Symmetric tensor approximation hierarchies for the completely positive cone

Hongbo Dong[†]

Revised: November 9, 2012

Abstract

In this paper we construct two approximation hierarchies for the completely positive cone based on symmetric tensors. We show that one hierarchy corresponds to dual cones of a known polyhedral approximation hierarchy for the copositive cone, and the other hierarchy corresponds to dual cones of a known semidefinite approximation hierarchy for the copositive cone. As an application, we consider a class of bounds on the stability number of a graph obtained from the polyhedral approximation hierarchy, and we construct a primal optimal solution with its tensor lifting for each of such linear programs.

Keywords: Relaxation, Completely Positive cone, Tensor

Mathematics Subject Classification: 90C05, 90C22, 90C25, 90C27

1 Introduction

An $n \times n$ real symmetric matrix X is called *completely positive* if and only if it has a factorization $X = \sum_i x_i x_i^T$ where $x_i \in \mathbb{R}^n_+$. An $n \times n$ real symmetric matrix A is called

[†]`Wisconsin Institutes for Discovery, University of Wisconsin-Madison. E-mail: hdong6@wisc.edu.`

copositive if $x^T Ax \geq 0$, $\forall x \in \mathbb{R}^n_+$, for $i = 1, ..., n$. We use $\mathcal{C}_n$ to denote the cone of $n \times n$ completely positive matrices. It is well known that its dual cone $\mathcal{C}_n^*$ is the cone of $n \times n$ copositive matrices. Both $\mathcal{C}_n$ and $\mathcal{C}_n^*$ are pointed, closed, and full-dimensional in the space of $n \times n$ real symmetric matrices $\mathcal{S}_n$.

The cones $\mathcal{C}_n$ and $\mathcal{C}_n^*$ have received much attention in recent years. One reason is that linear programs over these cones, usually referred as copositive programs, can be used to formulate many hard optimization problems exactly. Examples include (nonconvex) quadratic programs over the standard simplex [BDdK$^+$00, BK02], the stability number, the (fractional) chromatic number of a graph [dKP02, PVZ07, DR08, GL08], and the quadratic assignment problem [PR09]. Burer [Bur09] proved that a large class of deterministic (nonconvex) quadratic programming problems, possibly including binary variables and complementary constraints, have exact copositive representations. Another recent example is mixed binary linear programs with objective uncertainty (first two moments known) [NTZ11]. Also, copositive programs can be used to derive relaxations of the general quadratic constrained quadratic programming problem [BDdK$^+$00, Ans12]. Therefore any new knowledge about completely positive or copositive cone can be applied to all these problems. For a recent survey of copositive programming, see [Dür10]. Another survey emphasizing on copositive matrices is [HUS10].

The fact that many NP-hard problems have copositive reformulations implies that working directly with the copositive cone or completely positive cone must be extremely difficult. Even testing whether a given matrix is copositive is co-NP-complete [MK87], and the membership problem for the completely positive cone is recently shown to be NP-hard [DG11].

In the literature, several inner-approximation hierarchies exist for the copositive cone $\mathcal{C}_n^*$. These hierarchies are sequences of polyhedral or semidefinite representable cones. In other words, each of these approximation cones is computationally tractable. Usually these approximation hierarchies lead to polynomial-time approximation schemes or sequences of computable bounds for specific optimization problems. For example, by using this approach, in [BK02] the authors developed a polynomial-time approximation scheme for the standard quadratic programming problem. In [PVZ07], the authors studied properties of copositive motivated LP/SDP upper bounds for the stability number of a graph.

This paper proceeds as follows: After introducing standard notation in Section 2, we review some existing approximation hierarchies for $\mathcal{C}_n^*$ in Section 3. Then in Section 5, we construct two outer approximation hierarchies for $\mathcal{C}_n$ based on symmetric tensors. One of them is a sequence of polyhedral cones and the other a sequence of semidefinite representable cones. In section 6, we derive alternative characterizations of different hierarchies and establish the duality relations between our approximation hierarchies, which are outer-approximations of $\mathcal{C}_n$ and two known inner-approximation hierarchies for $\mathcal{C}_n^*$. Although the explicit characterization of the polyhedral hierarchy first appeared in [BK02], the tensor structure on the completely positive side have not been realized. As a result of the duality relation, any matrix which is not completely positive is excluded by our outer approximation hierarchies with a sufficiently large order. As an application, in Section 8 we consider the linear programming bounds for the stability number of a graph. We give a new combinatorial proof for a known result in [PVZ07]. We also explicitly construct a primal optimal solution with its tensor lifting corresponding to each of these bounds.

We comment that results in this paper not only describe the dual cones of two well-known outer-approximation hierarchies for $\mathcal{C}_n^*$, and provide a more intuitive way to understand these approximation hierarchies, but also provide new insights for copositive programming and its approximations. The primal (completely positive) formulations usually correspond to original decision variables in a more direct way. On the algorithmic side, this brings up an interesting question whether one could exploit this tensor structure to design more efficient algorithms to compute copositive programming motivated bounds for NP-hard problems.

2 Notation

We define the notation used throughout the rest of this paper. We use $\mathcal{M}_n$ to denote the set of all $n \times n$ real matrices and $\mathcal{S}_n$ to denote the set of all $n \times n$ symmetric matrices.

$\mathcal{N}_n \subseteq \mathcal{S}_n$ is the set of symmetric matrices with nonnegative entries. $\mathcal{S}_n^+$ represents the set of all $n \times n$ symmetric positive semidefinite matrices and $\mathcal{D}_n$ all $n \times n$ doubly nonnegative matrices, i.e., $\mathcal{D}_n = \mathcal{S}_n^+ \cap \mathcal{N}_n$. $\mathbb{Z}$ is the set of all integers. $\mathbb{Z}_+$ denotes all nonnegative integers, and $\mathbb{Z}_+^n$ all n-vectors with all entries in $\mathbb{Z}_+$. The inner product on $\mathcal{M}_n$ is defined as usual, for $A, B \in \mathcal{M}_n$, $\langle A, B \rangle = A \bullet B = \mathbf{trace}(AB^T)$. The Hadamard product on $\mathcal{M}_n$ is denoted by $\circ$, i.e., $(A \circ B)_{ij} = A_{ij} B_{ij}$.

To simplify our later analysis with polynomials, we use a pair of vectors, one as a superscript of the other, to represent a monomial; for example, x^m. Often we use x to denote a vector of variables used in a polynomial, whose dimension is usually n, and $m \in \mathbb{Z}_+^n$ is a vector of the same dimension used to represent the multiplicity of each variable. For example, if $x = [x_1, x_2, x_3]^T$ and $m = [0, 4, 2]^T$, then x^m means the monomial $x_2^4 x_3^2$. The *degree* of a monomial is the summation of the multiplicities of all variables. In the previous example, the degree of x^m (denoted by $|m|$) is $4 + 2 = 6$. The degree of a polynomial is the degree of its monomial term of highest degree, and a polynomial is *homogeneous* if each of its monomial terms has the same degree.

In later discussions, we are especially interested in monomials of a fixed degree. We use $\mathbb{I}^n(r)$ to denote all possible exponent vectors for n-variate monomials of degree r, i.e.,

$$\mathbb{I}^n(r) := \left\{ m \in \mathbb{Z}_+^n \;\middle|\; |m| = \sum_{i=1}^{n} m_i = r \right\}.$$

Usually we use the letter m or p to denote an element in $\mathbb{I}^n(r)$.

3 Approximation hierarchies for the copositive cone

In this section, we review some known approximation hierarchies for $\mathcal{C}_n^*$. Construction of such hierarchies is usually based on approximating nonnegative polynomials. A polynomial $P(x)$, where $x = [x_1, ..., x_n]^T$, is said to be *nonnegative* if $P(x) \geq 0, \forall x \in \mathbb{R}^n$. It is easy to see that a polynomial with odd degree is never nonnegative.

The starting point of constructing approximations for the copositive cone is usually the following observation, which relates the copositivity of a matrix with the nonnegativity of a homogeneous polynomial. Let

$$P^{(0)}(x) := (x \circ x)^T M(x \circ x) = \sum_{i=1}^{n} M_{ij} x_i^2 x_j^2 \tag{1}$$

Observation. An $n \times n$ symmetric matrix M is copositive if and only if its corresponding polynomial $P^{(0)}(x)$ is nonnegative.

An obvious sufficient condition for $P^{(0)}(x)$ to be nonnegative is that all its coefficients are nonnegative. This leads to the following inner approximation of the copositive cone $\mathcal{C}_n^*$:

$$\mathcal{C}_n^0 := \left\{ M \mid P^{(0)}(x) \text{ as in (1) has nonnegative coefficients} \right\}.$$

Since the coefficients in (1) are just entries in M, $\mathcal{C}_n^0 = \mathcal{N}_n$.

For a general homogeneous polynomial $p(x)$ of degree $2d$, a sufficient condition for its nonnegativity is that it can be represented as a sum of squares (s.o.s), i.e., $\exists$ polynomials $h_i(x)$, $i = 1, ..., l$ such that $p(x) = \sum_{i=1}^{l} h_i(x)^2$. It is known that we can assume without loss of generality that $h_i(x)$ is homogeneous and of degree d [BK02, Lemma 2.1].

By using this s.o.s. sufficient condition, another inner approximation of the copositive cone $\mathcal{C}_n^*$ can be constructed as follows:

$$\mathcal{K}_n^0 := \left\{ M \mid P^{(0)}(x) \text{ is an s.o.s., where } P^{(0)}(x) \text{ is defined in (1)} \right\}.$$

It was proved by Parrilo [Par00] that

$$\mathcal{K}_n^0 = \mathcal{D}_n^* = \mathcal{S}_n^+ + \mathcal{N}_n.$$

To obtain a full approximation hierarchy, for $r \in \mathbb{Z}_+$, define the polynomial $P^{(r)}(x)$ as follows:

$$P^{(r)}(x) := P^{(0)}(x) \left(\sum_{i=1}^{n} x_i^2 \right)^r = \sum_{i,j=1}^{n} M_{ij} x_i^2 x_j^2 \left(\sum_{i=1}^{n} x_i^2 \right)^r. \tag{2}$$

Then as in [Par00],[dKP02], two sequences of approximation sets may be defined as:

$$\begin{aligned} \mathcal{C}_n^r &:= \left\{ M \mid P^{(r)}(x) \text{ has nonnegative coefficients} \right\}, \\ \mathcal{K}_n^r &:= \left\{ M \mid P^{(r)}(x) \text{ is an s.o.s.} \right\}. \end{aligned}$$

It is easy to prove by definition that $\mathcal{C}_n^r \subseteq \mathcal{K}_n^r$, $\forall r \in \mathbb{Z}_+$. Further for fixed n, these two sequences provide two inner approximation hierarchies for the copositive cone:

$$\mathcal{N}_n = \mathcal{C}_n^0 \subseteq \mathcal{C}_n^1 \subseteq \cdots \subseteq \mathcal{C}_n^r \subseteq \mathcal{C}_n^{r+1} \subseteq \cdots \subseteq \mathcal{C}_n^*,$$
$$\mathcal{D}_n^* = \mathcal{K}_n^0 \subseteq \mathcal{K}_n^1 \subseteq \cdots \subseteq \mathcal{K}_n^r \subseteq \mathcal{K}_n^{r+1} \subseteq \cdots \subseteq \mathcal{C}_n^*.$$

Obviously $\mathcal{K}_n^r$ is a convex cone, and in fact also closed. To see this, we introduce the notation $\mathbb{R}[x]_{n,d}$ to denote the set of all n-variate homogeneous polynomials of degree d, and $\Sigma_{n,2d}$ to denote all s.o.s polynomials in $\mathbb{R}[x]_{n,2d}$. The topology on $\mathbb{R}[x]_{n,d}$ is defined by treating the coefficients of each polynomial as a real vector, with a fixed ordering of all monomials, then $\Sigma_{n,2d}$ is a closed convex cone [PS01, Sch05, Lau09]. First note the mapping $\mathcal{S}_n \longmapsto \mathbb{R}[x]_{n,2r+4}$ that maps M to $P^{(r)}(x)$ as in (2), is linear. Hence its image is a (finite dimensional) linear subspace. Now as $\Sigma_{n,2r+4}$ is closed in $\mathbb{R}[x]_{n,2r+4}$, $\mathcal{K}_n^r$ is the preimage of a closed set in $\mathbb{R}[x]_{n,2r+4}$, hence closed.

In [PVZ07], the authors introduced a new class of semidefinite inner approximations $\{\mathcal{Q}_n^r\}_{r=0}^\infty$ for the copositive cone $\mathcal{C}_n^*$. To simplify notation we use z to represent the vector $(x \circ x)$, i.e., for $m \in \mathbb{I}^n(r)$, $z^m = x^{2m}$. A set of homogeneous polynomials of degree $r+2$, $\mathcal{E}_{n,r}$, is defined as:

$$\mathcal{E}_{n,r} := \left\{ \sum_{p \in \mathbb{I}^n(r)} z^p z^T (P_p + N_p) z \,\middle|\, \forall p \in \mathbb{I}^n(r), P_p \in \mathcal{S}_n^+, N_p \in \mathcal{N}_n \right\}.$$

Then the approximation cone $\mathcal{Q}_n^r$ is:

$$\mathcal{Q}_n^r := \left\{ M \,\middle|\, \left(\sum_{i=1}^n z_i \right)^r z^T M z \in \mathcal{E}_{n,r} \right\}. \tag{3}$$

It is straightforward to verify that $\mathcal{Q}^r$ is a convex cone. In fact it is also closed. In [PVZ07] it was claimed without proof that the closedness of $\mathcal{Q}^r$ follows a limiting argument [1]. We outline an alternative proof here by using a theorem in convex analysis.

The key step is to show that $\mathcal{E}_{n,r}$ is a closed subset of $\mathbb{R}[z]_{n,r+2}$. Then $\mathcal{Q}_n^r$ is the preimage of $\mathcal{E}_{n,r}$ under a linear operator, and the closedness of $\mathcal{Q}_n^r$ follows. We prove the closedness of $\mathcal{E}_{n,r}$ by using the following theorem.

[1] However, after private communication with the authors, it appears the limiting proof is not trivial.

Theorem 1. *[Roc70, Corollary 9.1.3] Let $\{K_p\}_{p\in\mathcal{I}}$ be a finite collection of non-empty convex cones in a Euclidean space satisfying the following condition: if $\tau_p \in \mathbf{cl}(K_p), \forall p \in \mathcal{I}$ and $\sum_{p\in\mathcal{I}} \tau_p = 0$, then τ_p belongs to the lineality space of $\mathbf{cl}(K_p), \forall p \in \mathcal{I}$. Then*

$$\mathbf{cl}\left(\sum_{p\in\mathcal{I}} K_p\right) = \sum_{p\in\mathcal{I}} \mathbf{cl}\left(K_p\right).$$

To apply this theorem, let $K_p = \left\{z^p z^T(P+N)z : P \in \mathcal{S}_n^+, N \in \mathcal{N}_n\right\}$ and $\mathcal{I} = \mathbb{I}^n(r)$. Then $\mathcal{E}_{n,r} = \sum_{p\in\mathcal{I}} K_p$. Note each K_p is in a subspace of $\mathbb{R}[z]_{n,r+2}$ that is isomorphic to $\mathcal{S}_n$, and under this isomorphism K_p corresponds to $\mathcal{D}_n^*$, so K_p is closed. Now we verify the condition in Theorem 1. Take $\tau_p = z^p z^T(P_p + N_p)z \in K_p$ such that

$$\sum_{p\in\mathbb{I}^n(r)} z^p z^T(P_p + N_p)z = 0.$$

Then for *any* positive vector $z > 0$, $z^T(P_p + N_p)z = 0$, $\forall p \in \mathbb{I}^n(r)$. Further by continuity and convexity, $\langle P_p + N_p, X\rangle = 0, \forall X \in \mathcal{C}_n, \forall p \in \mathbb{I}^n(r)$. Since $\mathcal{C}_n$ is a full-dimensional cone in $\mathcal{S}_n$, $P_p + N_p = 0, \forall p$. Therefore the result of Theorem 1 implies $\mathbf{cl}\left(\mathcal{E}_{n,r}\right) = \sum_{p\in\mathcal{I}} K_p = \mathcal{E}_{n,r}$. Then by previous discussion $\mathcal{Q}_n^r$ is closed.

It is easy to verify by definitions that for any $r \in \mathbb{Z}_+$,

$$\mathcal{C}_n^r \subseteq \mathcal{Q}_n^r \subseteq \mathcal{K}_n^r. \tag{4}$$

Furthermore, as shown in [Par00] and [dKP02] by using a theorem of Pólya ([HLP67], Section 2.24), the linear programming approximation hierarchy is "almost exact" in the limit:

Theorem 2 ([Par00],[dKP02])**.** *Let M be a strictly copositive matrix, i.e., $y^T My > 0$ for all $y \in \mathbb{R}_+^n, e^T y = 1$. Then there exists a finite $R \in \mathbb{Z}_+$ such that $M \in \mathcal{C}_n^R$.*

It then follows from (4) and Theorem 2 that the hierarchies $\{Q_n^r\}_{r\in\mathbb{Z}_+}$ and $\{\mathcal{K}_n^r\}_{r\in\mathbb{Z}_+}$ are also "almost exact" in the limit.

4 Tensor Operators and Terminology

In this section we introduce some operators and terminology related to tensors. The term "tensor" we use throughout this paper really means hypermatrix. In each tensor of order r

and dimension n, one needs r indices to specify one entry, where each index takes integral value in $\{1, 2, ..., n\}$. In particular, a tensor of order 1 is a vector, and a tensor of order 2 is a square matrix. We say a tensor is *symmetric* if the values of its entries are independent of permutation of its indices. For example, a symmetric tensor of order 2 is simply a symmetric matrix. We use $\mathcal{M}_n^r$ to denote all tensors of order r and dimension n, and $\mathcal{S}_n^r$ to denote all such symmetric tensors. Note that for square matrices, $\mathcal{M}_n^2 = \mathcal{M}_n$ and $\mathcal{S}_n^2 = \mathcal{S}_n$.

Recall that we use an integral vector to represent the exponents of a monomial. Here we introduce another kind of multi-index notation to index entries in a tensor. In later analysis we will heavily use these two kinds of multi-index notations. It is important to understand the difference and connection between these notations.

To index entries in a tensor $T \in \mathcal{M}_n^r$, we use an r-dimension row vector whose coordinates are all integers between 1 and n, enclosed in a pair of box brackets. For example, for $T \in \mathcal{M}_n^3$, $T[1, 3, 2]$ is the entry at position 1 in the first dimension, position 3 in the second dimension, and position 2 in the last dimension. We use $\mathbb{N}^r(n)$ to denote all such indexing vectors:

$$\mathbb{N}^r(n) := \{\alpha \in \mathbb{Z}_+^r \mid 1 \leq \alpha_i \leq n, i = 1, ..., r\}.$$

Lower-case Greek letters are used to refer to elements in $\mathbb{N}^r(n)$. For example if $T \in \mathcal{M}_n^r$ and $\alpha \in \mathbb{N}^r(n)$, we use $T[\alpha]$ to refer an entry in T. The inner product over $\mathcal{M}_n^r$ is defined as

$$\langle T, Z \rangle := \sum_{\alpha \in \mathbb{N}^r(n)} T[\alpha] Z[\alpha]. \ \ \forall T, Z \in \mathcal{M}_n^r.$$

For $r > 0$, $\beta \in \mathbb{N}^r(n)$ and $T \in \mathcal{M}_n^{r+2}$, we use $T[\beta, :, :]$ to refer to an ordinary matrix obtained by fixing the first r indices of T as β. Since often the last two indices are special in our analysis, sometimes we use T^β to denote $T[\beta, :, :]$, and T_{ij}^β to denote $T[\beta, i, j]$, where $1 \leq i, j \leq n$.

For a symmetric tensor $T \in \mathcal{S}_n^r$, $T[\alpha] = T[\beta]$ if α and β are permutationally the same, i.e., there exists a permutation τ such that $\tau(\alpha) = \beta$. So we can define (permutational) equivalence classes in $\mathbb{N}^r(n)$ in this way, and we use $[\![\alpha]\!]$ to denote the equivalence class that includes α. One can define a bijection between all equivalence classes in $\mathbb{N}^r(n)$ and all elements in $\mathbb{I}^n(r)$. By abuse of notation, we represent this bijection by $[\![\cdot]\!]$. For $\alpha \in \mathbb{N}^r(n)$ and $m \in \mathbb{I}^n(r)$, we say $[\![\alpha]\!] = m$ if number i appears m_i times in α, $i = 1, ..., n$. For example, if $r = 4$ and $n = 5$, $\alpha = [4, 3, 1, 4] \in \mathbb{N}^r(n)$, then $m = [\![\alpha]\!] = [1, 0, 1, 2, 0] \in \mathbb{I}^n(r)$. Another way to interpret this relation is to identify an equivalence class in $\mathbb{N}^r(n)$, $[\![\alpha]\!]$, with the monomial x^m. In particular $x^m = \Pi_{i=1}^r x^{\beta(i)}, \forall \beta \; s.t. \; [\![\beta]\!] = [\![\alpha]\!]$. In the previous example, $x^{[1,0,1,2,0]} = x_4 x_3 x_1 x_4 = x_1 x_4 x_3 x_4 = \cdots$.

Remark 1. It is often beneficial to remember that the supscripts in $\mathbb{N}^r(n)$ and $\mathbb{I}^n(r)$ represent the length (number of dimensions) of vectors in these sets. In particular, an element in $\mathbb{N}^r(n)$ is a vector of length r and an element in $\mathbb{I}^n(r)$ is a vector of length n.

We next define several tensor operators that are convenient in our later analysis. First, given a tensor $T \in \mathcal{M}_n^{r+2}$, we define $\mathbf{Slices}(T)$ to be the set of $n \times n$ matrices which are "slices" of T obtained by fixing its first r indices, i.e., a matrix $P \in \mathbf{Slices}(T)$ iff $\exists \beta \in \mathbb{N}^r(n)$ such that $P = T^\beta = T[\beta, :, :]$.

Another operator $\mathbf{Collapse} : \mathcal{M}_n^{r+2} \longrightarrow \mathcal{M}_n^2$ is defined as:

$$\mathbf{Collapse}(T) := \sum_{P \in \mathbf{Slices}(T)} P \; and \; \mathbf{Collapse}(T)[i,j] = \sum_{\beta \in \mathbb{N}^r(n)} T[\beta, i, j] = \sum_{\beta \in \mathbb{N}^r(n)} T_{ij}^\beta.$$

Given a matrix $P \in \mathcal{M}_n^2$, we define an operator $\mathbf{Stack}_r(P)$ that "stacks" P for n layers along r dimensions; precisely, $\mathbf{Stack}_r(P) \in \mathcal{M}_n^{r+2}$ and

$$\mathbf{Stack}_r(P)[\beta, i, j] = P_{ij}, \quad \forall \beta \in \mathbb{N}^r(n).$$

Since the value r associated with the operator is usually clear in context, we usually omit the subscript and write $\mathbf{Stack}(\cdot)$ instead of $\mathbf{Stack}_r(\cdot)$.

Actually, $\mathbf{Collapse}(\cdot)$ and $\mathbf{Stack}(\cdot)$ are a pair of adjoint operators, i.e., $\forall T \in \mathcal{M}_n^{r+2}$, $\forall P \in \mathcal{M}_n^2$,

$$\langle \mathbf{Collapse}(T), P \rangle = \sum_{\beta \in \mathbb{N}^r(n)} \langle T[\beta, :, :], P \rangle = \langle T, \mathbf{Stack}(P) \rangle,$$

where the inner product in $\mathcal{M}_n^{r+2}$ is defined in the standard way: for $T, \tilde{T} \in \mathcal{M}_n^{r+2}$,

$$\left\langle T, \tilde{T} \right\rangle := \sum_{\alpha \in \mathbb{N}^{r+2}(n)} T[\alpha] \tilde{T}[\alpha].$$

For a square matrix $P \in \mathcal{M}_n^2$, one may symmetrize as $\frac{1}{2}(P + P^T)$. One can similarly symmetrize a tensor. The symmetrization operator $\mathbf{Sym} : \mathcal{M}_n^r \longrightarrow \mathcal{S}_n^r$ is defined as follows:

$$\mathbf{Sym}(T)[\beta] = \sum_{[\![\alpha]\!]=[\![\beta]\!]} T[\alpha] \Big/ c([\![\beta]\!]), \quad \forall \beta \in \mathbb{N}^r(n),$$

where $c([\![\beta]\!])$ is the cardinality of the equivalence class $[\![\beta]\!]$. Recall that an equivalence class $[\![\beta]\!]$ corresponds to an element in $\mathbb{I}^n(r)$, say m, we write $c([\![\beta]\!]) = c(m)$, which is defined by the following formula,

$$c(m) := |\{\alpha \in \mathbb{N}^r(n) : [\![\alpha]\!] = m\}| = \begin{cases} \frac{|m|!}{\prod_{i=1}^n (m_i)!} \ , & m \in \mathbb{I}^n(r); \\ 0 \ , & \textit{otherwise}. \end{cases} \tag{5}$$

Another place where $c(m)$ appears is when expanding the following polynomial:

$$\left(\sum_{i=1}^n z_i \right)^r = \sum_{m \in \mathbb{I}^n(r)} c(m) z^m.$$

For $m \in \mathbb{I}^n(r+2)$ and $r > 0$, we define $m(i,j) := m - e_i - e_j$, where e_i is a vector with 1 at the i-th position and 0's elsewhere, and similarly for e_j. Note that given $m \in \mathbb{I}^n(r+2)$, $m(i,j)$ is not always in $\mathbb{I}^n(r)$, since $m(i,j)$ might contain negative entries. If so, $c(m(i,j)) = 0$ by definition. As an example, let $\alpha = [1,3,3,5] \in \mathbb{N}^4(5)$, then $m = [\![\alpha]\!] = [1,0,2,0,1] \in \mathbb{I}^5(4)$, $c(m(1,2)) = 0$ and $c(m(1,3)) = 2$. Given $T \in \mathcal{M}_5^4$,

$$\mathbf{Sym}(T)[1,3,3,5] = \cdots = \mathbf{Sym}(T)[5,3,3,1] = \frac{1}{12} \left[T[1,3,3,5] + \cdots + T[5,3,3,1] \right].$$

Finally we use $T \simeq_{Sym} \tilde{T}$ to mean $\mathbf{Sym}(T) = \mathbf{Sym}(\tilde{T})$. If $T \simeq_{Sym} \tilde{T}$ and $Z \in \mathcal{S}_n^r$, then $\langle T, Z \rangle = \langle \tilde{T}, Z \rangle$.

5 Symmetric tensor approximation hierarchies for the completely positive cone

This section contains our main construction of two approximation hierarchies for the completely positive cone based on symmetric tensors. We first construct a polyhedral outer approximation hierarchy $\{\mathcal{T}_n^r\}_{r=1}^{\infty}$. Then by adding positive semidefinite conditions we obtain a semidefinite approximation hierarchy $\{\mathcal{TD}_n^r\}_{r=1}^{\infty}$. In the next section, we prove duality relations between these outer approximations and the inner approximations of $\mathcal{C}_n^*$ described in Section 3.

First consider the standard simplex in $\mathbb{R}_n$, $\Delta_n := \{x \in \mathbb{R}_n : e^T x = 1, x \geq 0\}$. For fixed $x \in \Delta_n$, we can construct an "outer-product" tensor $Z \in \mathcal{S}_n^{r+2}$ as:

$$Z[i_1, i_2, i_3, ..., i_{r+2}] := x_{i_1} x_{i_2} \cdots x_{i_{r+2}}, \ \forall \ (i_1, i_2, ..., i_{r+2}) \in \mathbb{N}_n^{r+2}. \tag{6}$$

Fixing $i_1, ..., i_r$, the corresponding slice of Z is $Z[i_1, ..., i_r, :, :] := (x_{i_1} x_{i_2} x_{i_r}) x x^T$, a nonnegative scaling of the matrix xx^T, and hence a completely positive matrix. Since $\sum_{i_1,...,i_r=1}^n x_{i_1} x_{i_2} \cdots x_{i_r} = (x_1 + \cdots + x_n)^r = (e^T x)^r = 1$, collapsing Z yields the matrix xx^T:

$$\textbf{Collapse}(Z) = \sum_{i_1,...,i_r=1}^{n} x_{i_1} \cdots x_{i_r} x x^T = x x^T. \tag{7}$$

Therefore we claim that the completely positive cone is equivalently written as

$$\mathcal{C}_n = \left\{ X \middle| \exists Z \in \mathcal{S}_n^{r+2}, \textbf{Slices}(Z) \subseteq \mathcal{C}_n, X = \textbf{Collapse}(Z) \right\}.$$

The right hand set is a subset of $\mathcal{C}_n$, as **Collapse**(Z) equals the sum of all matrices in **Slices**(Z). Previous discussion shows that all (normalized) extremal rays of $\mathcal{C}_n$ is in the right hand set, hence proves the equality. Then for the sake of computational tractability, we relax the complete positivity of each slice of the tensor Z to either nonnegativity or double nonnegativity, and define the following two sets for $r \in \mathbb{Z}_+$:

$$\mathcal{T}_n^r := \left\{ X \middle| \exists Z \in \mathcal{S}_n^{r+2}, Z \geq 0, X = \textbf{Collapse}(Z) \right\}, \tag{8}$$

$$\mathcal{TD}_n^r := \left\{ X \middle| \exists Z \in \mathcal{S}_n^{r+2}, \textbf{Slices}(Z) \subseteq \mathcal{D}_n, X = \textbf{Collapse}(Z) \right\}. \tag{9}$$

$\mathcal{T}_n^r$ is a closed convex polyhedron since it is the image of a linear transformation of a polyhedron of $\mathcal{S}_n^{r+2}$. $\mathcal{TD}_n^r$ is also a closed convex cone, whose closedness is proved as a by-product of Theorem 4 in the next section. Furthermore, we have the following inclusion relation:

Proposition 1. *For dimension $n \geq 1$,*

$$\mathcal{C}_n \subseteq \cdots \subseteq \mathcal{T}_n^{r+1} \subseteq \mathcal{T}_n^r \subseteq \cdots \subseteq \mathcal{T}_n^1 \subseteq \mathcal{T}_n^0 = \mathcal{N}_n,$$
$$\mathcal{C}_n \subseteq \cdots \subseteq \mathcal{TD}_n^{r+1} \subseteq \mathcal{TD}_n^r \subseteq \cdots \subseteq \mathcal{TD}_n^1 \subseteq \mathcal{TD}_n^0 = \mathcal{D}_n.$$

Proof. By (6) and (7), it is clear that $xx^T \in \mathcal{T}_n^r$ and $xx^T \in \mathcal{TD}_n^r$, $\forall x \in \Delta_n$. Note that both $\mathcal{T}_n^r$ and $\mathcal{TD}_n^r$ are convex cones, so $\mathcal{C}_n \subseteq \mathcal{T}_n^r$ and $\mathcal{C}_n \subseteq \mathcal{TD}_n^r$, $\forall r \in \mathbb{Z}_+$.

We show that $\mathcal{TD}_n^{r+1} \subseteq \mathcal{TD}_n^r$ for all $r > 0$. Suppose $X \in \mathcal{TD}_n^{r+1}$, then by definition $\exists Z \in \mathcal{S}_n^{r+3}$ such that $\mathbf{Slices}(Z) \subseteq \mathcal{D}_n$ and $X = \mathbf{Collapse}(Z)$. Now we define $\tilde{Z}$ by adding up all $(r+2)$-tensors generated by fixing the first index of Z:

$$\tilde{Z}[\alpha] = \sum_{i=1}^{n} Z[i, \alpha], \ \forall \alpha \in \mathbb{N}^{r+2}(n).$$

Since $Z \in \mathcal{S}_n^{r+3}$, it is easy to see $\tilde{Z} \in \mathcal{S}_n^{r+2}$. Further, every matrix in $\mathbf{Slices}(\tilde{Z})$ is a summation of n matrices in $\mathbf{Slices}(Z)$, so $\mathbf{Slices}(\tilde{Z}) \subseteq \mathcal{D}_n$. Finally,

$$\begin{aligned} \mathbf{Collapse}(\tilde{Z}) &= \sum_{\beta \in \mathbb{N}^r(n)} \tilde{Z}[\beta, :, :] = \sum_{\beta \in \mathbb{N}^r(n)} \sum_{i=1}^{n} Z[i, \beta, :, :] \\ &= \sum_{\alpha \in \mathbb{N}^{r+1}(n)} Z[\alpha, :, :] = \mathbf{Collapse}(Z) = X. \end{aligned}$$

Therefore $X \in \mathcal{TD}_n^r$. The proof of $\mathcal{T}_n^{r+1} \subseteq \mathcal{T}_n^r$ is similar. □

6 The duality relations

The main results in this section are theorems that establish duality relations between two pairs of convex cones:

1. $\mathcal{C}_n^r$ and $\mathcal{T}_n^r$ (Theorem 3),

2. $\mathcal{Q}_n^r$ and $\mathcal{TD}_n^r$ (Theorem 4).

First for $\mathcal{T}_n^r$ and $\mathcal{C}_n^r$, we derive explicit characterizations of $\mathcal{T}_n^r$ (Lemma 1) and $\mathcal{C}_n^r$ (Lemma 2). Then the duality between $\mathcal{C}_n^r$ and $\mathcal{T}_n^r$ is clear because these lemmas say that extreme rays of $\mathcal{T}_n^r$ define all the facets of $\mathcal{C}_n^r$.

Lemma 1. *For any nonnegative integer r,*

$$\mathcal{T}_n^r = \left\{ X \middle| X = \sum_{m \in \mathbb{I}^n(r+2)} \lambda_m F_m, \lambda_m \geq 0, \forall m \in \mathbb{I}^n(r+2) \right\},$$

where $F_m \in \mathcal{S}_n$ and $(F_m)_{ij} = c(m(i,j))$.

Proof. First we define a set of "elementary tensors" of order $r+2$: for every $m \in \mathbb{I}^n(r+2)$, define $E_m \in \mathcal{S}_n^{r+2}$ as follows:

$$(E_m)[\alpha] = \begin{cases} 1, & \text{if } [\![\alpha]\!] = m \\ 0, & \text{otherwise.} \end{cases}$$

Then it is easy to verify that $F_m = \textbf{Collapse}(E_m)$, $\forall m \in \mathbb{I}^n(r+2)$. Note $\forall Z \in \mathcal{S}_n^{r+2}$, $Z \geq 0$, Z can be decomposed as a conic combination of these "elementary tensors":

$$Z = \sum_{m \in \mathbb{I}^n(r+2)} \lambda_m E_m, \text{ where } \lambda_m = Z[\alpha] \text{ for } [\![\alpha]\!] = m.$$

Since $\textbf{Collapse}(\cdot)$ is a linear operator, we have

$$\begin{aligned} \mathcal{T}_n^r &= \left\{ X \middle| X = \sum_{m \in \mathbb{I}^n(r+2)} \lambda_m \textbf{Collapse}(E_m), \lambda_m \geq 0, \forall m \in \mathbb{I}^n(r+2) \right\} \\ &= \left\{ X \middle| X = \sum_{m \in \mathbb{I}^n(r+2)} \lambda_m F_m, \lambda_m \geq 0, \forall m \in \mathbb{I}^n(r+2) \right\}. \end{aligned}$$

□

Remark 2. Lemma 1 gives an explicit characterization of the polyhedral cone $\mathcal{T}_n^r$, i.e., every extreme ray of $\mathcal{T}_n^r$ is of form F_m where $m \in \mathbb{I}^n(r+2)$. Note m is a row vector, then some computation shows

$$F_m = \frac{c(m)}{(r+2)(r+1)} \left[m^T m - \textbf{Diag}(m) \right],$$

where $\textbf{Diag}(m)$ is a diagonal matrix with $\textbf{Diag}(m)_{ii} = m_i$.

The following explicit characterization of $\mathcal{C}_n^r$ first appeared as [BK02]. Also see [Yıl12, Section 3]. We include the derivation here for the sake of completeness.

Lemma 2. *[BK02, Theorem 2.4] For any nonnegative integer r,*

$$\mathcal{C}_n^r = \{M \in \mathcal{S}_n \mid \langle F_m, M\rangle \geq 0, \ \forall m \in \mathbb{I}^n(r+2)\}$$

where F_m is defined as in Lemma 1.

Proof.

$$\begin{aligned} P^{(r)}(x) &= \sum_{i,j=1}^n M_{ij}x_i^2x_j^2 \left(\sum_{i=1}^n x_i^2\right)^r &&= \sum_{i,j=1}^n M_{ij}x_i^2x_j^2 \sum_{m\in\mathbb{I}^n(r)} c(m)x^{2m} \\ &= \sum_{\substack{m\in\mathbb{I}^n(r)\\ i,j=1,\ldots,n}} c(m)M_{ij}x^{2m+2e_i+2e_j} &&= \sum_{m\in\mathbb{I}^n(r+2)} \left(\sum_{i,j=1}^n c(m(i,j))M_{ij}\right) x^{2m}. \end{aligned}$$

Hence, by the definition of $\mathcal{C}_n^r$, and $(F_m)_{ij} = c(m(i,j))$,

$$\mathcal{C}_n^r = \{M \in \mathcal{S}_n \mid \langle F_m, M\rangle \geq 0, \ \forall m \in \mathbb{I}^n(r+2)\}.$$

□

Now we are ready to state our first main result:

Theorem 3. *For any nonnegative integer r, the cones $\mathcal{C}_n^r$ and $\mathcal{T}_n^r$ are dual to one another:*

$$(\mathcal{C}_n^r)^* = \mathcal{T}_n^r, \quad (\mathcal{T}_n^r)^* = \mathcal{C}_n^r.$$

Proof. Since both $\mathcal{C}_n^r$ and $\mathcal{T}_n^r$ are finitely generated cones, the result follows immediately from Lemma 1 and Lemma 2. □

Next we derive the duality relation between $\mathcal{TD}_n^r$ and $\mathcal{Q}_n^r$. Again we start by deriving a tensor characterization of the cone $\mathcal{Q}_n^r$.

Lemma 3. *For any nonnegative integer r, $\mathcal{Q}_n^r$ has the following characterization:*

$$\mathcal{Q}_n^r = \left\{M \,\middle|\, \exists T \in \mathcal{M}_n^{r+2}, \mathbf{Stack}(M) \simeq_{Sym} T, \mathbf{Slices}(T) \subseteq \mathcal{D}_n^*\right\}.$$

Proof. First, we claim the following is an alternative characterization of $\mathcal{Q}_n^r$:

$$\mathcal{Q}_n^r = \left\{ M \;\middle|\; \exists T \in \mathcal{M}_n^{r+2}, \left(\sum_{i=1}^n z_i\right)^r z^T M z = \sum_{p \in \mathbb{I}^n(r)} z^p z^T \left(\sum_{[\![\alpha]\!]=p} T^\alpha\right) z, \mathbf{Slices}(T) \subseteq \mathcal{D}_n^* \right\}. \tag{10}$$

To see the equivalence with (3), we identify each matrix $P_p + N_p$ in (3) with the sum of slices of $T \in \mathcal{M}_n^{r+2}$ corresponding to an equivalence class in $\mathbb{N}^r(n)$. From the matrices $\{P_p + N_p\}_{p \in \mathbb{I}^n(r)}$ in (3), a tensor $T \in \mathcal{M}_n^{r+2}$ in (10) is constructed as follows:

$$T^\alpha = T[\alpha, :, :] = \frac{1}{c([\![\alpha]\!])}\left(P_{[\![\alpha]\!]} + N_{[\![\alpha]\!]}\right), \quad \forall \alpha \in \mathbb{N}^r(n).$$

Recall that $[\![\alpha]\!] \in \mathbb{I}^n(r)$ by abusing the notation of the bijection $[\![\cdot]\!]$ between equivalent classes in $\mathbb{N}^r(n)$ and elements in $\mathbb{I}^n(r)$. So T is properly defined. It then follows that $\mathbf{Slices}(T) \subseteq \mathcal{D}_n^*$ and $P_p + N_p = \sum_{[\![\alpha]\!]=p} T^\alpha$. So T satisfies conditions in (10). On the other hand, if we have $T \in \mathcal{M}_n^{r+2}$ with $\mathbf{Slices}(T) \subseteq \mathcal{D}_n^*$, we may construct P_p and N_p such that $P_p + N_p = \sum_{[\![\alpha]\!]=p} T^\alpha \in \mathcal{D}_n^*$.

Next we compare the coefficients of term z^m, for $m \in \mathbb{I}^n(r+2)$, on both sides of the equation in (10). Note that $c(m(i,j))$ is the coefficient of $z^{m(i,j)}$ when expanding $(\sum_{i=1}^n z_i)^r$, and $c(m(i,j)) = 0$ when $m(i,j) \notin \mathbb{I}^n(r)$. So

$$\sum_{i,j=1}^n c(m(i,j)) M_{ij} = \sum_{i,j=1}^n \left(\sum_{[\![\alpha]\!]=m(i,j)} T_{ij}^\alpha\right), \; \forall m \in \mathbb{I}^n(r+2). \tag{11}$$

Finally, (11) is equivalent to:

$$\mathbf{Stack}(M) \simeq_{Sym} T.$$

Actually, for every $m \in \mathbb{I}^n(r+2)$, because $c(m(i,j)) = |\{\alpha \in \mathbb{N}^r(n), [\![(\alpha, i, j)]\!] = m\}|$, the left hand side of (11) equals the summation of all entries in tensor $\mathbf{Stack}(M)$ with multi-index permutationally equivalent to m. Also the right hand side has the same interpretation about tensor T. This proves our lemma. □

We now state a technical lemma which will be useful later in the proof of Theorem 4.

Lemma 4. *Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a convex cone, and $\mathcal{H}$ be a closed convex cone such that $\mathcal{H} \subseteq \mathcal{K}^*$. If for any $X \notin \mathcal{K}$, $\exists H \in \mathcal{H}$, $\langle X, H \rangle < 0$, then $\mathcal{K}$ is closed and $\mathcal{H} = \mathcal{K}^*$.*

Proof. First by assumption, $\mathcal{K}$ is the intersection of (arbitrary many) closed half spaces defined by $\langle \cdot, H \rangle \geq 0$, $H \in \mathcal{H}$. So K is closed. Next we show $\mathcal{H} \supseteq \mathcal{K}^*$. Suppose there exists $\bar{Y} \in \mathcal{K}^* \setminus \mathcal{H}$, then by a standard separation theorem [Roc70, Corollary 11.4.2], there exists $\bar{X} \in \mathcal{H}^*$ such that $\langle \bar{X}, \bar{Y} \rangle < 0$ and $\langle \bar{X}, H \rangle > 0$ for all $H \in \mathcal{H}$. As $\bar{Y} \in \mathcal{K}^*$, $\bar{X} \notin \mathcal{K}$. Then by assumption there exists $\hat{H} \in \mathcal{H}$ such that $\langle \bar{X}, \hat{H} \rangle < 0$. This causes a contradiction. $\square$

Before stating the main theorem, we explain some additional notations. First we use $\mathcal{W}_n^{r+2}$ to denote the set of tensors with doubly nonnegative slices:

$$\mathcal{W}_n^{r+2} := \{Z \,\big|\, Z \in \mathcal{M}_n^{r+2}, \mathbf{Slices}(Z) \subseteq \mathcal{D}_n\}.$$

Note $\mathcal{W}_n^{r+2}$ is essentially a Cartesian product of n^r doubly nonnegative cones, and that the double nonnegativity on slices implies those tensors are symmetric with respect to the last two indices, but not necessarily symmetric in general. Since $\mathcal{S}_n^{r+2}$ is a linear subspace of $\mathcal{M}_n^{r+2}$, $\mathcal{S}_n^{r+2} \cap \mathcal{W}_n^{r+2}$ is a closed convex cone, and its dual cone in $\mathcal{M}_n^{r+2}$ is:

$$\left(\mathcal{S}_n^{r+2} \cap \mathcal{W}_n^{r+2}\right)^* = \mathbf{cl}\left(\left(\mathcal{S}_n^{r+2}\right)^{\perp} + \left(\mathcal{W}_n^{r+2}\right)^*\right),$$

where $(\mathcal{W}_n^{r+2})^* = \{T | T \in \mathcal{M}_n^{r+2}, \mathbf{Slices}(T) \subseteq \mathcal{D}_n^*\}$.

Theorem 4. *For any nonnegative integer r, $\mathcal{Q}_n^r$ and $\mathcal{TD}_n^r$ have the following duality relation:*

$$(\mathcal{TD}_n^r)^* = \mathcal{Q}_n^r, \quad (\mathcal{Q}_n^r)^* = \mathcal{TD}_n^r.$$

Proof. First we show that $(\mathcal{TD}_n^r)^* \supseteq \mathcal{Q}_n^r$ and $(\mathcal{Q}_n^r)^* \supseteq \mathcal{TD}_n^r$. It suffices to show that for any $M \in \mathcal{Q}_n^r$ and $X \in \mathcal{TD}_n^r$, $\langle M, X \rangle \geq 0$. This is because

$$\begin{aligned} \langle M, X \rangle &= \sum_{\alpha \in \mathbb{N}^r(n)} \langle M, Z^\alpha \rangle = \langle \mathbf{Stack}(M), Z \rangle \\ &= \langle T, Z \rangle = \sum_{\alpha \in \mathbb{N}^r(n)} \langle T^\alpha, Z^\alpha \rangle \geq 0, \end{aligned}$$

where Z is used as in (9), T is defined as in Lemma 3, and the third equality is because of $Z \in \mathcal{S}_n^{r+2}$ and $T \simeq_{Sym} \textbf{Stack}(M)$. The last inequality is because $T^\alpha \in \mathcal{D}_n^*$, $Z^\alpha \in \mathcal{D}_n$.

Next we show $(\mathcal{TD}_n^r)^* = \mathcal{Q}_n^r$. We use Lemma 4 and the following alternative system pair to prove this. Given a matrix $X \in \mathcal{M}_n$, we have the following alternative system pair in variables Z and M:

$$\textbf{Collapse}(Z) = X, \quad Z \in \mathcal{S}_n^{r+2} \cap \mathcal{W}_n^{r+2}, \tag{P}$$

$$\langle M, X \rangle < 0, \quad \textbf{Stack}(M) \in \textbf{cl}\left(\left(\mathcal{S}_n^{r+2}\right)^\perp + \left(\mathcal{W}_n^{r+2}\right)^* \right). \tag{D}$$

This pair is an alternative system because $\textbf{Collapse}(\cdot)$ and $\textbf{Stack}(\cdot)$ are adjoint operators, and the second condition in system (D) is strictly feasible (for example, let M be a matrix with all entries positive, then $\textbf{Stack}(M) \in \textbf{relint}\,(\mathcal{W}_n^{r+2})^*$).

For any $X \notin \mathcal{TD}_n^r$, (P) is infeasible, so there exists M such that (D) is feasible. By continuity, we can assume that $\textbf{Stack}(M) \in (\mathcal{S}_n^{r+2})^\perp + (\mathcal{W}_n^{r+2})^*$. So $\exists\, T \in (\mathcal{W}_n^{r+2})^*$ such that $\textbf{Stack}(M) \simeq_{Sym} T$. Therefore $M \in \mathcal{Q}_n^r$. Finally by applying Lemma 4, $(\mathcal{TD}_n^r)^* = \mathcal{Q}_n^r$ and $\mathcal{TD}_n^r$ is closed. Then $(\mathcal{Q}_n^r)^* = (\mathcal{TD}_n^r)^{**} = \mathcal{TD}_n^r$. This concludes our proof. □

A direct corollary of our duality results is that any matrix not in $\mathcal{C}_n$ is excluded from $\mathcal{T}_n^r$ with some finite r.

Corollary 1. *Suppose $Y \notin \mathcal{C}_n$, then there exists finite nonnegative integer R, such that $Y \notin \mathcal{T}_n^R$ and $Y \notin \mathcal{TD}_n^R$.*

Proof. For $Y \notin \mathcal{C}_n$, there exists a strictly copositive $M \in \mathcal{C}_n^*$, such that $\langle M, Y \rangle < 0$. Then the result follows readily from Theorem 2, Theorem 3 and Theorem 4. □

7 Relation with Gvozdenović-Laurent dual formulation

In [GL07], a formulation of $(\mathcal{K}_n^r)^*$ is derived by using the terminology of polynomials. They used this formulation to compare Lasserre's bounds and copositive based bounds for the stability number. In this section we reveal the connection between their formulation and our tensor terminology. Their key result is summarized in the following theorem.

Theorem 5. *[GL07, Proposition 1] The dual cone of $\mathcal{K}_n^r$ has the following formulation,*

$$(\mathcal{K}_n^r)^* = \left\{X \in \mathcal{S}_n \middle| \, \exists \; y \in \mathbb{R}^{|\mathbb{I}^n(2r+4)|}, N_{r+2}(y) \succeq 0, X = C(y)\right\},$$

where $\mathbb{R}^{|\mathbb{I}^n(2r+4)|}$ is the Euclidean space indexed by elements in $\mathbb{I}^n(2r+4)$; $N_{r+2}(y)$ and $C(y)$ are defined as follows. Given a vector $y = [y_m]_{m\in\mathbb{I}^n(2r+4)} \in \mathbb{R}^{|\mathbb{I}^n(2r+4)|}$, $N_{r+2}(y)$ is a symmetric matrix indexed by $\mathbb{I}^n(r+2)$, whose (m, m')-th entry is equal to $y_{m+m'}$, for $m, m' \in \mathbb{I}^n(r+2)$, and $C(y)$ is an $n \times n$ symmetric matrix constructed by

$$C(y) := \sum_{m\in\mathbb{I}^n(r)} c(m) N^m(y),$$

where $N^m(y)$ is the principal submatrix of $N_{r+2}(y)$ indexed by $m + 2e_1, ..., m + 2e_n$.

Their proof is based on the duality relation between s.o.s polynomials and positive semidefinite moment matrices. We remark that there exists an interesting relation between Theorem 5 and our tensor construction. Given a vector $y \in \mathbb{R}^{\mathbb{I}^n(2r+4)}$, one may construct a symmetric tensor $Z \in \mathcal{S}_n^{r+2}$ using only a subset of elements in y:

$$Z[\alpha] := y_{2p}, \;\; \forall\alpha \in \mathbb{N}^{r+2}(n), \;\; [\![\alpha]\!] = p \in \mathbb{I}^n(r+2).$$

Then the set $\{N^m(y)\}_{m\in\mathbb{I}^n(r)}$ are exactly **Slices**(Z), i.e.,

$$[N^m(y)]_{ij} = [N_{r+2}(y)]_{m+2e_i,m+2e_j} = y_{2m+2e_i+2e_j} = Z[\beta, i, j], for\ any\ \beta \in \mathbb{N}^r(n), [\![\beta]\!] = m.$$

Further, $C(y) = \textbf{Collapse}(Z)$. Indeed, for $1 \le i, j \le n$,

$$[C(y)]_{ij} = \sum_{m\in\mathbb{I}^n(r)} c(m) \left[N^m(y)\right]_{ij} = \sum_{m\in\mathbb{I}^n(r)} c(m) Z[\beta, i, j], \;\; for\ any\ \beta \in \mathbb{N}^r(n), [\![\beta]\!] = m.$$

Then by using (5) and the fact that Z is symmetric, we have $[C(y)]_{ij} = \sum_{\beta\in\mathbb{N}^r(n)} Z[\beta, i, j]$.

It is then clear that $N_{r+2}(y) \succeq 0$ is a stronger condition on Z than **Slices**$(Z) \subseteq \mathcal{D}_n$ that used in $\mathcal{TD}_n^r$. Indeed **Slices**(Z) are principal submatrices of $N_{r+2}(y)$, and $Z[\beta, i, j] = y_{2[\![[\beta,i,j]]\!]} = [N_{r+2}(y)]_{m+e_i+e_j,m+e_i+e_j} \ge 0$ by the construction of $N_{r+2}(y)$. It is also clear $(\mathcal{K}_n^r)^*$ uses more additional variables than $\mathcal{TD}_n^r$ and therefore is more expensive to compute in general.

8 An Application to the Stability Number and a Numerical Example

As an application of our results, we examine linear programming bounds of the stability number of a graph. In [PVZ07], a closed-form expression for these approximations was given. In this section, we use the tensor structure of $\mathcal{T}_n^r$ to give a new proof of a known sufficient and necessary condition under which these approximations are finite by using the pigeon-hole principle. Then we construct an explicit primal optimal solution as well as its tensor lifting for each of these tensor approximations.

Given a graph $\mathcal{G}$ with n vertices and edge set $\mathcal{E}(\mathcal{G})$, the Maximum Stable Set problem asks for the cardinality of the largest *stable set*, where a stable set is a subset of vertices that are pairwise unconnected. This cardinality is usually denoted by $\alpha(\mathcal{G})$. It has the following copositive formulation [dKP02]:

$$\alpha(\mathcal{G}) = \max_X \{E \bullet X : \mathbf{trace}(X) = 1, X_{ij} = 0,\ \forall (i,j) \in \mathcal{E}(\mathcal{G}),\ X \in \mathcal{C}_n\}. \tag{12}$$

Its dual problem is

$$\alpha(\mathcal{G}) = \min_\lambda \{\lambda : s.t.,\ \lambda(I + A_\mathcal{G}) - E \in \mathcal{C}_n^*\}, \tag{13}$$

where $A_\mathcal{G}$ is the incidence matrix such that $A_\mathcal{G}(i,j) = 1$ if $(i,j) \in \mathcal{E}(\mathcal{G})$, and $A_\mathcal{G}(i,j) = 0$ otherwise. Note that strong duality holds because (13) is strictly feasible, although (12) is not.

By replacing $\mathcal{C}_n$ with its polyhedral relaxations $\{\mathcal{T}_n^r\}$, a class of upper bounds of $\alpha(\mathcal{G})$ are defined as

$$\zeta^{(r)}(\mathcal{G}) := \max_X \{E \bullet X : \mathbf{trace}(X) = 1, X_{ij} = 0,\ \forall (i,j) \in \mathcal{E},\ X \in \mathcal{T}_n^r\}. \tag{14}$$

By Theorem 3, the dual problem of (14) reads:

$$\zeta^{(r)}(\mathcal{G}) = \min_\lambda \{\lambda : s.t.,\ \lambda(I + A_\mathcal{G}) - E \in \mathcal{C}_n^r\}. \tag{15}$$

Strong duality holds here because both (14) and (15) are linear programs.

Several properties about $\zeta^{(r)}(\mathcal{G})$ were proved in [PVZ07] using the dual form (15). One interesting result is that $\zeta^{(r)}(\mathcal{G})$ is finite if and only if $r \geq \alpha(\mathcal{G}) - 1$. Further $\zeta^{(r)}(\mathcal{G})$ can be written in closed form in terms of $\alpha(\mathcal{G})$ [PVZ07, Theorem 1]. Here we first give a combinatorial proof of the finite condition. Our proof uses the tensor structure and the pigeon-hole principle.

Proposition 2. *[PVZ07, Corollary 3]* $\zeta^{(r)}(\mathcal{G}) < \infty$ *if and only if* $r + 2 > \alpha(G)$.

Proof. Let $Z \in \mathcal{S}_n^{r+2}$ be the tensor corresponding to X as in (8). Note (14) is unbounded if and only if $\exists \beta \in \mathbb{N}^{r+2}(n)$, $Z[\beta]$ is unbounded. By examining the constraints in (14), every index in such a β appears at most once (otherwise by symmetry of Z, $Z[\beta]$ is bounded by $\textbf{trace}(X) = 1$ and elementwise nonnegativity), and β does not contain a pair of indices $(i, j) \in \mathcal{E}$ (otherwise bounded by $X_{ij} = 0$). This is equivalent to say that β corresponds to a stable set in $\mathcal{G}$. Note β is a vector of length $r + 2$. By pigeon-hole principle, there exists such an β if and only if $r + 2 \leq \alpha(G)$. □

It was shown in [PVZ07, Theorem 1] that $\zeta^{(r)}(\mathcal{G})$ has the following closed form:

Theorem 6. *Assume* $r + 2 > \alpha(\mathcal{G})$ *and* $r + 2 = u\alpha(\mathcal{G}) + v$*, where* u, v *are nonnegative integers with* $v < \alpha(\mathcal{G})$*. Then*

$$\zeta^{(r)}(\mathcal{G}) = \frac{\binom{r+2}{2}}{\binom{u}{2}\alpha(\mathcal{G}) + vu}.$$

In their analysis, the dual formulation (15) of $\zeta^{(r)}(\mathcal{G})$ was used. However it is interesting to analyze the structure of a primal optimal solution X in (14) as it is directly related to an indicator vector of a max stable set.

We define some terminology first. Let $V \subseteq \{1, 2, ..., n\}$. We say a multi-index $\beta \in \mathbb{N}^s(n)$ is *evenly distributed* over V, denoted by $\beta \curlywedge V$, if each element in V appears in β either $\left\lfloor \frac{s}{|V|} \right\rfloor$ or $\left\lceil \frac{s}{|V|} \right\rceil$ many times. For example, if $s = 5$, and $V = \{1, 2, 4\}$, then β is evenly distributed over V if β is a permutation of one of the following three cases

$$[1, 1, 2, 2, 4], \; [1, 2, 2, 4, 4], \; [1, 1, 2, 4, 4].$$

Now we do some counting. Let $u = \left\lfloor \frac{s}{|V|} \right\rfloor$, and $v = s - u|V|$. Then, for any $\beta \curlywedge V$, v elements in V are repeated $u+1$ times in β, and the other $|V| - v$ elements in V are repeated u times in β. In the previous example, $u = 1$, $v = 2$.

The number of multi-indices β that are evenly distributed over V is:

$$|\{\beta \in \mathbb{N}^s(n) : \beta \curlywedge V\}| = \binom{|V|}{v} \frac{s!}{[u!]^{|V|-v}\,[(u+1)!]^v}.$$

This is calculated by first choosing v elements from V, and then using the multinomial formula. Further for any fixed $\beta \in \mathbb{N}^s(n)$, if we random pick a pair of indices $1 \le i < j \le s$, the probability of $\beta_i = \beta_j$ is:

$$\begin{aligned} P_{pair} &= \left\{ \binom{u+1}{2} v + \binom{u}{2} [|V| - v] \right\} \Big/ \binom{s}{2} \\ &= \left\{ \binom{u}{2} |V| + uv \right\} \Big/ \binom{s}{2}. \end{aligned}$$

The set $\{\beta \in \mathbb{N}^s(n) : \beta \curlywedge V\}$ is closed under permutation. Assuming $s \ge 2$, the number of β that has **same last two entries** can be calculated by

$$|\{\beta \in \mathbb{N}^s(n) : \beta \curlywedge V, \beta_{s-1} = \beta_s\}| \;=\; |\{\beta \in \mathbb{N}^s(n) : \beta \curlywedge V\}| \cdot P_{pair}.$$

In the following theorem we construct an optimal solution X for the primal problem (14), as well as its corresponding tensor lifting Z. We will use the previous counting results with $s = r+2$, and $|V| = \alpha(\mathcal{G})$.

Theorem 7. *Assume V_{max} is a maximum stable set of graph $\mathcal{G}$, $\alpha(\mathcal{G}) = |V_{max}|$. Also assume $r+2 > \alpha(\mathcal{G})$, and $u = \left\lfloor \frac{r+2}{\alpha(\mathcal{G})} \right\rfloor$, $v = r+2 - \left\lfloor \frac{r+2}{\alpha(\mathcal{G})} \right\rfloor \alpha(\mathcal{G})$. Then an optimal solution to (14) is given by $X =$ **Collapse**(Z), where $Z \in \mathcal{S}_n^{r+2}$ with entries:*

$$Z[\beta] = \begin{cases} C^{-1} \cdot \zeta^{(r)}(\mathcal{G}) & ; \;\; \text{if } \beta \curlywedge V_{max}, \beta \in \mathbb{N}^{r+2}(n), \\ 0 & ; \;\; \text{otherwise}, \end{cases}$$

where $\beta \curlywedge V_{max}$ means β is evenly distributed over V_{max}, i.e., each element in V_{max} appears in β either u or $u+1$ times, and C is the following number

$$C = \binom{\alpha(\mathcal{G})}{v} \frac{(r+2)!}{[u!]^{\alpha(\mathcal{G})-v}[(u+1)!]^v}.$$

Proof. We verify that X is feasible.

$$\begin{aligned}\mathbf{trace}(X) &= \sum_{i=1}^{n} \sum_{\gamma \in \mathbb{N}^r(n)} Z[\gamma, i, i] = \sum_{\beta \curlywedge V_{max}, \beta_{r+1} = \beta_{r+2}} Z[\beta] \\ &= \frac{\zeta^{(r)}(\mathcal{G})}{C} \cdot |\{\beta \mid \beta \curlywedge V_{max}, \beta_{r+1} = \beta_{r+2}\}| = 1.\end{aligned}$$

For $(i,j) \in \mathcal{E}$, $X_{ij} = \sum_{\gamma \in \mathbb{N}^r(n)} Z[\gamma, i, j] = 0$. So X is feasible in (14). The objective is

$$\begin{aligned}E \bullet X &= \sum_{i,j=1}^{n} \sum_{\gamma \in \mathbb{N}^r(n)} Z[\gamma, i, j] = \sum_{\beta \curlywedge V_{max}} Z[\beta] \\ &= \frac{\zeta^{(r)}(\mathcal{G})}{C} \cdot \left|\{\beta \mid \beta \curlywedge V_{max}, \beta \in \mathbb{N}^{r+2}(n)\}\right| = \zeta^{(r)}(\mathcal{G}).\end{aligned}$$

□

Remark 3. Theorem 6 can also be proved by using the tensor structure and the primal form (14). The main idea is that for any feasible tensor lifting Z, one can continuously improve the objective by taking steps towards the direction that emphasizing $Z[\beta]$ with β evenly distributed over some maximal stable set of $\mathcal{G}$.

Next we provide some remarks with regard to using $\mathcal{T}_n^r$ and $\mathcal{TD}_n^r$ computationally. It is always beneficial to formulate $\mathcal{T}_n^r$ as in Lemma 1. Note one may normalize F_m since the constant $c(m)$ tends to be large when r is large. For $\mathcal{TD}_n^r$, the symmetry and $\mathbf{Collapse}(\cdot)$ can be formulated as a set of linear equations, while $\mathbf{Slices}(Z) \subseteq \mathcal{D}_n$ requires n^r number of semidefinite constraints.

The following is a small numerical example that illustrates different characteristics of linear and semidefinite relaxations hierarchies for $\mathcal{C}_n$.

Example 1. Let Q be the 7×7 Hoffman-Pereira matrix[HP73]:

$$Q := \begin{pmatrix} 1 & -1 & 1 & 0 & 0 & 1 & -1 \\ -1 & 1 & -1 & 1 & 0 & 0 & 1 \\ 1 & -1 & 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & -1 & 1 \\ 1 & 0 & 0 & 1 & -1 & 1 & -1 \\ -1 & 1 & 0 & 0 & 1 & -1 & 1 \end{pmatrix}.$$

and $D = \mathbf{Diag}[1, 2, 3, ..., 7]$. Since Q is known to be copositive, DQD is also copositive, and the following optimization problem has optimal value 0:

$$0 = \min_X \left\{ \langle DQD, X \rangle : \mathbf{trace}(X) \leq 1, X \in \mathcal{C}_7 \right\}. \tag{16}$$

Let $q(\mathcal{T}_7^r)$ and $q(\mathcal{TD}_7^r)$ be the optimal values of corresponding relaxations. We solve these relaxations using YALMIP [Löf04] as an interface for Gurobi [Gur12] (to solve LP relaxations) and CSDP [Bor99] (to solve SDP relaxations). Optimal values and solver running time are reported in table 1.

Table 1: Relaxations of (16) using $\mathcal{T}_7^r$ and $\mathcal{TD}_7^r$

r	≤ 7	8	9	10	12	16	18	20
$q(\mathcal{T}_7^r)$	$+\infty$	-23.250	-18.250	-15.000	-11.267	-7.697	-6.556	–5.764
Time(s)	-	0.05	0.10	0.18	0.47	2.19	4.21	7.75
r	0	1	2					
$q(\mathcal{TD}_7^r)$	-1.807	-0.013	-1.1E-9					
Time(s)	0.025	0.352	13.923					

9 Conclusion and Future work

In this paper we proposed two tensor relaxation hierarchies for the completely positive cone $\mathcal{C}_n$. Both hierarchies are "exact in the limit". We derived the duality relations between our relaxation hierarchies and two known inner-approximation hierarchies of $\mathcal{C}_n^*$.

There are several interesting open questions related to this work. One question is whether there exists efficiently computable barrier functions for cones $\mathcal{T}_n^r$ and $\mathcal{TD}_n^r$. In other words, what is the most effective way to exploit the tensor structures algorithmically? While an extremal characterization for $\mathcal{T}_n^r$ is available (Lemma 1), can one characterize all of its facets? Research on these directions could lead to more practical algorithms for solving relaxations of copositive programs.

Acknowledgements. The author is grateful to his PhD advisors Professor Kurt Anstreicher and Samuel Burer for discussions and comments on earlier drafts of this paper, which improved this paper significantly. The author also thank two anonymous reviewers for careful reading and constructive comments. Support from NSF grant CCF-0545514 is also appreciated.

References

[Ans12] Kurt Anstreicher, *On convex relaxations for quadratically constrained quadratic programming*, Mathematical Programming (Series B) (2012), Publised online: Oct, 2012.

[BDdK+00] I. M. Bomze, M. Dür, E. de Klerk, C. Roos, A. Quist, and T. Terlaky, *On copositive programming and standard quadratic optimization problems*, J. Global Optim. **18** (2000), 301–320.

[BK02] Immanuel M. Bomze and Etienne de Klerk, *Solving Standard Quadratic Optimization Problems via Linear, Semidefinite and Copositive Programming*, J. Global Optim. **24** (2002), 163–185.

[Bor99] B. Borchers, *CSDP, a C library for semidefinite programming*, Optimization Methods and Software **11** (1999), no. 1, 613–623.

[Bur09] S. Burer, *On the copositive representation of binary and continuous nonconvex quadratic programs*, Mathematical Programming **120** (2009), 479–495.

[DG11] Peter J.C. Dickinson and Luuk Gijben, *On the computational complexity of memebership problems for the completely positive cone and its dual*, Optimization Online. http://www.optimization-online.org/DB_HTML/2011/05/3041.html, April 2011.

[dKP02] Etienne de Klerk and Dmitrii V. Pasechnik, *Approximation of the stability number of a graph via copositive programming*, SIAM J. Optim. **12** (2002), no. 4, 875–892.

[DR08] Igor Dukanovic and Franz Rendl, *Copositive programming motivated bounds on the stability and the chromatic numbers*, Mathematical Programming **121** (2008), no. 2, 249–268.

[Dür10] Mirjam Dür, *Copositive Programming - a Survey*, Recent Advances in Optimization and its Applications in Engineering (M. Deihl, F. Glineur, E. Jarlebring, and W. Michiels, eds.), Springer, Nov. 2010, pp. 3–20.

[GL07] Nebojša Gvozdenović and Monique Laurent, *Semidefinite bounds for the stability number of a graph via sums of squares of polynomials*, Mathematical Programming **110** (2007), no. 1, 145–173.

[GL08] ———, *The operator Ψ for the chromatic number of a graph*, SIAM Journal on Optimization **19** (2008), no. 2, 572–591.

[Gur12] Gurobi Optimization, Inc., *Gurobi optimizer reference manual*, 2012.

[HLP67] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge University Press, 1967.

[HP73] A. J. Hoffman and F. Pereira, *On copositive matrices with -1, 0, 1 entries*, Journal of Combinatorial Theory (A) **14** (1973), 302–309.

[HUS10] J.-B. Hiriart-Urruty and A. Seeger, *A variational approach to copositive matrices*, SIAM Rev. **52** (2010), 593–629.

[Lau09] Monique Laurent, *Sums of squares, moment matrices and optimization over polynomials*, Emerging Applications of Algebraic Geometry (M. Putinar and S. Sullivant, eds.), IMA Volumes in Mathematics and its Applications, vol. 149, Springer, 2009, pp. 157–270.

[Löf04] J. Löfberg, *Yalmip: A toolbox for modeling and optimization in matalb*, Proceedings of the CACSD Conference (Taipei, Taiwan), 2004.

[MK87] Katta G. Murty and Santosh N. Kabadi, *Some NP-complete problems in quadratic and nonlinear programming*, Mathematical Programming **39** (1987), no. 2, 117–129.

[NTZ11] Karthik Natarajan, Chung-Piaw Teo, and Zhichao Zheng, *Mixed zero-one linear programs under objective uncertainty: A completely positive representation*, Operations Research **59** (2011), no. 3, 713–728.

[Par00] Pablo A. Parrilo, *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*, Ph.D. thesis, California Institute of Technology, May 2000.

[PR09] Janez Povh and Franz Rendl, *Copositive and semidefinite relaxations of the quadratic assignment problem*, Discrete Optim. **6** (2009), no. 3, 231–241.

[PS01] V. Powers and C. Scheiderer, *The moment problem for non-compact semialgebraic sets*, Adv. Geom. **1** (2001), 71–88.

[PVZ07] Javier Peña, Juan Vera, and Luis F. Zuluaga, *Computing the stability number of a graph via linear and semidefinite programming*, SIAM J. Optim. **18** (2007), no. 1, 87–105.

[Roc70] R. Tyrrell Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

[Sch05] Markus Schweighofer, *Optimization of Polynomials on Compact Semialgebraic Sets*, SIAM J. on Optimization **15** (2005), no. 3, 805–825.

[Yıl12] E. Alper Yıldırım, *On the accuracy of uniform polyhedral approximations of the copositive cone*, Optimization Methods and Software **27** (2012), no. 1, 155–173.