On the Accuracy of Uniform Polyhedral Approximations of the Copositive Cone

E. Alper Yıldırım*

March 11, 2010
Revised on June 17, 2010

Abstract

We consider linear optimization problems over the cone of copositive matrices. Such conic optimization problems, called *copositive programs*, arise from the reformulation of a wide variety of difficult optimization problems. We propose a hierarchy of increasingly better outer polyhedral approximations to the copositive cone. We establish that the sequence of approximations is exact in the limit. By combining our outer polyhedral approximations with the inner polyhedral approximations due to de Klerk and Pasechnik [SIAM J. Optim, 12 (2002), pp. 875–892], we obtain a sequence of increasingly sharper lower and upper bounds on the optimal value of a copositive program. Under primal and dual regularity assumptions, we establish that both sequences converge to the optimal value. For standard quadratic optimization problems, we derive tight bounds on the gap between the upper and lower bounds. We provide closed-form expressions of the bounds for the maximum stable set problem. Our computational results shed light on the quality of the bounds on randomly generated instances.

Key words: Copositive cone, completely positive cone, standard quadratic optimization, maximum stable set.

*Department of Industrial Engineering, Bilkent University, 06800 Bilkent, Ankara, Turkey (yildirim@bilkent.edu.tr). The author was suported in part by TÜBİTAK (Turkish Scientific and Technological Research Council) Grant 109M149.

AMS Subject Classifications: 90C25, 90C05, 90C20, 15A48, 05C69.

1 Introduction

In this paper, we consider linear optimization problems over the cone of copositive matrices, which is defined as

$$\mathcal{C} := \{X \in \mathcal{S} : u^T X u \geq 0 \text{ for all } u \in \mathbb{R}^n_+\}, \tag{1}$$

where $\mathcal{S}$ denotes the set of $n \times n$ real symmetric matrices and $\mathbb{R}^n_+$ denotes the nonnegative orthant in $\mathbb{R}^n$. Equipping $\mathcal{S}$ with the usual trace inner product given by $\langle A, B\rangle = \text{trace}(AB) = \sum_{i=1}^n \sum_{j=1}^n A_{ij} B_{ij}$ for all $A, B \in \mathcal{S}$, the dual cone of $\mathcal{C}$ with respect to this inner product is the cone of completely positive matrices, which is given by

$$\mathcal{C}^* = \{X \in \mathcal{S} : X = \sum_{i=1}^k v^i (v^i)^T,\ v^i \in \mathbb{R}^n_+,\ i = 1, \ldots, k\}. \tag{2}$$

Both of the cones $\mathcal{C}$ and $\mathcal{C}^*$ are closed, convex, pointed, full-dimensional, and nonpolyhedral. The interior of $\mathcal{C}$ is given by

$$\text{int}(\mathcal{C}) = \{X \in \mathcal{S}^n : u^T X u > 0 \text{ for all } u \in \mathbb{R}^n_+,\ u \neq 0\}. \tag{3}$$

We refer the reader to [9] for a characterization of $\text{int}(\mathcal{C}^*)$. Each extreme ray of $\mathcal{C}^*$ is given by a rank one matrix vv^T, where $v \in \mathbb{R}^n_+$.

A *completely positive program* is given by

$$\begin{array}{lll} \text{(CoP)} & \min & \langle C, X\rangle \\ & \text{s.t.} & \langle A_i, X\rangle = b_i, \quad i = 1, \ldots, m, \\ & & X \in \mathcal{C}^*, \end{array}$$

where $X \in \mathcal{S}$ is the decision variable, and $A_1, \ldots, A_m \in \mathcal{S}$, $b \in \mathbb{R}^m$, and $C \in \mathcal{S}$ constitute the data of the problem. The associated dual problem, called a *copositive program*, is given by

$$\begin{array}{lll} \text{(CoD)} & \max & b^T y \\ & \text{s.t.} & \sum_{i=1}^m y_i A_i + S = C, \\ & & S \in \mathcal{C}, \end{array}$$

where $y \in \mathbb{R}^m$ and $S \in \mathcal{S}$ are the decision variables. It follows from the conic duality theory that weak duality always holds between (CoP) and (CoD) and that strong duality is satisfied under regularity assumptions such as the Slater's condition.

Recently, Burer [6] established that every quadratic optimization problem with nonnegative and binary variables, linear equality constraints, and complementarity constraints on the pairs of variables can be reformulated as an instance of (CoP). The class of optimization problems that admits such a reformulation encompasses all binary integer programming problems, all quadratic programming problems, and specific problems such as the quadratic assignment problem. Therefore, despite the fact that (CoP) is a convex optimization problem, it follows from this reformulation that (CoP) is, in general, intractable. In fact, the problem of deciding whether $X \notin \mathcal{C}$ is NP-complete [13]. Therefore, the reformulation as a convex optimization problem by itself does not alter the complexity of the problem. However, it paves the way for new approximation results by replacing the intractable cone by various tractable inner or outer approximations.

It is well-known and easy to verify that

$$\mathcal{S}^+ + \mathcal{N} \subseteq \mathcal{C} \quad \text{and} \quad \mathcal{C}^* \subseteq \mathcal{S}^+ \cap \mathcal{N}, \tag{4}$$

where $\mathcal{S}^+$ and $\mathcal{N}$ denote the cone of positive semidefinite matrices and the cone of nonnegative matrices in $\mathcal{S}$, respectively. Therefore, a tractable relaxation of (CoP) can be obtained by replacing the cone of completely positive matrices by the intersection of the cones of semidefinite and nonnegative matrices. In fact, both inclusions (4) are satisfied with equality for $n \leq 4$ whereas they are known to be strict for $n \geq 5$ (see, e.g., [1]).

Recently, various hierarchies of tractable approximations of the cone of copositive matrices have been proposed. The main ingredient in most of these hierarchies is the observation that a matrix $M \in \mathcal{S}$ is copositive if and only if the polynomial

$$P_M(x) := \sum_{i=1}^{n} \sum_{j=1}^{n} M_{ij} x_i^2 x_j^2$$

is nonnegative for all $x \in \mathbb{R}^n$. Relying on the fact that any polynomial that admits a sum-of-squares decomposition is necessarily nonnegative, Parrilo [14] was the first to construct a hierarchy of convex cones satisfying $\mathcal{S}^+ + \mathcal{N} = \mathcal{K}^0 \subseteq \mathcal{K}^1 \subseteq \ldots \subseteq \mathcal{C}$ and $\text{int}(\mathcal{C}) \subseteq \cup_{r\in\mathbb{N}}\mathcal{K}^r$. Since each cone $\mathcal{K}^r$ can be represented using linear matrix inequalities, a linear optimization problem over $\mathcal{K}^r$ can be formulated as a semidefinite programming (SDP) problem.

Similarly, de Klerk and Pasechnik [8] exploited a weaker sufficient condition on the nonnegativity of a polynomial to propose another hierarchy of convex cones satisfying $\mathcal{N} = \mathcal{I}^0 \subseteq \mathcal{I}^1 \subseteq \ldots \subseteq \mathcal{C}$ and $\text{int}(\mathcal{C}) \subseteq \cup_{r\in\mathbb{N}}\mathcal{I}^r$. In contrast to Parrilo's hierarchy, each cone $\mathcal{I}^r$ is polyhedral. Therefore, a linear optimization problem over $\mathcal{I}^r$ is a linear programming (LP) problem.

More recently, Peña, Vera, and Zuluaga [15] developed yet another sufficient condition on the nonnegativity of a polynomial, which gave rise to a sequence of convex cones satisfying $\mathcal{Q}^0 \subseteq \mathcal{Q}^1 \subseteq \ldots \subseteq \mathcal{C}$. They also established that $\mathcal{I}^r \subseteq \mathcal{Q}^r \subseteq \mathcal{K}^r$ for each $r \in \mathbb{N}$ and that $\mathcal{Q}^r = \mathcal{K}^r$ for $r = 0, 1$. Since each $\mathcal{Q}^r$ can be represented by linear matrix inequalities, linear optimization over $\mathcal{Q}^r$ is equivalent to an SDP problem.

As noted in [5], each of these hierarchies provides a uniform inner approximation to the cone of copositive matrices. By duality, the dual cones in each hierarchy provide a uniform outer approximation to the cone of completely positive matrices. The sizes of the resulting tractable problems quickly reach beyond the current computational capabilities. Finally, with the exception of [3], there usually is not much information about the accuracy of the resulting approximation.

Motivated by these observations, Bundfuss and Dür [5] proposed two hierarchies of polyhedral cones that respectively provide inner and outer polyhedral approximations to the cone of copositive matrices. As such, their approximation scheme concurrently provides upper and lower bounds on the optimal value of an instance of (CoD), which leads to the exact information on the accuracy of the approximation. In contrast to the previously proposed hierarchies which uniformly approximate the copositive cone, Bundfuss and Dür adaptively improve their polyhedral approximations using the guidance of the objective function. In other words, their approximation scheme yields a finer approximation to the feasible region of (CoD) in the vicinity of the set of optimal solutions but only a coarse approximation in the remaining parts. They report very encouraging computational results on randomly generated standard quadratic optimization problems.

In this paper, we propose another hierarchy of *outer* polyhedral approximations to the cone of copositive matrices. We establish that our approximation is exact in the limit. Combining our hierarchy of outer polyhedral approximations with that of inner polyhedral approximations due to de Klerk and Pasechnik [8], we obtain a sequence of improving lower and upper bounds on the optimal value of an instance of (CoP). These bounds precisely reveal the duality gap arising from the inner and outer approximations. Under primal and dual regularity assumptions, we establish that the duality gap converges to zero.

For quadratic optimization over the unit simplex (also known as standard quadratic optimization), we provide tight bounds on the duality gap. For the special case of the stable set problem, we give closed-form expressions of the lower and upper bounds.

Our work is inspired by and closely related to the recent work of Bundfuss and Dür [5]. Similar to their approach, we also rely on inner and outer polyhedral approximations of the copositive cone in an attempt to quantify the quality of the resulting lower and upper bounds. In contrast to their adaptive approximations, we focus on uniform inner and outer approximations of the copositive cone. As such, our primary objective in this paper is to investigate and assess the accuracy of uniform approximations to the copositive cone.

This paper is organized as follows. We present a hierarchy of increasingly better outer polyhedral approximations that converges to the copositive cone in Section 2. By combining our hierarchy of outer polyhedral approximations with that of inner polyhedral approximations of [8], we discuss how to obtain sequences of increasingly sharper lower and upper bounds on the optimal value of an instance of (CoP) in Section 3. We establish that both sequences converge to the optimal value under primal and dual regularity assumptions. Section 4 is devoted to the specialization of our bounds to standard quadratic optimization problems. In particular, we derive a tight upper bound on the duality gap resulting from the inner and outer approximations. We also present closed-form expressions of the lower and upper bounds for the special case of the maximum stable set problem. Section 5 discusses the computational results. We conclude the paper in Section 6.

2 Outer Polyhedral Approximations of the Copositive Cone

In this section, we present a hierarchy of polyhedral cones that provide increasingly better outer approximations to the copositive cone.

Recall that a matrix $X \in \mathcal{S}$ is copositive if and only if $u^T X u \geq 0$ for all $u \in \mathbb{R}^n_+$. This condition is equivalent to

$$u^T X u \geq 0 \quad \text{for all } u \in \Delta_n, \tag{5}$$

where Δ_n denotes the $(n-1)$-dimensional unit simplex in $\mathbb{R}^n$ given by

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

where $e \in \mathbb{R}^n$ is the vector of all ones. The main idea behind our approximation scheme is to discretize the unit simplex and to enforce the condition (5) only on the discretized points as opposed to every point on the unit simplex.

For $r = 0, 1, 2, \ldots$, let us define the following regular grid of rational points on the unit simplex (see [3]):

$$\Delta(n, r) := \{x \in \Delta_n : (r+2)x \in \mathbb{N}\}. \tag{7}$$

The factor $(r+2)$ is chosen for consistency with the corresponding definition of the inner approximation scheme of [3]. For each r, $\Delta(n, r)$ provides a finite discretization of the unit simplex that consists only of rational points. It is easy to verify that

$$|\Delta(n, r)| = \binom{n+r+1}{r+2}, \quad r = 0, 1, 2, \ldots. \tag{8}$$

For $r = 0, 1, 2, \ldots$, let us define

$$\delta(n, r) := \bigcup_{k=0}^{r} \Delta(n, k). \tag{9}$$

For $n \geq 2$, it follows from (8) that

$$|\delta(n,r)| \leq \sum_{k=0}^{r} |\Delta(n,k)| = \sum_{k=0}^{r} \binom{n+k+1}{k+2} \leq \sum_{k=0}^{r} n^{k+2} = n^2 \left(\frac{n^{r+1}-1}{n-1} \right), \tag{10}$$

which is polynomial in n for a fixed value of r. Let us now define the following convex cones:

$$\mathcal{O}^r := \{X \in \mathcal{S}^n : d^T X d \geq 0 \text{ for all } d \in \delta(n,r)\}, \quad r = 0, 1, 2, \ldots. \tag{11}$$

Since $\delta(n,0) \subseteq \delta(n,1) \subseteq \ldots \subseteq \Delta_n$, it follows from (11) and (5) that $\mathcal{O}^0 \supseteq \mathcal{O}^1 \supseteq \ldots \supseteq \mathcal{C}$. Furthermore, $\mathcal{O}^r$ is a polyhedral cone for each r since $d^T X d = \langle X, dd^T \rangle \geq 0$ is a linear inequality constraint in $\mathcal{S}$ and there is a finite number of points in $\delta(n,r)$. For instance,

$$\mathcal{O}^0 = \{X \in \mathcal{S} : X_{ii} \geq 0,\ i = 1, \ldots, n; \quad X_{ii} + X_{jj} + 2X_{ij} \geq 0,\ 1 \leq i < j \leq n\},$$

and

$$\begin{aligned} \mathcal{O}^1 \ = \ & \{X \in \mathcal{S}^n : X_{ii} \geq 0,\ i = 1, \ldots, n; \quad X_{ii} + X_{jj} + 2X_{ij} \geq 0,\ 1 \leq i < j \leq n; \\ & 4X_{ii} + X_{jj} + 4X_{ij} \geq 0,\ 1 \leq i < j \leq n; \quad X_{ii} + 4X_{jj} + 4X_{ij} \geq 0,\ 1 \leq i < j \leq n; \\ & X_{ii} + X_{jj} + X_{kk} + 2X_{ij} + 2X_{ik} + 2X_{jk} \geq 0,\ 1 \leq i < j < k \leq n\}. \end{aligned}$$

It is easy to verify that each cone $\mathcal{O}^r$ is pointed and full-dimensional.

The next proposition establishes that the polyhedral cones $\mathcal{O}^r$ provide a hierarchy of outer approximations that converges to the cone of copositive matrices.

Theorem 2.1 *The polyhedral cones $\mathcal{O}^r$ satisfy $\mathcal{O}^0 \supseteq \mathcal{O}^1 \supseteq \ldots \supseteq \mathcal{C}$. Furthermore,*

$$\mathcal{C} = \bigcap_{r \in \mathbb{N}} \mathcal{O}^r. \tag{12}$$

Proof. Clearly, $\mathcal{C} \subseteq \bigcap_{r \in \mathbb{N}} \mathcal{O}^r$ since $\mathcal{C} \subseteq \mathcal{O}^r$ for each $r \in \mathbb{N}$. For the reverse inclusion, let $M \in \mathcal{S} \backslash \mathcal{C}$. Then, there exists $\bar{x} \in \Delta_n$ such that $\bar{x}^T M \bar{x} < 0$. By perturbing the zero components of $\bar{x}$ (if any) by a sufficiently small positive amount, we may assume that $\bar{x} > 0$. By continuity, there exists an $\bar{\epsilon} > 0$ such that $x^T M x < 0$ for all $x \in \mathbb{R}^n$ satisfying $\|x - \bar{x}\| < \bar{\epsilon}$. Let $\epsilon := \min\{\bar{\epsilon}, \min_{i=1,\ldots,n} \bar{x}_i\} > 0$. By the density of rational numbers in real numbers, there exists $\bar{w} \in \mathbb{Q}^n$ such that $\|\bar{w} - \bar{x}\| < \epsilon$. By the choice of ϵ, $\bar{w} > 0$. It follows that there exists $r_0 \in \mathbb{N}$ such that $\bar{d} := (1/(e^T \bar{w}))\bar{w} \in \delta(n,r)$ for all $r \geq r_0$. Since $\bar{d}^T M \bar{d} < 0$, we have $M \notin \mathcal{O}^r$ for all $r \geq r_0$ and hence, $M \notin \bigcap_{r \in \mathbb{N}} \mathcal{O}^r$. $\square$

The dual cone of $\mathcal{O}^r$ is given by

$$(\mathcal{O}^r)^* = \left\{ \sum_{d \in \delta(n,r)} \lambda_d d d^T : \lambda_d \geq 0 \text{ for all } d \in \delta(n,r) \right\}, \quad r = 0, 1, 2, \ldots. \tag{13}$$

Since $\mathcal{C} \subseteq \mathcal{O}^r$ for each $r \in \mathbb{N}$, it follows from duality that the dual cones satisfy $(\mathcal{O}^r)^* \subseteq \mathcal{C}^*$, i.e., each dual cone provides an inner approximation to the cone of completely positive matrices. The following theorem summarizes the relationships among these dual cones.

Theorem 2.2 *The polyhedral dual cones $(\mathcal{O}^r)^*$ satisfy $(\mathcal{O}^0)^* \subseteq (\mathcal{O}^1)^* \subseteq \ldots \subseteq \mathcal{C}^*$. Furthermore,*

$$int(\mathcal{C}^*) \subseteq \bigcup_{r \in \mathbb{N}} (\mathcal{O}^r)^* \subseteq \mathcal{C}^*, \tag{14}$$

and therefore

$$cl\left(\bigcup_{r \in \mathbb{N}} (\mathcal{O}^r)^* \right) = \mathcal{C}^*. \tag{15}$$

Proof. Since $\mathcal{C}^*$ is closed, (15) follows from (14). Therefore, it suffices to establish (14). By contradiction, suppose that there exists $M \in \text{int}(\mathcal{C}^*)$ but $M \notin \bigcup_{r \in \mathbb{N}} (\mathcal{O}^r)^*$. This implies that $M \notin (\mathcal{O}^r)^*$ for all $r \in \mathbb{N}$. Therefore, for each $r \in \mathbb{N}$, there exists $X_r \in \mathcal{O}^r$ such that $\langle X_r, M \rangle < 0$. Without loss of generality, we may assume that $\|X_r\| = \langle X_r, X_r \rangle^{1/2} = 1$ for each $r \in \mathbb{N}$. By passing to a subsequence if necessary, there exists $X^* \in \mathcal{S}$ such that $X_r \to X^*$. By Theorem 2.1, the subsequence $\{X_r, X_{r+1}, \ldots\} \in \mathcal{O}^r$ for all $r \in \mathbb{N}$. Since each $\mathcal{O}^r$ is closed, it follows that $X^* \in \mathcal{O}^r$ for each $r \in \mathbb{N}$. Therefore, $X^* \in \bigcap_{r \in \mathbb{N}} \mathcal{O}^r = \mathcal{C}$ by Theorem 2.1. Since $\langle X_r, M \rangle < 0$ for each $r \in \mathbb{N}$, we have $\langle X^*, M \rangle \leq 0$, which implies that $\langle X^*, M \rangle = 0$. However, this is a contradiction since $M \in \text{int}(\mathcal{C}^*)$ and $\langle X, M \rangle > 0$ for all $X \in \mathcal{C}$ such that $X \neq 0$. Therefore, $M \in \bigcup_{r \in \mathbb{N}} (\mathcal{O}^r)^*$. $\square$

3 Sequences of Improving Lower and Upper Bounds

In this section, we first review the hierarchy of inner polyhedral approximations to the copositive cone due to de Klerk and Pasechnik [8] (see also [3]). Then, we combine this hierarchy with our hierarchy of outer polyhedral approximations in order to obtain sequences of improving lower and upper bounds on the optimal value of an instance of (CoP). We establish that both sequences converge to the optimal value under primal and dual regularity assumptions. Furthermore, these bounds correspond to the duality gap and hence provide exact information on the quality of approximation.

Let us define

$$\Theta(n,r) := \left\{ z \in \mathbb{N}^n : \sum_{i=1}^{n} z_i = r+2 \right\}, \quad r = 0,1,2,\ldots. \tag{16}$$

By (7), it is easy to verify that

$$\Delta(n,r) = \{x \in \Delta_n : (r+2)x \in \Theta(n,r)\}, \tag{17}$$

which implies that

$$|\Theta(n,r)| = |\Delta(n,r)| = \binom{n+r+1}{r+2}, \quad r = 0,1,2,\ldots. \tag{18}$$

Consider the following convex cones:

$$\mathcal{I}^r := \{X \in \mathcal{S} : \langle zz^T - \mathrm{Diag}(z), X\rangle \geq 0 \text{ for all } z \in \Theta(n,r)\}, \quad r = 0,1,2,\ldots, \tag{19}$$

where $\mathrm{Diag}(z) \in \mathcal{S}$ is a diagonal matrix whose diagonal entries are given by $z \in \mathbb{R}^n$. By (18), $\mathcal{I}^r$ is a polyhedral cone for each $r \in \mathbb{N}$.

In [8], de Klerk and Pasechnik established that

$$\mathcal{N} = \mathcal{I}^0 \subseteq \mathcal{I}^1 \subseteq \ldots \subseteq \mathcal{C}, \quad \text{and} \quad \mathrm{int}(\mathcal{C}) \subseteq \bigcup_{r\in\mathbb{N}} \mathcal{I}^r \subseteq \mathcal{C}. \tag{20}$$

The dual cone of $\mathcal{I}^r$ is given by

$$(\mathcal{I}^r)^* = \left\{ \sum_{z\in\Theta(n,r)} \beta_z \left(zz^T - \mathrm{Diag}(z)\right) : \beta_z \geq 0 \text{ for all } z \in \Theta(n,r) \right\}, \quad r = 0,1,2,\ldots. \tag{21}$$

By duality, it follows that

$$\mathcal{C}^* \subseteq \ldots \subseteq (\mathcal{I}^1)^* \subseteq (\mathcal{I}^0)^* = \mathcal{N}, \quad \text{and} \quad \mathcal{C}^* = \bigcap_{r \in \mathbb{N}} (\mathcal{I}^r)^*. \tag{22}$$

Combining the relations (20) and (22) with Theorems 2.1 and 2.2, we obtain

$$\mathcal{N} = \mathcal{I}^0 \subseteq \mathcal{I}^1 \subseteq \ldots \subseteq \mathcal{C} \subseteq \ldots \subseteq \mathcal{O}^1 \subseteq \mathcal{O}^0, \tag{23}$$

and

$$(\mathcal{O}^0)^* \subseteq (\mathcal{O}^1)^* \subseteq \ldots \subseteq \mathcal{C}^* \subseteq \ldots \subseteq (\mathcal{I}^1)^* \subseteq (\mathcal{I}^0)^* = \mathcal{N}. \tag{24}$$

Therefore, we obtain a hierarchy of inner and outer polyhedral approximations to the copositive cone (respectively to the completely positive cone). Furthermore, each of these approximations is exact in the limit. We now discuss how these hierarchies can be used to obtain sequences of improving lower and upper bounds on the optimal value of an instance of a copositive programming problem.

Let us consider the following instance of (CoP):

$$\mu^* := \min\{\langle C, X\rangle : \langle A_i, X\rangle = b_i, \quad i = 1, \ldots, m, \quad X \in \mathcal{C}^*\}, \tag{25}$$

where $A_1, \ldots, A_m \in \mathcal{S}$, $b \in \mathbb{R}^m$, and $C \in \mathcal{S}$ are given and $X \in \mathcal{S}$ is the decision variable. Let us define

$$\mu_l^r := \min\{\langle C, X\rangle : \langle A_i, X\rangle = b_i, \quad i = 1, \ldots, m, \quad X \in (\mathcal{I}^r)^*\}, \quad r = 0, 1, \ldots, \tag{26}$$

and

$$\mu_u^r := \min\{\langle C, X\rangle : \langle A_i, X\rangle = b_i, \quad i = 1, \ldots, m, \quad X \in (\mathcal{O}^r)^*\}, \quad r = 0, 1, \ldots. \tag{27}$$

Since $(\mathcal{I}^r)^*$ and $(\mathcal{O}^r)^*$ are polyhedral cones for each $r \in \mathbb{N}$, it follows that the computation of each of μ_l^r and μ_u^r amounts to solving a linear programming problem. Furthermore, it follows from (24) that

$$\mu_l^0 \le \mu_l^1 \le \ldots \le \mu^* \le \ldots \le \mu_u^1 \le \mu_u^0. \tag{28}$$

Therefore, the sequence $\{\mu_u^r - \mu_l^r\}_{r=0}^{\infty}$ is nonincreasing and gives precise information about the accuracy of approximation with respect to the objective function value for each $r \in \mathbb{N}$. It is worth noticing that the number of constraints that define the inner and outer polyhedral cones is polynomial for each fixed value of r. However, the dependence on r is exponential, which implies that the cost of computing μ_l^r and μ_u^r rapidly increases as r increases. This is a common feature of all hierarchies that approximate the copositive cone uniformly. We refer the reader to [5] for an alternative and more effective approach.

In the next proposition, we establish that the sequence $\{\mu_u^r - \mu_l^r\}_{r=0}^{\infty}$ converges to zero under primal and dual regularity assumptions.

Theorem 3.1 *Let $\hat{X} \in \mathcal{S}$ be a strictly feasible solution of (CoP) and let $(\hat{y}, \hat{S}) \in \mathbb{R}^m \times \mathcal{S}$ be a strictly feasible solution of (CoD). Let μ^* denote the common optimal value of (CoP) and (CoD) and let μ_l^r and μ_u^r be defined as in (26) and (27), respectively. Then,*

$$\lim_{r \to \infty} \mu_l^r = \lim_{r \to \infty} \mu_u^r = \mu^*. \tag{29}$$

Proof. Our argument mimics the proof of [5, Theorem 4.2] but our hypotheses are slightly different. By the hypothesis, (CoP) has an optimal solution $X^* \in \mathcal{S}$ and (CoD) has an optimal solution $(y^*, S^*) \in \mathbb{R}^m \times \mathcal{S}$. Furthermore, strong duality holds between (CoP) and (CoD), i.e.,

$$\mu^* = \langle C, X^* \rangle = b^T y^*.$$

First, let us consider the sequence $\{\mu_u^r\}$ of upper bounds. By Theorem 2.2, there exists $r_0 \in \mathbb{N}$ such that $\hat{X} \in (\mathcal{O}^r)^*$ for all $r \geq r_0$. Therefore, $\hat{X}$ is a feasible solution of the linear programming problem in (27) and $\mu^* \leq \mu_u^r \leq \langle C, \hat{X} \rangle$ for all $r \geq r_0$. Since $\hat{X}$ is strictly feasible, $X_\lambda := \lambda X^* + (1 - \lambda)\hat{X}$ is a strictly feasible solution of (CoP) for all $\lambda \in (0, 1)$. For each $\lambda \in (0, 1)$, there exists $r_\lambda \in \mathbb{N}$ such that $X_\lambda \in (\mathcal{O}^r)^*$ for all $r \geq r_\lambda$ by Theorem 2.2. Therefore, X_λ is a feasible solution of (27) for all $r \geq r_\lambda$, which implies that $\mu^* \leq \mu_u^r \leq \langle C, X_\lambda \rangle = \lambda\mu^* + (1 - \lambda)\langle C, \hat{X} \rangle$ for $r \geq r_\lambda$. By taking the limit as λ goes to 1, we obtain that $\lim_{r \to \infty} \mu_u^r = \mu^*$.

Let us now focus on the the sequence $\{\mu_l^r\}$ of lower bounds. By linear programming duality,

$$\mu_l^r = \max \left\{ b^T y : \sum_{i=1}^m y_i A_i + S = C, \quad S \in \mathcal{I}^r \right\}, \quad r = 0, 1, \ldots, \tag{30}$$

By (20), there exists $r_1 \in \mathbb{N}$ such that $\hat{S} \in \mathcal{I}^r$ for each $r \geq r_1$. Therefore, $(\hat{y}, \hat{S})$ is a feasible solution of the linear programming problem in (30) for $r \geq r_1$, which implies that $b^T \hat{y} \leq \mu_l^r \leq \mu^*$ for all $r \geq r_1$. Let us define $(y_\lambda, S_\lambda) := \lambda(y^*, S^*) + (1-\lambda)(\hat{y}, \hat{S})$, which is a strictly feasible solution of (CoD) for each $\lambda \in (0,1)$. The convergence of the sequence of lower bounds is established by a similar limiting argument. □

Under the assumption that the feasible region of (CoP) (or (CoD)) is bounded and contains a strictly feasible solution, one can establish that the optimal solutions of (26) and (27) have an accumulation point and each accumulation point is an optimal solution using the proof technique of Bundfuss and Dür [5, Theorem 4.2].

Under the assumptions of Theorem 3.1, μ_l^r and μ_u^r are finite for all sufficiently large values of r. Let $X_u^r \in \mathcal{S}$ denote an optimal solution of (27) and let $(y_l^r, S_l^r) \in \mathbb{R}^m \times \mathcal{S}$ denote an optimal solution of (30). By (23) and (24), X_u^r is a feasible solution of (CoP) and (y_l^r, S_l^r) is a feasible solution of (CoD). It follows that the difference

$$\mu_u^r - \mu_l^r = \langle C, X_u^r \rangle - b^T y_l^r \tag{31}$$

precisely corresponds to the duality gap between these two feasible solutions for all sufficiently large values of r.

We close this section by discussing the relevance of the assumptions of Theorem 3.1. If (CoP) does not have a strictly feasible solution, then all the inner approximations given by (27) may remain infeasible for each $r \in \mathbb{N}$. For instance, if the feasible region of (CoP) is a subset of $\{\lambda u u^T : \lambda \geq 0\}$, where $u \in \mathbb{R}_+^n \backslash \mathbb{Q}_+^n$, then the feasible region is not contained in $(\mathcal{O}^r)^*$ for any $r \in \mathbb{N}$ (cf. (13)). If (CoP) is infeasible, then all inner approximations will necessarily be infeasible. However, by the previous example, we cannot conclude the infeasibility of (CoP) unless an outer approximation also happens to be infeasible. If the optimal solution set of (CoP) is empty (i.e., the optimal value is not attained) or unbounded, it follows that (CoD) cannot have a strictly feasible solution. Therefore, the inner approximations to the dual problem may remain infeasible for each $r \in \mathbb{N}$ as in the previous example. These discussions reveal that the assumptions of Theorem 3.1 are crucial in order to establish the convergence of the two sequences $\{\mu_l^r\}$ and $\{\mu_u^r\}$. Finally, if (CoP) is unbounded but contains a strictly feasible solution, then each outer approximation is necessarily unbounded. The inner approximations satisfy

$$\lim_{r \to \infty} \mu_u^r = -\infty,$$

which is easily proven by constructing a sequence of strictly feasible solutions whose objective function values tend to $-\infty$ and by using the fact that each such solution is feasible for the inner approximations for all sufficiently large values of r (see [5, Theorem 4.3]). Analogously, if (CoD) is unbounded but contains a strictly feasible solution, then

$$\lim_{r \to \infty} \mu_l^r = \infty.$$

4 Standard Quadratic Optimization

Let $Q \in \mathcal{S}$ be an arbitrary matrix. The standard quadratic optimization problem is given by

$$\mu^* := \min_{x \in \Delta_n} x^T Q x. \tag{32}$$

This optimization problem arises in many different application areas (see, e.g., [2]) and contains the maximum stable set problem as a special case (see Section 4.2). Therefore, it is in general an NP-hard problem.

Bomze et al. [4] established that the problem (32) can be reformulated as the following instance of (CoP):

$$\mu^* = \min_{x \in \Delta_n} x^T Q x = \min\{\langle Q, X \rangle : \langle E, X \rangle = 1, \quad X \in \mathcal{C}^*\}, \tag{33}$$

where $E = ee^T \in \mathcal{S}$ is the matrix of all ones. Let $A \in \mathbb{R}^{n \times n}$ be any nonsingular matrix with positive entries. Then, the matrix given by $(1/\|A^T e\|^2) A A^T$ is a strictly feasible solution of the copositive program (see [9]). Therefore,

$$\mu^* = \max\{y : yE + S = Q, \quad S \in \mathcal{C}\}. \tag{34}$$

It is also easy to verify that $\hat{S} = Q - \hat{y} E \in \mathrm{int}(\mathcal{C})$ for all $\hat{y} < \mu^*$. Therefore, the primal-dual pair of problems (33) and (34) satisfy the assumptions of Theorem 3.1.

Let us consider the specialization of the sequence of linear programming problems (26) and (27) to the copositive programming problem (33). By (21), $X \in (\mathcal{I}^r)^*$ if and only if

$$X = \sum_{z \in \Theta(n,r)} \beta_z (zz^T - \mathrm{Diag}(z)), \tag{35}$$

where $\beta_z \geq 0$ for all $z \in \Theta(n,r)$ and $\Theta(n,r)$ is given by (16). Together with (26), we have

$$\begin{aligned}
\mu_l^r &= \min \left\{ \sum_{z \in \Theta(n,r)} \beta_z (z^T Q z - z^T \mathrm{diag}(Q)) : \sum_{z \in \Theta(n,r)} \beta_z (z^T E z - z^T \mathrm{diag}(E)) = 1 \right\}, \\
&= \min \left\{ \sum_{z \in \Theta(n,r)} \beta_z (z^T Q z - z^T \mathrm{diag}(Q)) : \sum_{z \in \Theta(n,r)} \beta_z = \frac{1}{(r+1)(r+2)} \right\}, \\
&= \frac{1}{(r+1)(r+2)} \min_{z \in \Theta(n,r)} (z^T Q z - z^T \mathrm{diag}(Q)), \\
&= \frac{r+2}{r+1} \min_{x \in \Delta(n,r)} (x^T Q x - (1/(r+2)) x^T \mathrm{diag}(Q)), \quad r = 0, 1, \ldots,
\end{aligned} \tag{36}$$

where we used $z^T E z = (e^T z)^2 = (r+2)^2$ and $z^T \mathrm{diag}(E) = e^T z = r+2$ for each $z \in \Theta(n,r)$ in the second line and the relation (17) in the last one. The same characterization also appears in [3, Theorem 3.1].

Similarly, using the fact that $X \in (\mathcal{O}^r)^*$ if and only if

$$X = \sum_{d \in \delta(n,r)} \lambda_d d d^T, \tag{37}$$

where $\lambda_d \geq 0$ for all $d \in \delta(n,r)$ and $\delta(n,r)$ is given by (9), we obtain

$$\mu_u^r = \min_{d \in \delta(n,r)} d^T Q d, \quad r = 0, 1, \ldots. \tag{38}$$

It follows from (36) and (38) that μ_l^r and μ_u^r can simply be computed by evaluating the corresponding quadratic objective functions on a finite number of grid points and choosing the best one. Clearly, μ_l^r and μ_u^r are finite for each $r \in \mathbb{N}$. By Theorem 3.1, both sequences $\{\mu_l^r\}$ and $\{\mu_u^r\}$ converge to μ^*. We next establish explicit upper bounds on the terms of the sequence $\{\mu_u^r - \mu_l^r\}$.

Theorem 4.1 *Let μ^* be defined as in (33). Then, μ_l^r and μ_u^r given by (36) and (38) satisfy*

$$\mu_u^r - \mu_l^r \leq \frac{1}{r+1}\left(\max_{i=1,\ldots,n} Q_{ii} - \mu^*\right) \leq \frac{1}{r+1}(\bar{\mu} - \mu^*), \quad r = 0, 1, \ldots, \tag{39}$$

where $\bar{\mu} := \max_{x \in \Delta_n} x^T Q x$.

Proof. Let us fix $r \in \mathbb{N}$ and let $x^r \in \Delta(n, r)$ denote the point which achieves the smallest objective function value in (36). Since $\Delta(n, r) \subseteq \delta(n, r)$, we have

$$\begin{aligned} \mu_u^r - \mu_l^r &\leq (x^r)^T Q x^r - \left(\frac{r+2}{r+1}\right)(x^r)^T Q x^r + \left(\frac{1}{r+1}\right)(x^r)^T \mathrm{diag}(Q), \\ &= -\left(\frac{1}{r+1}\right)(x^r)^T Q x^r + \left(\frac{1}{r+1}\right)(x^r)^T \mathrm{diag}(Q), \\ &\leq \frac{1}{r+1}\left(\max_{i=1,\ldots,n} Q_{ii} - \mu^*\right), \end{aligned}$$

where we used $x^r \in \Delta_n$ to derive the inequality in the last line. The second inequality in (39) follows from the fact that $\bar{\mu} \geq \max_{i=1,\ldots,n} Q_{ii}$. $\square$

Theorem 4.1 establishes an upper bound on the sequence of duality gaps $\{\mu_u^r - \mu_l^r\}$. We remark that this upper bound can be used to compute the smallest value of r to obtain a prescribed relative accuracy provided that a lower bound on μ^* is available (see also [3]).

4.1 Relations to Previous Approximation Results

Bomze and de Klerk [3] study the implications of the sequence of inner polyhedral approximations to the cone of copositive matrices due to de Klerk and Pasechnik [8] in the context of standard quadratic optimization. As mentioned in the previous section, they obtain the same characterization (36) of μ_l^r. They establish that

$$\mu^* - \mu_l^r \leq \frac{1}{r+1}\left(\max_{i=1,\ldots,n} Q_{ii} - \mu^*\right) \leq \frac{1}{r+1}(\bar{\mu} - \mu^*), \quad r = 0, 1, \ldots, \tag{40}$$

where μ^* denotes the optimal value of the standard quadratic optimization problem and $\bar{\mu}$ is defined as in Theorem 4.1. Since $\mu_u^r \geq \mu^*$ for each $r = 0, 1, \ldots$, we remark that the upper bound (40) is already implied by our upper bound (39).

In an attempt to obtain a sequence of upper bounds on the optimal value μ^* of a standard quadratic optimization problem, they propose

$$\mu_{\Delta(n,r)} := \min_{d \in \Delta(n,r)} d^T Q d, \quad r = 0, 1, 2, \ldots, \tag{41}$$

and establish that $\mu_{\Delta(n,r)}$ satisfies

$$\mu_{\Delta(n,r)} - \mu^* \leq \frac{1}{r+2}\left(\max_{i=1,\ldots,n} Q_{ii} - \mu^*\right) \leq \frac{1}{r+2}(\bar{\mu} - \mu^*), \quad r = 0, 1, \ldots. \tag{42}$$

By (9) and (38), we obtain

$$\mu_u^r = \min_{k=0,1,\ldots,r} \mu_{\Delta(n,k)}, \quad r = 0, 1, 2, \ldots, \tag{43}$$

which implies that $\mu_u^r \leq \mu_{\Delta(n,r)}$ for each $r = 0, 1, \ldots$. Therefore, we readily obtain

$$\mu_u^r - \mu^* \leq \frac{1}{r+2}\left(\max_{i=1,\ldots,n} Q_{ii} - \mu^*\right) \leq \frac{1}{r+2}(\bar{\mu} - \mu^*), \quad r = 0, 1, \ldots. \tag{44}$$

Bomze and de Klerk use these bounds to establish polynomial-time approximation schemes for standard quadratic optimization. Since our upper bounds are at least as good as theirs, it follows that our bounds yield at least the same approximation guarantees.

We remark that the sequence of the upper bounds $\{\mu_u^r\}$ is monotone nonincreasing by construction (cf. (43)). In contrast, the sequence of upper bounds $\{\mu_{\Delta(n,r)}\}$ may not be monotone in general as illustrated by the following simple example. Consider an instance of (33) with

$$Q = \begin{bmatrix} 3 & -1 \\ -1 & 1 \end{bmatrix}.$$

The unique global minimizer is $x^* = [1/3, 2/3]^T$ with $\mu^* = 1/3$. It is easy to verify that $\mu_{\Delta(n,0)} = 1/2; \mu_{\Delta(n,1)} = 1/3 = \mu^*; \mu_{\Delta(n,2)} = 3/8 > \mu^*$. Note that $\mu_{\Delta(n,r)} > \mu^*$ as long as $r+2$ is not a multiple of 3. In contrast, $\mu_u^0 = 1/2$ and $\mu_u^r = \mu^* = 1/3$ for each $r \geq 1$.

We remark that the idea of taking the unions $\delta(n,r)$ is the key that allows us to construct the hierarchy of outer polyhedral approximations to the copositive cone and consequently to obtain the sequence of monotone nonincreasing upper bounds. In a more recent paper by de Klerk, Laurent, and Parrilo [7], the authors employ the same idea of using the rational grid

$\Delta(n,r)$ to construct polynomial-time approximation schemes for the more general problem of minimizing a polynomial of fixed degree over the unit simplex, of which standard quadratic optimization is a special case. It follows from our discussion that using the "union grid" $\delta(n,r)$ will yield a hierarchy of polynomially computable bounds that are at least as sharp as $\mu_{\Delta(n,r)}$. This observation may lead to new insights for polynomial optimization.

4.2 The Stable Set Problem

Let $G = (V, E)$ be a simple, undirected graph, where $V = \{1, 2, \ldots, n\}$ denotes the set of vertices and E denotes the set of edges. A set $S \subseteq V$ is called a stable set if no two vertices in S are connected by an edge in E. The maximum stable set problem is that of finding the stable set with the largest cardinality in G. The size of the largest stable set, denoted by $\alpha(G)$, is called the stability number of G. The stability number cannot be approximated within a factor of $n^{1/2} - \epsilon$ for any $\epsilon > 0$ unless P = NP [10, Theorem 5.3], and within a factor of $n^{1-\epsilon}$ unless any problem in NP admits a probabilistic polynomial-time algorithm [10, Theorem 5.2].

Motzkin and Straus [12] established that the stability number satisfies

$$\frac{1}{\alpha(G)} = \min_{x \in \Delta_n} x^T (I + A_G) x, \tag{45}$$

where $A_G \in \mathcal{S}$ denotes the vertex adjacency matrix of G. In addition, for any maximum stable set $S^* \subseteq V$, $x^* := (1/|S^*|)\chi^{S^*} \in \mathbb{R}^n$ is an optimal solution of (45), where χ^{S^*} is the incidence vector of S^*. By (33),

$$\mu^* := \frac{1}{\alpha(G)} = \min \left\{ \langle (I + A_G), X \rangle : \langle E, X \rangle = 1, \quad X \in \mathcal{C}^* \right\}. \tag{46}$$

The reader is also referred to [8] for a derivation of a different but equivalent copositive programming reformulation of the stability number.

As in Section 4, let us define

$$\mu_l^r = -\frac{1}{r+1} + \frac{r+2}{r+1} \min_{x \in \Delta(n,r)} x^T (I + A_G) x, \quad r = 0, 1, \ldots, \tag{47}$$

where we used the fact that $x^T \mathrm{diag}(I + A_G) = e^T x = 1$, and

$$\mu_u^r = \min_{d \in \delta(n,r)} d^T (I + A_G) d, \quad r = 0, 1, \ldots. \tag{48}$$

The next proposition establishes closed-form solutions of μ_l^r and μ_u^r for each $r \in \mathbb{N}$.

Theorem 4.2 *Let $G = (V, E)$ be a graph and let μ_l^r and μ_u^r be as defined in (47) and (48), respectively. Then $\mu_l^r = 0$ if $r \leq \alpha(G) - 2$. If $r > \alpha(G) - 2$, then*

$$\mu_l^r = \frac{\binom{s}{2}\alpha(G) + st}{\binom{r+2}{2}}, \tag{49}$$

where $s \in \mathbb{N}$ and $t \in \mathbb{N}$ satisfy $r + 2 = s\alpha(G) + t$ and $0 \leq t < \alpha(G)$ (i.e., s is the quotient and t is the remainder obtained by dividing $r + 2$ to $\alpha(G)$) with the convention that $\binom{s}{2} = 0$ if $s < 2$. Similarly, μ_u^r satisfies

$$\mu_u^r = \begin{cases} \frac{1}{r+2} & \text{if } r \leq \alpha(G) - 2, \\ \frac{1}{\alpha(G)} & \text{otherwise,} \end{cases} \tag{50}$$

Proof. Let us define (see [8, 15])

$$\zeta^r(G) := \min\{\lambda : \lambda(I + A_G) - E \in \mathcal{I}^r\}, \quad r = 0, 1, \ldots. \tag{51}$$

This definition was introduced by de Klerk and Pasechnik [8], who proved that $\zeta^0(G) \geq \zeta^1(G) \geq \ldots \geq \alpha(G)$. Peña et al. [15, Theorem 1] established that

$$\zeta^r(G) = \frac{\binom{r+2}{2}}{\binom{s}{2}\alpha(G) + st}, \quad r = 0, 1, \ldots, \tag{52}$$

where s and t are nonnegative integers satisfying $r + 2 = s\alpha(G) + t$ and $0 \leq t < \alpha(G)$, with the conventions that $a/0 = +\infty$ for $a > 0$ and $\binom{a}{2} = 0$ for $a = 0$ and $a = 1$.

We establish (49) by showing that $\mu_u^r = 1/\zeta^r(G)$. Let us replace the decision variable λ in (51) by $y = 1/\lambda$. Then,

$$\begin{aligned} \zeta^r(G) &= \min\{(1/y) : (1/y)(I + A_G) - E \in \mathcal{I}^r\}, \\ &= \min\{(1/y) : I + A_G - yE \in \mathcal{I}^r\}, \\ &= \min\{(1/y) : yE + S = I + A_G, \quad S \in \mathcal{I}^r\}, \end{aligned}$$

which implies that

$$\frac{1}{\zeta^r(G)} = \max\{y : yE + S = I + A_G, \quad S \in \mathcal{I}^r\} = \mu_l^r,$$

where the second equality follows from the dual formulation (cf. (30)). This establishes our claim.

By [15, Corollary 3], $\zeta^r(G) = \infty$ if and only if $r \leq \alpha(G) - 2$. Therefore, $\mu_l^r = 0$ if and only if $r \leq \alpha(G) - 2$. For $r > \alpha(G) - 2$, the relation (49) follows from (52).

Let us now focus on μ_u^r. If $r \leq \alpha(G) - 2$, then there exists a stable set $S \subseteq V$ such that $|S| = r + 2$. Note that $\bar{d} := (1/(r+2))\chi^S \in \delta(n, r)$, where $\chi^S \in \mathbb{R}^n$ denotes the incidence vector of S. Therefore,

$$\mu_u^r \leq \frac{1}{(r+2)^2}(\chi^S)^T(I + A_G)\chi^S = \frac{1}{r+2}.$$

In order to establish the reverse inequality, consider any $d \in \delta(n, r)$. Let $P := \{i \in \{1, \ldots, n\} : d_i > 0\}$. Clearly, $|P| \leq r + 2 \leq \alpha(G) \leq n$. Let $d_P \in \mathbb{R}^{|P|}$ denote the restriction of d to its positive entries and let $G(P)$ denote the subgraph of G induced by $P \subseteq V$. We have

$$\frac{1}{r+2} \leq \frac{1}{|P|} \leq \frac{1}{\alpha(G(P))} = \min_{u \in \Delta_{|P|}} u^T(I + A_{G(P)})u \leq (d_P)^T(I + A_{G(P)})d_P = d^T(I + A_G)d, \quad (53)$$

where the third inequality follows from the fact that $d_P \in \Delta_{|P|}$. Therefore,

$$\mu_u^r \geq \frac{1}{r+2},$$

which, together with the previous inequality, implies that $\mu_u^r = 1/(r+2)$.

Finally, if $r > \alpha(G) - 2$, then $\bar{d} := (1/\alpha(G))\chi^{S^*} \in \delta(n, r)$, where $S^* \subseteq V$ is a stable set with the maximum cardinality. It follows that $\mu_u^r \leq 1/\alpha(G)$. Since $\mu_u^r \geq \mu^* = 1/\alpha(G)$, we obtain $\mu_u^r = 1/\alpha(G)$, which establishes (50). $\square$

We remark that Bomze and de Klerk [3] established that $\lfloor 1/\mu_l^r \rfloor = \alpha(G)$ if and only if $r \geq \alpha^2(G) - 1$. Similarly, Peña et al. [15] showed that $\lfloor \zeta^r(G) \rfloor = \alpha(G)$ if and only if $r \geq \alpha^2(G) - 1$. These two results are equivalent since $\zeta^r(G) = 1/\mu_l^r$. In addition, the exact characterization (52) implies that $1/\mu_l^r = \zeta^r(G) > \alpha(G)$ for each $r \in \mathbb{N}$ if $\alpha(G) > 1$ (see [15, Corollary 1]). In contrast, Theorem 4.2 implies that $\lfloor 1/\mu_u^r \rfloor = 1/\mu_u^r = r+2$ for each $r < \alpha(G) - 2$ and $\lfloor 1/\mu_u^r \rfloor = 1/\mu_u^r = \alpha(G)$ for all $r \geq \alpha(G) - 2$.

We close this section by the following immediate corollary.

Corollary 4.1 *Let $G = (V, E)$ be a graph and let μ_l^r and μ_u^r be as defined in (47) and (48), respectively. We have*

$$\mu_u^r - \mu_l^r = \begin{cases} \frac{1}{r+2}, & \text{if } r \leq \alpha(G) - 2, \\ \frac{1}{\alpha(G)} - \frac{\binom{s}{2}\alpha(G) + st}{\binom{r+2}{2}}, & \text{otherwise,} \end{cases} \tag{54}$$

where s and t are integers satisfying $r + 2 = s\alpha(G) + t$ and $0 \leq t < \alpha(G)$.

When specialized to the stable set problem, the first inequality in (39) in Proposition 4.1 is given by

$$\mu_u^r - \mu_l^r \leq \frac{1}{r+1}\left(1 - \frac{1}{\alpha(G)}\right) = \frac{\alpha(G) - 1}{(r+1)\alpha(G)}. \tag{55}$$

By Corollary 4.1, if $r > \alpha(G) - 2$ and $\alpha(G)$ divides $r + 2$, then $s = (r+2)/\alpha(G)$ and $t = 0$. Therefore,

$$\mu_u^r - \mu_l^r = \frac{1}{\alpha(G)} - \frac{s(s-1)\alpha(G)}{(r+2)(r+1)} = \frac{1}{\alpha(G)} - \frac{r + 2 - \alpha(G)}{(r+1)\alpha(G)} = \frac{\alpha(G) - 1}{(r+1)\alpha(G)}.$$

It follows from (55) that the upper bound (39) is tight and cannot, in general, be improved.

4.3 Error Bounds for Other Classes of Problems

We discuss the extensions of the error bound of Theorem 4.1 to other classes of optimization problems that can be reformulated as an instance of (CoP) or (CoD).

Consider the following general quadratic optimization problem over the unit simplex:

$$\text{(QP1)} \quad \min_{x \in \Delta_n} \; x^T Q x + 2c^T x,$$

where $Q \in \mathcal{S}^n$ and $c \in \mathbb{R}^n$ are given and $x \in \mathbb{R}^n$ constitutes the decision variables. Despite the fact that (QP1) seems to be a more general problem than the standard quadratic optimization problem, it turns out that (QP1) can be reformulated as the following instance of (32) (see, e.g., [4]):

$$\text{(SQP1)} \quad \min_{x \in \Delta_n} \; x^T \tilde{Q} x,$$

where $\tilde{Q} := Q + ec^T + ce^T \in \mathcal{S}$ and $e \in \mathbb{R}^n$ denotes the vector of all ones. It is easy to verify that the objective function values of (QP1) and (SQP1) coincide on the unit simplex. It follows from this reformulation that the error bound of Theorem 4.1 applies to any quadratic optimization problem over the unit simplex.

Let us now consider the more general problem of quadratic optimization over a polytope. Such a problem can be formulated as

$$\text{(QP2)} \quad \min\{x^T Q x + 2c^T x : x \in \text{conv}(\{v^1, v^2, \ldots, v^k\})\},$$

where $Q \in \mathcal{S}^n$ and $c \in \mathbb{R}^n$ are given, $x \in \mathbb{R}^n$ constitutes the decision variables, $v^1, v^2, \ldots, v^k \in \mathbb{R}^n$ denote the vertices of the feasible region and $\text{conv}(\cdot)$ denotes the convex hull. Since every feasible solution can be represented as a convex combination of the vertices $v^1, v^2, \ldots, v^k$, (QP2) can be reformulated as the following instance of quadratic optimization over the unit simplex:

$$\text{(QP3)} \quad \min\{u^T V^T Q V u + 2c^T V u : e^T u = 1, \quad u \geq 0\},$$

where $V = [v^1, v^2, \ldots, v^k] \in \mathbb{R}^{n \times k}$ and $u \in \mathbb{R}^k$ corresponds to the weights used in the convex combination. By using the aforementioned transformation, (QP3) can be reformulated as an instance of the standard quadratic optimization problem. Therefore, the error bound of Theorem 4.1 encompasses all quadratic optimization problems over a polytope.

We remark that quadratic optimization over a polytope subsumes several classes of well-known optimization problems such as the box-constrained quadratic optimization. However, the transformation of (QP2) into (QP3) requires the explicit information about each vertex of the feasible region. For box-constrained quadratic optimization, there are 2^n vertices, which implies that there is an exponential number of variables in (QP3). Therefore, despite the theoretical equivalence, the aforementioned transformation may not be useful in practice if there is a large number of vertices.

It is an interesting open problem whether similar error bounds can be constructed for other classes of optimization problems that can be cast as an instance of (CoP) or (CoD). We leave this as a future research problem.

5 Computational Results

In this section, we present and discuss our computational results. We set up and solved the linear programming formulations arising from the inner and outer approximations in MATLAB using the YALMIP [11] interface and the MATLAB Optimization Toolbox with MATLAB's linear programming solver `linprog`. The computational tests were conducted using MATLAB version 2008b on an AMD Athlon 64 X2 6000+ Dual Core Processor with 2 GB of RAM running under Linux.

We first report our computational results on several instances of standard quadratic optimization (33) taken from the literature. Let us consider the following examples from [3]:

$$Q_1 = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}, \quad \text{and} \quad Q_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The problem (33) corresponds to the computation of the stability number in a pentagon for $Q = Q_1$ and in the complement of an icosahedron for $Q = Q_2$. For the first example,

$\mu^* = 1/2$. We obtain $\mu_l^0 = 0, \mu_l^1 = 1/3, \mu_l^2 = 1/3$, and $\mu_l^3 = 2/5$ as the first few lower bounds. For the upper bound, we have $\mu_u^0 = 1/2$, which is already exact. For the second example, $\mu^* = 1/3$. Our computations reveal that the lower bounds are given by $\mu_l^0 = \mu_l^1 = 0, \mu_l^2 = 1/6, \mu_l^3 = 1/5$ while the upper bounds are $\mu_u^0 = 1/2$ and $\mu_u^1 = \mu^* = 1/3$. Observe that the upper bounds quickly match the stability number since $\alpha(G)$ is small for both examples (cf. Theorem 4.2).

The third example taken from [3] arises from a problem in population genetics:

$$Q_3 = \begin{pmatrix} -14 & -15 & -16 & 0 & 0 \\ -15 & -14 & -12.5 & -22.5 & -15 \\ -16 & -12.5 & -10 & -26.5 & -16 \\ 0 & -22.5 & -26.5 & 0 & 0 \\ 0 & -15 & -16 & 0 & -14 \end{pmatrix}.$$

The optimal value is given by $\mu^* = -16\frac{1}{3}$. The lower bounds are given by $\mu_l^0 = -26.5, \mu_l^1 = -21, \mu_l^2 = -19\frac{1}{3}$, and $\mu_l^3 = -18.9$ while the upper bounds are $\mu_u^0 = -15.75$ and $\mu_u^1 = -16\frac{1}{3} = \mu^*$, which is already exact.

The next example, also taken from [3] (see also [5]), corresponds to a portfolio optimization problem:

$$Q_4 = \begin{pmatrix} 0.9044 & 0.1054 & 0.5140 & 0.3322 & 0 \\ 0.1054 & 0.8715 & 0.7385 & 0.5866 & 0.9751 \\ 0.5140 & 0.7385 & 0.6936 & 0.5368 & 0.8086 \\ 0.3322 & 0.5866 & 0.5368 & 0.5633 & 0.7478 \\ 0 & 0.9751 & 0.8086 & 0.7478 & 1.2932 \end{pmatrix}.$$

For this example, $\mu^* = 0.4839$. The lower bounds are given by $\mu_l^0 = 0, \mu_l^1 = 0.3015, \mu_l^2 = 0.3484$, and $\mu_l^3 = 0.4005$. The upper bounds are $\mu_u^0 = 0.4967, \mu_u^1 = 0.4875, \mu_u^2 = 0.4875$, and $\mu_u^3 = 0.4867$.

Each of these examples illustrates that the upper bounds μ_u^r provide an accurate approximation of the optimal value μ^* already for small values of r.

By (18) and (19), the number of inequality constraints that define $\mathcal{I}^r$ is given by $\binom{n+r+1}{r+2}$. Similarly, an upper bound on the number of constraints that define $\mathcal{O}^r$ is given by (10). The exact numbers of constraints for $r = 0, 1, 2, 3$, respectively, are given by $\binom{n}{2}$, $n(n^2+6n-1)/6$, $n(n+5)(n^2+5n-2)/24$, and $n(n^4+15n^3+85n^2+165n-146)/120$. Therefore, the number of constraints quickly increases with r.

In an attempt to assess the accuracy of the bounds, we generated random instances of the quadratic optimization problem for different values of n. We used $n = 25$ and $n = 50$ in our experiments. For each choice of n, we generated 100 instances in which each entry of $Q \in \mathcal{S}$ was generated uniformly in $[0, 1]$. For each instance, we computed μ_l^r and μ_u^r and the approximation ratio μ_l^r/μ_u^r for the first few choices of r. Note that we have $0 \le \mu_l^r \le \mu^* \le \mu_u^r$ for each $r = 0, 1, \ldots$. Therefore, the ratio μ_l^r/μ_u^r is nondecreasing as r increases.

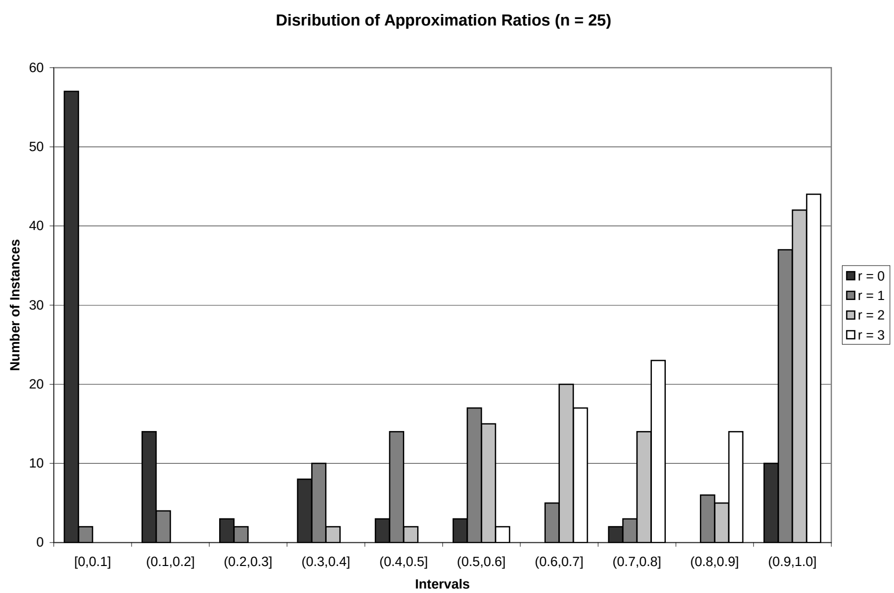

Figure 1: Distribution of μ_l^r/μ_u^r for $r = 0, 1, 2, 3$

Figure 1 illustrates the distribution of the approximation ratios μ_l^r/μ_u^r over 100 instances for $r = 0, 1, 2, 3$ using $n = 25$. The horizontal axis denotes the interval [0,1] divided into ten equal subintervals and the vertical axis indicates the number of instances whose approximation ratio falls into the corresponding interval. Note that the weight shifts towards larger ratios as r increases. It is worth noticing that the number of instances whose approximation ratio is equal to 1 is 10, 34, 39, and 39 for $r = 0, 1, 2, 3$, respectively. Therefore, the polyhedral approximations yield the exact solution on these instances.

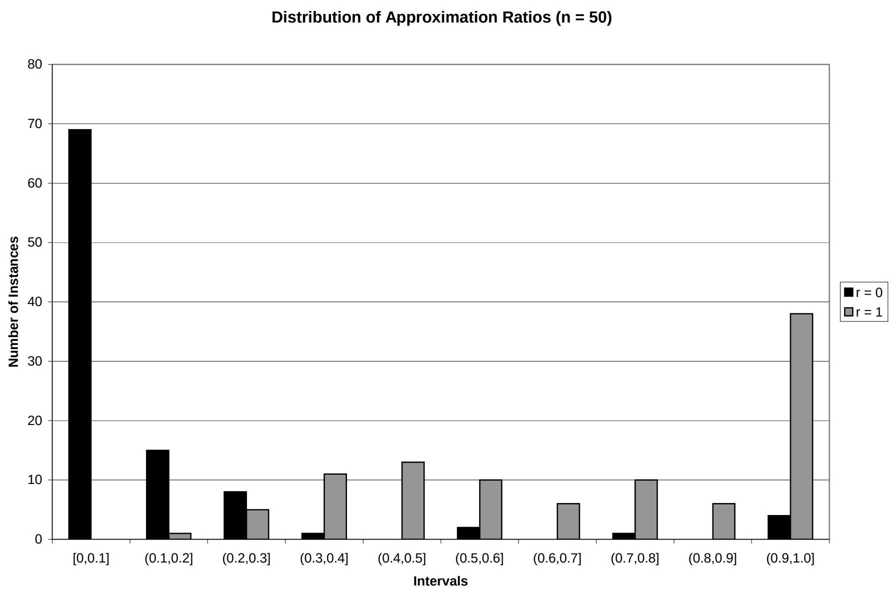

Figure 2: Distribution of μ_l^r/μ_u^r for $r = 0, 1$

In Figure 2, which is organized similarly to Figure 1, we present the distribution of approximation ratios μ_l^r/μ_u^r over 100 instances using $n = 50$ for $r = 0, 1$. For larger values of r, we ran into memory problems. Similarly, the number of instances shifts towards the larger ratios as r increases from 0 to 1. The approximation was exact on 4 and 34 instances for $r = 0$ and $r = 1$, respectively.

Next, we present some statistics in an attempt to shed light onto the average behavior of the approximation ratios. In Table 1, we report the average approximation ratios for all combinations of n and r. As expected, the approximation ratios improve as r increases. The average ratios indicate the quality of approximation. It is worth noting that the solutions returned by the polyhedral approximations are already within 15% of the optimal solution for $n = 25$ and $r = 3$ on the average.

	μ_l^r/μ_u^r			
	$r = 0$	$r = 1$	$r = 2$	$r = 3$
$n = 25$	0.2238	0.6754	0.7966	0.8497
$n = 50$	0.1255	0.7095	-	-

Table 1: Average approximation ratios μ_l^r/μ_u^r

Finally, we report the average computation times in CPU seconds for all combinations of n and r in Table 2. Observe that the cost of computing the bounds quickly increases as n and r increase. We briefly discuss how the computation times can be improved in Section 6.

We ran into memory problems for instances with $n > 50$ and $r > 1$. We therefore did not include computational experiments for larger instances. We remark, however, that the computational efficiency can potentially be improved by using a state-of-the-art solver such as CPLEX. An efficient way of computing the bounds for larger values of r even for small values of n still remains a challenge.

	μ_l^0	μ_u^0	μ_l^1	μ_u^1	μ_l^2	μ_u^2	μ_l^3	μ_u^3
$n = 25$	0.66	0.74	1.99	2.31	34.42	52.03	1098.3	1760.4
$n = 50$	30.75	31.25	68.53	79.71	-	-	-	-

Table 2: Average computation times in CPU seconds

Our computational results reveal that the polyhedral bounds are fairly accurate even for small values of r on randomly generated standard quadratic optimization problems. However, the cost of computing the bounds increases drastically as the value of r increases.

6 Concluding Remarks

In this paper, we proposed a hierarchy of increasingly better outer polyhedral approximations of the copositive cone that is exact in the limit. By combining our hierarchy with a previously proposed hierarchy of inner polyhedral approximations, we obtained two sequences of improving upper and lower bounds on the optimal value of a copositive program. We established that both of these sequences converge to the optimal value under primal and dual regularity. For standard quadratic optimization problems, we derived tight bounds on the duality gap resulting from these sequences. We provided closed-form solutions for the upper and lower bounds for the stable set problem. Our computational experiments revealed the quality of the bounds on randomly generated standard quadratic optimization problems.

In our experiments, we included all of the constraints that define the polyhedral cones $(\mathcal{I}^r)^*$ and $(\mathcal{O}^r)^*$. Similar to the approach of Bundfuss and Dür [5], the inner and outer approximations can be adaptively guided using the objective function. For instance, rather than adding all the inequalities that define the inner and polyhedral cones, one may include only (a subset of) the violated constraints as in a cutting plane scheme. We believe that such an approach may considerably increase the value of r for which the bounds μ_l^r and μ_u^r can be computed. We intend to pursue this direction in the near future.

Acknowledgments

The author gratefully acknowledges the valuable comments from his colleagues and two anonymous referees on an earlier version of this manuscript.

References

[1] A. Berman and N. Shaked-Monderer. *Completely Positive Matrices.* World Scientific, River Edge, New Jersey, 2003.

[2] I. M. Bomze. On standard quadratic optimization problems. *Journal of Global Optimization*, 13(4):369–387, 1998.

[3] I. M. Bomze and E. de Klerk. Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. *Journal of Global Optimization*, 24:163–185, 2002.

[4] I. M. Bomze, M. Dür, E. de Klerk, C. Roos, A. Quist, and T. Terlaky. On copositive programming and standard quadratic optimization problems. *Journal of Global Optimization*, 18(4):301–320, 2000.

[5] S. Bundfuss and M. Dür. An adaptive linear approximation algorithm for copositive programs. *SIAM Journal on Optimization*, 20(1):30–53, 2009.

[6] S. Burer. On the copositive representation of binary and continuous nonconvex quadratic programs. *Mathematical Programming*, 120(2):479–495, 2009.

[7] E. de Klerk, M. Laurent, and P. Parrilo. A PTAS for the minimization of polynomials of fixed degree over the simplex. *Theoretical Computer Science*, 361:210–225, 2006.

[8] E. de Klerk and D. V. Pasechnik. Approximation of the stability number of a graph via copositive programming. *SIAM Journal on Optimization*, 12(4):875–892, 2002.

[9] M. Dür and G. Still. Interior points of the completely positive cone. *Electronic Journal of Linear Algebra*, 17:48–53, 2008.

[10] J. Håstad. Clique is hard to approximate within $n^{1-\epsilon}$. *Acta Mathematica*, 182(1):105–142, 1999.

[11] J. Löfberg. YALMIP: A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[12] T. S. Motzkin and E. G. Straus. Maxima for graphs and a new proof of a theorem of Turán. *Canadian Journal of Mathematics*, 17:533–540, 1965.

[13] K. G. Murty and S. N. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Mathematical Programming*, 39(2):117–129, 1987.

[14] P. A. Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization.* PhD thesis, California Institute of Technology, Pasadena, CA, 2000.

[15] J. Peña, J. Vera, and L. F. Zuluaga. Computing the stability number of a graph via linear and semidefinite programming. *SIAM Journal on Optimization*, 18(1):87–105, 2007.