


Article

# Approximation Hierarchies for the Copositive Tensor Cone and Their Application to the Polynomial Optimization over the Simplex

Muhammad Faisal Iqbal<sup>1</sup> and Faizan Ahmed<sup>2,\*</sup> 

<sup>1</sup> Department of Applied Mathematics and Statistics, Institute of Space Technology, Islamabad 44000, Pakistan; faisal.iqbal@aiou.edu.pk

<sup>2</sup> Formal Methods and Tools Group, University of Twente, 7522 NB Enschede, The Netherlands

\* Correspondence: faizan.ahmed@utwente.nl

**Abstract:** In this paper, we discuss the cone of copositive tensors and its approximation. We describe some basic properties of copositive tensors and positive semidefinite tensors. Specifically, we show that a non-positive tensor (or  $Z$ -tensor) is copositive if and only if it is positive semidefinite. We also describe cone hierarchies that approximate the copositive cone. These hierarchies are based on the sum of squares conditions and the non-negativity of polynomial coefficients. We provide a compact representation for the approximation based on the non-negativity of polynomial coefficients. As an immediate consequence of this representation, we show that the approximation based on the non-negativity of polynomial coefficients is polyhedral. Furthermore, these hierarchies are used to provide approximation results for optimizing a (homogeneous) polynomial over the simplex.

**Keywords:** copositive tensor; positive semidefinite tensor; sum of squares; approximation hierarchies; polynomial optimization; simplex

**MSC:** 15A69; 15B48; 52A27; 11E25; 90C23



**Citation:** Iqbal, M.F.; Ahmed, F. Approximation Hierarchies for the Copositive Tensor Cone and Their Application to the Polynomial Optimization over the Simplex. *Mathematics* **2022**, *10*, 1683. <https://doi.org/10.3390/math10101683>

Academic Editor: Ferenc Hartung

Received: 1 April 2022

Accepted: 10 May 2022

Published: 14 May 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Multidimensional arrays or tensors arise naturally as an extension of matrices. They occur in applications where one needs to represent multidimensional data such as in signal processing [1–3], machine learning [2,4,5], material science [6], and speech recognition [7]. For example, any homogeneous polynomial of degree  $d$  in  $n$ -variables is associated with a symmetric tensor of order  $d$  and dimension  $n$ . In polynomial optimization and, likewise, in control theory, checking the non-negativity of a polynomial is a fundamental problem [8,9]. Nonnegativity of a polynomial over the real space or its non-negative orthant results in positive semidefinite and copositive tensors, respectively.

The copositive and completely positive cones of matrices, which are tensors of order two, are very well explored (see, e.g., [10–12] and also [13] for a list of open problems). Therefore, it seems natural to study similar results for copositive tensors. The generalization from matrix to tensor is not trivial since a higher dimension usually destroys the nice structure present at a lower dimension.

The current research on copositive tensor is focused on describing properties that can be generalized from the quadratic case to higher dimensional case. The area is not very well explored. Similar to its matrix analog a characterization of copositive tensors using eigenvectors of principal sub-tensors is described in [14]. Moreover, Qi and co-authors have discussed several basic properties of copositive tensors in a series of papers (see e.g., [15] and the book [16]).

The set of copositive tensors forms the copositive cone. The copositive cone is used to reformulate hard combinatorial optimization problem as a linear conic optimization

problem. Reformulation of polynomial optimization as copositive program does not reduce its complexity. However, the complexity is packaged in the copositivity constraint, which is known to be NP-hard. To approximate copositive cone several tractable approximation hierarchies are developed. These approximation hierarchies are based on sum-of-squares conditions, non-negativity of polynomial coefficients, simplicial partition, or rational gridding of the simplex. For instance, Parrilo [17] had provided a hierarchy of linear and semi-definite inner approximations for copositive cone (see also [18]). Moreover, Bomze and de Klerk [19] developed a hierarchy for copositive matrices based on non-negativity of polynomial coefficients.

In this paper, we describe approximation hierarchies for the cone of copositive tensors. We extend the hierarchies presented by Parrilo [17] (represented by  $\mathcal{K}_{n,d}^r$ ) an Bomze and de Klerk [19] (represented by  $\mathcal{C}_{n,d}^r$ ) for higher dimension. Earlier work was focused on extending polynomial approximation schemes from lower to higher dimension. For example, Bomze and de Klerk [19] focused on improving the approximation result as presented by Nesterov [20]. They first presented a polyhedral representation of  $\mathcal{C}_{n,2}^r$  and then used this representation to derive an approximation scheme for polynomial optimization over the simplex. The approximation scheme was extended to a higher-degree fixed-degree polynomial by De Klerk, Laurent and Parrilo [21]. However, they used rational gridding of the simplex to derive this polynomial time approximation scheme. The results was further refined using the Bernstein approximation in [22]. Since these approximation schemes relied on rational gridding, they have therefore also presented an error analysis based on multivariate hyper-geometric distribution (see [23] and also [24] for a note on the convergence rate).

Furthermore, we develop a compact representation of  $\mathcal{C}_{n,d}^{(r)}$  with an aim to extend the initial results as presented in Bomze and de Klerk [19]. To the best of our knowledge, the closest attempt in this direction is by [25]. However, their results are restricted to the tensor of order four. Secondly, the results rely heavily on breaking the tensor of dimension four into tensors of lower dimensions. The representation process was tedious and had no obvious generalization to higher dimension cases. Our representation is general and holds for all dimensions. These hierarchies can be used to develop polynomial time approximation schemes for polynomial optimization over the simplex. Moreover, we have provided results in this direction (see Section 5).

The main contributions of this paper are as follows. (a) To discuss basic properties of the copositive tensor cone. Moreover, we show that every Z-tensor is copositive if and only if it is positive semidefinite. (b) To describe approximation hierarchies for the copositive cone of tensors based on sum of square decomposition and non-negativity of the polynomial coefficients. Moreover, we present the polyhedral representation of the approximation hierarchical cone  $\mathcal{C}_{n,d}^{(r)}$  based on non-negative coefficients. The representation is compact and has not appeared in the literature. (c) application of approximation hierarchies to the polynomial optimization over the simplex.

The article is arranged as follows. Section 2 is comprised of the basic definitions and notations. In Section 3, we define tensor cones and related results. In Section 4, we present approximation hierarchies for copositive cone of tensors. We discuss special cases and characterization of these hierarchies. Section 5 provides approximation results based on these hierarchies. In Section 6, we provide a conclusion and future directions.

## 2. Preliminaries

Throughout this article, the  $n$ -dimensional Euclidean space and its non-negative orthant are denoted by  $\mathbb{R}^n$  and  $\mathbb{R}_+^n$ , respectively. The set of natural numbers is denoted by  $\mathbb{N}$  and the set of first  $n$  natural number is denoted by  $[1 : n]$ . The set of whole numbers is denoted by  $\mathbb{N}_0 = \{0, 1, 2, \dots\}$  while the set of first  $n + 1$  whole numbers is denoted by  $[0 : n]$ . For any  $\alpha \in [0 : d]^n$  we define  $|\alpha| := \sum_{i=1}^n \alpha_i$ . We also define the index set

$$\mathbb{I}^n(d) = \{\alpha \in [0 : d]^n : |\alpha| = d\}. \quad (1)$$

The cardinality of  $\mathbb{I}^n(d)$  is  $|\mathbb{I}^n(d)| = \binom{n+d-1}{d}$  (see, e.g., [22]).

A *tensor* is a multi-dimensional array of real numbers. Specifically, an  $n$ -dimensional  $d$ th order tensor is given by  $\mathcal{A} = (a_{i_1 i_2 \dots i_d})_{1 \leq i_1, \dots, i_d \leq n}$ . Moreover, a tensor  $\mathcal{A}$  is said to be symmetric if

$$a_{i_1 i_2 i_3 \dots i_d} = a_{i_{\sigma(1)} i_{\sigma(2)} i_{\sigma(3)} \dots i_{\sigma(d)}} \text{ for all permutations } \sigma \text{ on } \{1, 2, \dots, d\}.$$

The set of all symmetric tensors is denoted by  $\mathcal{S}_{n,d}$ . For brevity of notation, if some index  $i_j$  of an element  $a_{i_1 i_2 \dots i_d} \in \mathcal{A}$  is repeated  $k$ -times, we write it as  $(i_j)^k$  i.e.,

$$a_{i_1 i_2 \dots (i_j)^k \dots i_{d-k}} = a_{i_1 i_2 \dots \underbrace{i_j i_j \dots i_j}_{k\text{-times}} \dots i_{d-k}}.$$

Using above notation the  $i^{\text{th}}$  diagonal element of the tensor  $\mathcal{A} \in \mathcal{S}_{n,d}$  is denoted by  $a_{(i)^d}$ . The inner product of two tensors  $\mathcal{A}, \mathcal{B} \in \mathcal{S}_{n,d}$  is defined as follows

$$\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{i_1, i_2, \dots, i_d=1}^n a_{i_1, i_2, \dots, i_d} b_{i_1, i_2, \dots, i_d}. \tag{2}$$

For  $\alpha \in \mathbb{I}^n(d)$  and  $x \in \mathbb{R}^n$ ,  $x^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n} = \prod_{i=1}^n x_i^{\alpha_i}$  represent a monomial. The maximum degree among all the monomials of  $p(x)$  is called the degree of a polynomial. A polynomial whose all monomials have the same degree is termed as a ‘form’ or a ‘homogeneous polynomial’. Moreover, for  $x \in \mathbb{R}^n$ ,  $x^d$  denotes a symmetric tensor of dimension  $n$  and degree  $d$  and is given below,

$$x^d = \underbrace{x \odot x \odot \dots \odot x}_{d\text{-times}} = (x_{i_1} x_{i_2} \dots x_{i_d})_{1 \leq i_1, \dots, i_d \leq n} \in \mathcal{S}_{n,d}.$$

Notice that, the entries of tensor  $x^d$  are the monomials of degree  $d$  in  $n$ -variables. Thus for any symmetric tensor  $\mathcal{A} \in \mathcal{S}_{n,d}$  its associated form can be written as

$$h_{\mathcal{A}}(x) := \mathcal{A}x^d = \sum_{i_1, \dots, i_d=1}^n a_{i_1, \dots, i_d} x_{i_1} \dots x_{i_d} = \sum_{\alpha \in \mathbb{I}^n(d)} c(\alpha) \mathcal{A}_\alpha x^\alpha. \tag{3}$$

where  $\mathcal{A}_\alpha = a_{i_1, \dots, i_d}$  denote the coefficient of monomial  $x^\alpha$  in  $h_{\mathcal{A}}(x)$  and

$$c(\alpha) = \begin{cases} \frac{|\alpha|!}{\prod_i (\alpha_i!)} & \text{if } \alpha \in \mathbb{N}_0^n \\ 0 & \text{if } \alpha \in \mathbb{R}^n \setminus \mathbb{N}_0^n \end{cases} \tag{4}$$

and  $\alpha_i!$  denotes the factorial of  $\alpha_i$ .

For  $\mathcal{A} \in \mathcal{S}_{n,d}$  a subset  $K \subseteq \mathcal{S}_{n,d}$  is said to be a cone if for each tensor  $\mathcal{A} \in K$  the scalar product  $\lambda \mathcal{A} \in K$  for all  $\lambda \geq 0$ . Moreover, the cone  $K$  is said to be a convex cone if for  $\mathcal{A}, \mathcal{B} \in K$  and for non-negative scalars  $\lambda_1, \lambda_2 \in \mathbb{R}$  we have,  $\lambda_1 \mathcal{A} + \lambda_2 \mathcal{B} \in K$ . The dual  $K^*$  of the cone  $K$  is defined as

$$K^* := \{ \mathcal{U} \in \mathcal{S}_{n,d} : \langle \mathcal{U}, \mathcal{V} \rangle \geq 0, \text{ for all } \mathcal{V} \in K \}. \tag{5}$$

A convex cone  $K$  is said to be pointed if  $\{K\} \cap \{-K\} = \{0\}$ , and  $K$  is said to be solid if its interior is nonempty. A convex cone which is closed, pointed, and solid is termed as *proper cone*. A convex cone  $K$  is said to be a *polyhedral cone* if it is finitely generated.

The cone of entry-wise non-negative tensors is denoted by  $\mathcal{N}_{n,d}$ . Finally, a tensor with non-positive off-diagonal entries is termed as *Z-tensor* (it is also called *essentially non-positive*

tensor see e.g., [26]). Finally, the standard simplex  $\Delta_n$  in  $n$ -dimensional Euclidean space  $\mathbb{R}^n$  is

$$\Delta_n := \{x \in \mathbb{R}^n : e^T x = 1\} \tag{6}$$

where  $e^T = [1, 1, \dots, 1] \in \mathbb{R}^n$ .

### 3. Positive Semidefinite Tensors, Copositive Tensors and Their Duals

The cone of  $d$ th-order (with  $d$  even)  $n$ -dimensional, positive semidefinite tensors is denoted by  $\mathcal{S}_{n,d}^+$  and is given below

$$\mathcal{S}_{n,d}^+ := \{A \in \mathcal{S}_{n,d} : h_A(x) = Ax^d \geq 0, \forall x \in \mathbb{R}^n\}. \tag{7}$$

For a tensor  $A \in \mathcal{S}_{n,d}^+$ , the polynomial  $h_A(x)$  is called PSD polynomial. The dual of PSD cone  $\mathcal{S}_{n,d}^+$  is the cone of completely positive semidefinite tensors, denoted by  $\mathcal{S}_{n,d}^{+*}$ , which is defined as .

$$\mathcal{S}_{n,d}^{+*} := \{X \in \mathcal{S}_{n,d} : \langle A, X \rangle \geq 0, \forall A \in \mathcal{S}_{n,d}^+\} = \left\{ \sum_{k=1}^N (x_k)^d : x_k \in \mathbb{R}^n, N \in \mathbb{N} \right\}. \tag{8}$$

For  $d = 2$ , PSD cone is self dual i.e.,  $\mathcal{S}_{n,2}^+ = \mathcal{S}_{n,2}^{+*}$  (cf. [16]). However, for  $d \geq 4$  we have  $\mathcal{S}_{n,d}^+ \neq \mathcal{S}_{n,d}^{+*}$  in general (see [27], Example 4.5).

It is well known that a function is convex if and only if its Hessian matrix is positive semidefinite (see, e.g., [28], Theorem 4.5). Therefore the convexity of homogeneous polynomial defined in (3) amounts to checking if  $\nabla^2 h_A(x) \in \mathcal{S}_{n,2}^+$ . It has been shown that if a polynomial  $h_A(x)$  is convex then its associated tensor  $A$  is positive semidefinite (see [27], Proposition 5.10). However, the converse need not to be true in general (see [27] (Example 5.11) and [29]).

A tensor  $A \in \mathcal{S}_{n,d}$  is said to be *copositive* if  $h_A(x) = Ax^d \geq 0$  for all  $x \in \mathbb{R}_+^n$ , whereas a tensor is called *strictly copositive* if  $h_A(x) = Ax^d > 0$  for all  $x \in \mathbb{R}_+^n \setminus \{0\}$ . The set of  $n$ -dimensional,  $d^{\text{th}}$ -order copositive tensors defines a cone given below,

$$\mathcal{C}_{n,d} := \{A \in \mathcal{S}_{n,d} : Ax^d \geq 0, \forall x \in \mathbb{R}_+^n\}. \tag{9}$$

It is well known that, a tensor  $A \in \text{int}(\mathcal{C}_{n,d})$  if and only if it is strictly copositive, where  $\text{int}(\mathcal{C}_{n,d})$  denotes the interior of the set  $\mathcal{C}_{n,d}$ . The dual of copositive cone  $\mathcal{C}_{n,d}$  is completely positive cone denoted by  $\mathcal{C}_{n,d}^*$  (see e.g., [16] (Theorem 6.9), [30]), which is defined below,

$$\mathcal{C}_{n,d}^* = \{x^d \in \mathcal{S}_{n,d} : \langle A, x^d \rangle \geq 0, \forall A \in \mathcal{C}_{n,d} \text{ and } x \in \mathbb{R}_+^n\} \tag{10}$$

$$= \left\{ \sum_{k=1}^N (x_k)^d : x_k \in \mathbb{R}_+^n, N \in \mathbb{N} \right\}. \tag{11}$$

It is clear that if a tensor is positive semidefinite then it is copositive also. A copositive tensor need not to be positive semidefinite (cf. Example 1). The question arises in which case a copositive tensor is also positive semidefinite. In the following theorem, we describe one such case (for  $d = 2$  see [31], Lemma 2.6).

**Theorem 1.** *Let  $A \in \mathcal{S}_{n,d}$  be a Z-tensor, then  $A$  is copositive if and only if it is positive semidefinite, where  $d$  is even.*

**Proof.** Let us take  $A \in \mathcal{C}_{n,d}$  and since  $A$  is Z-tensor, we can write  $A = \mathcal{P} - \mathcal{Q}$  where  $\mathcal{P}, \mathcal{Q} \in \mathcal{N}_{n,d}$  such that,

$$p_{i_1 \dots i_d} := \begin{cases} a_{(i)}^d & \text{if } i_1 = \dots = i_d = i \\ 0 & \text{otherwise} \end{cases} \quad q_{i_1 \dots i_d} := \begin{cases} 0 & \text{if } i_1 = \dots = i_d = i \\ |a_{i_1 \dots i_d}| & \text{otherwise} \end{cases} \quad (12)$$

To show that  $\mathcal{A} \in \mathcal{S}_{n,d}^+$  take  $x \in \mathbb{R}^n$  and consider,

$$h_{\mathcal{A}}(x) = \mathcal{A}x^d = (\mathcal{P} - \mathcal{Q})x^d = \mathcal{P}x^d - \mathcal{Q}x^d \quad \forall x \in \mathbb{R}^n. \quad (13)$$

Since  $d$  is even, we have  $h_{\mathcal{P}}(x) := \mathcal{P}x^d = \sum_i^n a_{(i)}^d x_i^d \geq 0$  for all  $x \in \mathbb{R}^n$ . However,  $h_{\mathcal{Q}}(x) := \mathcal{Q}x^d$  can be positive or negative. Note that if  $h_{\mathcal{Q}}(x) \leq 0$  for some  $x \in \mathbb{R}^n$  then clearly from (13) we have,  $h_{\mathcal{A}}(x) = \mathcal{P}x^d + |\mathcal{Q}x^d| \geq 0$ .

So, the only case left is when  $h_{\mathcal{Q}}(x) \geq 0$  for some  $x \in \mathbb{R}^n$ . To show that  $h_{\mathcal{A}}(x) \geq 0$  in this case also, we define  $x^+ \in \mathbb{R}_+^n$  such that  $x_i^+ = \begin{cases} x_i & x_i \geq 0 \\ -x_i & x_i < 0 \end{cases}$ . To show that  $h_{\mathcal{Q}}(x) \leq h_{\mathcal{Q}}(x^+)$  consider,

$$S := \{\alpha \in \mathbb{I}^n(d) : x^\alpha < 0\}$$

and note that for  $\alpha \in S$  we have  $x^\alpha = -(x^+)^{\alpha}$ . Clearly, we have

$$\begin{aligned} 0 \leq h_{\mathcal{Q}}(x) &= \sum_{\alpha \in \mathbb{I}^n(d)} c(\alpha) \mathcal{Q}_{\alpha} x^{\alpha} = \sum_{\alpha \in \mathbb{I}^n(d) \setminus S} c(\alpha) \mathcal{Q}_{\alpha} x^{\alpha} + \sum_{\alpha \in S} c(\alpha) \mathcal{Q}_{\alpha} x^{\alpha} \\ &= \underbrace{\sum_{\alpha \in \mathbb{I}^n(d) \setminus S} c(\alpha) |\mathcal{A}_{\alpha}| (x^+)^{\alpha}}_{\geq 0} - \underbrace{\sum_{\alpha \in S} c(\alpha) |\mathcal{A}_{\alpha}| (x^+)^{\alpha}}_{\geq 0} \\ &\leq \sum_{\alpha \in \mathbb{I}^n(d) \setminus S} c(\alpha) |\mathcal{A}_{\alpha}| (x^+)^{\alpha} + \sum_{\alpha \in S} c(\alpha) |\mathcal{A}_{\alpha}| (x^+)^{\alpha} = h_{\mathcal{Q}}(x^+). \end{aligned} \quad (14)$$

Since  $d$  is even therefore we have  $h_{\mathcal{P}}(x) = h_{\mathcal{P}}(x^+)$ . Furthermore,  $\mathcal{A}$  is copositive and  $x^+ \in \mathbb{R}_+^n$  implying  $h_{\mathcal{P}}(x^+) \geq h_{\mathcal{Q}}(x^+)$ . Thus, we have

$$h_{\mathcal{P}}(x) = h_{\mathcal{P}}(x^+) \geq h_{\mathcal{Q}}(x^+) \geq h_{\mathcal{Q}}(x). \quad (15)$$

From (13) and (15) we deduce that  $h_{\mathcal{A}}(x) \geq 0$  in this case also. Hence,  $\mathcal{A}$  is positive semidefinite.

Converse is obvious since every positive semidefinite tensor is also copositive.  $\square$

Note that the above result also appeared in Zhang et al. [32] (Theorem 3.5(e) and Theorem 3.12), where the proof is constructed based on the spectral properties of the so called M-tensors. The proof given above is self-contained and does not require any extra structure.

#### 4. Approximation Hierarchies for the Copositive Cone

Recall that a tensor  $\mathcal{A} \in \mathcal{S}_{n,d}$  is copositive if  $h_{\mathcal{A}}(x) := \sum_{\alpha \in \mathbb{I}^n(d)} \mathcal{A}_{\alpha} x^{\alpha} \geq 0$  for all  $x \in \mathbb{R}_+^n$ . Notice that for any  $x \in \mathbb{R}_+^n$ , we can write  $x = z \circ z \in \mathbb{R}_+^n$  for some  $z \in \mathbb{R}^n$ , where  $\circ$  indicates the component wise (Hadamard) product, giving

$$h_{\mathcal{A}}(z \circ z) := \sum_{\alpha \in \mathbb{I}^n(d)} \mathcal{A}_{\alpha} (z \circ z)^{\alpha} = \sum_{i_1, \dots, i_d=1}^n a_{i_1 i_2 \dots i_d} z_{i_1}^2 z_{i_2}^2 \dots z_{i_d}^2. \quad (16)$$

Thus, the copositivity condition translates to  $h_{\mathcal{A}}(z \circ z) \geq 0$  for all  $z \in \mathbb{R}^n$ , for which a sufficient condition is that if (16) can be written as a sum of squares (SOS). Let us illustrate this with an example:

**Example 1.** Consider a tensor  $\mathcal{A} \in \mathcal{S}_{2,4}$  such that  $a_{1111} = 0, a_{1112} = 1/4, a_{1122} = 1/6, a_{1222} = -1/2, a_{2222} = 1$ . Then for  $x \in \mathbb{R}^2$ , we have

$$h_{\mathcal{A}}(x) := (x_1 - x_2)^2 x_2^2 + x_1^3 x_2 = x_1^3 x_2 + x_1^2 x_2^2 - 2x_1 x_2^3 + x_2^4.$$

Clearly,  $\mathcal{A}$  is not positive semidefinite since  $h_{\mathcal{A}}([-5, 2]^T) = -54$ . Let's consider  $x = z \circ z \in \mathbb{R}_+^2$  where  $z \in \mathbb{R}^2$  with  $x_1 = z_1^2, x_2 = z_2^2$ , then we have,

$$h_{\mathcal{A}}(z \circ z) = z_1^6 z_2^2 + z_1^4 z_2^4 - 2z_1^2 z_2^6 + z_2^8 = \left( (z_1^2 - z_2^2) z_2^2 \right)^2 + \left( z_1^3 z_2 \right)^2. \tag{17}$$

Thus,  $\mathcal{A}$  is copositive.

In the above example  $h_{\mathcal{A}}(z \circ z)$  can be written as sum of squares. However, this is not the case in general (see Example 2). Therefore, to develop higher order sufficient conditions, the following polynomial, introduced by Parrilo [17], is most often used.

$$P^{(r)}(z) := h_{\mathcal{A}}(z \circ z) \left( \sum_{k=1}^n z_k^2 \right)^r \text{ for all } z \in \mathbb{R}^n. \tag{18}$$

Clearly,  $P^{(r)}(z)$  is a polynomial of degree  $2(r + d)$ . Based on (18), one can define two cone approximations for the copositive cone (as we will see),

$$\mathcal{K}_{n,d}^{(r)} = \left\{ \mathcal{A} \in \mathcal{S}_{n,d} : P^{(r)}(z) \text{ has an SOS decomposition} \right\} \tag{19}$$

$$\mathcal{C}_{n,d}^{(r)} = \left\{ \mathcal{A} \in \mathcal{S}_{n,d} : P^{(r)}(z) \text{ has non-negative coefficients} \right\}. \tag{20}$$

For the tensor  $\mathcal{A} \in \mathcal{S}_{2,4}$  given in Example 1, we have  $\mathcal{A} \in \mathcal{K}_{2,4}^{(0)}$ . It is clear from (17) that  $h_{\mathcal{A}}(z \circ z)$  has a negative coefficient implying  $\mathcal{A} \notin \mathcal{C}_{n,d}^{(0)}$ . However, one can show that  $\mathcal{A} \in \mathcal{C}_{2,4}^{(1)}$ .

Inclusions  $\mathcal{K}_{n,d}^{(r)} \subseteq \mathcal{K}_{n,d}^{(r+1)}$  and  $\mathcal{C}_{n,d}^{(r)} \subseteq \mathcal{C}_{n,d}^{(r+1)}$  for  $r \in \mathbb{N}_0$  are evident from the following,

$$P^{(r+1)}(z) = h_{\mathcal{A}}(z \circ z) \left( \sum_{k=1}^n z_k^2 \right)^{r+1} = \left( \sum_{k=1}^n z_k^2 \right) \left( h_{\mathcal{A}}(z \circ z) \left( \sum_{k=1}^n z_k^2 \right)^r \right) \tag{21}$$

$$= \left( \sum_{k=1}^n z_k^2 \right) P^{(r)}(z). \tag{22}$$

The cone hierarchies  $\mathcal{K}_{n,d}^{(r)}$  and  $\mathcal{C}_{n,d}^{(r)}$  approximates the copositive cone from the inside. For instance if  $\mathcal{A} \in \mathcal{K}_{n,d}^{(r)}$  then  $P^{(r)}(z)$  being sum of squares is non-negative i.e.,  $P^{(r)}(z) \geq 0$  for all  $z \in \mathbb{R}^n$ , which in turn imply that  $h_{\mathcal{A}}(z \circ z) \geq 0$  for all  $z \in \mathbb{R}^n$ , hence  $\mathcal{A} \in \mathcal{C}_{n,d}$ . Similarly, one can show that if  $\mathcal{A} \in \mathcal{C}_{n,d}^{(r)}$  for some  $r \in \mathbb{N}_0$  then  $\mathcal{A}$  is copositive. Thus, we have

$$\bigcup_{r=0}^{\infty} \mathcal{K}_{n,d}^{(r)} \subseteq \mathcal{C}_{n,d} \text{ and } \bigcup_{r=0}^{\infty} \mathcal{C}_{n,d}^{(r)} \subseteq \mathcal{C}_{n,d}. \tag{23}$$

Referring to Polya's theorem (see e.g., [25], Theorem 2.1), which states that for a tensor  $\mathcal{A} \in \text{int}(\mathcal{C}_{n,d})$  there exists a large enough  $r$  such that  $\mathcal{A} \in \mathcal{C}_{n,d}^{(r)}$ . This further implies that, for some  $r \in \mathbb{N}_0$  the strictly copositive tensor  $\mathcal{A} \in \text{int}(\mathcal{C}_{n,d})$  allows  $P^{(r)}(z)$  to have sum of squares decomposition, that is  $\mathcal{A} \in \mathcal{K}_{n,d}^{(r)}$ . Therefore, the infinite union of these cones contains the interior of copositive cone i.e.,

$$\text{int}(\mathcal{C}_{n,d}) \subseteq \bigcup_{r=0}^{\infty} \mathcal{C}_{n,d}^{(r)} \text{ and } \text{int}(\mathcal{C}_{n,d}) \subseteq \bigcup_{r=0}^{\infty} \mathcal{K}_{n,d}^{(r)}. \tag{24}$$

For the tensor  $\mathcal{A} \in \mathcal{S}_{2,4}$  as given in Example 1 it holds  $\mathcal{A} \notin \text{int}(\mathcal{C}_{2,4})$  since  $h_{\mathcal{A}}([1, 0]^T) = 0$ . Thus, neither  $\cup_{r=0}^{\infty} \mathcal{K}_{n,d}^{(r)} \not\subseteq \text{int}(\mathcal{C}_{n,d})$  nor  $\cup_{r=0}^{\infty} \mathcal{C}_{n,d}^{(r)} \not\subseteq \text{int}(\mathcal{C}_{n,d})$  holds.

4.1. The Case  $R = 0$

The case  $r = 0$  is interesting and require further exploration. It is clear that  $\mathcal{C}_{n,d}^{(0)} = \mathcal{N}_{n,d}$ , which require no further exploration. For  $d = 2$ , the tensor  $\mathcal{A} \in \mathcal{K}_{n,2}^{(0)}$  is often characterized in terms of the decomposition  $\mathcal{A} = \mathcal{S} + \mathcal{T}$ , where  $\mathcal{S} \in \mathcal{S}_{n,2}^+$  and  $\mathcal{T} \in \mathcal{N}_{n,2}$ . However, for  $d > 4$  only one direction is possible, as shown below.

**Theorem 2.** Let  $\mathcal{A} \in \mathcal{S}_{n,d}$  and if  $\mathcal{A} \in \mathcal{K}_{n,d}^{(0)}$  then  $\mathcal{A} = \mathcal{S} + \mathcal{T}$ , where  $\mathcal{S} \in \mathcal{S}_{n,d}^+$  and  $\mathcal{T} \in \mathcal{N}_{n,d}$  i.e.,  $\mathcal{K}_{n,d}^{(0)} \subseteq \mathcal{S}_{n,d}^+ + \mathcal{N}_{n,d}$ .

**Proof.** The proof for the matrix case can be easily generalized to accommodate for higher order tensors see e.g., ([19], Theorem 2.1). □

The converse of the above theorem is not true in general. The following is a counter example,

**Example 2.** Let  $\mathcal{A} \in \mathcal{S}_{3,6}$  be such that

$$a_{i_1 i_2 i_3 i_4 i_5 i_6} = \begin{cases} 1 & i_1 = i_2 = \dots = i_6 \in \{1, 2, 3\} \\ 1/30 & i_1 = i_2 = i, i_3 = i_4 = j, i_5 = i_6 = k, i \neq j \neq k \in \{1, 2, 3\} \\ -1/15 & i_1 = i_2 = i_3 = i_4 = i, i_5 = i_6 = j, i \neq j \in \{1, 2, 3\} \\ 0 & \text{otherwise} \end{cases} \quad (25)$$

The associated polynomial is given by (cf. (3))

$$h_{\mathcal{A}}(\mathbf{x}) = x_1^6 + x_2^6 + x_3^6 + 3x_1^2 x_2^2 x_3^2 - (x_1^4 x_2^2 + x_1^4 x_3^2 + x_1^2 x_2^4 + x_1^2 x_3^4 + x_2^4 x_3^2 + x_2^2 x_3^4) \quad (26)$$

The polynomial given in (26) is the well known Robinson’s polynomial [33]. It is well known that  $h_{\mathcal{A}}(\mathbf{x}) \geq 0$  for all  $\mathbf{x} \in \mathbb{R}^3$ . Therefore, trivially  $\mathcal{A}$  can be written as sum of positive semidefinite and non-negative (actually zero) tensor. It is also well known that Robinson polynomial cannot be written as SOS (see [33] for a proof) i.e.,  $\mathcal{A} \notin \mathcal{K}_{3,6}^{(0)}$ .

To find special cases in which the converse of Theorem 2 holds note that from  $\mathcal{A} = \mathcal{S} + \mathcal{T}$ , we have

$$h_{\mathcal{A}}(\mathbf{z} \circ \mathbf{z}) = \langle (\mathcal{S} + \mathcal{T}), (\mathbf{z} \circ \mathbf{z})^d \rangle = \underbrace{\langle \mathcal{S}, (\mathbf{z} \circ \mathbf{z})^d \rangle}_{=h_{\mathcal{S}}(\mathbf{z} \circ \mathbf{z})} + \langle \mathcal{T}, (\mathbf{z} \circ \mathbf{z})^d \rangle \quad \forall \mathbf{z} \in \mathbb{R}^n \quad (27)$$

Clearly,  $\langle \mathcal{T}, (\mathbf{z} \circ \mathbf{z})^d \rangle$  is SOS since  $\mathcal{T}$  is non-negative. Thus, if  $h_{\mathcal{S}}(\mathbf{z} \circ \mathbf{z})$  can be written as SOS the converse of the Theorem 2 holds. It is well known that a matrix (i.e.,  $d = 2$ ) is positive semidefinite if and only if its associated polynomial can be written as sum of squares. Therefore, the converse of Theorem 2 holds for  $d = 2$  (see [19] (Theorem 2.1) for a proof). Moreover, for  $n = 2$  and  $d = 4$  a tensor is positive semidefinite if and only if its associated form is sum-of-squares, see, e.g., [33]. Therefore, the converse of Theorem 2 holds in this case also. We close this discussion by showing that the converse of Theorem 2 holds for the Z-tensors.

**Theorem 3.** Let  $\mathcal{A} \in \mathcal{S}_{n,d}$  be a Z-tensor and  $\mathcal{A} = \mathcal{S} + \mathcal{T}$ , where  $\mathcal{S} \in \mathcal{S}_{n,d}^+$  and  $\mathcal{T} \in \mathcal{N}_{n,d}$  then  $\mathcal{A} \in \mathcal{K}_{n,d}^{(0)}$ .



**Proof.** Note that since  $\mathcal{A}$  is a  $Z$ -tensor and  $\mathcal{T}$  is non-negative therefore  $\mathcal{S}$  is also a  $Z$ -tensor. It is well-known that a  $Z$ -tensor is positive semidefinite if and only if it is sum of squares. (see e.g., [26] (Proposition 2.1), [34] (Theorem 11)). Hence (27) can be written as sum of square.  $\square$

#### 4.2. Characterization of $\mathcal{K}_{n,d}^{(r)}$

In this subsection, we formulate a characterization for  $\mathcal{K}_{n,d}^{(r)}$ . Before presenting this characterization, we provide bounds on the value of coefficients for the polynomial  $P^{(r)}(z)$ . For this consider,

$$P^{(r)}(z) = h_{\mathcal{A}}(z \circ z) \left( \sum_{k=1}^n z_k^2 \right)^r = \left( \sum_{i_1, \dots, i_d=1}^n a_{i_1 i_2 \dots i_d} z_{i_1}^2 z_{i_2}^2 \dots z_{i_d}^2 \right) \left( \sum_{\alpha \in \mathbb{I}^n(r)} c(\alpha) z^{2\alpha} \right) \quad (28)$$

Note that  $z_{i_j}^2$  can be written as  $z^{2\mathbf{e}_j}$  where  $\mathbf{e}_j$  is a unit vector and  $1 \leq j \leq n$ . Using this notation (28) can be written as,

$$P^{(r)}(z) = \left( \sum_{i_1, \dots, i_d=1}^n a_{i_1 \dots i_d} z^{2\mathbf{e}_{i_1}} z^{2\mathbf{e}_{i_2}} \dots z^{2\mathbf{e}_{i_d}} \right) \left( \sum_{\alpha \in \mathbb{I}^n(r)} c(\alpha) z^{2\alpha} \right) \quad (29)$$

Denote by  $\theta := (\mathbf{e}_{i_1} + \mathbf{e}_{i_2} \dots + \mathbf{e}_{i_d}) \in \mathbb{I}^n(d)$ , then  $\omega := \alpha + \theta \in \mathbb{I}^n(d+r)$ , and with these notations we can write (29) as follows,

$$P^{(r)}(z) = \sum_{\omega \in \mathbb{I}^n(r+d)} [\mathcal{B}_{\omega}(\mathcal{A})] z^{2\omega} \text{ where } \mathcal{B}_{\omega}(\mathcal{A}) := \sum_{\theta \in \mathbb{I}^n(d)} \left( \frac{r!}{\prod_{i=1}^n (\omega_i - \theta_i)!} \right) c(\theta) \mathcal{A}_{\theta} \quad (30)$$

where  $\omega_i - \theta_i \geq 0$  for all  $i = 1, \dots, n$ .

Note that the maximum value of  $\prod_{i=1}^n (\omega_i - \theta_i)!$  is  $r!$  (Observe that  $\sum_{i=1}^n (\omega_i - \theta_i) = r + d - d = r$ . Take  $r_i = \omega_i - \theta_i$  then maximum value occurs when  $r_i = r$  and  $r_j = 0$  for all  $i \neq j$ ). Moreover, the minimum value of  $\prod_{i=1}^n (\omega_i - \theta_i)!$  occurs when  $\omega_i - \theta_i \in \{0, 1\}$  for all  $i \in [1 : n]$  and the minimum value is 1. Thus, we have

$$1 \leq \frac{r!}{\prod_{i=1}^n (\omega_i - \theta_i)!} \leq r!, \text{ where } \omega_i - \theta_i \geq 0 \forall i \in [1 : n]. \quad (31)$$

To show that the upper bound in (31) is sharp take  $n = r + d$  and  $\omega = (1, 1, \dots, 1)^T \in \mathbb{I}^n(d+r)$  then for  $\theta_i = i, 1 \leq i \leq d$ , we have  $(\omega_i - \theta_i) \in \{0, 1\}$ , thus the denominator reduces to 1 and the upper bound is exactly  $r!$ . In the following theorem, we describe a characterization of  $\mathcal{K}_{n,d}^{(r)}$ .

**Theorem 4** (cf. [19] for  $d = 2$ ). *The tensor  $\mathcal{A} \in \mathcal{K}_{n,d}^{(r)}$  if and only if there exists a PSD matrix  $\tilde{M} \in S_{\tilde{n}_r, 2}^+$  associated with the polynomial (18)*

$$P^{(r)}(z) = \tilde{z}^T \tilde{M} \tilde{z} \text{ where } \tilde{z} = [z^{\alpha}]_{\alpha \in \mathbb{I}^n(d+r)} \in \mathbb{R}^{\tilde{n}_r} \text{ with } \tilde{n}_r = \binom{n+r+d-1}{r+d}$$



such that

$$\sum_{\substack{(\beta_i, \beta_j) \in \mathbb{I}^n(d+r) \times \mathbb{I}^n(d+r) : \beta_i + \beta_j = 2\omega \\ \omega \in \mathbb{I}^n(d+r)}} \tilde{m}_{i,j} := \mathcal{B}_\omega(\mathcal{A}) \text{ for all } \omega \in \mathbb{I}^n(d+r) \tag{32}$$

$$\sum_{\substack{(\beta_i, \beta_j) \in \mathbb{I}^n(d+r) \times \mathbb{I}^n(d+r) : \beta_i + \beta_j = u \\ u \in \mathbb{I}^n(2d+2r) \setminus 2\mathbb{I}^n(d+r)}} \tilde{m}_{i,j} := 0 \text{ for all } u \in \mathbb{I}^n(2d+2r) \setminus 2\mathbb{I}^n(d+r). \tag{33}$$

**Proof.** The proof is an easy generalization of matrix case (Theorem 2.2, [19]), which is presented here for the sake of completeness. Let us consider the polynomial  $P^{(r)}(z)$  described in (30) and using its matrix formulation as follows

$$P^{(r)}(z) = \tilde{z}^T \tilde{M} \tilde{z} \text{ where } \tilde{z} = [z^\alpha]_{\alpha \in \mathbb{I}^n(d+r)} \in \mathbb{R}^{\tilde{n}_r} \text{ with } \tilde{n}_r = \binom{n+d+r-1}{d+1} \tag{34}$$

Taking  $\mathcal{A} \in \mathcal{K}_{n,d}^{(r)}$  implies that its associate polynomial  $P^{(r)}(z)$  allows a sum of squares decomposition, which further implies that the matrix  $\tilde{M}$  is PSD. For converse we assume that, the matrix  $\tilde{M}$  in (34) is PSD. So the following decomposition is evident

$$\tilde{M} = \sum_{k=1}^{m_r} \tilde{u}_k \tilde{u}_k^T \text{ where } m_r \text{ is the rank of } \tilde{M}. \tag{35}$$

Combining (34) and (35) gives,

$$P^{(r)}(z) = \sum_{k=1}^{m_r} (\tilde{u}_k^T \tilde{z})^T (\tilde{u}_k^T \tilde{z}) \text{ where } \tilde{z} = [z^\alpha]_{\alpha \in \mathbb{I}^n(d+r)} \in \mathbb{R}^{\tilde{n}_r} \tag{36}$$

$$= \sum_{k=1}^{m_r} (\tilde{u}_k^T \tilde{z})^2 \text{ where } \tilde{z} = [z^\alpha]_{\alpha \in \mathbb{I}^n(d+r)} \in \mathbb{R}^{\tilde{n}_r} \tag{37}$$

By comparing (30) and (34), we obtain the following

$$\tilde{z}^T \tilde{M} \tilde{z} = \sum_{\omega \in \mathbb{I}^n(d+r)} [\mathcal{B}_\omega(\mathcal{A})] z^{2\omega} \tag{38}$$

Comparing the coefficient of  $z^{2\omega}$  in (38) gives,

$$\sum_{\substack{(\beta_i, \beta_j) \in \mathbb{I}^n(d+r) \times \mathbb{I}^n(d+r) : \beta_i + \beta_j = 2\omega \\ \omega \in \mathbb{I}^n(d+r)}} \tilde{m}_{i,j} = \mathcal{B}_\omega(\mathcal{A}) \text{ for all } \omega \in \mathbb{I}^n(d+r). \tag{39}$$

However, for  $u \in \mathbb{I}^n(2d+2r) \setminus 2\mathbb{I}^n(d+r)$  the coefficients of  $z^u$  on R.H.S. of (38) are all zero, thus we have

$$\sum_{\substack{(\beta_i, \beta_j) \in \mathbb{I}^n(d+r) \times \mathbb{I}^n(d+r) : \beta_i + \beta_j = u \\ u \in \mathbb{I}^n(2d+2r) \setminus 2\mathbb{I}^n(d+r)}} \tilde{m}_{i,j} = 0 \text{ for all } u \in \mathbb{I}^n(2d+2r) \setminus 2\mathbb{I}^n(d+r).$$

Hence the proof is complete.  $\square$

From Theorem 4, it is clear that the matrix  $\tilde{M}$  is sparse. To deal with the sparsity Ahmadi and Majumdar [35] has introduced polyhedral approximations using the observation that a diagonally dominant matrix is positive semidefinite.

### 4.3. Polyhedral Characterization of $\mathcal{C}_{n,d}^{(r)}$

In this section, we present a compact representation for the cone  $\mathcal{C}_{n,d}^{(r)}$ . The representation is useful in deducing polyhedral characterization of  $\mathcal{C}_{n,d}^{(r)}$ . For this, recall that for any  $\omega \in \mathbb{I}^n(r+d)$  (30) can be re-written as follows

$$\begin{aligned}
 \mathcal{B}_\omega(\mathcal{A}) &= \sum_{\theta \in \mathbb{I}^n(d)} c(\theta)c(\omega - \theta)\mathcal{A}_\theta \\
 &= \sum_{\theta \in \mathbb{I}^n(d)} c(\theta) \frac{(|\omega - \theta|)!}{(\omega_1 - \theta_1)!(\omega_2 - \theta_2)! \cdots (\omega_n - \theta_n)!} \mathcal{A}_\theta \\
 &= r! \sum_{\theta \in \mathbb{I}^n(d)} c(\theta) \frac{[\omega_1(\omega_1 - 1) \cdots (\omega_1 - (\theta_1 - 1))] \cdots [\omega_n(\omega_n - 1) \cdots (\omega_n - (\theta_n - 1))]}{(\omega_1)! \cdots (\omega_n)!} \mathcal{A}_\theta \\
 &= r! \sum_{\theta \in \mathbb{I}^n(d)} c(\theta) \prod_{k=1}^n \frac{[\omega_k(\omega_k - 1) \cdots (\omega_k - (\theta_k - 1))]}{(\omega_k)!} \mathcal{A}_\theta \\
 &= \frac{r!}{\prod_{k=1}^n (\omega_k)!} \sum_{\theta \in \mathbb{I}^n(d)} c(\theta)\mathcal{A}_\theta \left( \prod_{k=1}^n [\omega_k(\omega_k - 1) \cdots (\omega_k - (\theta_k - 1))] \right) \tag{40}
 \end{aligned}$$

Note that the first equality follows by using the definition of polynomial coefficient (see (4)). Recognize that the product  $[\omega_k(\omega_k - 1) \cdots (\omega_k - (\theta_k - 1))]$  is linked with the falling factorials. The falling factorials can be represented using the Stirling number of the first kind as follows

$$\omega_k(\omega_k - 1) \cdots (\omega_k - (\theta_k - 1)) = \sum_{m=0}^{\theta_k} s(\theta_k, m)\omega_k^m \tag{41}$$

where  $s(\theta_k, m) := (-1)^{\theta_k - m} \begin{bmatrix} \theta_k \\ m \end{bmatrix}$  is well known Stirling number of the first kind (see e.g., [36], Chapter 6.1). Using (41) in (40) gives

$$\mathcal{B}_\omega(\mathcal{A}) = \frac{r!}{\prod_{k=1}^n (\omega_k)!} \sum_{\theta \in \mathbb{I}^n(d)} c(\theta)\mathcal{A}_\theta \left( \prod_{k=1}^n \sum_{m=0}^{\theta_k} s(\theta_k, m)\omega_k^m \right) \tag{42}$$

For  $\theta \in \mathbb{I}^n(d)$  and  $\alpha \in \mathbb{I}^n(t)$  the inequality  $\alpha \leq \theta$  is defined to hold element wise i.e., the inequality is true if  $\alpha_i \leq \theta_i$  for all  $1 \leq i \leq n$ . Furthermore,  $s(\theta, \alpha) = s(\theta_1, \alpha_1)s(\theta_2, \alpha_2) \cdots s(\theta_n, \alpha_n)$ . Observe that, if  $\theta_k = m$  then  $\begin{bmatrix} \theta_k \\ m \end{bmatrix} = 1$  and also  $\begin{bmatrix} \theta_k \\ m \end{bmatrix} = 0$  if  $\theta_k < m$ . The observation leads to the following simplification for each  $\theta \in \mathbb{I}^n(d)$  and  $\alpha \in \mathbb{I}^n(t)$  for all  $t \in [0 : d]$

$$\begin{aligned}
 \prod_{k=1}^n \sum_{m=0}^{\theta_k} s(\theta_k, m)\omega_k^m &= \sum_{m_1, m_2, \dots, m_n=0}^{\theta_1, \theta_2, \dots, \theta_n} s(\theta_1, m_1)s(\theta_2, m_2) \cdots s(\theta_n, m_n)\omega_1^{m_1}\omega_2^{m_2} \cdots \omega_n^{m_n} \\
 &= \sum_{\alpha \leq \theta} s(\theta, \alpha)\omega^\alpha \tag{43}
 \end{aligned}$$

Combining (42) and (43) leads to

$$\mathcal{B}_\omega(\mathcal{A}) = \frac{r!}{\prod_{k=1}^n (\omega_k)!} \sum_{\theta \in \mathbb{I}^n(d)} c(\theta)\mathcal{A}_\theta \left( \sum_{\substack{\alpha \in \mathbb{I}^n(t), t \in [0:d] \\ \alpha \leq \theta}} s(\theta, \alpha)\omega^\alpha \right) \tag{44}$$

We define the tensors  $\mathcal{S}(t)$  and  $\mathcal{W}(t)$  of order  $|\alpha| = t$  and dimension  $n$  as follows,

$$\mathcal{S}(t) = \left( s(\boldsymbol{\theta}, \boldsymbol{\alpha}) \right)_{\boldsymbol{\alpha} \in \mathbb{I}^n(t)} \quad \text{and} \quad \mathcal{W}(t) = \left( \omega^\alpha \right)_{\boldsymbol{\alpha} \in \mathbb{I}^n(t)} \quad \text{for all } t \in [0 : d]. \tag{45}$$

**Remark 1.** Interestingly for each  $\boldsymbol{\alpha} \in \mathbb{I}^n(d)$  we have

$$(\mathcal{S}(d))_\alpha := \begin{cases} 1 & \text{if } \boldsymbol{\alpha} = \boldsymbol{\theta} \\ 0 & \text{otherwise} \end{cases}. \tag{46}$$

Consequently, the above notations lead to the following simplified representation of (44)

$$\mathcal{B}_\omega(\mathcal{A}) = \frac{r!}{\prod_{k=1}^n (\omega_k)!} \sum_{\boldsymbol{\theta} \in \mathbb{I}^n(d)} c(\boldsymbol{\theta}) \mathcal{A}_\theta \left( \sum_{t=0}^d \langle \mathcal{S}(t), \mathcal{W}(t) \rangle_\theta \right). \tag{47}$$

Finally, we define the notation  $\mathcal{Y}_\theta^t := \langle \mathcal{S}(t), \mathcal{W}(t) \rangle_\theta$ . The notation leads to define the tensor

$$\mathcal{Y}^t := \left( \mathcal{Y}_\theta^t \right)_{\boldsymbol{\theta} \in \mathbb{I}^n(d)} \tag{48}$$

of order  $d$  and dimension  $n$ , whose entries are either forms of degree  $t$  or zero that is,

$$\mathcal{Y}_\theta^t := \begin{cases} \sum_{\boldsymbol{\alpha} \in \mathbb{I}^n(t)} s(\boldsymbol{\theta}, \boldsymbol{\alpha}) \omega^\alpha & \boldsymbol{\alpha} \leq \boldsymbol{\theta} \\ 0 & \text{otherwise} \end{cases} \tag{49}$$

Note that  $\mathcal{Y}_\theta^t \in \mathbb{R}$  and  $\mathcal{Y}^t \in \mathcal{S}_{n,d}$ . Moreover, from (49) one can easily assert that  $\mathcal{Y}_\theta^d = \omega^d$ .

**Remark 2.** Recall that  $\boldsymbol{\theta} \in \mathbb{I}^n(d)$  implying  $\boldsymbol{\theta}$  can be written as a linear combination of unit vectors as follows  $\boldsymbol{\theta} = \beta_1 \mathbf{e}_1 + \dots + \beta_n \mathbf{e}_n$ ,  $\boldsymbol{\beta} \in [0 : d]^n$ . Since,  $\boldsymbol{\theta} \in \mathbb{I}^n(d)$  the maximum number of non-zero elements of  $\boldsymbol{\beta}$  will be  $\min\{d, n\}$ . Similarly, for  $\boldsymbol{\alpha} \in \mathbb{I}^n(t)$ , we can have at most  $\min\{n, t\}$  non-zero coefficients in the linear combination of the basis. Based on this observation we can obtain an explicit representation of the tensor  $\mathcal{Y}^t \in \mathcal{S}_{n,d}$  for all  $t \in [1 : d - 1]$ . Thus, for each  $\boldsymbol{\theta} \in \mathbb{I}^n(d)$  the entries of tensor  $\mathcal{Y}^t$  are described as

$$\mathcal{Y}_\theta^t := \begin{cases} s(d, t) \omega_i^t & \text{if } t\mathbf{e}_i = \boldsymbol{\alpha} \leq \boldsymbol{\theta} = d\mathbf{e}_i \\ & \forall i \in [1 : n] \\ \sum_{\substack{\alpha_{i_1}, \alpha_{i_2}=1 \\ \text{s.t. } \alpha_{i_1} + \alpha_{i_2} = t}} s(\theta_{i_1}, \alpha_{i_1}) s(\theta_{i_2}, \alpha_{i_2}) \omega_{i_1}^{\alpha_{i_1}} \omega_{i_2}^{\alpha_{i_2}} & \text{if } \left\{ \begin{array}{l} \boldsymbol{\alpha} \leq \boldsymbol{\theta} \\ \boldsymbol{\alpha} = \alpha_{i_1} \mathbf{e}_{i_1} + \alpha_{i_2} \mathbf{e}_{i_2} \\ \boldsymbol{\theta} = \theta_{i_1} \mathbf{e}_{i_1} + \theta_{i_2} \mathbf{e}_{i_2} \\ \forall i_1 \neq i_2 \in [1 : n] \end{array} \right\} \\ \sum_{\substack{\alpha_{i_1}, \alpha_{i_2}, \alpha_{i_3}=1 \\ \text{s.t. } \alpha_{i_1} + \alpha_{i_2} + \alpha_{i_3} = t}} s(\theta_{i_1}, \alpha_{i_1}) s(\theta_{i_2}, \alpha_{i_2}) s(\theta_{i_3}, \alpha_{i_3}) \omega_{i_1}^{\alpha_{i_1}} \omega_{i_2}^{\alpha_{i_2}} \omega_{i_3}^{\alpha_{i_3}} & \text{if } \left\{ \begin{array}{l} \boldsymbol{\alpha} \leq \boldsymbol{\theta} \\ \boldsymbol{\alpha} = \alpha_{i_1} \mathbf{e}_{i_1} + \alpha_{i_2} \mathbf{e}_{i_2} + \alpha_{i_3} \mathbf{e}_{i_3} \\ \boldsymbol{\theta} = \theta_{i_1} \mathbf{e}_{i_1} + \theta_{i_2} \mathbf{e}_{i_2} + \theta_{i_3} \mathbf{e}_{i_3} \\ \forall i_1 \neq i_2 \neq i_3 \in [1 : n] \end{array} \right\} \\ \vdots & \vdots \\ \sum_{\substack{\alpha_{i_1}, \dots, \alpha_{i_m}=1 \\ \text{s.t. } \alpha_{i_1} + \dots + \alpha_{i_m} = t \\ \text{where } m = \min\{t, n\}}} s(\theta_{i_1}, \alpha_{i_1}) \dots s(\theta_{i_m}, \alpha_{i_m}) \omega_{i_1}^{\alpha_{i_1}} \dots \omega_{i_m}^{\alpha_{i_m}} & \text{if } \left\{ \begin{array}{l} \boldsymbol{\alpha} \leq \boldsymbol{\theta} \\ \boldsymbol{\alpha} = \alpha_{i_1} \mathbf{e}_{i_1} + \dots + \alpha_{i_m} \mathbf{e}_{i_m} \\ \boldsymbol{\theta} = \theta_{i_1} \mathbf{e}_{i_1} + \dots + \theta_{i_m} \mathbf{e}_{i_m} \\ \forall i_1 \neq \dots \neq i_m \in [1 : n] \end{array} \right\} \\ 0 & \text{otherwise} \end{cases} \tag{50}$$

Thus, from (45) and (48) one obtains the following notionally convenient formulation of (44),

$$\begin{aligned} \mathcal{B}_\omega(\mathcal{A}) &= \frac{r!}{\prod_{k=1}^n (\omega_k)!} \left( \left\langle \mathcal{A}, \mathcal{Y}^0 + \mathcal{Y}^1 + \dots + \mathcal{Y}^{d-1} + \mathcal{Y}^d \right\rangle \right) \\ &= \frac{r!}{\prod_{k=1}^n (\omega_k)!} \left( \left\langle \mathcal{A}, \omega^d \right\rangle + \sum_{t=0}^{(d-1)} \left\langle \mathcal{A}, \mathcal{Y}^t \right\rangle \right). \end{aligned} \tag{51}$$

**Remark 3.** For sanity check we consider a special case, the tensor of all ones  $\mathcal{A} = \mathcal{E}$ , that is  $a_{i_1, \dots, i_d} = 1$  for all  $i_1, i_2, \dots, i_d \in [1 : n]$ . Note that  $\langle \mathcal{E}, \omega^d \rangle = |\omega|^d$  and from (51) we have

$$\begin{aligned} \mathcal{B}_\omega(\mathcal{E}) &= \frac{r!}{\prod_{k=1}^n (\omega_k)!} \left( \left\langle \mathcal{E}, \omega^d \right\rangle + \left\langle \mathcal{E}, \mathcal{Y}^{d-1} \right\rangle + \dots + \left\langle \mathcal{E}, \mathcal{Y}^1 \right\rangle + \left\langle \mathcal{E}, \mathcal{Y}^0 \right\rangle \right) \\ &= \frac{r!}{\prod_{k=1}^n (\omega_k)!} \left( |\omega|^d + s(d, d-1)|\omega|^{d-1} + \dots + s(d, 1)|\omega|^1 \right) \\ &= \frac{r!}{\prod_{k=1}^n (\omega_k)!} \left( (r+d)(r+(d-1)) \dots (r+2)(r+1) \right) \\ &= \frac{r!}{\prod_{k=1}^n (\omega_k)!} \frac{(r+d)!}{r!} = \frac{(r+d)!}{\prod_{k=1}^n (\omega_k)!} = c(\omega) \end{aligned} \tag{52}$$

The above representation leads to a polyhedral representation of the cone hierarchy  $\mathcal{C}_{n,d}^{(r)}$  and is presented in the following theorem (cf. [19], Theorem 2.4).

**Theorem 5.** For all  $r \in \mathbb{N}_0$  and  $n \in \mathbb{N}$  the polyhedral representation of cones  $\mathcal{C}_{n,d}^{(r)}$  is given as

$$\mathcal{C}_{n,d}^{(r)} = \left\{ \mathcal{A} \in \mathcal{S}_{n,d} \mid \left\langle \mathcal{A}, \omega^d + \sum_{t=0}^{(d-1)} \mathcal{Y}^t \right\rangle \geq 0 \text{ for all } \omega \in \mathbb{I}^n(r+d) \right\}$$

where  $\mathcal{Y}^t := (\langle \mathcal{S}(t), \mathcal{W}(t) \rangle)_{\theta \in \mathbb{I}^n(d)}$  for all  $t \in [0 : d-1]$ .

**Proof.** The proof follows immediately from (20) and (51).  $\square$

### 5. Approximating Polynomial Optimization over the Simplex

In this section, we consider the homogeneous polynomial optimization over the simplex

$$\min h_{\mathcal{A}}(x) \text{ subject to : } h_{\mathcal{B}}(x) = 1 \text{ for } x \in \mathbb{R}_+^n \tag{53}$$

where  $\mathcal{A}, \mathcal{B} \in \mathcal{S}_{n,d}$ . It is well-known (see e.g., [18,19,25]) that (53) can be equivalently reformulated as a copositive program over the cone of completely positive tensors. The reformulation is as follows,

$$\min \left\langle \mathcal{A}, x^d \right\rangle \text{ subject to : } \left\langle \mathcal{B}, x^d \right\rangle = 1, x^d \in \mathcal{C}_{n,d}^*. \tag{54}$$

The dual formulation of (54) is also a conic program over the cone of copositive tensors, which is given below.

$$\max_{\lambda \in \mathbb{R}} \lambda \text{ subject to : } (\mathcal{A} - \lambda \mathcal{B}) \in \mathcal{C}_{n,d} \tag{55}$$

We consider a special case where  $\mathcal{B} = \mathcal{E}$ . Obviously, the feasible set  $\{\mathcal{E}x^d = 1, x \in \mathbb{R}_+^n\}$  is precisely the simplex  $\Delta_n$ . Thus, the minimum (maximum) value of (53) in this special case, that is optimization of the homogeneous polynomial over the simplex  $\Delta_n$  is

$$h_{\mathcal{A}}^{\min(\Delta_n)} := \min_{x \in \Delta_n} h_{\mathcal{A}}(x) \quad \left( h_{\mathcal{A}}^{\max(\Delta_n)} := \max_{x \in \Delta_n} h_{\mathcal{A}}(x) \right) \tag{56}$$

As mentioned before testing if a tensor is copositive is co-NP-hard. To find an approximate solution we replace the cone  $\mathcal{C}_{n,d}$  (for the special case  $\mathcal{B} = \mathcal{E}$ ) in (55) by its approximation  $\mathcal{C}_{n,d}^{(r)}$  where  $r \in \mathbb{N}_0$ ,

$$v_c^{(r)} := \max \left\{ \lambda \mid \mathcal{A} - \lambda \mathcal{E} \in \mathcal{C}_{n,d}^{(r)}, \lambda \in \mathbb{R} \right\} \tag{57}$$

We are interested to compute the bound on the difference of approximate solution (to the dual program)  $v_c^{(r)}$  and the exact solution  $h_{\mathcal{A}}^{\min(\Delta_n)}$ . For this we use rational girding of the simplex  $\Delta_n$  i.e., for non-negative integer  $r \in \mathbb{N}_0$  we have,

$$\Delta_n(r) := \{x \in \Delta_n : (r + d)x \in \mathbb{N}_0^n\} = \frac{\mathbb{I}^n(r + d)}{(r + d)}. \tag{58}$$

The rational grid  $\Delta_n(r)$  is a discretization of  $\Delta_n$ , which leads to a natural approximation of (56), i.e.,

$$h_{\mathcal{A}}^{\min(\Delta_n(r))} := \min_{x \in \Delta_n(r)} \{h_{\mathcal{A}}(x) := \langle \mathcal{A}, x^d \rangle\} \tag{59}$$

Note that  $v_c^{(r)}$  approximates the dual while  $h_{\mathcal{A}}^{\min(\Delta_n(r))}$  approximates the primal as given in (56). It is interesting to investigate the connection between these two approximations. The connection is given below (cf. [19] (Theorem 3.1), for  $d = 2$  and [25] (Theorem 3.1) for  $d = 4$ ),

**Theorem 6.** Let  $\Delta_n(r)$  be a rational discretization of simplex  $\Delta_n$  as given in (58) for any  $r \in \mathbb{N}_0$ , then for  $\mathcal{Q} \in \mathcal{S}_{n,d}$  we have

$$v_c^{(r)} = \frac{r!(r + d)^d}{(r + d)!} \min_{x \in \Delta_n(r)} \left\langle \mathcal{Q}, x^d + \sum_{t=1}^{(d-1)} \frac{\mathcal{X}^t}{(r + d)^{d-t}} \right\rangle \tag{60}$$

where  $\mathcal{X}^t \in \mathcal{S}_{n,d}$ .

**Proof.** Let  $\mathcal{A} := \mathcal{Q} - \lambda \mathcal{E}$  be a feasible point of the program given in (57) then for  $\omega \in \mathbb{I}^n(r + d)$  it follows, from (51) and the definition of  $\mathcal{C}_{n,d}^{(r)}$ , that,

$$0 \leq \mathcal{B}_{\omega}(\mathcal{A}) = \mathcal{B}_{\omega}(\mathcal{Q}) - \lambda \mathcal{B}_{\omega}(\mathcal{E}) \tag{61}$$

$$= \frac{r!}{\prod_{k=1}^n (\omega_k)!} \left( \langle \mathcal{Q}, \omega^d \rangle + \sum_{t=1}^{(d-1)} \langle \mathcal{Q}, \mathcal{Y}^t \rangle \right) - \lambda c(\omega) \tag{62}$$

$$= \frac{(r + d)(r + d - 1) \cdots (r + 1)r!}{\prod_{k=1}^n (\omega_k)!} \left( \frac{\langle \mathcal{Q}, \omega^d \rangle + \sum_{t=1}^{(d-1)} \langle \mathcal{Q}, \mathcal{Y}^t \rangle}{(r + d)(r + d - 1) \cdots (r + 1)} \right) - \lambda c(\omega) \tag{63}$$

$$= c(\omega) \left( \frac{\langle \mathcal{Q}, \omega^d \rangle + \sum_{t=1}^{(d-1)} \langle \mathcal{Q}, \mathcal{Y}^t \rangle}{(r + d)(r + d - 1) \cdots (r + 1)} - \lambda \right). \tag{64}$$

Giving,

$$\frac{\langle \mathcal{Q}, \omega^d \rangle + \sum_{t=1}^{(d-1)} \langle \mathcal{Q}, \mathcal{Y}^t \rangle}{(r+d)(r+d-1)\dots(r+1)} \geq \lambda.$$

The above imply that the maximum value of  $\lambda$  in (57) is attained at the minimum value of  $\left(\frac{\langle \mathcal{Q}, \omega^d \rangle + \sum_{t=1}^{(d-1)} \langle \mathcal{Q}, \mathcal{Y}^t \rangle}{(r+d)(r+d-1)\dots(r+1)}\right)$ . Thus, (57) can be equivalently written as follows,

$$\begin{aligned} v_c^{(r)} &= \min_{\omega \in \mathbb{I}^n(d+r)} \left\{ \frac{(r+d)^{d-1}}{(r+d-1)\dots(r+1)} \left( \frac{\langle \mathcal{Q}, \omega^d \rangle + \sum_{t=1}^{(d-1)} \langle \mathcal{Q}, \mathcal{Y}^t \rangle}{(r+d)^d} \right) \right\} \\ &= \frac{r!(r+d)^d}{(r+d)!} \min_{\omega \in \mathbb{I}^n(d+r)} \left\langle \mathcal{Q}, \frac{\omega^d}{(r+d)^d} + \sum_{t=1}^{(d-1)} \frac{\mathcal{Y}^t}{(r+d)^d} \right\rangle. \end{aligned} \tag{65}$$

A mere change of variable  $\omega = (r+d)x$  where  $x \in \Delta_n(r)$  and  $\mathcal{Y}^t = (r+d)^t \mathcal{X}^t$  in (65) yields the required expression, i.e.,

$$v_c^{(r)} = \frac{r!(r+d)^d}{(r+d)!} \min_{x \in \Delta_n(r)} \left\langle \mathcal{Q}, x^d + \sum_{t=1}^{(d-1)} \frac{\mathcal{X}^t}{(r+d)^{d-t}} \right\rangle. \tag{66}$$

□

Note that, for any  $r \in \mathbb{N}_0$  we have the relation  $h_Q^{\min(\Delta_n(r))} \geq h_Q^{\min(\Delta_n)}$ . However, in (66) a correction term  $\sum_{t=0}^{d-1} \frac{\langle \mathcal{Q}, \mathcal{X}^t \rangle}{(r+d)^{d-t}}$  is deducted from the actual objective  $\langle \mathcal{Q}, x^d \rangle$  for obtaining a closer approximation to  $h_Q^{\min(\Delta_n)}$ . Clearly, for increasing  $r \in \mathbb{N}_0$  the value  $v_c^{(r)}$  surpass the value  $h_{\Delta_n}^{(r)}$  invariably. However, one has to compensate with the factor  $\frac{r!(r+d)^d}{(r+d)!} = 1 + \mathcal{O}\left(\frac{1}{r^{d-1}}\right) > 1$ . It would be interesting to find bounds on the difference between two approximations namely  $v_c^{(r)}$  and  $h_Q^{\min(\Delta_n(r))}$ . To compute the bound we define some notations. First recall that  $\begin{bmatrix} \theta \\ \alpha \end{bmatrix}$  denotes Stirling number. In addition, for  $x \in \Delta_n, t \in [0 : d]$  we define a function  $q_t(x)$  as follows,

$$q_t(x) := \frac{1}{(r+d)^{d-t}} \sum_{\substack{\alpha \leq \theta \\ \theta \in \mathbb{I}^n(d), \alpha \in \mathbb{I}^n(t)}} \begin{bmatrix} \theta \\ \alpha \end{bmatrix} c(\theta) \mathcal{Q}_\theta x^\alpha \tag{67}$$

$$:= \frac{1}{(r+d)^{d-t}} \left\{ \sum_{\substack{\alpha \leq \theta, \mathcal{Q}_\theta \geq 0 \\ \theta \in \mathbb{I}^n(d), \alpha \in \mathbb{I}^n(t)}} \begin{bmatrix} \theta \\ \alpha \end{bmatrix} c(\theta) |\mathcal{Q}_\theta| x^\alpha - \sum_{\substack{\alpha \leq \theta, \mathcal{Q}_\theta < 0 \\ \theta \in \mathbb{I}^n(d), \alpha \in \mathbb{I}^n(t)}} \begin{bmatrix} \theta \\ \alpha \end{bmatrix} c(\theta) |\mathcal{Q}_\theta| x^\alpha \right\} \tag{68}$$

$$:= q_t(x)^+ - q_t(x)^- \tag{69}$$

If there is no dependence on the variable  $x$  we simply write  $q_t$  that is,

$$q_t^+ := \frac{1}{(r+d)^{d-t}} \sum_{\substack{\alpha \leq \theta, \mathcal{Q}_\theta \geq 0 \\ \theta \in \mathbb{I}^n(d), \alpha \in \mathbb{I}^n(t)}} \begin{bmatrix} \theta \\ \alpha \end{bmatrix} c(\theta) |\mathcal{Q}_\theta|$$

One can define  $q_t^+$  analogously.

**Theorem 7.** Let  $\Delta_n(r)$  be a rational discretization of simplex  $\Delta_n$  as given in (58) for any  $r \in \mathbb{N}_0$ , then for  $\mathcal{Q} \in \mathcal{S}_{n,d}$  we have

$$-\left( \sum_{t \in [0:(d-1)] \cap \mathbb{N}_{Even}} q_t^- + \sum_{t \in [0:(d-1)] \cap \mathbb{N}_{Odd}} q_t^+ \right) \leq \frac{(r+d)!}{r!(r+d)^d} v_c^{(r)} - h_{\mathcal{Q}}^{\min(\Delta_n(r))} \tag{70}$$

$$\leq \left( \sum_{t \in [0:(d-1)] \cap \mathbb{N}_{Even}} q_t^+ + \sum_{t \in [0:(d-1)] \cap \mathbb{N}_{Odd}} q_t^- \right) \tag{71}$$

**Proof.** From the expression given in Equation (66),

$$\frac{(r+d)!}{r!(r+d)^d} v_c^{(r)} = \min_{x \in \Delta_n(r)} \left\{ \langle \mathcal{Q}, x^d \rangle + \left\langle \mathcal{Q}, \sum_{t=0}^{(d-1)} \frac{x^t}{(r+d)^{d-t}} \right\rangle \right\} \tag{72}$$

$$= \min_{x \in \Delta_n(r)} \left\{ \langle \mathcal{Q}, x^d \rangle + \sum_{t=0}^{(d-1)} \frac{\left\{ \sum_{\theta \in \mathbb{I}^n(d)} c(\theta) \mathcal{Q}_{\theta} x_{\theta}^t \right\}}{(r+d)^{d-t}} \right\} \tag{73}$$

$$= \min_{x \in \Delta_n(r)} \left\{ \langle \mathcal{Q}, x^d \rangle + \sum_{t=0}^{(d-1)} \frac{\left\{ \sum_{\substack{\alpha \leq \theta \\ \theta \in \mathbb{I}^n(d) \ \& \ \alpha \in \mathbb{I}^n(t)}} c(\theta) \mathcal{Q}_{\theta} x^{\alpha} \right\}}{(-1)^t (r+d)^{d-t}} \right\} \tag{74}$$

Notice that  $x \in \Delta_n$  then for all  $\alpha \in \mathbb{I}^n(d)$ , we have  $x^{\alpha} \leq 1$ . From this observation, we have that  $q_t(x)^{\pm} \leq q_t^{\pm}$ . This observation leads to the following,

$$\frac{(r+d)!}{r!(r+d)^d} v_c^{(r)} = \min_{x \in \Delta_n(r)} \left\{ \langle \mathcal{Q}, x^d \rangle + \sum_{t=0}^{(d-1)} q_t(x) \right\} \tag{75}$$

$$= \min_{x \in \Delta_n(r)} \left\{ \langle \mathcal{Q}, x^d \rangle + \left( \begin{array}{l} \sum_{t \in [0:(d-1)] \cap \mathbb{N}_{Even}} q_t(x)^+ + \sum_{t \in [0:(d-1)] \cap \mathbb{N}_{Odd}} q_t(x)^- \\ - \left( \sum_{t \in [0:(d-1)] \cap \mathbb{N}_{Even}} q_t(x)^- + \sum_{t \in [0:(d-1)] \cap \mathbb{N}_{Odd}} q_t(x)^+ \right) \end{array} \right) \right\} \tag{76}$$

$$\leq \min_{x \in \Delta_n(r)} \left\{ \langle \mathcal{Q}, x^d \rangle + \left( \sum_{t \in [0:(d-1)] \cap \mathbb{N}_{Even}} q_t(x)^+ + \sum_{t \in [0:(d-1)] \cap \mathbb{N}_{Odd}} q_t(x)^- \right) \right\} \tag{77}$$

$$\leq \min_{x \in \Delta_n(r)} \left\{ \langle \mathcal{Q}, x^d \rangle + \left( \sum_{t \in [0:(d-1)] \cap \mathbb{N}_{Even}} q_t^+ + \sum_{t \in [0:(d-1)] \cap \mathbb{N}_{Odd}} q_t^- \right) \right\} \tag{78}$$

$$\leq h_{\mathcal{Q}}^{\min(\Delta_n(r))} + \sum_{t \in [0:(d-1)] \cap \mathbb{N}_{Even}} q_t^+ + \sum_{t \in [0:(d-1)] \cap \mathbb{N}_{Odd}} q_t^- \tag{79}$$

Thus, we have

$$\frac{(r+d)!}{r!(r+d)^d} v_c^{(r)} - h_{\mathcal{Q}}^{\min(\Delta_n(r))} \leq \sum_{t \in [0:(d-1)] \cap \mathbb{N}_{Even}} q_t^+ + \sum_{t \in [0:(d-1)] \cap \mathbb{N}_{Odd}} q_t^- \tag{80}$$

The lower bound could be done similar manner.  $\square$

### 6. Conclusions

The paper was focused on describing the copositive tensor cone and its approximations. We have shown that every Z-tensor is copositive if and only if it is positive definite. The result has appeared already in [32] (Theorem 3.5(e) and Theorem 3.12). However, the proof given by Zhang et al. relies heavily on the notion of M-tensor and convex analysis. The proof we have provided is simpler and self-contained. We had discussed some approximation hierarchies for the copositive cone, focusing on providing a compact



representation of these hierarchies. For the Parrilo cone,  $\mathcal{K}_{n,d}^r$  the proof techniques are a straightforward generalization to a high dimensional case. For the cone,  $\mathcal{C}_{n,d}^r$  a more rigorous approach is used to derive the representation. Most notions used are unique and have not appeared in the literature. We have illustrated this by applying our compact representation to the polynomial optimization over the simplex. We have compared the approximation obtained by our representation  $\mathcal{C}_{n,d}^r$  with the approximation based on the rational gridding. The bounds are proved between the two approximations. Moreover, the characterization helped to simplify the proofs and results related to approximating polynomial optimization over the simplex. In future it would be interesting to investigate the convergence rate of these approximations.

In the future, we work towards utilizing these hierarchies for providing approximation results related to copositive optimization, especially to recover approximation results for polynomial optimization over the simplex as obtained by De Klerk and co-authors [21–24]. Furthermore, our aim is to use these approximation hierarchies to develop numerical algorithms for application domains such as approximating clique numbers for uniform hypergraphs (see, e.g., [37,38]).

**Author Contributions:** Writing—original draft, M.F.I.; Writing—review & editing, F.A. All authors have read and agreed to the published version of the manuscript

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Maricic, B.; Luo, Z.; Davidson, T. Blind constant modulus equalization via convex optimization. *IEEE Trans. Signal Process.* **2003**, *51*, 805–818. [\[CrossRef\]](#)
2. Sidiropoulos, N.D.; De Lathauwer, L.; Fu, X.; Huang, K.; Papalexakis, E.E.; Faloutsos, C. Tensor decomposition for signal processing and machine learning. *IEEE Trans. Signal Process.* **2017**, *65*, 3551–3582. [\[CrossRef\]](#)
3. Weiland, S.; van Belzen, F. Singular Value Decompositions and Low Rank Approximations of Tensors. *Signal Process. IEEE Trans.* **2010**, *58*, 1171–1182. [\[CrossRef\]](#)
4. Cohen, N.; Sharir, O.; Shashua, A. On the expressive power of deep learning: A tensor analysis. In Proceedings of the Conference on Learning Theory, New York, NY, USA, 23–26 June 2016; pp. 698–728.
5. Kolda, T.G.; Bader, B.W. Tensor decompositions and applications. *SIAM Rev.* **2009**, *51*, 455–500. [\[CrossRef\]](#)
6. Soare, S.; Yoon, J.W.; Cazacu, O. On the use of homogeneous polynomials to develop anisotropic yield functions with applications to sheet forming. *Int. J. Plast.* **2008**, *24*, 915–944. [\[CrossRef\]](#)
7. Michelli, C.A.; Olsen, P. Penalized maximum-likelihood estimation, the Baum-Welch algorithm, diagonal balancing of symmetric matrices and applications to training acoustic data. *J. Comput. Appl. Math.* **2000**, *119*, 301–331. [\[CrossRef\]](#)
8. Hamadneh, T.; Ali, M.; AL-Zoubi, H. Linear optimization of polynomial rational functions: Applications for positivity analysis. *Mathematics* **2020**, *8*, 283. [\[CrossRef\]](#)
9. Henrion, D.; Garulli, A. *Positive Polynomials in Control*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2005; Volume 312.
10. Berman, A.; Shaked-Monderer, N. *Completely Positive Matrices*; World Scientific Publishing Company Pte Limited: Singapore, 2003.
11. Dickinson, P.J.C. The Copositive Cone, the Completely Positive Cone and Their Generalisations. Ph.D. Thesis, Groningen University, Groningen, The Netherlands, 2013.
12. Kostyukova, O.; Tchemisova, T. Structural Properties of Faces of the Cone of Copositive Matrices. *Mathematics* **2021**, *9*, 2698. [\[CrossRef\]](#)
13. Avi Berman, M.D.; Shaked-monderer, N. Open Problems in the Theory of Completely Positive and Copositive Matrices. *Electron. J. Linear Algebra* **2015**, *29*, 46–58. [\[CrossRef\]](#)
14. Song, Y.; Qi, L. Necessary and sufficient conditions for copositive tensors. *Linear Multilinear Algebra* **2015**, *63*, 120–131. [\[CrossRef\]](#)
15. Qi, L. Symmetric nonnegative tensors and copositive tensors. *Linear Algebra Its Appl.* **2013**, *439*, 228–238. [\[CrossRef\]](#)
16. Qi, L.; Luo, Z. *Tensor Analysis: Spectral Theory and Special Tensors*; SIAM: Philadelphia, PA, USA, 2017.
17. Parrilo, P.A. Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization. Ph.D. Thesis, California Institute of Information Technology, Pasadena, CA, USA, 2000.
18. de Klerk, E.; Pasechnik, D.V. Approximation of the Stability Number of a Graph via Copositive Programming. *SIAM J. Optim.* **2002**, *12*, 875–892. [\[CrossRef\]](#)
19. Bomze, I.M.; De Klerk, E. Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. *J. Glob. Optim.* **2002**, *24*, 163–185. [\[CrossRef\]](#)

20. Nesterov, Y. Global quadratic optimization on the sets with simplex structure. In *LIDAM Discussion Papers CORE 1999015*; Université Catholique de Louvain, Center for Operations Research and Econometrics (CORE): Louvain-la-Neuve, Belgium, 1999.
21. de Klerk, E.; Laurent, M.; Parrilo, P.A. A PTAS for the minimization of polynomials of fixed degree over the simplex. *Theor. Comput. Sci.* **2006**, *361*, 210–225. [[CrossRef](#)]
22. de Klerk, E.; Laurent, M.; Sun, Z. An alternative proof of a PTAS for fixed-degree polynomial optimization over the simplex. *Math. Program.* **2015**, *151*, 433–457. [[CrossRef](#)]
23. de Klerk, E.; Laurent, M.; Sun, Z. An Error Analysis for Polynomial Optimization over the Simplex Based on the Multivariate Hypergeometric Distribution. *SIAM J. Optim.* **2015**, *25*, 1498–1514. [[CrossRef](#)]
24. de Klerk, E.; Laurent, M.; Sun, Z.; Vera, J.C. On the convergence rate of grid search for polynomial optimization over the simplex. *Optim. Lett.* **2017**, *11*, 597–608. [[CrossRef](#)]
25. Ling, C.; He, H.; Qi, L. Improved approximation results on standard quartic polynomial optimization. *Optim. Lett.* **2017**, *11*, 1767–1782. [[CrossRef](#)]
26. Hu, S.; Li, G.; Qi, L. A tensor analogy of Yuan’s theorem of the alternative and polynomial optimization with sign structure. *J. Optim. Theory Appl.* **2016**, *168*, 446–474. [[CrossRef](#)]
27. Luo, Z.; Qi, L.; Ye, Y. Linear operators and positive semidefiniteness of symmetric tensor spaces. *Sci. China Math.* **2015**, *58*, 197–212. [[CrossRef](#)]
28. Rockafellar, R. *Convex Analysis*; Princeton Landmarks in Mathematics and Physics; Princeton University Press: Princeton, NJ, USA, 1970.
29. Ahmadi, A.A.; Parrilo, P.A. A convex polynomial that is not SOS-convex. *Math. Program.* **2012**, *135*, 275–292. [[CrossRef](#)]
30. Peña, J.; Vera, J.C.; Zuluaga, L.F. Completely positive reformulations for polynomial optimization. *Math. Program.* **2015**, *151*, 405–431. [[CrossRef](#)]
31. Ahmed, F. Copsitive Programming and Related Problems. Ph.D. Thesis, University of Twente, Twente, The Netherlands, 2014.
32. Zhang, L.; Qi, L.; Zhou, G. M-tensors and some applications. *SIAM J. Matrix Anal. Appl.* **2014**, *35*, 437–452. [[CrossRef](#)]
33. Reznick, B. Uniform denominators in Hilbert’s seventeenth problem. *Math. Z.* **1995**, *220*, 75–97. [[CrossRef](#)]
34. Chen, H.; Wang, Y.; Zhou, G. High-order sum-of-squares structured tensors: Theory and applications. *Front. Math. China* **2020**, *15*, 255–284. [[CrossRef](#)]
35. Ahmadi, A.A.; Majumdar, A. DSOS and SDSOS optimization: More tractable alternatives to sum of squares and semidefinite optimization. *SIAM J. Appl. Algebra Geom.* **2019**, *3*, 193–230. [[CrossRef](#)]
36. Graham, R.L.; Knuth, D.E.; Patashnik, O. *Concrete Mathematics: A Foundation for Computer Science*; Addison-Wesley: Boston, MA, USA, 1994.
37. Ahmed, F.; Still, G. Two methods for the maximization of homogeneous polynomials over the simplex. *Comput. Optim. Appl.* **2021**, *80*, 523–548. [[CrossRef](#)]
38. Rota Bulò, S.; Pelillo, M. A generalization of the Motzkin–Straus theorem to hypergraphs. *Optim. Lett.* **2009**, *3*, 287–295. [[CrossRef](#)]