$(\boldsymbol{\lambda}^{(i_0)}, \mathbf{f}(\boldsymbol{\lambda}^{(i_0)})) = \sum_{i=1}^{S_L} c_i \cdot (\boldsymbol{\lambda}^{(i)}, \mathbf{f}(\boldsymbol{\lambda}^{(i)})). \quad (138)$$

Assume without loss of generality that $\boldsymbol{\lambda}^{(i_0)} \in \mathcal{G}_L^0$. Since we assume at the beginning of Appendix E that $\zeta \geq 2$ for $\boldsymbol{\lambda}^{(i_0)}$, we can see from (38) that $(\lambda_2^{(i_0)}, \lambda_3^{(i_0)}, \dots, \lambda_L^{(i_0)}) \in \mathcal{G}_{L-1}^0$ by construction. Let $\mathcal{G}_{L-1} = \{\boldsymbol{\lambda}_{L-1}^{(1)}, \boldsymbol{\lambda}_{L-1}^{(2)}, \dots, \boldsymbol{\lambda}_{L-1}^{(S_{L-1})}\}$. Then there exists a unique $j_0 \in \{1, 2, \dots, S_{L-1}\}$ such that

$$\boldsymbol{\lambda}_{L-1}^{(j_0)} = (\lambda_2^{(i_0)}, \lambda_3^{(i_0)}, \dots, \lambda_L^{(i_0)}). \quad (139)$$

By Lemma 11, (138) implies that

$$\begin{aligned} & (\boldsymbol{\lambda}_{L-1}^{(j_0)}, \mathbf{f}(\boldsymbol{\lambda}_{L-1}^{(j_0)})) \\ &= \sum_{i=1}^{S_L} c_i \cdot ((\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}), \mathbf{f}(\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)})). \end{aligned} \quad (140)$$

Let $\mathcal{K}_L = \{1, 2, \dots, S_L\}$, $\mathcal{I}_L = \{i \in \mathcal{K}_L : c_i \neq 0\}$, $\mathcal{K}_L^{(j_0)} = \{i \in \mathcal{K}_L : (\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}) = \boldsymbol{\lambda}_{L-1}^{(j_0)}\}$, and

$$d_0 = \sum_{i \in \mathcal{K}_L^{(j_0)}} c_i. \quad (141)$$

In the proof of Theorem 2, we have shown that any vector in \mathcal{F}_L^1 is a conic combination of the vectors in \mathcal{F}_L^2 . Then for any $i \in \mathcal{K}_L \setminus \mathcal{K}_L^{(j_0)}$, there exists $(t_1^{(i)}, t_2^{(i)}, \dots, t_{S_{L-1}}^{(i)}) \in \mathbb{R}_+^{S_{L-1}}$ such that

$$\begin{aligned} & ((\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}), \mathbf{f}(\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)})) \\ &= \sum_{j=1}^{S_{L-1}} t_j^{(i)} (\boldsymbol{\lambda}_{L-1}^{(j)}, \mathbf{f}(\boldsymbol{\lambda}_{L-1}^{(j)})). \end{aligned} \quad (142)$$

Substitute (141) and (142) into (140), we have

$$\begin{aligned} & (\boldsymbol{\lambda}_{L-1}^{(j_0)}, \mathbf{f}(\boldsymbol{\lambda}_{L-1}^{(j_0)})) \\ &= d_0 (\boldsymbol{\lambda}_{L-1}^{(j_0)}, \mathbf{f}(\boldsymbol{\lambda}_{L-1}^{(j_0)})) \\ &+ \sum_{i \in \mathcal{K}_L \setminus \mathcal{K}_L^{(j_0)}} c_i \left[\sum_{j=1}^{S_{L-1}} t_j^{(i)} (\boldsymbol{\lambda}_{L-1}^{(j)}, \mathbf{f}(\boldsymbol{\lambda}_{L-1}^{(j)})) \right] \\ &= \left(d_0 + \sum_{i \in \mathcal{K}_L \setminus \mathcal{K}_L^{(j_0)}} c_i t_{j_0}^{(i)} \right) (\boldsymbol{\lambda}_{L-1}^{(j_0)}, \mathbf{f}(\boldsymbol{\lambda}_{L-1}^{(j_0)})) \\ &+ \sum_{i \in \mathcal{K}_L \setminus \mathcal{K}_L^{(j_0)}} c_i \left[\sum_{j \in \mathcal{K}_{L-1} \setminus \{j_0\}} t_j^{(i)} (\boldsymbol{\lambda}_{L-1}^{(j)}, \mathbf{f}(\boldsymbol{\lambda}_{L-1}^{(j)})) \right] \\ &= \left(d_0 + \sum_{i \in \mathcal{K}_L \setminus \mathcal{K}_L^{(j_0)}} c_i t_{j_0}^{(i)} \right) (\boldsymbol{\lambda}_{L-1}^{(j_0)}, \mathbf{f}(\boldsymbol{\lambda}_{L-1}^{(j_0)})) \\ &+ \sum_{j \in \mathcal{K}_{L-1} \setminus \{j_0\}} \left(\sum_{i \in \mathcal{K}_L \setminus \mathcal{K}_L^{(j_0)}} c_i t_j^{(i)} \right) (\boldsymbol{\lambda}_{L-1}^{(j)}, \mathbf{f}(\boldsymbol{\lambda}_{L-1}^{(j)})). \end{aligned}$$

Thus,

$$\begin{aligned} & \left(1 - d_0 - \sum_{i \in \mathcal{K}_L \setminus \mathcal{K}_L^{(j_0)}} c_i t_{j_0}^{(i)} \right) (\boldsymbol{\lambda}_{L-1}^{(j_0)}, \mathbf{f}(\boldsymbol{\lambda}_{L-1}^{(j_0)})) \\ &= \sum_{j \in \mathcal{K}_{L-1} \setminus \{j_0\}} \left(\sum_{i \in \mathcal{K}_L \setminus \mathcal{K}_L^{(j_0)}} c_i t_j^{(i)} \right) (\boldsymbol{\lambda}_{L-1}^{(j)}, \mathbf{f}(\boldsymbol{\lambda}_{L-1}^{(j)})). \end{aligned} \quad (143)$$

Proposition 2. *There exists $i \in \mathcal{I}_L$ such that $(\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}) \neq \boldsymbol{\lambda}_{L-1}^{(j_0)}$.*

The proof of Proposition 2 is given in Appendix G. The proposition implies that

$$\mathcal{I}_L \cap (\mathcal{K}_L \setminus \mathcal{K}_L^{(j_0)}) \neq \emptyset. \quad (144)$$

For any $i \in \mathcal{I}_L \cap (\mathcal{K}_L \setminus \mathcal{K}_L^{(j_0)})$, we can rewrite (142) as follows:

$$\begin{aligned} & ((\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}), \mathbf{f}(\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)})) \\ &= t_{j_0}^{(i)} (\boldsymbol{\lambda}_{L-1}^{(j_0)}, \mathbf{f}(\boldsymbol{\lambda}_{L-1}^{(j_0)})) + \sum_{j \in \mathcal{K}_{L-1} \setminus \{j_0\}} t_j^{(i)} (\boldsymbol{\lambda}_{L-1}^{(j)}, \mathbf{f}(\boldsymbol{\lambda}_{L-1}^{(j)})). \end{aligned}$$

Since $(\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}) \neq \boldsymbol{\lambda}_{L-1}^{(j_0)}$, there must exist $j \in \mathcal{K}_{L-1} \setminus \{j_0\}$ such that

$$t_j^{(i)} > 0. \quad (145)$$

For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^{L-1}$, define a binary relation ‘>’ by $\mathbf{x} > \mathbf{y}$ if and only if $(\mathbf{x} - \mathbf{y}) \in \mathbb{R}_+^{L-1}$, i.e., \mathbf{x} is strictly greater than

\mathbf{y} in at least one component (cf. (3)). Then for the RHS of (143), we have

$$\begin{aligned}
& \sum_{j \in \mathcal{K}_{L-1} \setminus \{j_0\}} \left(\sum_{i \in \mathcal{K}_L \setminus \mathcal{K}_L^{(j_0)}} c_i t_j^{(i)} \right) (\lambda_{L-1}^{(j)}, \mathbf{f}(\lambda_{L-1}^{(j)})) \\
&= \sum_{i \in \mathcal{K}_L \setminus \mathcal{K}_L^{(j_0)}} c_i \left[\sum_{j \in \mathcal{K}_{L-1} \setminus \{j_0\}} t_j^{(i)} (\lambda_{L-1}^{(j)}, \mathbf{f}(\lambda_{L-1}^{(j)})) \right] \\
&= \sum_{i \in \mathcal{I}_L \cap (\mathcal{K}_L \setminus \mathcal{K}_L^{(j_0)})} c_i \left[\sum_{j \in \mathcal{K}_{L-1} \setminus \{j_0\}} t_j^{(i)} (\lambda_{L-1}^{(j)}, \mathbf{f}(\lambda_{L-1}^{(j)})) \right] \\
&> 0, \tag{146}
\end{aligned}$$

where the inequality follows from (144), (145) and the fact that $\lambda_{L-1}^{(j)} > 0$ for all $j \in \mathcal{K}_{L-1} \setminus \{j_0\}$. Then we can see from (143) that

$$\left(1 - d_0 - \sum_{i \in \mathcal{K}_L \setminus \mathcal{K}_L^{(j_0)}} c_i t_{j_0}^{(i)} \right) (\lambda_{L-1}^{(j_0)}, \mathbf{f}(\lambda_{L-1}^{(j_0)})) > 0,$$

which implies that

$$1 - d_0 - \sum_{i \in \mathcal{K}_L \setminus \mathcal{K}_L^{(j_0)}} c_i t_{j_0}^{(i)} > 0.$$

For each $j \in \mathcal{K}_{L-1} \setminus \{j_0\}$, let

$$d_j = \frac{1}{1 - d_0 - \sum_{i \in \mathcal{K}_L \setminus \mathcal{K}_L^{(j_0)}} c_i t_{j_0}^{(i)}} \sum_{i \in \mathcal{K}_L \setminus \mathcal{K}_L^{(j_0)}} c_i t_j^{(i)}. \tag{147}$$

It is easy to see that $d_j \geq 0$ for all $j \in \mathcal{K}_{L-1} \setminus \{j_0\}$. By (146), there exists $j \in \mathcal{K}_{L-1} \setminus \{j_0\}$ such that $d_j > 0$. Upon letting $d_j = 0$ for $j = j_0$, by (143) and (147), we have

$$(\lambda_{L-1}^{(j_0)}, \mathbf{f}(\lambda_{L-1}^{(j_0)})) = \sum_{j=1}^{S_{L-1}} d_j (\lambda_{L-1}^{(j)}, \mathbf{f}(\lambda_{L-1}^{(j)})).$$

This means that Lemma 13 is false for $L-1$. The proposition is proved.

APPENDIX G PROOF OF PROPOSITION 2

Since $\lambda^{(i_0)} \in \mathcal{G}_L^0$, there exists a unique $\eta_{i_0} \in \{0, 1, \dots, L-1\}$ such that

$$\lambda_1^{(i_0)} = \frac{1}{\eta_{i_0}} \sum_{j=2}^L \lambda_j^{(i_0)}. \tag{148}$$

Recall from (139) that $\lambda_{L-1}^{(j_0)} = (\lambda_2^{(i_0)}, \lambda_3^{(i_0)}, \dots, \lambda_L^{(i_0)})$. Since we assume at the beginning of Appendix E that $\zeta \geq 2$ for $\lambda^{(i_0)}$, we see that

$$\lambda_{L-1}^{(j_0)} \neq \mathbf{0}, \tag{149}$$

which implies that $\sum_{j=2}^L \lambda_j^{(i_0)} > 0$. Then we have $\eta_{i_0} \neq 0$, otherwise $\lambda_1^{(i_0)} = \infty$ in (148). Thus, $\eta_{i_0} \in \{1, 2, \dots, L-1\}$. By Remark 4 following Lemma 7, we have

$$f_{\eta_{i_0}+1}(\lambda^{(i_0)}) = f_{\eta_{i_0}}(\lambda_{L-1}^{(j_0)}) = \frac{1}{\eta_{i_0}+1} \sum_{j=1}^L \lambda_j^{(i_0)}. \tag{150}$$

We now prove the proposition by contradiction. Assume that for all $i \in \mathcal{I}_L$,

$$(\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}) = \lambda_{L-1}^{(j_0)}. \tag{151}$$

This means that for each $i \in \mathcal{I}_L$,

$$\begin{aligned}
\lambda^{(i)} &= (\lambda_1^{(i)}, \lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}) \\
&= (\lambda_1^{(i)}, \lambda_2^{(i_0)}, \lambda_3^{(i_0)}, \dots, \lambda_L^{(i_0)}). \tag{152}
\end{aligned}$$

Furthermore, since $\lambda^{(i_0)} \in \mathcal{G}_L^0$, we see from (38) that $\lambda_{L-1}^{(j_0)} \in \mathcal{G}_{L-1}^0$ by construction, so that $(\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}) \in \mathcal{G}_{L-1}^0$. Then by Lemma 12, we have $\lambda_1^{(i)} = 0$ or $\lambda_{\pi_i(1)}^{(i)}$.

Let \mathcal{I}_L^0 be the subset of \mathcal{I}_L such that $i \in \mathcal{I}_L^0$ if and only if $\lambda_1^{(i)} = 0$. For $i \in \mathcal{I}_L^0$, it is easy to see that $\lambda_{\pi_i(L)}^{(i)} = 0$. Then upon noting that

$$(\lambda_{\pi_i(1)}^{(i)}, \lambda_{\pi_i(2)}^{(i)}, \dots, \lambda_{\pi_i(L-1)}^{(i)}) = (\lambda_2^{(i_0)}, \lambda_3^{(i_0)}, \dots, \lambda_L^{(i_0)}),$$

by Theorem 1 we have

$$\begin{aligned}
f_{\eta_{i_0}+1}(\lambda^{(i)}) &= \min_{\beta \in \{0, 1, \dots, \eta_{i_0}\}} \frac{1}{(\eta_{i_0}+1) - \beta} \sum_{j=\beta+1}^{L-1} \lambda_{\pi_i(j)}^{(i)} \\
&= \min_{\beta \in \{0, 1, \dots, \eta_{i_0}\}} \frac{1}{(\eta_{i_0}+1) - \beta} \sum_{j=\beta+2}^L \lambda_j^{(i_0)} \\
&= f_{\eta_{i_0}+1}(\lambda_{L-1}^{(j_0)}) \\
&< f_{\eta_{i_0}}(\lambda_{L-1}^{(j_0)}),
\end{aligned}$$

where the inequality follows from Lemma 5 in [5]. By (150), this implies that

$$f_{\eta_{i_0}+1}(\lambda^{(i)}) < \frac{1}{\eta_{i_0}+1} \sum_{j=1}^L \lambda_j^{(i_0)}. \tag{153}$$

For $i \in \mathcal{I}_L \setminus \mathcal{I}_L^0$, we have $\lambda_1^{(i)} = \lambda_{\pi_i(1)}^{(i)}$. Since $\lambda^{(i)} \in \mathcal{G}_L$, there exists a unique $\eta_i \in \{0, 1, \dots, L-1\}$ such that

$$\lambda_1^{(i)} = \frac{1}{\eta_i} \sum_{j=2}^L \lambda_j^{(i)}. \tag{154}$$

From (149) and the assumption in (151), we have $(\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}) \neq \mathbf{0}$. Then from (154), we have $\eta_i \neq 0$, and thus $\eta_i \in \{1, 2, \dots, L-1\}$. Since $c_{i_0} = 0$ in (138), we have $\lambda^{(i)} \neq \lambda^{(i_0)}$ for all $i \in \mathcal{I}_L$ and hence for all $i \in \mathcal{I}_L \setminus \mathcal{I}_L^0$. In light of (152), $\lambda^{(i)} \neq \lambda^{(i_0)}$ implies $\lambda_1^{(i)} \neq \lambda_1^{(i_0)}$, and upon comparing (148) and (154), we see that

$$\eta_i \neq \eta_{i_0}. \tag{155}$$

Let $\mathcal{I}_L^1 = \{i \in \mathcal{I}_L \setminus \mathcal{I}_L^0 : \eta_i > \eta_{i_0}\}$ and $\mathcal{I}_L^2 = \{i \in \mathcal{I}_L \setminus \mathcal{I}_L^0 : \eta_i < \eta_{i_0}\}$. Then we can see from (155) that

$$\mathcal{I}_L^0 \cup \mathcal{I}_L^1 \cup \mathcal{I}_L^2 = \mathcal{I}_L. \tag{156}$$

For $i \in \mathcal{I}_L^1$, we have $\lambda_1^{(i)} < \lambda_1^{(i_0)}$, which is equivalent to

$$\lambda_1^{(i)} < \frac{1}{\eta_{i_0}} \sum_{j=2}^L \lambda_j^{(i_0)}.$$

Thus, by Lemma 7 (i), we have

$$f_{\eta_{i_0}+1}(\lambda^{(i)}) = \frac{1}{\eta_{i_0}+1} \sum_{j=1}^L \lambda_j^{(i)} < \frac{1}{\eta_{i_0}+1} \sum_{j=1}^L \lambda_j^{(i_0)}. \quad (157)$$

On the other hand, for $i \in \mathcal{I}_L^2$, we have $\lambda_1^{(i)} > \lambda_1^{(i_0)}$, which is equivalent to

$$\lambda_1^{(i)} > \frac{1}{\eta_{i_0}} \sum_{j=2}^L \lambda_j^{(i)}.$$

Thus, by Lemma 7 (ii), we have

$$f_{\eta_{i_0}+1}(\lambda^{(i)}) = f_{\eta_{i_0}}(\lambda_{L-1}^{(j_0)}),$$

which by (150) implies that

$$f_{\eta_{i_0}+1}(\lambda^{(i)}) = \frac{1}{\eta_{i_0}+1} \sum_{j=1}^L \lambda_j^{(i_0)}. \quad (158)$$

From (138) and (156), we have

$$\begin{aligned} f_{\eta_{i_0}+1}(\lambda^{(i_0)}) &= \sum_{i=1}^{S_L} c_i f_{\eta_{i_0}+1}(\lambda^{(i)}) \\ &= \sum_{i \in \mathcal{I}_L} c_i f_{\eta_{i_0}+1}(\lambda^{(i)}) \\ &= \sum_{i \in \mathcal{I}_L^0} c_i f_{\eta_{i_0}+1}(\lambda^{(i)}) + \sum_{i \in \mathcal{I}_L^1} c_i f_{\eta_{i_0}+1}(\lambda^{(i)}) \\ &\quad + \sum_{i \in \mathcal{I}_L^2} c_i f_{\eta_{i_0}+1}(\lambda^{(i)}). \end{aligned} \quad (159)$$

Comparing (150) for $f_{\eta_{i_0}+1}(\lambda^{(i_0)})$ and (153), (157), and (158) for $f_{\eta_{i_0}+1}(\lambda^{(i)})$, we see that both \mathcal{I}_L^0 and \mathcal{I}_L^1 must be empty in order for the equality in (159) to hold, and hence

$$\mathcal{I}_L = \mathcal{I}_L^2. \quad (160)$$

For any $i \in \mathcal{I}_L^2$, since $\eta_i < \eta_{i_0}$ and $\eta_i \geq 1$, we see that $\eta_{i_0} \geq 2$. Thus from (148), we have

$$\begin{aligned} \frac{1}{\eta_{i_0}} \sum_{j=1}^L \lambda_j^{(i_0)} &= \frac{1}{\eta_{i_0}} \left[\frac{1}{\eta_{i_0}} \sum_{j=2}^L \lambda_j^{(i_0)} + \sum_{j=2}^L \lambda_j^{(i_0)} \right] \\ &= \frac{1}{\eta_{i_0}} \left(\frac{1}{\eta_{i_0}} + 1 \right) \sum_{j=2}^L \lambda_j^{(i_0)} \\ &< \frac{1}{\eta_{i_0}-1} \sum_{j=2}^L \lambda_j^{(i_0)}. \end{aligned}$$

Then by Lemma 7 (i), (148) implies that

$$f_{\eta_{i_0}}(\lambda^{(i_0)}) = \frac{1}{\eta_{i_0}} \sum_{j=1}^L \lambda_j^{(i_0)} < \frac{1}{\eta_{i_0}-1} \sum_{j=2}^L \lambda_j^{(i_0)}. \quad (161)$$

Since $\lambda^{(i_0)}$ is ordered, by (148), we have

$$\lambda_2^{(i_0)} \leq \lambda_1^{(i_0)} = \frac{1}{\eta_{i_0}} \sum_{j=2}^L \lambda_j^{(i_0)},$$

which implies that

$$\lambda_2^{(i_0)} \leq \frac{1}{\eta_{i_0}-1} \sum_{j=3}^L \lambda_j^{(i_0)}.$$

Then by Lemma 7 (i), we have

$$f_{\eta_{i_0}-1}(\lambda_{L-1}^{(j_0)}) = \frac{1}{\eta_{i_0}-1} \sum_{j=2}^L \lambda_j^{(i_0)}. \quad (162)$$

It follows from (161) and (162) that

$$f_{\eta_{i_0}}(\lambda^{(i_0)}) < f_{\eta_{i_0}-1}(\lambda_{L-1}^{(j_0)}). \quad (163)$$

For $i \in \mathcal{I}_L^2$, we have $\eta_{i_0} \geq \eta_i + 1$. Then by Lemma 7 (ii), (154) implies that

$$f_{\eta_{i_0}}(\lambda^{(i)}) = f_{\eta_{i_0}-1}(\lambda_{L-1}^{(j_0)}). \quad (164)$$

Following (138), we have

$$\begin{aligned} &(\lambda_2^{(i_0)}, \lambda_3^{(i_0)}, \dots, \lambda_L^{(i_0)}) \\ &= \sum_{i \in \mathcal{I}_L} c_i (\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}) \\ &= \sum_{i \in \mathcal{I}_L^2} c_i (\lambda_2^{(i)}, \lambda_3^{(i)}, \dots, \lambda_L^{(i)}) \end{aligned} \quad (165)$$

$$= \left(\sum_{i \in \mathcal{I}_L^2} c_i \right) (\lambda_2^{(i_0)}, \lambda_3^{(i_0)}, \dots, \lambda_L^{(i_0)}), \quad (166)$$

where (165) follows from (160) and (166) follows from the assumption in (151). Thus we have

$$\sum_{i \in \mathcal{I}_L^2} c_i = 1. \quad (167)$$

Then from (164) and (167), we see that

$$\sum_{i \in \mathcal{I}_L^2} c_i f_{\eta_{i_0}}(\lambda^{(i)}) = \left(\sum_{i \in \mathcal{I}_L^2} c_i \right) f_{\eta_{i_0}-1}(\lambda_{L-1}^{(j_0)}) = f_{\eta_{i_0}-1}(\lambda_{L-1}^{(j_0)}),$$

and it follows from (163) that

$$\sum_{i \in \mathcal{I}_L^2} c_i f_{\eta_{i_0}}(\lambda^{(i)}) > f_{\eta_{i_0}}(\lambda^{(i_0)}).$$

This is a contradiction to (138). Therefore, the assumption in (151) is false and the proposition is proved.

APPENDIX H PROOF OF LEMMA 14

For any permutation ω on $\{1, 2, \dots, L\}$ and any $\lambda \in \mathbb{R}_+^L$, recall from the beginning of Section III that

$$\omega(\lambda) = (\lambda_{\omega(1)}, \lambda_{\omega(2)}, \dots, \lambda_{\omega(L)}).$$

Then for the ordered permutation π , we have $\pi(\lambda) = (\lambda_{\pi(1)}, \lambda_{\pi(2)}, \dots, \lambda_{\pi(L)})$.

If $\lambda = \pi(\lambda)$, the lemma is immediate. Otherwise, let $\omega_0(i) = i$ for all $i \in \mathcal{L}$ so that $\omega_0(\lambda) = \lambda$. Set $t = 1$ and we sort λ in descending order by iteration as follows:

- (i) Let $i_t = \min\{i \in \mathcal{L} : \omega_{t-1}(i) \neq \pi(i)\}$. Let k_t, j_t be any indexes in \mathcal{L} such that

$$\pi(k_t) = \omega_{t-1}(i_t) \quad (168)$$

and

$$\omega_{t-1}(j_t) = \pi(i_t). \quad (169)$$

It is easy to check that $k_t > i_t$ and $j_t > i_t$, which implies

$$\lambda_{\pi(i_t)} - \lambda_{\pi(k_t)} \geq 0 \quad (170)$$

and

$$R_{j_t} - R_{i_t} \geq 0. \quad (171)$$

Let $\omega_t(\lambda) = (\lambda_{\omega_t(1)}, \lambda_{\omega_t(2)}, \dots, \lambda_{\omega_t(L)})$ be a permutation of $\omega_{t-1}(\lambda)$ where we switch $\lambda_{\omega_{t-1}(j_t)}$ and $\lambda_{\omega_{t-1}(i_t)}$, i.e.,

$$\omega_t(i) = \begin{cases} \pi(i_t), & \text{if } i = i_t \\ \pi(k_t), & \text{if } i = j_t \\ \omega_{t-1}(i), & \text{otherwise.} \end{cases} \quad (172)$$

Then we have

$$\begin{aligned} & \sum_{i=1}^L \lambda_{\omega_{t-1}(i)} R_i - \sum_{i=1}^L \lambda_{\omega_t(i)} R_i \\ &= (\lambda_{\omega_{t-1}(i_t)} R_{i_t} + \lambda_{\omega_{t-1}(j_t)} R_{j_t}) \\ & \quad - (\lambda_{\omega_t(i_t)} R_{i_t} + \lambda_{\omega_t(j_t)} R_{j_t}) \\ &= (\lambda_{\pi(k_t)} R_{i_t} + \lambda_{\pi(i_t)} R_{j_t}) - (\lambda_{\pi(i_t)} R_{i_t} + \lambda_{\pi(k_t)} R_{j_t}) \\ &= R_{i_t} (\lambda_{\pi(k_t)} - \lambda_{\pi(i_t)}) + R_{j_t} (\lambda_{\pi(i_t)} - \lambda_{\pi(k_t)}) \\ &= (\lambda_{\pi(i_t)} - \lambda_{\pi(k_t)}) (R_{j_t} - R_{i_t}) \\ &\geq 0, \end{aligned}$$

where the second equality follows from (168), (169) and (172), and the inequality follows from (170) and (171).

- (ii) If $\omega_t(\lambda) = \pi(\lambda)$, return $T = t$ and stop. Otherwise, let $t = t + 1$ and go back to step (i).

At the end of the iteration, $\omega_T(\lambda)$ is sorted in the same order as $\pi(\lambda)$, and we have

$$\begin{aligned} & \sum_{i=1}^L \lambda_i R_i - \sum_{i=1}^L \lambda_{\pi(i)} R_i \\ &= \sum_{i=1}^L \lambda_{\omega_0(i)} R_i - \sum_{i=1}^L \lambda_{\omega_T(i)} R_i \\ &= \sum_{t=1}^T \left(\sum_{i=1}^L \lambda_{\omega_{t-1}(i)} R_i - \sum_{i=1}^L \lambda_{\omega_t(i)} R_i \right) \\ &\geq 0. \end{aligned}$$

This proves the lemma.

REFERENCES

- [1] T. Guo and R. W. Yeung, "The explicit coding rate region of symmetric multilevel diversity coding," in *2018 IEEE International Symposium on Information Theory (ISIT)*, (Vail, CO, USA), June 2018.
- [2] R. W. Yeung, "Multilevel diversity coding with distortion," *IEEE Trans. Inf. Theory*, vol. 41, pp. 412–422, Mar. 1995.
- [3] J. R. Roche, R. W. Yeung, and K. P. Hau, "Multilevel diversity coding with symmetrical connectivity," in *IEEE Int. Symp. Inf. Theory*, (Whistler, BC, Canada), Sep. 1995.
- [4] J. R. Roche, R. W. Yeung, and K. P. Hau, "Symmetrical multilevel diversity coding," *IEEE Trans. Inf. Theory*, vol. 43, pp. 1059–1064, May 1997.
- [5] R. W. Yeung and Z. Zhang, "On symmetrical multilevel diversity coding," *IEEE Trans. Inf. Theory*, vol. 45, pp. 609–621, Mar. 1999.
- [6] J. R. Roche, *distributed information storage*. PhD thesis, Stanford University, Stanford, CA, Mar. 1992.
- [7] J. R. Roche, A. Dembo, and A. Nobel, "Distributed information storage," in *IEEE Int. Symp. Inf. Theory*, (Kobe, Japan), Jun. 1988.
- [8] A. Shamir, "How to share a secret," *Commun. ACM*, vol. 22, pp. 612–613, Nov. 1979.
- [9] G. R. Blakley, "Safeguarding cryptographic keys," in *Proceedings of the AFIPS National Computer Conference*, pp. 313–317, Feb. 1979.
- [10] G. R. Blakley and C. Meadows, "Security of ramp type," *Proc. Adv. Cryptology*, vol. LNCS-196, pp. 242–269, 1985.
- [11] E. Ayanoglu, C.-L. I, R. D. Gitlin, and J. E. Mazo, "Diversity coding for transparent self-healing and fault-tolerant communication networks," *IEEE Trans. Commun.*, vol. 41, pp. 1677–1686, Nov. 1993.
- [12] T. T. Lee and S. C. Liew, "Parallel communications for atm network control and management," in *GLOBECOM '93*, (Houston, TX, USA), pp. 442–446, Nov. 1993.
- [13] R. W. Yeung and Z. Zhang, "Distributed source coding for satellite communications," *IEEE Trans. Inf. Theory*, vol. 45, pp. 1111–1120, May 1999.
- [14] R. Ahlswede, N. Cai, S. Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, pp. 1204–1216, Jul. 2000.
- [15] L. Song, R. W. Yeung, and N. Cai, "Zero-error network coding for acyclic networks," *IEEE Trans. Inf. Theory*, vol. 49, pp. 3129–3139, Dec. 2003.
- [16] A. El Gamal and T. M. Cover, "Achievable rates for multiple descriptions," *IEEE Trans. Inf. Theory*, vol. IT-28, pp. 851–857, Nov. 1982.
- [17] Z. Zhang and T. Berger, "New results in binary multiple descriptions," *IEEE Trans. Inf. Theory*, vol. 33, pp. 502–521, Jul. 1987.
- [18] R. Venkataramani, G. Kramer, and V. K. Goyal, "Multiple description coding with many channels," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2106–2114, Sep. 2003.
- [19] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge: Cambridge University Press, 2011.
- [20] C. Tian, S. Mohajer, and S. N. Diggavi, "Approximating the Gaussian multiple description rate region under symmetric distortion constraints," *IEEE Trans. Inf. Theory*, vol. 55, pp. 3869–3891, Aug. 2009.
- [21] T. Guo and R. W. Yeung, "Extended multiple descriptions with reconstruction consistency constraints," in *GLOBECOM*, (Washington DC, USA), Dec. 2016.
- [22] A. Albanese, J. Blömer, J. Edmonds, M. Luby, and M. Sudan, "Priority encoding transmission," *IEEE Trans. Inf. Theory*, vol. 42, pp. 1737–1744, Nov. 1996.
- [23] A. Balasubramanian, H. D. Ly, S. Li, T. Liu, and S. L. Miller, "Secure symmetrical multilevel diversity coding," *IEEE Trans. Inf. Theory*, vol. 59, pp. 3572–3581, Jun. 2013.
- [24] J. Jiang, N. Marukala, and T. Liu, "Symmetrical multilevel diversity coding and subset entropy inequalities," *IEEE Trans. Inf. Theory*, vol. 60, pp. 84–103, Jan. 2014.
- [25] Z. Xiao, J. Chen, Y. Li, and J. Wang, "Distributed multilevel diversity coding," *IEEE Trans. Inf. Theory*, vol. 61, pp. 6368–6384, Nov. 2015.
- [26] C. Tian and T. Liu, "Multilevel diversity coding with regeneration," *IEEE Trans. Inf. Theory*, vol. 62, pp. 4833–4847, Sep. 2016.
- [27] S. Mohajer, C. Tian, and S. N. Diggavi, "Asymmetric multilevel diversity coding and asymmetric gaussian multiple descriptions," *IEEE Trans. Inf. Theory*, vol. 56, pp. 4367–4387, Sep. 2010.
- [28] C. Li and J. W. S. Weber, "Multilevel diversity coding systems: Rate regions, codes, computation, & forbidden minors," *IEEE Trans. Inf. Theory*, vol. 63, pp. 230–251, Nov. 2016.
- [29] A. Schrijver, *Theory of linear and integer programming*. Wiley-Interscience series in discrete mathematics and optimization, Wiley, 1998.
- [30] O. L. Mangasarian, "Pseudo-convex functions," *Journal of the society for industrial and applied math., series a control*, vol. 3, no. 2, pp. 281–290, 1965.
- [31] J. M. Steel, *Cauchy-Schwarz master class - an introduction to the art of mathematical inequalities*. Cambridge University Press, 2004.
- [32] T. S. Han, "Nonnegative entropy measures of multivariate symmetric correlations," *Inf. Control*, vol. 36, pp. 133–156, Feb. 1978.
- [33] M. Madiman and P. Tetali, "Information inequalities for joint distributions, with interpretations and applications," *IEEE Trans. Inf. Theory*, vol. 56, pp. 2699–2713, Jun. 2010.

APPENDIX I

TABLES OF NON-REDUNDANT λ

For $L = 1, 2, \dots, 5$, the vectors $\lambda \in \mathcal{G}_L^0$ and the corresponding $f_\alpha(\lambda)$ are listed in the following tables. The parameter θ is the integer such that $\lambda_1 = \frac{1}{\theta} \sum_{i=2}^L \lambda_i$.

TABLE I: non-redundant constraint for $L = 1$.

λ	$f_1(\lambda)$
1	1

TABLE II: non-redundant constraints for $L = 2$.

suffix	λ	$f_1(\lambda)$	$f_2(\lambda)$	θ
-	(1, 0)	1	0	0
(1)	(1, 1)	2	1	1

TABLE III: non-redundant constraints for $L = 3$.

suffix	λ	$f_1(\lambda)$	$f_2(\lambda)$	$f_3(\lambda)$	θ
-	(1, 0, 0)	1	0	0	0
(1, 0)	(1, 1, 0)	2	1	0	1
(1, 1)	(1, 1, 1)	3	$\frac{3}{2}$	1	2
	(2, 1, 1)	4	2	1	1

TABLE IV: non-redundant constraints for $L = 4$.

suffix	λ	$f_1(\lambda)$	$f_2(\lambda)$	$f_3(\lambda)$	$f_4(\lambda)$	θ
-	(1, 0, 0, 0)	1	0	0	0	0
(1, 0, 0)	(1, 1, 0, 0)	2	1	0	0	1
(1, 1, 0)	(1, 1, 1, 0)	3	$\frac{3}{2}$	1	0	2
	(2, 1, 1, 0)	4	2	1	0	1
(1, 1, 1)	(1, 1, 1, 1)	4	2	$\frac{4}{3}$	1	3
	($\frac{3}{2}$, 1, 1, 1)	$\frac{9}{2}$	$\frac{9}{4}$	$\frac{3}{2}$	1	2
	(3, 1, 1, 1)	6	3	$\frac{3}{2}$	1	1
(2, 1, 1)	(2, 2, 1, 1)	6	3	2	1	2
	(4, 2, 1, 1)	8	4	2	1	1

TABLE V: non-redundant constraints for $L = 5$.

suffix	λ	$f_1(\lambda)$	$f_2(\lambda)$	$f_3(\lambda)$	$f_4(\lambda)$	$f_5(\lambda)$	θ
-	(1, 0, 0, 0, 0)	1	0	0	0	0	0
(1, 0, 0, 0)	(1, 1, 0, 0, 0)	2	1	0	0	0	1
(1, 1, 0, 0)	(1, 1, 1, 0, 0)	3	$\frac{3}{2}$	1	0	0	2
	(2, 1, 1, 0, 0)	4	2	1	0	0	1
(1, 1, 1, 0)	(1, 1, 1, 1, 0)	4	2	$\frac{4}{3}$	1	0	3
	($\frac{3}{2}$, 1, 1, 1, 0)	$\frac{9}{2}$	$\frac{9}{4}$	$\frac{3}{2}$	1	0	2
	(3, 1, 1, 1, 0)	6	3	$\frac{3}{2}$	1	0	1
(2, 1, 1, 0)	(2, 2, 1, 1, 0)	6	3	2	1	0	2
	(4, 2, 1, 1, 0)	8	4	2	1	0	1
(1, 1, 1, 1)	(1, 1, 1, 1, 1)	5	$\frac{5}{2}$	$\frac{5}{3}$	$\frac{5}{4}$	1	4
	($\frac{4}{3}$, 1, 1, 1, 1)	$\frac{16}{3}$	$\frac{8}{3}$	$\frac{16}{9}$	$\frac{4}{3}$	1	3
	(2, 1, 1, 1, 1)	6	3	2	$\frac{4}{3}$	1	2
	(4, 1, 1, 1, 1)	8	4	2	$\frac{4}{3}$	1	1
(3/2, 1, 1, 1)	($\frac{3}{2}$, $\frac{3}{2}$, 1, 1, 1)	6	3	2	$\frac{3}{2}$	1	3
	($\frac{9}{4}$, $\frac{3}{2}$, 1, 1, 1)	$\frac{27}{4}$	$\frac{27}{8}$	$\frac{9}{4}$	$\frac{3}{2}$	1	2
	($\frac{5}{2}$, $\frac{3}{2}$, 1, 1, 1)	9	$\frac{9}{2}$	$\frac{9}{4}$	$\frac{3}{2}$	1	1
(3, 1, 1, 1)	(3, 3, 1, 1, 1)	9	$\frac{9}{2}$	3	$\frac{3}{2}$	1	2
	(6, 3, 1, 1, 1)	12	6	3	$\frac{3}{2}$	1	1
(2, 2, 1, 1)	(2, 2, 2, 1, 1)	8	4	$\frac{8}{3}$	2	1	3
	(3, 2, 2, 1, 1)	9	$\frac{9}{2}$	3	2	1	2
	(6, 2, 2, 1, 1)	12	6	3	2	1	1
(4, 2, 1, 1)	(4, 4, 2, 1, 1)	12	6	4	2	1	2
	(8, 4, 2, 1, 1)	16	8	4	2	1	1