

Copositive Programming via Semi-Infinite Optimization

Faizan Ahmed · Mirjam Dür · Georg Still

Received: 19 June 2012 / Accepted: 18 May 2013 / Published online: 8 June 2013
© Springer Science+Business Media New York 2013

Abstract Copositive programming (CP) can be regarded as a special instance of linear semi-infinite programming (SIP). We study CP from the viewpoint of SIP and discuss optimality and duality results. Different approximation schemes for solving CP are interpreted as discretization schemes in SIP. This leads to sharp explicit error bounds for the values and solutions in dependence on the mesh size. Examples illustrate the structure of the original program and the approximation schemes.

Keywords Copositive programming · Semi-infinite programming · Optimality and duality · Discretization method · Order of maximizer

1 Introduction

During the last years copositive programming has attracted much attention due to the fact that many difficult (NP-hard) quadratic and integer programs can be reformulated equivalently as CP (see, e.g., [1–3]). This in particular implies that copositive programming is NP-hard.

However, CP represents a specially structured convex program. So, the hope is that this structure allows the construction of new methods for approximately solving integer programs.

Communicated by Johannes Jahn.

F. Ahmed · G. Still (✉)
Department of Mathematics, University of Twente, 7500 AE Enschede, the Netherlands
e-mail: g.j.still@utwente.nl

F. Ahmed
e-mail: f.ahmed@utwente.nl

M. Dür
Department of Mathematics, University of Trier, 54286 Trier, Germany
e-mail: duer@uni-trier.de

CP can also be regarded as a special instance of linear semi-infinite programming (SIP). The aim of this paper is to take advantage of this relation.

The article is organized as follows. Section 2 provides an introduction into copositive programming and linear semi-infinite optimization. In Sect. 3, first order optimality conditions and duality results of SIP are applied to CP leading to known results but also to a new insight. In Sect. 4, we reinterpret approximation schemes for solving CP (in [1, 4, 5]) as discretization methods in SIP. This leads to new explicit error bounds between the approximate and the original problem. For the schemes in [5], the approximation error for the feasible sets and the value functions are shown to behave like $\mathcal{O}(d^2)$ when the mesh-size d goes to zero. Some examples compare the structure of the original program with that of the approximation schemes.

In Sect. 5, it is proven that for another scheme (in [1]) the (sharp) order of convergence is $\mathcal{O}(d)$. Section 6 gives error bounds for the maximizers in dependence on the order of the maximizer of the original program. We also show by examples that maximizers of any order can occur in copositive programming. Section 7 comments briefly on the fact that the results can easily be extended from copositive to set-copositive programming.

2 Preliminaries

In this paper, we consider the pair of primal/dual copositive programs (CP):

$$(P) \quad \max_{x \in \mathbb{R}^n} c^T x \quad \text{s.t.} \quad B - \sum_{i=1}^n x_i A_i \in \mathcal{C}_m,$$

$$(D) \quad \min_{Y \in \mathcal{S}_m} \langle Y, B \rangle \quad \text{s.t.} \quad \langle Y, A_i \rangle = c_i \quad (i = 1, \dots, n), \quad Y \in \mathcal{C}_m^*,$$

with $c \in \mathbb{R}^n$ and $A_i, B \in \mathcal{S}_m$. \mathcal{S}_m denotes the set of symmetric $m \times m$ -matrices, and $\langle Y, B \rangle = \sum_{ij} y_{ij} b_{ij}$ the inner product of $Y = (y_{ij})$ and $B = (b_{ij})$. We assume throughout that the matrices A_i ($i = 1, \dots, n$) are linearly independent. \mathcal{C}_m denotes the cone of copositive matrices, and \mathcal{C}_m^* its dual, the cone of completely positive matrices:

$$\mathcal{C}_m := \left\{ A \in \mathcal{S}_m : z^T A z \geq 0 \text{ for all } z \in \mathbb{R}_+^m \right\},$$

$$\mathcal{C}_m^* := \left\{ A \in \mathcal{S}_m : A = \sum_{j=1}^k z_j z_j^T \text{ with } z_j \in \mathbb{R}_+^m, k \in \mathbb{N} \right\}.$$

In this definition, \mathbb{R}_+^m denotes the non-negative orthant. It will be convenient to use the fact [6] that the interior $\text{int} \mathcal{C}_m$ of the cone \mathcal{C}_m is the cone of so-called strictly copositive matrices:

$$\text{int} \mathcal{C}_m = \left\{ A \in \mathcal{C}_m : z^T A z > 0 \text{ for all } z \in \mathbb{R}_+^m, z \neq 0 \right\}. \tag{1}$$

For a characterization of the interior of C_m^* we refer to [7, 8]. Linear semi-infinite programs are of the form:

$$(SIP_P) \quad \max_{x \in \mathbb{R}^n} c^T x \quad \text{s.t.} \quad b(z) - a(z)^T x \geq 0 \quad \forall z \in Z,$$

with an (infinite) compact index set $Z \subset \mathbb{R}^m$ and continuous functions $a : Z \rightarrow \mathbb{R}^n$ and $b : Z \rightarrow \mathbb{R}$. The (Haar-) dual reads:

$$(SIP_D) \quad \min_{y_z} \sum_{z \in Z} y_z b(z) \quad \text{s.t.} \quad \sum_{z \in Z} y_z a(z) = c, \quad y_z \geq 0,$$

where the min is taken over all finite sums. For an introduction to (linear) SIP we refer, e.g., to [9, 10]. Note that the condition $A \in C_m$ can be equivalently expressed by either of the conditions:

$$\begin{aligned} z^T A z \geq 0 \quad \forall z \in \mathcal{B}_m &:= \{z \in \mathbb{R}_+^m : \|z\| = 1\} \text{ (unit orthant),} \\ z^T A z \geq 0 \quad \forall z \in \Delta_m &:= \left\{ z \in \mathbb{R}_+^m : \sum_{i=1}^m z_i = 1 \right\} \text{ (unit simplex).} \end{aligned}$$

In view of this relation, the primal CP can be written as a SIP with

$$a(z) = (z^T A_1 z, \dots, z^T A_n z)^T, \quad b(z) = z^T B z, \quad \text{and} \quad Z \in \{\mathcal{B}_m, \Delta_m\}. \quad (2)$$

In this paper, we always take $Z = \Delta_m$, and defining $F(x) := (B - \sum_{i=1}^n x_i A_i)$, we write the copositive primal problem (P) in SIP form:

$$(P) \quad \max_{x \in \mathbb{R}^n} c^T x \quad \text{s.t.} \quad z^T F(x) z \geq 0 \quad \forall z \in Z := \Delta_m. \quad (3)$$

In view of (2), the feasibility condition for (SIP_D) becomes

$$c_i = \sum_{z \in Z} y_z \langle z z^T, A_i \rangle \quad (i = 1, \dots, n), \quad y_z \geq 0,$$

and with $Y := \sum_{z \in Z} y_z z z^T \in C_m^*$ this coincides with the feasibility condition $c_i = \langle Y, A_i \rangle$ ($i = 1, \dots, n$) of (D). Moreover,

$$\sum_{z \in Z} y_z b(z) = \sum_{z \in Z} y_z \langle z z^T, B \rangle = \langle Y, B \rangle.$$

So, the dual (SIP_D) of (P) in SIP form (3) is equivalent to the CP dual (D) and we simply denote both versions by (D).

3 Optimality Conditions and Duality for CP via SIP

From the SIP form of CP it follows that any result for linear SIP can directly be translated to CP. We will do this for some optimality conditions and duality results.

Optimality conditions for SIP are usually presented in terms of Karush–Kuhn–Tucker (KKT) conditions for a feasible candidate maximizer \bar{x} . Denoting the *active index-set* at a point \bar{x} by $Z(\bar{x}) := \{z \in Z : z^T F(\bar{x})z = 0\}$ and the Lagrange multipliers by y_j , the KKT conditions for the general linear semi-infinite problem (SIP_P) read:

$$\sum_{j=1}^k y_j a(z_j) = c \quad \text{for } z_j \in Z(\bar{x}) \text{ and } y_j \geq 0 \ (j = 1, \dots, k).$$

Using (2), for the copositive problem in SIP-form (3), this condition translates to

$$c = \sum_{j=1}^k y_j a(z_j) = \sum_{j=1}^k y_j \begin{pmatrix} z_j^T A_1 z_j \\ \vdots \\ z_j^T A_n z_j \end{pmatrix}, \quad z_j \in Z(\bar{x}), \ y_j \geq 0 \ (j = 1, \dots, k). \quad (4)$$

It is important to note that any solution of the KKT system with feasible \bar{x} automatically yields a minimizer \bar{Y} of the dual program (D):

$$\bar{Y} := \sum_{j=1}^k y_j z_j z_j^T \in C_m^*. \quad (5)$$

Observe that, by Carathéodory’s Lemma for cones (see, e.g., [11]), we can assume that

$$\text{the KKT condition (4) is satisfied with } k \leq n \text{ active points } z_j \in Z(\bar{x}). \quad (6)$$

This implies that the dual minimizer \bar{Y} allows a representation (5) with $k \leq n$, i.e., $\bar{Y} \in C_m^*$ has CP-rank $\leq n$. The CP-rank of a matrix $Y \in C_m^*$ is defined as the smallest number r possible in a factorization $Y = \sum_{j=1}^r v_j v_j^T$ with $v_j \in \mathbb{R}_+^n$. Determining the CP-rank of an arbitrary matrix in C_m^* is an open problem, see [6].

Before applying the standard results of SIP to copositive programming, we have to translate the primal/dual constraint qualification (Slater condition) from SIP to the copositive terminology.

Lemma 3.1 *Consider the copositive problem in its SIP-formulation (3). The primal SIP constraint qualification*

$$(CQ_P) : \quad z^T F(x_0)z \geq \sigma_0 > 0, \quad \text{for all } z \in Z \text{ and for some } \sigma_0 > 0 \quad (7)$$

is satisfied for $x_0 \in \mathbb{R}^n$ if and only if $F(x_0) \in \text{int} C_m$. The dual SIP constraint qualification

$$(CQ_D) : \quad c \in \text{int } M, \quad \text{with } M := \text{cone}\{a(z) : z \in Z\}$$

holds if and only if there exists Y_0 feasible for (D) such that $Y_0 \in \text{int} C_m^$.*

Proof The fact that (CQ_P) implies $F(x_0) \in \text{int}C_m$ follows immediately from (1).

For the converse we use the (Frobenius) matrix-norm $\|A\| = (\sum_{ij} a_{ij}^2)^{1/2}$. Let $F(x_0) \in \text{int}C_m$. Then there exists $\varepsilon > 0$ such that $F \in C_m$ for all F with $\|F - F(x_0)\| \leq \varepsilon$. Define $F := F(x_0) - \frac{\varepsilon}{\sqrt{m}}I$. Then $\|F - F(x_0)\| \leq \varepsilon$ and thus $F \in C_m$. Consequently,

$$0 \leq z^T Fz = z^T F(x_0)z - \frac{\varepsilon}{\sqrt{m}}z^T z \quad \text{for all } z \in Z.$$

Using $z^T z \geq \frac{1}{m}$ for $z \in Z$, we obtain $z^T F(x_0)z \geq \frac{\varepsilon}{\sqrt{mm}} =: \sigma_0 > 0$ for all $z \in Z$.

To prove the equivalence of the dual constraint qualifications we define the mapping $c(Y) := (\langle A_1, Y \rangle, \dots, \langle A_n, Y \rangle)^T$. We first show that

$$Y \in C_m^* \quad \Rightarrow \quad c(Y) \in M. \tag{8}$$

To see this, note that $Y \in C_m^*$ has a rank-one representation $Y = \sum_{j=1}^k v_j v_j^T$ with $0 \neq v_j \geq 0$ for all $j = 1, \dots, k$. Define $z_j := v_j / (v_j^T e)$ to obtain $z_j \in Z$, and $y_j := (v_j^T e)^2 > 0$. Then $Y = \sum_{j=1}^k y_j z_j z_j^T$. Therefore we get

$$c(Y) = \sum_{j=1}^k y_j (\langle A_1, z_j z_j^T \rangle, \dots, \langle A_n, z_j z_j^T \rangle)^T = \sum_{j=1}^k y_j a(z_j) \in M,$$

and (8) is proved. Now let $Y_0 \in \text{int}C_m^*$ be feasible for (D), i.e., $c(Y_0) = c$. To prove $c \in \text{int}M$ we assert that there exists some $\varepsilon > 0$ such that, for any $\gamma \in \mathbb{R}$, $|\gamma| < \varepsilon$, the relation

$$c + \gamma e_k \in M \quad \text{holds for all (standard basis) vectors } e_k \ (k = 1, \dots, n). \tag{9}$$

To show this we note that, since the A_i 's are linearly independent, for any k the linear system $c(Y_k) = (\langle A_1, Y_k \rangle, \dots, \langle A_n, Y_k \rangle)^T = e_k$ has a solution $\tilde{Y}_k \in S_m$. Since $Y_0 \in \text{int}C_m^*$, there exists some $\varepsilon > 0$ such that for all γ , $|\gamma| < \varepsilon$:

$$Y_k := Y_0 + \gamma \tilde{Y}_k \in C_m^* \quad \text{for all } k = 1, \dots, n.$$

Using (8) and $c(Y_0) = c$ we get $M \ni c(Y_k) = c(Y_0) + \gamma c(\tilde{Y}_k) = c + \gamma e_k$, which proves (9).

We finally show that (CQ_D) yields some Y_0 feasible for (D) with $Y_0 \in \text{int}C_m^*$. To do so, choose any $Y_* \in \text{int}C_m^*$ and define $b := c(Y_*) = (\langle A_1, Y_* \rangle, \dots, \langle A_n, Y_* \rangle)^T$. Since $c \in \text{int}M$, we have for some $\varepsilon > 0$ that $c - \varepsilon b \in M$, which means that for some $y_j \geq 0, z_j \in Z$ ($j = 1, \dots, k$) we have

$$c - \varepsilon b = \sum_{j=1}^k y_j a(z_j) = \sum_{j=1}^k y_j c(z_j z_j^T).$$

Defining $Y := \sum_{j=1}^k y_j z_j z_j^T \in C_m^*$, we find that $c(Y) = c - \varepsilon b$ by construction. Next, define $Y_0 := Y + \varepsilon Y_*$. Then $Y_0 \in \text{int}C_m^*$ because $Y \in C_m^*, Y_* \in \text{int}C_m^*$ and C_m^* is a

convex cone. Moreover, $c(Y_0) = c(Y) + \varepsilon c(Y_*) = c - \varepsilon b + \varepsilon b = c$, which means that Y_0 is feasible for (D) . This completes the proof. \square

We mention the standard first order optimality conditions for (the convex program) SIP (see, e.g., [10, Theorem 3] and [12, Theorem 2(b)] for a proof in SIP context):

If a feasible point \bar{x} satisfies the KKT condition (4) then \bar{x} is a (global) maximizer of (P) . On the other hand, under (CQ_P) a maximizer \bar{x} of (P) must satisfy the KKT conditions.

We emphasize that relation (6) implies that, under (CQ_P) to any maximizer \bar{x} of (P) , there always exists a corresponding (complementary) optimal solution \bar{Y} of (D) that has CP-rank $\leq n$.

Finally, we apply some standard duality and existence results in SIP to the copositive problem. Let $v(P)$, $v(D)$ and $\mathcal{F}(P)$, $\mathcal{F}(D)$ denote the optimal values and feasible sets of (P) , (D) , respectively. We introduce the *upper/lower level sets*

$$\mathcal{F}_\alpha(P) = \{x \in \mathcal{F}(P) : c^T x \geq \alpha\} \quad \text{and} \quad \mathcal{F}_\alpha(D) = \{Y \in \mathcal{F}(D) : \langle Y, B \rangle \leq \alpha\}.$$

Let $\mathcal{S}(P)$ denote the set of maximizers of (P) . Note that, in general, for SIP (and CP) strong duality need not hold and solutions of (P) and/or (D) need not exist (see [13] for many examples). However, the following is true for linear SIP.

Theorem 3.1 *We have:*

- (a) *If either (CQ_P) or (CQ_D) holds, then $v(D) = v(P)$.*
- (b) *Let $\mathcal{F}(P)$ be nonempty. Then*

$$(CQ_D) \text{ holds} \iff \forall \alpha \in \mathbb{R}: \mathcal{F}_\alpha(P) \text{ is compact} \iff \emptyset \neq \mathcal{S}(P) \text{ compact.}$$

Thus, if one of these equivalent conditions holds, then a solution of (P) exists.

A result as in (b) also holds for the dual problem.

Proof See, e.g., [12, Theorems 6.9, 6.11] and [10, Theorem 4] for the second equivalence in (b). \square

4 Application of SIP: Discretization Methods for CP

Due to the SIP representation of CP, any solution method of SIP can directly be applied to CP. In this paper, we only consider discretization methods. Recently an inner and outer approximation algorithm for CP has been proposed and analyzed by Bundfuss and Dür [5]. We re-analyze this approach in the light of discretization methods in SIP as outlined in [14]. This will lead to additional insight and explicit error bounds.

We start with the CP in SIP-form (3) with $Z = \Delta_m$. The approach in [5] is based on the following partition of Δ_m .

Definition 4.1 We partition the unit simplex $Z = \Delta_m$ into finitely many sub-simplices $\Delta^1, \dots, \Delta^k$ of Δ_m such that

$$\Delta_m = \bigcup_{l=1}^k \Delta^l \quad \text{and} \quad \text{int } \Delta^l \cap \text{int } \Delta^p = \emptyset \quad \text{for } l \neq p.$$

This partition defines a mesh-size d , a discretization Z_d and a set E_d of “edges” (pairs of vertices):

$$\begin{aligned} Z_d &= \{v_j : v_j \text{ is a vertex of } \Delta^l \text{ for some } l\}, \\ E_d &= \{(v_i, v_j) : v_i, v_j \text{ are vertices in the same } \Delta^l \text{ for some } l \text{ (possibly } i = j)\}, \\ d &= \max\{\|v_i - v_j\| : (v_i, v_j) \in E_d\}. \end{aligned}$$

In [5], the following outer and inner approximation schemes for (3) are given:

$$\begin{aligned} (P_d) \quad & \max_{x \in \mathbb{R}^n} c^T x \quad \text{s.t.} \quad z^T F(x)z \geq 0 \quad \forall z \in Z_d, \\ (\tilde{P}_d) \quad & \max_{x \in \mathbb{R}^n} c^T x \quad \text{s.t.} \quad u^T F(x)v \geq 0 \quad \forall (u, v) \in E_d. \end{aligned}$$

Note that (P_d) represents a special instance of a discretization scheme in SIP. (\tilde{P}_d) provides feasible points for the original copositive problem (P) , see [5] and below. Observe that both (P_d) and (\tilde{P}_d) are LPs.

Remark 4.1 Note that any point $z \in Z = \Delta_m$ is contained in one of the sub-simplices Δ^l and thus $z \in \Delta^l$ can be written as a convex combination $z = \sum_v \lambda_v v_v$, with $\sum_v \lambda_v = 1, \lambda_v \geq 0$ of vertices v_v of Δ^l . Consequently, for any $z \in Z$, the inequality $\min_{z_j \in Z_d} \|z - z_j\| \leq d$ holds so that d above really defines a mesh-size:

$$d \geq \max_{z \in Z} \min_{z_j \in Z_d} \|z - z_j\|.$$

In the following, $\mathcal{F}(P), \mathcal{F}(P_d), \mathcal{F}(\tilde{P}_d)$ and $v(P), v(P_d), v(\tilde{P}_d)$ denote the feasible sets and the maximum values of $(P), (P_d), (\tilde{P}_d)$, respectively. The vector \bar{x} is always a maximizer of (P) and \bar{x}_d, \tilde{x}_d are feasible points (possibly maximizers) of $(P_d), (\tilde{P}_d)$. We are now going to discuss some of the convergence results of [14] for our special program (P) in terms of the mesh-size d in an explicit form. The proofs are independent and mainly based on the following two relations: For any $F \in \mathcal{S}_m$ and $z, u \in \mathbb{R}^m$ we have

$$z^T F u = \frac{1}{2} [z^T F z + u^T F u - (z - u)^T F (z - u)]. \tag{10}$$

Moreover, as mentioned earlier, for every $z \in \Delta^l \subseteq Z$ we have the representation $z = \sum_v \lambda_v v_v$ with v_v the vertices of $\Delta^l, \lambda_v \geq 0$, and $\sum_v \lambda_v = 1$. This gives:

$$v_v^T F v_\mu \geq \gamma, \quad \forall (v_v, v_\mu) \in E_d \quad \Rightarrow \quad z^T F z = \sum_{v,\mu} \lambda_v \lambda_\mu v_v^T F v_\mu \geq \gamma, \quad \forall z \in \Delta_m. \tag{11}$$

Clearly, $\mathcal{F}(P) \subset \mathcal{F}(P_d)$ holds, and using (11) for $\gamma = 0$ we obtain the relations

$$\mathcal{F}(\tilde{P}_d) \subset \mathcal{F}(P) \subset \mathcal{F}(P_d) \quad \text{and thus} \quad v(\tilde{P}_d) \leq v(P) \leq v(P_d). \tag{12}$$

We are interested in accurate bounds, e.g., for $v(P_d) - v(P)$ and $v(P) - v(\tilde{P}_d)$, depending explicitly on the mesh-size d . From [14], we know that even for nonlinear SIP under a constraint qualification, the approximation error between $\mathcal{F}(P), v(P)$ and $\mathcal{F}(P_d), v(P_d)$ behaves like $\mathcal{O}(d^2)$ in the mesh-size d , provided that the discretization Z_d of Z “covers all boundary parts of Z of all dimensions.” In the above discretization scheme this is automatically fulfilled.

The next lemma shows that the inner approximation (\tilde{P}_d) yields points feasible for the original program (P) and the outer approximation (P_d) generates points with an infeasibility error of order $\mathcal{O}(d^2)$. In the following, $\|z\|$ denotes the 2-norm in \mathbb{R}^m and $\|F\|$, the Frobenius norm in \mathcal{S}_m .

Lemma 4.1 *Let \bar{x}_d, \tilde{x}_d be feasible for $(P_d), (\tilde{P}_d)$. Then for all $z \in Z$ and for all d we have:*

- (a) $z^T F(\bar{x}_d)z \geq -\frac{1}{2}\|F(\bar{x}_d)\| \cdot d^2.$
- (b) $z^T F(\tilde{x}_d)z \geq 0.$

So \tilde{x}_d is feasible for (P) , and \bar{x}_d is feasible up to an error of order $\mathcal{O}(d^2)$.

Proof Let $F = F(\bar{x}_d)$. Using $z^T Fz \geq 0$ for all $z \in Z_d$, we find from (10) that for all $(z, u) \in E_d$

$$z^T Fu \geq -\frac{1}{2}(z - u)^T F(z - u) \geq -\frac{1}{2}\|F\|\|z - u\|^2 \geq -\frac{1}{2}\|F\| \cdot d^2.$$

The second inequality follows from the fact that with the 2-norms the relation $\|Fz\| \leq \|F\|\|z\|$ holds. In view of (11), this shows (a). Letting $F := F(\tilde{x}_d)$, (b) follows from (11) with $\gamma = 0$. □

Assuming a strictly feasible point x_0 we show that small perturbations of any feasible point \bar{x}_d for (P_d) leads to points in $\mathcal{F}(P)$ or even $\mathcal{F}(\tilde{P}_d)$.

Lemma 4.2 *Let (CQ_P) be satisfied for $x_0 \in \mathcal{F}(P)$ with $\sigma_0 > 0$ (see (7)). Then, for any \bar{x}_d feasible for (P_d) and d small enough, we have:*

- (a) $\bar{x}_d^* := \bar{x}_d + \rho d^2(x_0 - \bar{x}_d) \in \mathcal{F}(P)$ for $\rho \geq \frac{\|F(\bar{x}_d)\|}{2\sigma_0}$ and $0 < \rho d^2 < 1.$
- (b) $\tilde{x}_d^* := \bar{x}_d + \tau d^2(x_0 - \bar{x}_d) \in \mathcal{F}(\tilde{P}_d)$ for $\tau \geq \frac{\|F(\bar{x}_d)\|}{2\sigma_0 + d^2(\|F(\bar{x}_d)\| - \|F(x_0)\|)}$ and $0 < \tau d^2 < 1.$

Recall that $\mathcal{F}(\tilde{P}_d) \subset \mathcal{F}(P)$ holds, cf. (12).

- (c) *If \bar{x}_d is a solution of (P_d) , i.e., $c^T \bar{x}_d = v(P_d)$, it follows that $0 \leq v(P_d) - v(\tilde{P}_d) \leq \tau[c^T(\bar{x}_d - x_0)] \cdot d^2$ for τ satisfying the bound in (b).*

Proof Recall that (CQ_P) means that $z^T F(x_0)z \geq \sigma_0 > 0$ for all $z \in Z$. Using this, the fact that $F(\bar{x}_d^*) = (1 - \rho d^2)F(\bar{x}_d) + \rho d^2 F(x_0)$, and Lemma 4.1, we see that for

any $z \in Z$ and $0 \leq 1 - \rho d^2$,

$$\begin{aligned} z^T F(\bar{x}_d^*)z &= (1 - \rho d^2)z^T F(\bar{x}_d)z + \rho d^2 z^T F(x_0)z \\ &\geq -\frac{1}{2}(1 - \rho d^2) \|F(\bar{x}_d)\| \cdot d^2 + \rho d^2 \sigma_0 \\ &\geq d^2 \left(\rho \sigma_0 - \frac{1}{2} \|F(\bar{x}_d)\| \right), \end{aligned}$$

which shows (a). Part (b) is proven similarly. Here, observing

$$F(\tilde{x}_d^*) = (1 - \tau d^2)F(\bar{x}_d) + \tau d^2 F(x_0),$$

for any pair $(z, u) \in E_d$, we find using (10), $z^T F(x_0)z \geq \sigma_0$ and $z^T F(\bar{x}_d)z \geq 0$:

$$\begin{aligned} z^T F(\tilde{x}_d^*)u &= (1 - \tau d^2) \frac{1}{2} [z^T F(\bar{x}_d)z + u^T F(\bar{x}_d)u - (z - u)^T F(\bar{x}_d)(z - u)] \\ &\quad + \tau d^2 \frac{1}{2} [z^T F(x_0)z + u^T F(x_0)u - (z - u)^T F(x_0)(z - u)] \\ &\geq -(1 - \tau d^2) \frac{1}{2} \|F(\bar{x}_d)\| d^2 + \tau d^2 \left(\sigma_0 - d^2 \frac{\|F(x_0)\|}{2} \right) \\ &= d^2 \left(-\frac{\|F(\bar{x}_d)\|}{2} + \tau \left[\sigma_0 + \frac{d^2}{2} (\|F(\bar{x}_d)\| - \|F(x_0)\|) \right] \right) \geq 0 \end{aligned}$$

if τ is chosen as stated (assuming $\|F(x_0)\|d^2 \leq \sigma_0$, implying $\tau > 0$). The inequality (c) for the maximum values is deduced easily using that \tilde{x}_d^* is feasible for (\tilde{P}_d) :

$$0 \leq v(P_d) - v(\tilde{P}_d) \leq c^T (\bar{x}_d - \tilde{x}_d^*) = [c^T (\bar{x}_d - x_0)\tau] \cdot d^2. \quad \square$$

Observe that the bounds in Lemma 4.2(c) depend on the actual solutions \bar{x}_d of P_d . In order to use these bounds (a-priori) we must assure that the solutions \bar{x}_d exist and that they are bounded. As we shall see below, the key assumption here is a dual constraint qualification. We define the distance between a point x and the set $\mathcal{S}(P)$ of maximizers of (P) ,

$$\delta(x, \mathcal{S}(P)) := \min \{ \|x - \bar{x}\| : \bar{x} \in \mathcal{S}(P) \}.$$

Under feasibility of (P) , the existence of solutions \bar{x}_d of (P_d) and the convergence towards $\mathcal{S}(P)$ follow by only assuming the dual constraint qualification (CQ_D) , or equivalently, the boundedness of the level sets $\mathcal{F}_\alpha(P)$ (or the condition $\emptyset \neq \mathcal{S}(P)$ compact), see Theorem 3.1.

Theorem 4.1 *Let (P) be feasible and let (CQ_D) be satisfied. Then, for any mesh-size d small enough, the sets $\mathcal{S}(P_d)$ of optimal solutions of (P_d) are nonempty and compact. Moreover, for any sequence of solutions $\bar{x}_d \in \mathcal{S}(P_d)$ we have $\delta(\bar{x}_d, \mathcal{S}(P)) \rightarrow 0$ for $d \rightarrow 0$.*

Proof See [10, Theorem 9] for a proof. See also [5, Theorem 4.2(b), (c)] for a proof under slightly stronger assumptions. \square

Since the feasible set $\mathcal{F}(P)$ may consist of a single point, it is clear that, in order to ensure the existence of a feasible point of the inner approximation (\tilde{P}_d) , we have to assume that $\mathcal{F}(P)$ has interior points (see also [5, Theorem 4.2(a), (c)]).

Theorem 4.2 *Let (CQ_P) and (CQ_D) hold. Then for any mesh-size d small enough the sets $\mathcal{S}(\tilde{P}_d)$ of optimal solutions of (\tilde{P}_d) are nonempty and compact. Moreover, for any sequence of solutions $\tilde{x}_d \in \mathcal{S}(\tilde{P}_d)$ we have $\delta(\tilde{x}_d, \mathcal{S}(P)) \rightarrow 0$ for $d \rightarrow 0$.*

Proof If (CQ_P) holds for x_0 , then we find from (10) that for all $(u, v) \in E_d$ and d small enough,

$$\begin{aligned} u^T F(x_0)v &= \frac{1}{2} [u^T F(x_0)u + v^T F(x_0)v - (u - v)^T F(x_0)(u - v)] \\ &\geq \sigma_0 - \frac{1}{2} \|F(x_0)\| \cdot d^2 \geq 0. \end{aligned}$$

Hence, $x_0 \in \mathcal{F}(\tilde{P}_d)$ if d is small. By (CQ_D) the level sets $\mathcal{F}_\alpha(P)$ are bounded (compact) (see Theorem 3.1). Since $\mathcal{F}_\alpha(\tilde{P}_d) \subset \mathcal{F}_\alpha(P)$ (see (12)), also the level sets $\mathcal{F}_\alpha(\tilde{P}_d)$ are bounded. Therefore, solutions \tilde{x}_d of the linear programs (\tilde{P}_d) exist and the sets $\mathcal{S}(\tilde{P}_d)$ of maximizers are nonempty and compact.

Suppose now that a sequence \tilde{x}_{d_k} of such solutions does not satisfy

$$\delta(\tilde{x}_{d_k}, \mathcal{S}(P)) \rightarrow 0 \quad \text{for } k \rightarrow \infty.$$

Then there exists $\varepsilon > 0$ and a subsequence $\tilde{x}_{d_{k_v}}$ such that

$$\delta(\tilde{x}_{d_{k_v}}, \mathcal{S}(P)) \geq \varepsilon \quad \forall v. \tag{13}$$

Since the minimizers $\tilde{x}_{d_{k_v}}$ are elements of a compact set $\mathcal{F}_\alpha(P)$, we can select a convergent subsequence and w.l.o.g. we can assume $\tilde{x}_{d_{k_v}} \rightarrow \hat{x} \in \mathcal{F}_\alpha(P)$ for $v \rightarrow \infty$. In view of Lemma 4.2(c) we have $v(P_d) - v(\tilde{P}_d) \rightarrow 0$ and thus, by (12), $v(\tilde{P}_d) \rightarrow v(P)$, $d \rightarrow 0$. This yields $c^T \tilde{x}_{d_{k_v}} = v(P_{d_{k_v}}) \rightarrow c^T \hat{x} = v(P)$, $v \rightarrow \infty$ and since $\hat{x} \in \mathcal{F}_\alpha(P)$ is feasible for (P) , we obtain $\hat{x} \in \mathcal{S}(P)$ contradicting (13). \square

The next example shows that it may happen that every program (P_d) and (\tilde{P}_d) has a solution while no solution of the original program (P) exists.

Example 4.1 Consider the copositive program (based on [13, Theorem 3.1]) with $c = (1, 1, 0)^T$ and

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then (P) becomes:

$$\max x_1 + x_2 \quad \text{s.t.} \quad F(x_1, x_2, x_3) := \begin{pmatrix} 1 - x_1 & 0 & 0 \\ 0 & -x_2 & -1 \\ 0 & -1 & -x_3 \end{pmatrix} \in \mathcal{C}_3.$$

The feasibility conditions for this program read:

$$x_1 \leq 1, \quad x_2 \leq 0, \quad x_3 \leq 0, \quad x_2 x_3 \geq 1.$$

Obviously, $x_1 + x_2 \leq 1$ holds for any feasible x and for any $\epsilon > 0$ the point $x = (1, -\epsilon, -1/\epsilon)^T$ is feasible with objective value $x_1 + x_2 = 1 - \epsilon$. On the other hand, no feasible \bar{x} exists with objective $\bar{x}_1 + \bar{x}_2 = 1$ ($\bar{x}_2 = 0$ is excluded). So, the sup value of (P) is $v(P) = 1$ but a maximizer does not exist. Now, consider the program (P_d) :

$$(P_d) \quad \max x_1 + x_2 \quad \text{s.t.} \quad z^T F(x_1, x_2, x_3) z \geq 0 \quad \forall z \in Z_d,$$

where Z_d is any (finite) discretization of Δ_3 containing the basis vectors $z = e_i \in \mathbb{R}^3$, $i = 1, 2, 3$. Then (P_d) in particular contains the constraints

$$e_i^T F(x) e_i \geq 0, \quad i = 1, 2, 3 \quad \text{or} \quad 1 - x_1 \geq 0, \quad x_2 \leq 0, \quad x_3 \leq 0.$$

This implies $x_1 + x_2 \leq 1$. So, the linear program (P_d) is bounded and a solution exists. In fact, any program (P_d) has a solution $\bar{x}_d = (1, 0, \bar{x}_3(d))^T$ with objective value $v(P_d) = 1$ (and $\bar{x}_3(d) \rightarrow -\infty$ for $d \rightarrow 0$).

Note that also the inner LP-approximations (\tilde{P}_d) have solutions. Indeed, since the feasible sets $\mathcal{F}(\tilde{P}_d)$ are contained in $\mathcal{F}(P)$, the values $v(\tilde{P}_d)$ are bounded by 1. Moreover, the feasible sets are nonempty. To see this, take e.g. the (CQ_P) -point $x_0 = (0, -2, -2)^T$ in the interior of $\mathcal{F}(P)$. Then, as in the proof of Theorem 4.2, it follows that $x_0 \in \mathcal{F}(\tilde{P}_d)$, provided d is small enough.

We finish this section with some remarks. Note that for any solution \bar{x}_d of the standard linear program (P_d) the KKT condition holds:

$$c = \sum_{j=1}^k y_j \cdot (z_j^T A_1 z_j, \dots, z_j^T A_n z_j), \quad \text{for some } y_j \geq 0, \quad z_j \in Z_d(\bar{x}_d), \quad (14)$$

where $Z_d(\bar{x}_d) := \{z \in Z_d : z^T F(\bar{x}_d) z = 0\}$. Again, any such solution \bar{x}_d generates a dual feasible matrix

$$\bar{Y}_d := \sum_{j=1}^k y_j z_j z_j^T \in \mathcal{F}(D) \quad \text{such that} \quad (\bar{Y}_d, B) = v(P_d) \geq v(D) \geq v(P).$$

Remark 4.2 Any solution \tilde{x}_d of (\tilde{P}_d) also satisfies the KKT condition

$$c = \sum_{j=1}^s \tilde{y}_j \cdot (u_j^T A_1 v_j, \dots, u_j^T A_n v_j), \quad \tilde{y}_j \geq 0, \quad (u_j, v_j) \in E_d(\tilde{x}_d), \quad s \in \mathbb{N},$$

where $E_d(\tilde{x}_d) := \{(u, v) \in E_d : u^T F(\tilde{x}_d)v = 0\}$. Such a solution \tilde{x}_d generates the matrix $\tilde{Y}_d := \sum_{j=1}^s \tilde{y}_j \cdot \frac{1}{2}(u_j v_j^T + v_j u_j^T)$ which satisfies the constraints $\langle \tilde{Y}_d, A_i \rangle = c_i$ for all i . However, in general, $\tilde{Y} \notin \mathcal{C}_m^*$, so \tilde{Y} is not necessarily feasible for (D) . Using (10), we see (under the assumption of Theorem 4.1) that \tilde{Y}_d is in \mathcal{C}_m^* up to an error of order $\mathcal{O}(d^2)$.

5 Comparison with an Inner Approximation via Sets \mathcal{C}_m^r

In this section, we consider a special discretization scheme first considered in [2] which is connected to an inner approximation of \mathcal{C}_m by subsets $\mathcal{C}_m^r \subset \mathcal{C}_m$. For $r \in \mathbb{N}$, let us define

$$\mathcal{C}_m^r := \left\{ A = (a_{ij}) \in \mathcal{S}_m : \sum_{i,j=1}^m a_{ij} x_i^2 x_j^2 \left(\sum_{k=1}^m x_k^2 \right)^r \text{ has non-negative coefficients} \right\}.$$

The following is shown in [2] :

$$\mathcal{C}_m^r \subset \mathcal{C}_m^{r+1} \subset \dots \subset \mathcal{C}_m \quad \text{and} \quad \text{cl}\left(\lim_{r \rightarrow \infty} \mathcal{C}_m^r\right) = \mathcal{C}_m.$$

The interesting connection with the discretization approach above is based on the following description of the sets \mathcal{C}_m^r (see[1]) (with $\text{diag}(A) := (a_{11}, \dots, a_{mm}) \in \mathbb{R}^m$):

$$\mathcal{C}_m^{r-2} = \left\{ A \in \mathcal{S}_m : v^T A v - v^T \text{diag}(A) \geq 0 \text{ for all } v \in I_m^r \right\}, \tag{15}$$

where I_m^r is the grid $I_m^r = \{v \in \mathbb{N}_0^m : \sum_{j=1}^m v_j = r\}$. By (15), we can write

$$\mathcal{C}_m^{r-2} = \left\{ A \in \mathcal{S}_m : z^T A z - \frac{1}{r} z^T \text{diag}(A) \geq 0 \text{ for all } z \in Z_d^0 := \frac{1}{r} I_m^r \right\}. \tag{16}$$

It is not difficult to see that the set $Z_d^0 := \frac{1}{r} I_m^r$ defines a uniform discretization of the simplex $Z = \Delta_m$ with mesh-size of Z_d^0 given by

$$d = \max_{z_j \in \frac{1}{r} I_m^r} \min_{z_i \in \frac{1}{r} I_m^r, z_i \neq z_j} \|z_j - z_i\| = \frac{\sqrt{2}}{r}.$$

So it is natural to compare the outer and inner approximations (P_d) , (\tilde{P}_d) of (P) in Sect. 4 with the following approximations, where $d = \sqrt{2}/r$, $r \in \mathbb{N}$:

$$(\hat{P}_d) \quad \max_{x \in \mathbb{R}^n} c^T x \quad \text{s.t.} \quad z^T F(x) z - \frac{d}{\sqrt{2}} z^T \text{diag}(F(x)) \geq 0 \quad \forall z \in Z_d^0. \tag{17}$$

Denote the solution, the optimal value and the feasible set of (\widehat{P}_d) by \widehat{x}_d , $v(\widehat{P}_d)$ and $\mathcal{F}(\widehat{P}_d)$, respectively. Note that, by (16), a point x is feasible for (\widehat{P}_d) if and only if $F(x) \in \mathcal{C}_m^{r-2}$. So (\widehat{P}_d) provides an inner approximation, i.e., $\mathcal{F}(\widehat{P}_d) \subset \mathcal{F}(P)$ and $v(\widehat{P}_d) \leq v(P)$. Similarly to Lemma 4.2 we obtain.

Lemma 5.1 *Let (CQP) be satisfied for $x_0 \in \mathcal{F}(P)$. Then with the solutions \bar{x}_d of (P_d) (with discretization $Z_d = Z_d^0$) the following holds for all $d = \frac{\sqrt{2}}{r}$, $r \in \mathbb{N}$, d small enough:*

$$\widehat{x}_d^* = \bar{x}_d + \tau d(x_0 - \bar{x}_d) \in \mathcal{F}(\widehat{P}_d) \subset \mathcal{F}(P)$$

and

$$0 \leq v(P_d) - v(\widehat{P}_d) \leq \tau [c^T (\bar{x}_d - x_0)] \cdot d$$

if $\tau \geq \frac{\|\text{diag}(F(\bar{x}_d))\|}{\sqrt{2}\sigma_0 - d\|\text{diag}(F(x_0))\|}$ and $0 < \tau d < 1$.

Proof We use the relation $F(\widehat{x}_d^*) = (1 - \tau d)F(\bar{x}_d) + \tau dF(x_0)$ and proceed as in the proof of Lemma 4.2. By Lemma 4.1, using the relation $\|z\| \leq 1$ for $z \in Z_d^0$ and $z^T F(\bar{x}_d)z \geq 0$ for $z \in Z_d = Z_d^0$, we obtain for any $z \in Z_d^0$:

$$\begin{aligned} & z^T F(\widehat{x}_d^*)z - \frac{d}{\sqrt{2}} z^T \text{diag}(F(\widehat{x}_d^*)) \\ &= \left[(1 - \tau d)z^T F(\bar{x}_d)z + \tau dz^T F(x_0)z \right. \\ &\quad \left. - \frac{d}{\sqrt{2}}(1 - \tau d)z^T \text{diag}(F(\bar{x}_d)) - \tau \frac{d^2}{\sqrt{2}} z^T \text{diag}(F(x_0)) \right] \\ &\geq \tau d\sigma_0 - \frac{d}{\sqrt{2}}(1 - \tau d)\|\text{diag}(F(\bar{x}_d))\| - \tau \frac{d^2}{\sqrt{2}}\|\text{diag}(F(x_0))\| \\ &\geq d \left[\tau \left(\sigma_0 - \frac{d}{\sqrt{2}}\|\text{diag}(F(x_0))\| \right) - \frac{\|\text{diag}(F(\bar{x}_d))\|}{\sqrt{2}} \right] \geq 0 \end{aligned}$$

for any $d > 0$ (small enough) if τ is as given above. This shows the first relation. The inequality for the maximum values follows again easily using that \widehat{x}_d^* is feasible for (\widehat{P}_d) :

$$0 \leq v(P_d) - v(\widehat{P}_d) \leq c^T (\bar{x}_d - \widehat{x}_d^*) = [c^T (\bar{x}_d - x_0)\tau] \cdot d. \quad \square$$

According to the analysis above, under the assumption that the sequence \bar{x}_d , $d \rightarrow 0$, is bounded (cf. Theorem 4.1), we have established the following error bounds (the last bound holds for $Z_d^0 = \frac{1}{r}I_m^r$ with $d = \sqrt{2}/r$, $r \in \mathbb{N}$):

$$\begin{aligned} 0 \leq v(P_d) - v(P) &\leq \mathcal{O}(d^2), & 0 \leq v(P) - v(\widetilde{P}_d) &\leq \mathcal{O}(d^2), \\ 0 \leq v(P) - v(\widehat{P}_d) &\leq \mathcal{O}(d). \end{aligned}$$

The next example shows that the bound $\mathcal{O}(d)$ for (\widehat{P}_d) is sharp.

Example 5.1 We consider the program

$$(P) \quad \max_{x \in \mathbb{R}} x \quad \text{s.t.} \quad F(x) := \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + x \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \in \mathcal{C}_2.$$

The maximizer of (P) is $\bar{x} = 0$ with $v(P) = \bar{x} = 0$. The corresponding unique active index is $\bar{z} = (\frac{1}{2}, \frac{1}{2})^T$. For odd $r = 2l + 1$ and $d = \sqrt{2}/r$, the discretization Z_d^0 of $Z = \Delta_2 = \{z \in \mathbb{R}_+^2 : z_1 + z_2 = 1\}$ is given by

$$Z_d^0 = \left\{ z(\lambda) := \lambda \begin{pmatrix} 1 \\ 0 \end{pmatrix} + (1 - \lambda) \begin{pmatrix} 0 \\ 1 \end{pmatrix} : \lambda = \frac{i}{r}, i = 0, \dots, r \right\}.$$

It is not difficult to see that the optimal solutions of $(P_d), (\widehat{P}_d)$ are given by the solutions \bar{x}_d, \widehat{x}_d of the equations

$$\begin{aligned} z \begin{pmatrix} l \\ 2l+1 \end{pmatrix}^T F(x) z \begin{pmatrix} l \\ 2l+1 \end{pmatrix} &= 0, \\ z \begin{pmatrix} l \\ 2l+1 \end{pmatrix}^T F(x) z \begin{pmatrix} l \\ 2l+1 \end{pmatrix} - \frac{d}{\sqrt{2}} z \begin{pmatrix} l \\ 2l+1 \end{pmatrix}^T \text{diag } F(x) &= 0, \end{aligned}$$

respectively. After some calculations we obtain $v(P_d) = \bar{x}_d = \frac{1}{2} \frac{1}{l(l+1)} = \mathcal{O}(d^2)$ and

$$v(\widehat{P}_d) = \widehat{x}_d = -\frac{\sqrt{2}d}{2} \left[1 + \frac{2l^2 + 1}{2l(l+1)} \right] + \frac{1}{2l(l+1)} = -\sqrt{2}d + \mathcal{O}(d^2) = \mathcal{O}(d).$$

Let us further compare the inner approximations (\widetilde{P}_d) and (\widehat{P}_d) . It is not difficult to show that the number of points in the discretization Z_d^0 for $d = \frac{\sqrt{2}}{(r+2)}$ (approximation by \mathcal{C}_m^r see (16)) are given by $N := \binom{m+r-1}{r}$ (cf. [4]). To obtain a corresponding inner approximation (\widetilde{P}_d) one could think of the so-called Delaunay triangulation (by simplices) of the point set Z_d^0 . The number of edges in such a triangulation is “much smaller” than N^2 (edge from each point to each other, instead of edges only to “neighboring points”). So (for fixed m) the same order of approximation $\mathcal{O}(\frac{1}{r^2})$ (w.r.t. r) would require “much less” than $N^2 = \binom{m+r-1}{r}^2$ constraints in (\widetilde{P}_d) and $\binom{m+r^2-1}{r^2}$ constraints in (\widehat{P}_d) . This can be seen to be in favor of the scheme (\widetilde{P}_d) .

Interested in an inner approximation, one could also avoid both inner approximations $(\widetilde{P}_d), (\widehat{P}_d)$ and only make use of (P_d) . Indeed, the a posteriori error bound of Lemma 4.2 allows us to construct a feasible point $\bar{x}_d^* = \bar{x}_d + \mathcal{O}(d^2)$ from the “outer approximation” \bar{x}_d , if a strictly feasible point x_0 is available.

Note however that with regard to the “exploding number” of constraints in the above methods, a practical algorithm should avoid most of the constraints. This means that in practice, the pure discretization methods have to be modified to a so-called exchange method where (as in [5]) during the computation only those grid

points in Z_d are kept in the discretization which still play a role as candidates for the active points $z_j \in Z(\bar{x})$ of a solution \bar{x} of (P) (see also [10]). We emphasize that for such exchange methods the order bounds obtained above remain valid.

6 Order of Convergence for the Maximizers

In this section, we shortly discuss error bounds for $\|\bar{x} - \bar{x}_d\|, \|\bar{x} - \tilde{x}_d\|, \|\bar{x} - \hat{x}_d\|$ for the maximizers of $(P_d), (\tilde{P}_d), (\hat{P}_d)$, respectively. These bounds are based on the concept of the order of a maximizer. A feasible point $\bar{x} \in \mathcal{F}(P)$ is a maximizer of (P) of order $p > 0$, iff with some $\gamma > 0, \varepsilon > 0$,

$$c^T \bar{x} \geq c^T x + \gamma \|x - \bar{x}\|^p \quad \text{for all } x \in \mathcal{F}(P), \|x - \bar{x}\| < \varepsilon \tag{18}$$

holds (and such an inequality does not hold for a smaller $\bar{p} < p$). Note that, if \bar{x} is a maximizer of order $0 < p$, in particular, $\mathcal{S}(P) = \{\bar{x}\}$ is nonempty and compact. So, by Theorem 3.1 the condition (CQ_D) is satisfied and we can apply Theorem 4.1.

Corollary 6.1 *Let (CQ_P) be satisfied and let \bar{x} be a maximizer of (P) of order $p \geq 1$. Then for the maximizers $\bar{x}_d, \tilde{x}_d, \hat{x}_d$ of $(P_d), (\tilde{P}_d), (\hat{P}_d)$, respectively, we have:*

$$\|\bar{x} - \bar{x}_d\| = \mathcal{O}(d^{2/p}), \quad \|\bar{x} - \tilde{x}_d\| = \mathcal{O}(d^{2/p}), \quad \|\bar{x} - \hat{x}_d\| = \mathcal{O}(d^{1/p}).$$

Proof Recall from Lemma 4.2(a) that $\bar{x}_d^* := \bar{x}_d + \rho d^2(x_0 - \bar{x}_d) \in \mathcal{F}(P)$ for ρ large enough. Using (18) and $c^T(\bar{x} - \bar{x}_d) \leq 0$ we get (\bar{x}_d^* is feasible for (P))

$$\begin{aligned} \|\bar{x} - \bar{x}_d^*\|^p &\leq \frac{1}{\gamma} c^T(\bar{x} - \bar{x}_d^*) = \frac{1}{\gamma} c^T(\bar{x} - \bar{x}_d) - \frac{\rho}{\gamma} d^2 c^T(x_0 - \bar{x}_d) \\ &\leq \frac{\rho}{\gamma} d^2 c^T(\bar{x}_d - x_0) \leq \mathcal{O}(d^2) \end{aligned}$$

or $\|\bar{x} - \bar{x}_d^*\| \leq \mathcal{O}(d^{2/p})$. We thus find using $1 \leq p$,

$$\|\bar{x} - \bar{x}_d\| \leq \|\bar{x} - \bar{x}_d^*\| + \|\bar{x}_d^* - \bar{x}_d\| \leq \mathcal{O}(d^{2/p}) + \mathcal{O}(d^2) = \mathcal{O}(d^{2/p}).$$

The other bounds are proven in the same way. For \hat{x}_d , e.g., we obtain using Lemma 5.1,

$$\|\bar{x} - \hat{x}_d\| \leq \|\bar{x} - \hat{x}_d^*\| + \|\hat{x}_d^* - \hat{x}_d\| = \mathcal{O}(d^{1/p}) + \mathcal{O}(d) = \mathcal{O}(d^{1/p}). \quad \square$$

According to this corollary, the smaller the order p of the maximizer \bar{x} , the faster is the convergence. The following examples show that for copositive programs (P) (unique) maximizer of orders 1, 2 and of arbitrarily large order can occur.

Example 6.1 Obviously, in Example 4.1 the maximizer $\bar{x} = 0$ is of order $p = 1$. Considering the copositive program

$$(P) \quad \max x_1 \quad \text{s.t.} \quad F(x) := \begin{pmatrix} -x_1 & x_2 & 0 \\ x_2 & 1 & 0 \\ 0 & 0 & -x_2 \end{pmatrix} \in \mathcal{C}_3$$

we see that x is feasible if and only if $-x_1 \geq 0, -x_2 \geq 0$ and $-x_1 - x_2^2 \geq 0$ hold, or

$$x_1 \leq 0, \quad x_2 \leq 0, \quad x_1 \leq -x_2^2.$$

The maximum value is $x_1 = 0$ implying $x_2 = 0$. So $\bar{x} = (0, 0)^T$ is the unique maximizer. For the feasible points $x = (-x_2^2, x_2)^T, x_2 < 0$ ($|x_2|$ small) we find with $\|x\|_\infty := \max\{|x_1|, |x_2|\}$:

$$c^T \bar{x} - c^T x = x_2^2 = \|x\|_\infty^2,$$

and \bar{x} is a maximizer of order 2. Now, we take the program

$$(P) \quad \max x_1 \quad \text{s.t.} \quad F(x) := \begin{pmatrix} -x_1 & x_2 & 0 & 0 & 0 \\ x_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & -x_2 & x_3 & 0 \\ 0 & 0 & x_3 & 1 & 0 \\ 0 & 0 & 0 & 0 & -x_3 \end{pmatrix} \in \mathcal{C}_5.$$

In view of the block structure of $F(x)$, a vector $x \in \mathbb{R}^3$ is feasible iff:

$$x_1 \leq 0, \quad x_2 \leq 0, \quad x_3 \leq 0, \quad x_1 \leq -x_2^2, \quad x_2 \leq -x_3^2.$$

Thus $\bar{x} = (0, 0, 0)^T$ is the (unique) maximizer and with feasible vectors $x = (-x_3^4, -x_3^2, x_3)^T, x_3 < 0$ ($|x_3|$ small) we find

$$c^T \bar{x} - c^T x = x_3^4 = \|x\|_\infty^4,$$

showing that \bar{x} is maximizer of order 4. Similarly, we can construct copositive programs with maximizer of arbitrarily large order.

Remark 6.1 In [1, Sect. 3], approximation results have been established for the values $v^* = \min_{z \in \Delta_m} z^T A z$ with $A \in \mathcal{S}_m$. We briefly show that these bounds appear in our result above as special instances. Obviously v^* is the value of

$$(P) \quad \max_{x \in \mathbb{R}} x \quad \text{s.t.} \quad z^T (A - xI) z \geq 0 \quad \forall z \in Z := \Delta_m,$$

with dual

$$(D) \quad \min_{Y \in \mathcal{S}_m} \langle Y, A \rangle \quad \text{s.t.} \quad \langle Y, I \rangle = 1, \quad Y \in \mathcal{C}_m^*.$$

Obviously, (P) satisfies (CQ_P) with some x_0 (small enough) and also (D) has strictly feasible matrices Y_0 (with any $Y \in \text{int } \mathcal{C}_m^*$ take $Y_0 = Y / \langle Y, I \rangle$). Let \bar{x} be the solution of (P) and consider the approximations $(P_d), (\tilde{P}_d)$ defined by the grids Z_d^0 ($d = \sqrt{2}/r$) with corresponding values v_d, \tilde{v}_d and solutions \bar{x}_d, \tilde{x}_d . It is easy to see that these solutions must be unique, satisfy $x_0 \leq \tilde{x}_d \leq \bar{x} \leq \bar{x}_d$ and are monotonic, i.e., $\tilde{x}_d \uparrow \bar{x}, \bar{x}_d \downarrow \bar{x}$ for $d \rightarrow 0$. Then by Lemma 4.2(a) we obtain the bound

$$0 \leq v_d - v^* = \bar{x}_d - \bar{x} \leq \bar{x}_d - \bar{x}_d^* = \frac{\|F(\bar{x}_d)\|}{2\sigma_0} (\bar{x}_d - x_0) d^2,$$

and Lemma 5.1 yields

$$0 \leq v^* - \tilde{v}_d \leq v_d - \tilde{v}_d \leq \bar{x}_d - \tilde{x}_d^* \leq \tau(\bar{x}_d - x_0)d.$$

The latter gives (up to a constant factor) the bound in [1] and the first bound yields a $\mathcal{O}(d^2)$ error instead of a rate $\mathcal{O}(d)$ in [1].

7 Set-Copositive Programming

Instead of studying copositivity with respect to \mathbb{R}_+^m , it is also possible to consider copositivity with respect to an arbitrary set $Z \subset \mathbb{R}^m$, meaning that one requires $z^T A z \geq 0$ for all $z \in Z$. This concept was studied in detail in [15]. It plays an important role in modeling quadratically constrained quadratic problems, cf. [16].

We can easily extend the analysis above, and consider programs with a feasibility condition

$$F(x) \in \mathcal{C}_m(Z, \alpha) := \{A \in \mathcal{S}_m : z^T A z - \alpha z^T \text{diag}(A) \geq 0 \forall z \in Z\}, \tag{19}$$

where $\alpha \in \mathbb{R}$ and a compact set $Z \subset \mathbb{R}^m$ is given. Here again, $\text{diag}(A)$ denotes the vector with components a_{ii} , and later for $z \in \mathbb{R}^m$, $\text{Diag}(z)$ will denote the diagonal matrix with diagonal elements z_i . So we consider

$$(P) \quad \max_{x \in \mathbb{R}^n} c^T x \quad \text{s.t.} \quad F(x) := B - \sum_{i=1}^n x_i A_i \in \mathcal{C}_m(Z, \alpha).$$

The cases $Z = \Delta_m, \alpha = 0$ and $Z = \Delta_m, \alpha = \frac{1}{r}$ have been discussed in Sect. 5. Here again, the SIP-form of (P) will lead us to the dual program (D) and to the dual cone $\mathcal{C}_m^*(Z, \alpha)$ as follows. In SIP-form (P) reads:

$$(P_{\text{SIP}}) \quad \max_{x \in \mathbb{R}^n} c^T x \quad \text{s.t.} \quad b(z) - a(z)^T x \geq 0 \quad \forall z \in Z,$$

with vector $a(z)$ having components $a_i(z) = z^T A_i z - \alpha z^T \text{diag}(A_i)$ and $b(z) = z^T B z - \alpha z^T \text{diag}(B)$. The dual of (P_{SIP}) is (with finite sums \sum_j , see Sect. 2):

$$(D_{\text{SIP}}) \quad \min \sum_j y_j b(z_j) \quad \text{s.t.} \quad \sum_j y_j a(z_j) = c, \quad y_j \geq 0, z_j \in Z.$$

To convert this SIP-dual to the conic dual of (P) we make use of the relation $z^T \text{diag}(B) = \langle B, \text{Diag}(z) \rangle$ and find

$$\begin{aligned} \sum_j y_j b(z_j) &= \left\langle B, \sum_j y_j (z_j z_j^T - \alpha \text{Diag}(z_j)) \right\rangle \\ c_i &= \left\langle A_i, \sum_j y_j (z_j z_j^T - \alpha \text{Diag}(z_j)) \right\rangle \quad (i = 1, \dots, n). \end{aligned}$$

So (D_{SIP}) is equivalent to the program

$$(D) \quad \min_Y \langle Y, B \rangle \quad \text{s.t.} \quad \langle Y, A_i \rangle = c_i, \quad (i = 1, \dots, n), \quad Y \in C_m^*(Z, \alpha),$$

where

$$C_m^*(Z, \alpha) := \left\{ A = \sum_{j=1}^k y_j (z_j z_j^T - \alpha \text{Diag}(z_j)) : y_j \geq 0, z_j \in Z, k \in \mathbb{N} \right\}.$$

It is easily shown that this set is the dual of $C_m(Z, \alpha)$. So, under the assumption that Z is compact, all results in the preceding sections can be extended to this more general problem (P) .

8 Conclusions

Copositive programs have been treated as special cases of linear semi-infinite problems (LSIP). We have applied known results from LSIP to cone programming and re-obtained known results but also gained new insight.

We have interpreted different approximation schemes for solving copositive programs as discretization schemes. The behavior of the approximation error for the optimal values in dependence of the discretization mesh-size d has been analyzed. The concept of order of maximizers allows to analyse the behavior of the error for the maximizers in the approximation schemes. It also has been shown that maximizer of arbitrary large order may appear in copositive programming.

Acknowledgements The authors would like to thank the anonymous referees for many valuable comments and suggestions.

This work was done while Mirjam Dür was working at Johann Bernoulli Institute of Mathematics and Computer Science, University of Groningen, The Netherlands. She was supported by the Netherlands Organization for Scientific Research (NWO) through Vici Grant No. 639.033.907.

References

1. Bomze, I.M., De Klerk, E.: Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. *J. Glob. Optim.* **24**(2), 163–185 (2002)
2. de Klerk, E., Pasechnik, D.V.: Approximation of the stability number of a graph via copositive programming. *SIAM J. Optim.* **12**(4), 875–892 (2002)
3. Burer, S.: On the copositive representation of binary and continuous nonconvex quadratic programs. *Math. Program.* **120**(2), 479–495 (2009)
4. Alper Yıldırım, E.: On the accuracy of uniform polyhedral approximations of the copositive cone. *Optim. Methods Softw.* **27**(1), 155–173 (2012)
5. Bundfuss, S., Dür, M.: An adaptive linear approximation algorithm for copositive programs. *SIAM J. Optim.* **20**(1), 30–53 (2009)
6. Berman, A., Shaked-Monderer, N.: *Completely Positive Matrices*. World Scientific, Singapore (2003)
7. Dickinson, P.J.: An improved characterisation of the interior of the completely positive cone. *Electron. J. Linear Algebra* **20**, 723–729 (2010)
8. Dür, M., Still, G.: Interior points of the completely positive cone. *Electron. J. Linear Algebra* **17**, 48–53 (2008)

9. Goberna, M.A., López, M.A.: *Linear Semi-Infinite Optimization*. Wiley, Chichester (1998)
10. López, M., Still, G.: Semi-infinite programming (invited review). *Eur. J. Oper. Res.* **180**(2), 491–518 (2007)
11. Faigle, U., Kern, W., Still, G.: *Algorithmic Principles of Mathematical Programming*. Springer, Berlin (2002)
12. Hettich, R., Kortanek, K.O.: Semi-infinite programming: theory, methods, and applications. *SIAM Rev.* **35**(3), 380–429 (1993)
13. Bomze, I.M., Schachinger, W., Uchida, G.: Think co (mpletely) positive! Matrix properties, examples and a clustered bibliography on copositive optimization. *J. Glob. Optim.* **52**(3), 423–445 (2012)
14. Still, G.: Discretization in semi-infinite programming: the rate of convergence. *Math. Program.* **91**(1), 53–69 (2001)
15. Eichfelder, G., Jahn, J.: Set-semidefinite optimization. *J. Convex Anal.* **15**(4), 767–801 (2008)
16. Burer, S., Dong, H.: Representing quadratically constrained quadratic programs as generalized copositive programs. *Oper. Res. Lett.* **40**(3), 203–206 (2012)