Copositive relaxation beats Lagrangian dual bounds in quadratically and linearly constrained QPs

Immanuel M. Bomze

ISOR, University of Vienna, Austria

Abstract

We study non-convex quadratic minimization problems under (possibly non-convex) quadratic and linear constraints, and characterize both Lagrangian and Semi-Lagrangian dual bounds in terms of conic optimization. While the Lagrangian dual is equivalent to the SDP relaxation, the Semi-Lagrangian dual we study is equivalent to a natural copositive relaxation. This way, we arrive at a full hierarchy of tractable conic bounds tighter than the usual Lagrangian dual (and thus than the SDP) bounds. In particular, the usual zero-order approximation by doubly nonnegative matrices improves upon the Lagrangian dual bounds. We also relate the new relaxation with an alternative, still tighter one which was earlier introduced by Burer who showed that his formulation is indeed tight in an important subclass of the problem type studied here. Further we specify sufficient conditions for tightness of the Semi-Lagrangian relaxation and show that copositivity of the slack matrix guarantees global optimality for KKT points of this problem. Motivated by these observations, we propose a seemingly new approximation hierarchy based on LMIs on matrices of basically the original order plus additional linear constraints, in contrast to more familiar sum-of-squares or moment approximation hierarchies. This approach may have merits in particular for large instances where it is important to employ only a few inequality constraints for the conic problems.

Key words: Copositive matrices, non-convex optimization, polynomial optimization, quadratically constrained problem, approximation hierarchies, global optimality condition

November 20, 2013

1 Introduction and basic concepts

As is well known, the effectiveness of Lagrangian relaxation – and optimization methods in general – heavily depends on the formulation of the problem, and of the treatment of constraints. For instance, if the ground set is not the full space but rather incorporates some (simpler) constraints, we arrive at Semi-Lagrangian relaxation yielding tighter bounds than the classic Lagrangian relaxation which uses the full Euclidean space $\mathbb{R}^n$ as the ground set. However, Semi-Lagrangian dual bounds cannot always be calculated efficiently.

Here we study non-convex quadratic minimization problems under (possibly non-convex) quadratic and linear constraints, and characterize both duals in terms of conic optimization. Due to their pivotal role for applications, bounds for such type of problems receive currently much interest in the optimization community, a for sure non-exhaustive list is [2, 12, 13, 19, 20, 21, 22].

In absence of linear constraints, the full Lagrangian dual problem is equivalent to the direct semidefinite relaxation. Under additional linear constraints, we arrive at an LMI description of the Lagrangian dual which is an extension thereof, while the Semi-Lagrangian dual can be shown to result from a natural copositive relaxation. This way, we arrive at a full hierarchy of tractable conic bounds tighter than the usual Lagrangian dual (and thus than the SDP) bounds. In particular, the usual zero-order approximation by doubly nonnegative matrices improves upon the Lagrangian dual bounds. Therefore we manage a tractable approximation tightening towards Semi-Lagrangian dual bounds.

The resulting approximation hierarchy is apparently new, and based on LMIs on matrices of basically the original order plus relatively few additional linear constraints, in contrast to more familiar sum-of-squares hierarchies or moment approximation hierarchies. We also relate the new relaxation with an alternative, still tighter, relaxation earlier introduced by Burer who showed that his formulation is indeed tight in an important subclass of the problem type studied here, including all mixed-binary QPs satisfying the so-called key condition. Further we study strong duality of the resulting conic problems, and also specify sufficient conditions for tightness of the Semi-Lagrangian (i.e. copositive) relaxation. We also show that copositivity of the slack matrix guarantees global optimality for KKT points of this problem. Finally, we address an alternative to replace all linear constraints by one convex quadratic. Similar approaches have been tried recently along different roads [2, 16, 20].

The paper is organized as follows: first, after recapitulating shortly the principles of copositive optimization, we introduce a new approximation hierarchy in Section 2, which may be of particular interest in large instances, i.e., in regimes where every additional linear inequality constraint "hurts" in the conic problem, so the focus is to employ as few of them as possible. Then we discuss several variants of (Semi-)Lagrangian relaxations in Section 3. Section 4 presents a new perspective on the full Lagrangian duals as SDPs, first for all-quadratic problems without any linear constraints, and then with explicit treatment of linear constraints. Thereafter we incorporate, in the central Section 5, the sign constraints into the ground set, and show that the resulting Semi-Lagrangian bounds exactly lead to the natural copositive relaxation of the all-quadratic problem with linear constraints. In this section, we also briefly explain how to tighten Lagrangian bounds by the approximation hierarchies previously developed, and thereafter relate our construction to Burer's [10].

Further, under widely used strict feasibility conditions, we establish full strong duality of the primal-dual pair of copositive problems in Section 6. Section 7 contains conditions which guarantee that the Semi-Lagrangian relaxation (and thus the copositive relaxation) is tight, and discusses global optimality conditions for a KKT point of the original problem. Finally we address an alternative formulation which replaces all linear constraints by one convex quadratic, and add some observations on the relation to the previous findings, including Burer's relaxation.

1.1 Notation and terminology

We abbreviate by $[m:n] := \{m, m+1, \dots, n\}$ the integer range between two integers m, n with $m \leq n$. By bold-faced lower-case letters we denote vectors in n-dimensional Euclidean space $\mathbb{R}^n$, by bold-faced upper case letters matrices, and by $^\top$ transposition. The positive orthant is denoted by $\mathbb{R}^n_+ := \{\mathsf{x} \in \mathbb{R}^n : x_i \geq 0 \text{ for all } i \in [1:n]\}$. I_n is the $n \times n$ identity matrix with columns e_i, $i \in [1:n]$, while $\mathsf{e} := \sum_{i=1}^{n} \mathsf{e}_i = [1, \dots, 1] \in \mathbb{R}^n$ and the compact *standard simplex* is

$$\Delta := \left\{ \mathsf{y} \in \mathbb{R}^n_+ : \mathsf{e}^\top \mathsf{x} = 1 \right\} ,$$

which of course satisfies $\mathbb{R}_+ \Delta = \mathbb{R}^n_+$. The letters o and O stand for zero vectors, and zero matrices, respectively, of appropriate orders.

For a given symmetric matrix $\mathsf{H} = \mathsf{H}^\top$, we denote the fact that H is positive-semidefinite by $\mathsf{H} \succeq \mathsf{O}$. Sometimes we write instead "H is psd." Linear forms in symmetric matrices X will play an important role in this paper; they are expressed by Frobenius duality $\langle \mathsf{S}, \mathsf{X}\rangle = \mathrm{trace}(\mathsf{SX})$, where $\mathsf{S} = \mathsf{S}^\top$ is another symmetric matrix of the same order as X.

Given any cone $\mathcal{C}$ of symmetric $n \times n$ matrices,

$$\mathcal{C}^\star := \left\{ \mathsf{S} = \mathsf{S}^\top \; n \times n : \langle \mathsf{S}, \mathsf{X}\rangle \geq 0 \text{ for all } \mathsf{X} \in \mathcal{C} \right\}$$

denotes the dual cone of $\mathcal{C}$. For instance, if $\mathcal{C} = \left\{ \mathsf{X} = \mathsf{X}^\top \; n \times n : \mathsf{X} \succeq \mathsf{O} \right\}$, then $\mathcal{C}^\star = \mathcal{C}$ itself, an example of a *self-dual cone*. Trusting the sharp eyes of my readers, I chose a notation with subtle differences between the five-star denoting a dual cone, e.g., $\mathcal{C}^\star$, and the six-star, e.g. z^*, denoting optimality.

The key notion used below is that of *copositivity*. Given a symmetric $n \times n$ matrix Q, we say that

$$\begin{aligned} \mathsf{Q} \text{ is copositive if} \quad & \mathsf{v}^\top \mathsf{Q} \mathsf{v} \geq 0 \text{ for all } \mathsf{v} \in \mathbb{R}^n_+ , \quad \text{and that} \\ \mathsf{Q} \text{ is strictly copositive if} \quad & \mathsf{v}^\top \mathsf{Q} \mathsf{v} > 0 \text{ for all } \mathsf{v} \in \mathbb{R}^n_+ \setminus \{\mathsf{o}\} . \end{aligned}$$

Strict copositivity generalizes positive-definiteness (all eigenvalues strictly positive) and copositivity generalizes positive-semidefiniteness (no eigenvalue strictly negative) of a symmetric matrix. Contrasting to positive-semidefiniteness, checking copositivity is NP-hard, see [15, 23].

The set of all copositive matrices form a closed, convex cone, the *copositive cone*

$$\mathcal{C}^\star = \left\{ \mathsf{Q} = \mathsf{Q}^\top \; n \times n : \mathsf{Q} \text{ is copositive} \right\}$$

with non-empty interior which exactly consists of all strictly copositive matrices. However, the cone $\mathcal{C}^\star$ is not self-dual. Rather one can show that it is the dual cone of

$$\mathcal{C} = \left\{ \mathsf{X} = \mathsf{F}\mathsf{F}^\top : \mathsf{F} \text{ has } \tfrac{n(n+1)-8}{2} \text{ columns in } \mathbb{R}^n_+ \right\} ,$$

the cone of *completely positive* matrices. Note that the factor matrix F has many more columns than rows. The quadratic upper bound on the necessary number of columns was recently established by [25], reducing an old bound by 3. It is not assumed to be tight, however; a lower bound for the largest minimal number of columns required to describe $\mathcal{C}$, is $\lfloor \frac{n^2}{4} \rfloor$ for all n, which is exact for $n = 5$ [26]. Anyhow, a probably more amenable representation is

$$\mathcal{C} = \mathrm{conv} \left\{ \mathsf{x}\mathsf{x}^\top : \mathsf{x} \in \mathbb{R}^n_+ \right\} ,$$

where conv S stands for the convex hull of a set $S \subset \mathbb{R}^n$. Caratheodory's theorem then elucidates the quadratic character of above discussed bound.

2 Basics of Copositive Optimization

2.1 A key lemma and its consequences

We start with a key observation which involves bordering of $n \times n$ matrices (in which context we always address the first row/column as the zeroth one). To this end, we denote by $\mathsf{e}_0 = [1, 0, \dots, 0]^\top \in \mathbb{R}^{n+1}$, and by

$$\mathsf{J}_0 := \mathsf{e}_0 \mathsf{e}_0^\top = \begin{bmatrix} 1 & \mathsf{o}^\top \\ \mathsf{o} & \mathsf{O} \end{bmatrix} .$$

Lemma 2.1 *Consider a quadratic function* $q(\mathsf{x}) = \mathsf{x}^\top \mathsf{H} \mathsf{x} - 2\mathsf{d}^\top \mathsf{x} + \gamma$ *defined on* $\mathbb{R}^n$*, with* $q(\mathsf{o}) = \gamma$*,* $\nabla q(\mathsf{o}) = -2\mathsf{d}$ *and* $D^2 q(\mathsf{o}) = 2\mathsf{H}$ *(the factors* 2 *being here just for ease of later notation). Define the* Shor relaxation matrix [27]

$$\mathsf{M}(q) := \begin{bmatrix} \gamma & -\mathsf{d}^\top \\ -\mathsf{d} & \mathsf{H} \end{bmatrix} . \tag{1}$$

Then for any $\mu \in \mathbb{R}$*, we have*

(a) $q(\mathsf{x}) \geq \mu$ *for all* $\mathsf{x} \in \mathbb{R}^n$ *if and only if* $\mathsf{M}(q - \mu) = \mathsf{M}(q) - \mu \mathsf{J}_0 \succeq \mathsf{O}$.

(b) $q(\mathsf{x}) \geq \mu$ *for all* $\mathsf{x} \in \mathbb{R}^n_+$ *if and only if* $\mathsf{M}(q - \mu) = \mathsf{M}(q) - \mu \mathsf{J}_0 \in \mathcal{C}^\star$.

Proof. The identity $\mathsf{M}(q - \mu) = \mathsf{M}(q) - \mu \mathsf{J}_0$ is evident. Assertion (a) is proved, e.g., in [19, Lemma 1]. The argument for claim (b) is completely analogous, but for the readers' convenience we provide a proof. Suppose that $q(\mathsf{x}) \geq \mu$ for all $\mathsf{x} \in \mathbb{R}^n_+$. Then H must be copositive. Indeed, otherwise consider a $\mathsf{y} \in \mathbb{R}^n_+$ such that $\mathsf{y}^\top \mathsf{H} \mathsf{y} < 0$ and look at $\mathsf{x} = t\mathsf{y}$. For large enough $t > 0$, we get

$$q(\mathsf{x}) = q(t\mathsf{y}) = t^2 \mathsf{y}^\top \mathsf{H} \mathsf{y} - 2t\mathsf{d}\mathsf{y} + \gamma < \mu ,$$

contradicting the hypothesis. So we have $[0, \mathsf{x}^\top] \mathsf{M}(q - \mu) [0, \mathsf{x}^\top]^\top = \mathsf{x}^\top \mathsf{H} \mathsf{x} \geq 0$ for all $\mathsf{x} \in \mathbb{R}^n_+$. On the other hand, we get

$$[1, \mathsf{x}^\top] \mathsf{M}(q - \mu) \begin{bmatrix} 1 \\ \mathsf{x} \end{bmatrix} = [1, \mathsf{x}^\top] \begin{bmatrix} \gamma - \mu & -\mathsf{d}^\top \\ -\mathsf{d} & \mathsf{H} \end{bmatrix} \begin{bmatrix} 1 \\ \mathsf{x} \end{bmatrix} = q(\mathsf{x}) - \mu , \tag{2}$$

and the latter is nonnegative for all $\mathsf{x} \in \mathbb{R}^n_+$, by hypothesis. By homogeneity, we arrive at $\mathsf{z}^\top \mathsf{M}(q - \mu) \mathsf{z} \geq 0$ for all $\mathsf{z} \in \mathbb{R}^{n+1}$ and one implication is shown. The converse follows readily from (2). □

This observation implies the following identities with a duality flavor:

Corollary 2.1 *For a quadratic function* $q(\mathsf{x}) = \mathsf{x}^\top \mathsf{H}\mathsf{x} - 2\mathsf{d}^\top \mathsf{x} + \gamma$*,*

(a) $\inf\{q(\mathsf{x}) : \mathsf{x} \in \mathbb{R}^n\} = \sup\{\mu \in \mathbb{R} : \mathsf{M}(q) - \mu \mathsf{J}_0 \succeq \mathsf{O}\}$*; and*

(b) $\inf\left\{q(\mathsf{x}) : \mathsf{x} \in \mathbb{R}^n_+\right\} = \sup\{\mu \in \mathbb{R} : \mathsf{M}(q) - \mu \mathsf{J}_0 \in \mathcal{C}^\star\}$*.*

Note that above equalities hold, by the usual convention ($\sup \emptyset = -\infty$), also if $q(\mathsf{x})$ is unbounded from below on $\mathbb{R}^n$ or $\mathbb{R}^n_+$.

So quite naturally we are led to our first SDP, in (a), or copositive optimization problem, in (b): optimize a linear function of a variable μ under the constraint that a matrix affine-linear in μ is either psd or copositive. More generally, in a *copositive optimization problem,* for surveys see, e.g. [4, 6, 11, 17], we are given $\mathsf{r} \in \mathbb{R}^m$ as well as $m+1$ symmetric matrices $\{\mathsf{M}_0, \dots \mathsf{M}_m\}$ of same order, and we have to maximize a linear function of m variables $y_i \geq 0$ such that the affine combination $\mathsf{M}_0 + \sum_{i=1}^m y_i \mathsf{M}_i \in \mathcal{C}^\star$:

$$z^*_{CD} := \sup \left\{ \mathsf{r}^\top \mathsf{y} : \mathsf{y} \in \mathbb{R}^m_+ \,,\ \mathsf{M}_0 + \sum_{i=1}^m y_i \mathsf{M}_i \in \mathcal{C}^\star \right\} . \tag{3}$$

This convex program has no local, non-global solutions, and the formulation shifts complexity from global optimization towards sheer feasibility questions (is $\mathsf{S} \in \mathcal{C}^\star$?). On the other hand, there are several hard non-convex programs which can be formulated as copositive problems, among them mixed-binary QPs or Standard QPs. The copositive formulation offers a unified view on some key classes of (mixed) continuous and discrete optimization problems. Applications range from machine learning to several combinatorial problems, including the maximum-clique problem or the maximum-cut problem.

Unlike the more popular SDP case, problem (3) is the conic dual of a problem involving a different matrix cone $\mathcal{C}$. Here we have to minimize a linear function $\langle \mathsf{M}_0, \mathsf{X}\rangle$ in a completely positive matrix variable X subject to linear constraints $\langle \mathsf{M}_i, \mathsf{X}\rangle \leq r_i$, $i \in [1\!:\!m]$:

$$z^*_{CP} := \inf\left\{\langle \mathsf{M}_0, \mathsf{X}\rangle : \langle \mathsf{M}_i, \mathsf{X}\rangle \leq r_i \,,\ i \in [1\!:\!m] \,,\ \mathsf{X} \in \mathcal{C}\right\} . \tag{4}$$

The reason why we consider (4) as the primal problem will be clear immediately.

Consider, for ease of exposition only, an all-quadratic optimization problem over the positive orthant,

$$z^*_+ := \inf\left\{q_0(\mathsf{x}) : q_i(\mathsf{x}) \leq 0 \,,\ i \in [1\!:\!m] \,,\ \mathsf{x} \in \mathbb{R}^n_+\right\} , \tag{5}$$

where all q_i are quadratic functions. Then $\mathsf{z} = [1, \mathsf{x}^\top]^\top \in \mathbb{R}^{n+1}_+$ and $\mathsf{X} = \mathsf{z}\mathsf{z}^\top$ is completely positive. Further, for $\mathsf{M}_i = \mathsf{M}(q_i)$ as defined in (1), we get

$q_i(\mathsf{x}) = \mathsf{z}^\top \mathsf{M}_i \mathsf{z}$ for all $i \in [0\!:\!m]$ by (2). Therefore, and by weak conic duality, we get

$$z^*_{CD} \le z^*_{CP} \le z^*_+ \, .$$

Dropping the sign constraints, we arrive at the problem

$$z^* := \inf \{ q_0(\mathsf{x}) : q_i(\mathsf{x}) \le 0 \, , \; i \in [1\!:\!m] \, , \; \mathsf{x} \in \mathbb{R}^n \} \, , \tag{6}$$

with its familiar SDP relaxation [27, 24]

$$z^*_{SD} \le z^*_{SP} \le z^* \, ,$$

where

$$z^*_{SD} := \sup \left\{ y_0 : (y_0, \mathsf{u}) \in \mathbb{R} \times \mathbb{R}^m_+ \, , \; \mathsf{M}_0 - y_0 \mathsf{J}_0 + \sum_{i=1}^m u_i \mathsf{M}_i \succeq \mathsf{O} \right\} \tag{7}$$

which is very similar to (3), and which is the dual of the SDP

$$z^*_{SP} := \inf \{ \langle \mathsf{M}_0, \mathsf{X} \rangle : \langle \mathsf{J}_0, \mathsf{X} \rangle = 1, \; \langle \mathsf{M}_i, \mathsf{X} \rangle \le 0, \; i \in [1\!:\!m] \, , \; \mathsf{X} \succeq \mathsf{O} \} \, , \tag{8}$$

the counterpart of (4). In [24] it is also shown (for the first time to the author's belief), that z^*_{SD} coincides with the Lagrangian dual for z^*.

2.2 Duality and approximation hierarchies

Strong duality for the pair (3) and (4) follow in a way standard for convex problems: strict feasibility of (4) implies attainability of z^*_{CD}, and strict feasibility of (3) implies attainability of z^*_{CP}. In either of these cases we have zero duality gap, $z^*_{CD} = z^*_{CP}$. We will investigate, and formally define, strict feasibility of these conic problems in more detail in Section 6 below. Here let us assume, for brevity of exposition, that the duality gap is zero. Both cones $\mathcal{C}$ and $\mathcal{C}^\star$ involved in this primal-dual pair are intractable. So we need to approximate them by so-called hierarchies, i.e., a sequence of tractable cones $\mathcal{D}^\star_d$ such that $\mathcal{D}^\star_d \subset \mathcal{D}^\star_{d+1} \subset \mathcal{C}^\star$ where d is the level of the hierarchy, and $\bigcup_{d=0}^\infty \mathcal{D}^\star_d = [\mathcal{C}^\star]^\circ$, i.e., every strictly copositive matrix is contained in $\mathcal{D}^\star_d$ for some d. On the dual side, $\mathcal{D}_d$ are also tractable, $\mathcal{D}_{d+1} \subset \mathcal{D}_d$, and $\bigcap_{d=0}^\infty \mathcal{D}_d = \mathcal{C}$ contains no matrix which is not completely positive. Again for brevity, assume that strong duality also holds for the approximation:

$$\begin{array}{rcl} z^*_{\mathcal{D},d} & := & \min \{ \langle \mathsf{M}_0, \mathsf{X} \rangle : \langle \mathsf{M}_i, \mathsf{X} \rangle \le r_i \, , \; i \in [1\!:\!m] \, , \; \mathsf{X} \in \mathcal{D}_d \} \\ & = & \max \left\{ \mathsf{r}^\top \mathsf{y} : \mathsf{y} \in \mathbb{R}^m_+ \, , \; \mathsf{M}_0 + \sum_{i=1}^m y_i \mathsf{M}_i \in \mathcal{D}^\star_d \right\} \, . \end{array}$$

Then by above we get $z^*_{\mathcal{D},d} \to z^*_{CD} = z^*_{CP}$ as $d \to \infty$. By now, there are many possibilities explored for hierarchies $(\mathcal{D}_d)_d$, for a concise survey see [6]. Many of these involve linear or psd constraints of matrices of order n^{d+2}, which in particular for LMIs pose a serious memory problem for algorithmic implementations even for moderate d if n is large. LP-based hierarchies suffer less from this curse of dimensionality, and therefore we will follow a compromise between LP-based and SDP-based hierarchies. We start with the usual zero-order approximation by the cone of *doubly nonnegative (DNN) matrices*

$$\mathcal{D}_0 = \{\mathsf{X} \text{ is psd} : \mathsf{X} \text{ has no negative entries}\} \ . \tag{9}$$

For the dual cone

$$\mathcal{D}_0^\star = \{\mathsf{P} + \mathsf{N} : \mathsf{P} \text{ is psd. and } \mathsf{N} \text{ has no negative entries}\} \tag{10}$$

Florian Jarre (personal communication) very recently has coined the term *nonnegative decomposable (NND)* for matrices in $\mathcal{D}_0^\star$, using the duality calculus pun $(DNN)^\star = NND$. Anyhow, based upon this construct, we may add valid linear inequalities, e.g., as done in [8, 9], yielding polyhedral inner approximations $\mathcal{L}_d^\star$ of the copositive cone, and, on the dual side, polyhedral outer approximations $\mathcal{L}_d$ for the completely positive cone, and finally define

$$\mathcal{D}_d := \mathcal{D}_0 \cap \mathcal{L}_d \, , \; d \in \{0, 1, 2, \ldots\} \ , \tag{11}$$

or, by duality, $\mathcal{D}_d^\star := \mathcal{D}_0^\star + \mathcal{L}_d^\star$ using the Minkowski sum. Of course, this approximation satisfies above properties of exhaustivity, and involves LMIs only for matrices of order linear in n; in fact, we only employ the matrices $\mathsf{M}_i = \mathsf{M}(q_i)$ of order $n + 1$.

A similar yet different approach is taken in [21] where a conic *exact reformulation* of problem (6) is proposed, using another untractable cone, and constructing tractable approximation hierarchies for this cone. The examples specified in [21] reduce again to the NND cone $\mathcal{D}_0^\star$ or its dual, the DNN cone $\mathcal{D}_0$. However for large n, even $\mathcal{D}_0$ may involve too many (namely $\frac{(n+1)n}{2}$) linear inequalities to allow for efficient computation. This problem can be overcome by warmstarting as in [18], identifying or separating valid linear inequalities on the fly, or by the recently proposed tightening and acceleration method in [20].

The following proposal is an alternative: suppose that we only employ, say, n inequalities, e.g., by forbidding negative entries only in the first row of a matrix, to proxy for complete positivity:

$$\mathcal{D}_\diamond := \{\mathsf{X} \text{ is psd} : X_{0j} \geq 0 \text{ for all } j \in [1:n]\} \ .$$

This cone can be seen as a sub-zero level approximation of $\mathcal{C}$ in light of above discussion. Its dual cone is given by

$$\mathcal{D}_\diamond^\star := \left\{ \mathsf{P} + \begin{bmatrix} 0 & \mathsf{v}^\top \\ \mathsf{v} & \mathsf{O} \end{bmatrix} : \mathsf{P} \text{ is psd} \, , \, \mathsf{v} \in \mathbb{R}_+^n \right\} \ , \tag{12}$$

and both will play a prominent role in Section 4 below.

3 Lagrangian duality for quadratic problems

3.1 Different problems and different formulations

Consider two problems with quadratic constraints:

$$z^* := \inf \{q_0(\mathsf{x}) : \mathsf{x} \in F\} \qquad \text{with } F := \{\mathsf{x} \in \mathbb{R}^n : q_i(\mathsf{x}) \le 0\,,\ i\in[1:m]\} \quad (13)$$

where all $q_i(\mathsf{x}) = \mathsf{x}^\top \mathsf{Q}_i \mathsf{x} - 2\mathsf{b}_i^\top \mathsf{x} + c_i$ are quadratic functions (we may assume $c_0 = 0$) for $i\in[0:m]$; and

$$z_+^* := \inf \{q_0(\mathsf{x}) : \mathsf{x} \in F \cap P\} \qquad \text{with } P := \left\{\mathsf{x} \in \mathbb{R}^n_+ : \mathsf{A}\mathsf{x} = \mathsf{a}\right\}\,. \qquad (14)$$

where $\mathsf{a} \in \mathbb{R}^p$ and A is a $p \times n$ matrix of full row rank p. We further assume that

$$\text{for all } j\in[1:n] \text{ there is an } \mathsf{x}^{(j)} \in P \text{ such that } x_j^{(j)} > 0\,. \qquad (15)$$

This is no restriction of generality, since we can test this condition by solving, in a preprocessing step, for all $j\in[1:n]$, the n LPs $z_j^* := \sup\{x_j : \mathsf{x} \in P\}$, and discard the variable x_j if $z_j^* = 0$.

Neither of the optimal values z^* of (13), or z_+^* of (14) need be attained, and they could also equal to $-\infty$ (in the unbounded case) or to $+\infty$ (in the infeasible case). Of course, we have $z^* \le z_+^*$ due to the additional sign constraints. Considering $\mathsf{Q}_i = \mathsf{O}$ would also allow for linear inequalities in the constraints. However, it is often advisable to discriminate the functional form of constraints, and we will adhere to this principle in what follows. Structural linear *inequality* constraints, however, are cast into above form by use of slack variables.

Note that defining $\mathsf{Q}_{m+1} = \mathsf{A}^\top\mathsf{A}$, $\mathsf{b}_{m+1} = \mathsf{A}^\top\mathsf{a}$ and $c_{m+1} = \mathsf{a}^\top\mathsf{a}$, we may rephrase the m linear constraints $\mathsf{A}\mathsf{x} = \mathsf{a}$ into one homogeneous quadratic constraint $\mathsf{z}^\top\mathsf{M}_{m+1}\mathsf{z} = \|\mathsf{A}\mathsf{x} - \mathsf{a}\|^2 = 0$. We will return later to this formulation. Still, the resulting feasible set is not of the form of F, the difference being the sign constraints $x_j \ge 0$.

Finally note that binarity constraints $x_j \in \{0, 1\}$ can be recast into two inequality constraints of the form $x_j \le 1$ (this constraint would ensure Burer's key condition [7, 10]) and $x_j - x_j^2 \le 0$. This fits into above formulation, but then one has to be careful with strict feasibility assumptions; also, introducing slacks for $x_j \le 1$ will double the number of variables. We will address an alternative later in Subsection 5.2.

3.2 The Lagrangian (dual) functions

Now consider multipliers $\mathsf{u} \in \mathbb{R}^m_+$ of the inequality constraints $q_i(\mathsf{x}) \leq 0$, $2\mathsf{v} \in \mathbb{R}^n_+$ for the sign constraints $\mathsf{x} \in \mathbb{R}^n_+$, and $2\mathsf{w} \in \mathbb{R}^p$ for the linear equality constraints $\mathsf{A}\mathsf{x} = \mathsf{a}$ (again, the factors two are introduced for notational convenience only). Then the full Lagrangian function

$$L(\mathsf{x};\mathsf{u},\mathsf{v},\mathsf{w}) := q_0(\mathsf{x}) + \sum_i u_i q_i(\mathsf{x}) - 2\mathsf{v}^\top\mathsf{x} + 2\mathsf{w}^\top(\mathsf{a} - \mathsf{A}\mathsf{x})$$

and its first two derivatives w.r.t. x are given by

$$\begin{aligned} L(\mathsf{x};\mathsf{u},\mathsf{v},\mathsf{w}) &= \mathsf{x}^\top \mathsf{H}_u \mathsf{x} - 2(\mathsf{d}_\mathsf{u} + \mathsf{v} + \mathsf{A}^\top\mathsf{w})^\top\mathsf{x} + \mathsf{c}^\top\mathsf{u} + 2\mathsf{w}^\top\mathsf{a}\,, \\ \nabla_\mathsf{x} L(\mathsf{x};\mathsf{u},\mathsf{w}) &= 2[\mathsf{H}_\mathsf{u}\mathsf{x} - (\mathsf{d}_\mathsf{u} + \mathsf{v} + \mathsf{A}^\top\mathsf{w})] \quad \text{and} \\ D^2_\mathsf{x} L(\mathsf{x};\mathsf{u},\mathsf{w}) &= 2\mathsf{H}_\mathsf{u} \qquad \text{for all } (\mathsf{x};\mathsf{u},\mathsf{v},\mathsf{w}) \in \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}^p \times \mathbb{R}^n_+\,. \end{aligned}$$

Here we denote by $\mathsf{H}_\mathsf{u} = \mathsf{Q}_0 + \sum_{i=1}^m u_i \mathsf{Q}_i$, by $\mathsf{d}_\mathsf{u} = \mathsf{b}_0 + \sum_{i=1}^m u_i \mathsf{b}_i$ and by $\mathsf{c} = [c_1, \ldots, c_m]^\top$. Abbreviating $L_0(\mathsf{x};\mathsf{u}) = L(\mathsf{x};\mathsf{u},\mathsf{o},\mathsf{o})$, the Lagrangian dual function for problem (13) reads

$$\Theta_0(\mathsf{u}) := \inf\left\{L_0(\mathsf{x};\mathsf{u}) : \mathsf{x} \in \mathbb{R}^n\right\}\,, \tag{16}$$

and the dual optimal value is

$$z^*_{LD} := \sup\left\{\Theta_0(\mathsf{u}) : \mathsf{u} \in \mathbb{R}^m_+\right\}\,. \tag{17}$$

Standard weak duality implies $z^*_{LD} \leq z^*$.

The full Lagrangian dual for problem (14) with additional linear constraints reads instead

$$\Theta(\mathsf{u},\mathsf{v},\mathsf{w}) := \inf\left\{L(\mathsf{x};\mathsf{u},\mathsf{v},\mathsf{w}) : \mathsf{x} \in \mathbb{R}^n\right\}\,, \tag{18}$$

with dual optimal value

$$z^*_{LD,+} := \sup\left\{\Theta(\mathsf{u},\mathsf{v},\mathsf{w}) : (\mathsf{u},\mathsf{v},\mathsf{w}) \in \mathbb{R}^m_+ \times \mathbb{R}^n_+ \times \mathbb{R}^p\right\}\,. \tag{19}$$

The idea to incorporate some of the constraints defining $F \cap P$ into the ground set, or equivalently, to relax only some of the constraints, leads to the corresponding Semi-Lagrangian (sometimes also called *partial Lagrangian*) dual and is not new, see, e.g. [19] and references therein. However, previous work has concentrated to do this with linear equality constraints, which then leads to an SDP formulation similar to those treated in the previous section. Here, we take an alternative path, incorporating the sign (i.e., inequality) constraints into the ground set, and relax all other constraints.

So we arrive at the Semi-Lagrangian variant

$$\Theta_{\rm semi}(\mathsf{u},\mathsf{w}) := \inf \left\{ L(\mathsf{x};\mathsf{u},\mathsf{o},\mathsf{w})) : \mathsf{x} \in \mathbb{R}^n_+ \right\} , \tag{20}$$

with dual optimal value

$$z^*_{\rm semi} := \sup \left\{ \Theta_{\rm semi}(\mathsf{u},\mathsf{w}) : (\mathsf{u},\mathsf{w}) \in \mathbb{R}^m_+ \times \mathbb{R}^p \right\} . \tag{21}$$

The relation between full and Semi-Lagrangian bounds is a general principle. For ease of reference, we repeat the argument here: for any $\mathsf{v} \in \mathbb{R}^n_+$,

$$\begin{aligned} \Theta(\mathsf{u},\mathsf{v},\mathsf{w}) &= \inf \{ L(\mathsf{x};\mathsf{u},\mathsf{v},\mathsf{w}) : \mathsf{x} \in \mathbb{R}^n \} \\ &\leq \inf \left\{ L(\mathsf{x};\mathsf{u},\mathsf{v},\mathsf{w}) : \mathsf{x} \in \mathbb{R}^n_+ \right\} \\ &= \inf \left\{ L(\mathsf{x};\mathsf{u},\mathsf{o},\mathsf{w}) - 2\mathsf{v}^\top \mathsf{x} : \mathsf{x} \in \mathbb{R}^n_+ \right\} \\ &\leq \inf \left\{ L(\mathsf{x};\mathsf{u},\mathsf{o},\mathsf{w}) : \mathsf{x} \in \mathbb{R}^n_+ \right\} \;=\; \Theta_{\rm semi}(\mathsf{u},\mathsf{w}) , \end{aligned}$$

as $\mathsf{v}^\top \mathsf{x} \geq 0$ for all $\mathsf{x} \in \mathbb{R}^n_+$. So we arrive at the following chain of inequalities

$$z^*_{LD,+} \leq z^*_{\rm semi} \leq z^*_+ ,$$

where the last inequality above follows, again, from standard weak duality.

We also have $z^*_{LD} \leq z^*_{LD,+}$ as $\Theta_0(\mathsf{u}) = \Theta(\mathsf{u},\mathsf{o},\mathsf{o})$, but as z^*_{LD} and $z^*_{LD,+}$ refer to different problems, their relation cannot be seen as a tightening, but rather as a reflection of the relation $z^* \leq z^*_+$ of the optimal (primal) values of (13) and (14), respectively.

4 A new perspective on SDP relaxations

4.1 SDP and Lagrangian dual in absence of linear constraints

We have $\Theta_0(\mathsf{u}) > -\infty$ if and only if (a) $\mathsf{H_u} \succeq \mathsf{O}$; and (b) the linear equation system $\mathsf{H_u x} = \mathsf{d_u}$ has a solution. In this case $\Theta_0(\mathsf{u}) = L_0(\mathsf{x};\mathsf{u})$ for any x with $\mathsf{H_u x} = \mathsf{d_u}$, or

$$\Theta_0(\mathsf{u}) = L_0(\mathsf{x};\mathsf{u}) = \mathsf{x}^\top \mathsf{d}_u - 2\mathsf{d}_\mathsf{u}^\top \mathsf{x} + \mathsf{c}^\top \mathsf{u} = \mathsf{c}^\top \mathsf{u} - \mathsf{d}_\mathsf{u}^\top \mathsf{x} .$$

So the Lagrangian dual problem can be written as a Wolfe dual with an additional psd constraint, namely as

$$z^*_{LD} = \sup \left\{ L_0(\mathsf{x};\mathsf{u}) : (\mathsf{x},\mathsf{u}) \in \mathbb{R}^n \times \mathbb{R}^m_+, \mathsf{H_u} \succeq \mathsf{O} \,,\; \mathsf{H_u x} = \mathsf{d_u} \right\} .$$

Unfortunately, the condition $\Theta_{\rm semi}(\mathsf{u},\mathsf{w}) > -\infty$ does not allow for nice conditions similar to requiring $\mathsf{H_u} \succeq \mathsf{O}$ and solvability of $\mathsf{H_u x} = \mathsf{d}_u + \mathsf{v} + \mathsf{A}^\top \mathsf{w}$, which would now be the first-order condition $\nabla L(\mathsf{x};\mathsf{u},\mathsf{v},\mathsf{w}) = \mathsf{o}$. However, for $\Theta_0(\mathsf{u})$ these conditions played a key role for the equivalence result $z^*_{LD} = z^*_{SD}$, established in [24]. Here we will pass, also in light of the difficulties with $\Theta_{\mathrm{semi}}(\mathsf{u},\mathsf{w})$, to a different formulation of this semidefinite relaxation for the problem (13) which immediately follows from Corollary 2.1:

Theorem 4.1 *Consider problem (13) and its Lagrangian dual function as defined in (16). Then*

$$\Theta_0(\mathsf{u}) = \sup\left\{\mu : \mu \in \mathbb{R}\,,\ \mathsf{M}(L_0(\cdot;\mathsf{u})) - \mu \mathsf{J}_0 \succeq \mathsf{O}\right\}$$

and

$$z^*_{LD} = \sup\left\{\mu : (\mu,\mathsf{u}) \in \mathbb{R}\times\mathbb{R}^m_+,\ \mathsf{M}(L_0(\cdot;\mathsf{u})) - \mu \mathsf{J}_0\mathsf{J}_0 \succeq \mathsf{O}\right\}\ .$$

Further, we have $z^*_{LD} = z^*_{SD}$ as defined in (7); so a zero duality gap $z^*_{LD} = z^*$ occurs if and only if (a) the SDP relaxation has itself no positive duality gap, and (b) the SDP relaxation is tight.

Proof. The first equation follows directly from Corollary 2.1(a), and the second equation is then immediate. But obviously

$$\mathsf{M}(L_0(\cdot;\mathsf{u})) - \mu\mathsf{J}_0 = \mathsf{M}(q_0) - y_0\mathsf{J}_0 + \sum_{i=1}^m u_i \mathsf{M}(q_i)$$

when $y_0 = \mu$. Now, considering the equality constraint $\langle \mathsf{J}_0, \mathsf{X}\rangle = 1$ with multiplier $y_0 \in \mathbb{R}$ and the inequality constraints $\langle \mathsf{M}(q_i), \mathsf{X}\rangle \le 0$ with multiplier $u_i \ge 0$, all $i \in [1\!:\!m]$, we arrive at the dual SDP (7), exactly as required. So we arrive at

$$z^*_{LD} = z^*_{SD} \le z^*_{SP} \le z^*$$

wherefrom the last assertion follows. □

Thus the slack matrix of the conic relaxation for (13) is

$$\mathsf{Z}(\mathsf{y}) := \mathsf{M}_0 - y_0\mathsf{J}_0 + \sum_{i=1}^m u_i\mathsf{M}_i = \begin{bmatrix} \mathsf{c}^\top\mathsf{u} - y_0 & -\mathsf{d}_\mathsf{u}^\top \\ -\mathsf{d}_\mathsf{u} & \mathsf{H}_\mathsf{u} \end{bmatrix}, \tag{22}$$

where $\mathsf{y} = (y_0,\mathsf{u}) \in \mathbb{R}\times\mathbb{R}^m_+$ collects all dual variables. We will encounter updates of these slack matrices in the sequel.

4.2 Full Lagrangian dual with linear constraints

Now consider the full Lagrangian function; perhaps somewhat unexpectedly, we are lead to a conic constraint where the sub-zero level approximation cone $\mathcal{D}_\diamond$ occurs:

Theorem 4.2 *Consider problem (14) and its Lagrangian dual function as defined in (18). Then for all* $(\mathsf{u},\mathsf{v},\mathsf{w}) \in \mathbb{R}^m_+ \times \mathbb{R}^n_+ \times \mathbb{R}^p$

$$\Theta(\mathsf{u},\mathsf{v},\mathsf{w}) = \sup\left\{\mu : \mu \in \mathbb{R}\,,\ \mathsf{M}(L(\cdot;\mathsf{u},\mathsf{v},\mathsf{w})) - \mu \mathsf{J}_0 \succeq \mathsf{O}\right\}$$

and the full Lagrangian dual problem of (14) can be written as

$$z^*_{LD,+} = \sup\left\{\mu : (\mu,\mathsf{u},\mathsf{w}) \in \mathbb{R}\times\mathbb{R}^m_+\times\mathbb{R}^p,\ \mathsf{M}(L(\cdot;\mathsf{u},\mathsf{o},\mathsf{w})) - \mu\mathsf{J}_0 \in \mathcal{D}^\star_\diamond\right\}\,. \tag{23}$$

Proof. The first equation is again a direct consequence of Corollary 2.1(a). For the second, observe that

$$\mathsf{M}(L(\cdot;\mathsf{u},\mathsf{v},\mathsf{w})) - \mu\mathsf{J}_0 = \mathsf{M}(L(\cdot;\mathsf{u},\mathsf{o},\mathsf{w})) - \mu\mathsf{J}_0 - \begin{bmatrix} 0 & -\mathsf{v}^\top \\ -\mathsf{v} & \mathsf{O}\end{bmatrix},$$

so that $\mathsf{M}(L(\cdot;\mathsf{u},\mathsf{o},\mathsf{w})) - \mu\mathsf{J}_0 \in \mathcal{D}^\star_\diamond$ if and only if $\mathsf{M}(L(\cdot;\mathsf{u},\mathsf{v},\mathsf{w})) - \mu\mathsf{J}_0 \succeq \mathsf{O}$ for some $\mathsf{v}\in\mathbb{R}^n_+$, by (12). The result follows. □

Hence we can characterize also the full Lagrangian dual for (14) as an SDP, namely the dual of the natural SDP relaxation of (14): to this end, let us express the p linear equality constraints as $\mathsf{r}_k^\top\mathsf{x} = a_k$ with $\mathsf{r}_k \in \mathbb{R}^n$ for all $k\in[1\!:\!p]$. So $\mathsf{A}^\top = [\mathsf{r}_1,\dots,\mathsf{r}_p]^\top$ with $\mathsf{r}_k^\top$ the kth row of A. For all $k\in[1\!:\!p]$, we define the symmetric matrices of order $n+1$

$$\mathsf{A}_k := \begin{bmatrix} 2a_k & -\mathsf{r}_k^\top \\ -\mathsf{r}_k & \mathsf{O}\end{bmatrix}. \tag{24}$$

Theorem 4.3 *Consider the full Lagrangian dual* $z^*_{LD,+}$ *as defined in (19) and expressed in Theorem 4.2. Then this is the conic dual of the SDP*

$$z_{SP,+} := \inf\left\{\langle\mathsf{M}_0,\mathsf{X}\rangle : \mathsf{X} = \begin{bmatrix}1 & \mathsf{x}^\top \\ \mathsf{x} & \mathsf{Y}\end{bmatrix} \succeq \mathsf{O}, \mathsf{x}\in P,\ \langle\mathsf{M}_i,\mathsf{X}\rangle \le 0, i\in[1\!:\!m]\right\}, \tag{25}$$

which can be easily seen as the natural SDP relaxation of (14). Therefore we have

$$z^*_{LD,+} = z^*_{SD,+} \le z^*_{SP,+} \le z^*_+\,,$$

and the full Lagrangian relaxation is tight, $z^*_{LD,+} = z^*_+$*, if and only if (a) the SDP relaxation has zero duality gap,* $z^*_{SD,+} = z^*_{SP,+}$*, and (b) the primal SDP relaxation (25) is tight.*

Proof. Whenever the top (zeroth) row of X reads $\mathsf{z}^\top = [1, \mathsf{x}^\top]$, we have, due to (24), $2(\mathsf{r}_k^\top \mathsf{x} - a_k) = \mathsf{z}^\top \mathsf{A}_k \mathsf{z} = \langle \mathsf{A}_k, \mathsf{X}\rangle$. Hence the constraint $\langle \mathsf{A}_k, \mathsf{X}\rangle = 0$ is equivalent to $\mathsf{r}_k^\top \mathsf{x} = a_k$. So $\mathsf{x} \in P$ is equivalent to $\mathsf{x} \in \mathbb{R}^n_+$ and $\langle \mathsf{A}_k, \mathsf{X}\rangle = 0$ for all $k \in [1:p]$. Now choose multipliers $w_k \in \mathbb{R}$ for these equality constraints. Then, if we dualize the SPD (25) by the standard procedure, we arrive at the new slack matrix

$$\mathsf{Z}_+(\mathsf{y}, \mathsf{w}) := \mathsf{Z}(\mathsf{y}) + \sum_{k=1}^{p} w_k \mathsf{A}_k = \begin{bmatrix} \mathsf{c}^\top \mathsf{u} - y_0 + 2\mathsf{w}^\top \mathsf{a} & -\mathsf{d}_\mathsf{u}^\top - \mathsf{w}^\top \mathsf{A} \\ -\mathsf{d}_\mathsf{u} - \mathsf{A}^\top \mathsf{w} & \mathsf{H}_\mathsf{u} \end{bmatrix}, \quad (26)$$

where $\mathsf{Z}(\mathsf{y})$ is defined as in (22). Now notice that for $y_0 = \mu$, we have

$$\mathsf{M}(L(\cdot, \mathsf{u}, \mathsf{o}, \mathsf{w})) - \mu \mathsf{J}_0 = \mathsf{Z}_+(\mathsf{y}, \mathsf{w}) \quad \text{if } \mathsf{y} = (\mu, \mathsf{u}).$$

Hence the result follows by (23). $\square$

There are several, a priori different, SDP formulations for the full Lagrangian dual of (14), some adapted to special subclasses; see, e.g. [19] and references therein.

5 Semi-Lagrangian dual and copositive relaxation

5.1 A two-fold characterization of Semi-Lagrangian dual

Before we proceed to the Semi-Lagrangian case, we introduce the natural copositive relaxation of (14), in analogy to (25). Consider therefore A_k as in (24) and form the problem

$$z^*_{CP} := \inf_{\mathsf{X} \in \mathcal{C}} \left\{ \langle \mathsf{M}_0, \mathsf{X}\rangle : \begin{array}{rcll} \langle \mathsf{M}_i, \mathsf{X}\rangle & \leq & 0\,, & i \in [1:m], \\ \langle \mathsf{A}_k, \mathsf{X}\rangle & = & 0\,, & k \in [1:p], \\ \langle \mathsf{J}_0, \mathsf{X}\rangle & = & 1 & \end{array} \right\} \quad (27)$$

and its dual

$$z^*_{CD} := \sup \left\{ y_0 : \mathsf{Z}_+(\mathsf{y}, \mathsf{w}) \in \mathcal{C}^\star, (\mathsf{y}, \mathsf{w}) = (y_0, \mathsf{u}, \mathsf{w}) \in \mathbb{R} \times \mathbb{R}^m_+ \times \mathbb{R}^p \right\} \quad (28)$$

with the slack matrix $\mathsf{Z}_+(\mathsf{y}, \mathsf{w})$ as defined in (26).

Theorem 5.1 *Consider problem (14) and its Semi-Lagrangian dual function as defined in (20), the dual z^*_{semi} as defined in (21), as well as the copositive relaxation (27) and (28). Then*

$$\Theta_{\text{semi}}(\mathsf{u}, \mathsf{w}) = \sup \{ \mu : \mu \in \mathbb{R}\,,\ \mathsf{M}(L(\cdot; \mathsf{u}, \mathsf{o}, \mathsf{w})) - \mu \mathsf{J}_0 \in \mathcal{C}^\star \}$$

and the Semi-Lagrangian dual problem of (14) can be written as

$$z^*_{\rm semi} = \sup \left\{ \mu : (\mu, \mathsf{u}, \mathsf{w}) \in \mathbb{R} \times \mathbb{R}^m_+ \times \mathbb{R}^p,\ \mathsf{M}(L(\cdot; \mathsf{u}, \mathsf{o}, \mathsf{w})) - \mu \mathsf{J}_0 \in \mathcal{C}^\star \right\} .$$

Further, we have

$$z^*_{LD,+} \le z^*_{\rm semi} = z^*_{CD} \le z^*_{CP} \le z^*_+ \,, \tag{29}$$

and the Semi-Lagrangian relaxation is tight, $z^*_{\rm semi} = z^*$*, if and only if if and only if (a) the copositive relaxation has no positive duality gap,* $z^*_{CD} = z^*_{CP}$*, and (b) the copositive primal relaxation (27) is tight.*

Proof. The first equation is now a direct consequence of Corollary 2.1(b). The remainder is as an immediate generalization of Theorem 4.3. □

So we have characterized the Semi-Lagrangian dual in two ways: (a) as the dual of the natural (primal) copositive relaxation for the problem (14); and (b) as the natural extension of the (dual) SDP relaxation for the same problem. But we can say more, in particular regarding computational consequences, see the next subsection.

5.2 Approximate copositive bounds dominate Lagrangian dual bounds even at (sub)zero level

The fact that every positive-semidefinite matrix lies in $\mathcal{D}^\star_\diamond$ is another reflection of the relation $z^*_{LD} \le z^*_{LD,+}$. On the other side, we by now can easily see that even at zero level of approximation, the resulting tractable bound tightens the Lagrangian bound:

Theorem 5.2 *Consider any approximation hierarchy* $\mathcal{D}_d$ *starting with* $\mathcal{D}_0$ *as defined in (9), e.g. the one defined in (11), together with their bounds* $z^*_{\mathcal{D},d}$*. Then*

$$z^*_{LD,+} \le z^*_{\mathcal{D},d} \quad \text{for all } d \in \{0, 1, \ldots\} \,,$$

and $z^*_{\mathcal{D},d} \to z^*_{\rm semi}$ *as* $d \to \infty$*.*

Proof. Follows from $\mathcal{D}^\star_\diamond \subset \mathcal{D}^\star_0 \subset \mathcal{D}^\star_d \subset \mathcal{C}^\star$ for all $d \ge 0$; cf. (10) and (12). □

At the end of this section, we pass to an even tighter copositive relaxation put forward by Burer in his seminal paper [10], although this is not made explicit there in full generality; but see the more recent papers [12, 13]. Basically, he proposed in [10] to complement the condition $\langle \mathsf{A}_k, \mathsf{X}\rangle = 0$ by another one resulting from squaring the linear constraint $\mathsf{r}_k^\top \mathsf{x} = a_k$: again, with $\mathsf{X} = [1, \mathsf{x}^\top]^\top [1, \mathsf{x}^\top]$, we have

$$\langle \mathsf{r}_k \mathsf{r}_k^\top, \mathsf{x}\mathsf{x}^\top \rangle = (\mathsf{r}_k^\top \mathsf{x})^2 = a_k^2 \quad \iff \quad \langle \mathsf{B}_k, \mathsf{X}\rangle = 0 \text{ with } \mathsf{B}_k := \begin{bmatrix} -a_k^2 & \mathsf{o}^\top \\ \mathsf{o} & \mathsf{r}_k \mathsf{r}_k^\top \end{bmatrix}.$$

So we arrive at another copositive relaxation for (14),

$$\left.\begin{array}{rcl} z^*_{\mathrm{Burer},P} & := & \inf\limits_{\mathsf{X}\in\mathcal{C}} \left\{ \langle \mathsf{M}_0, \mathsf{X}\rangle : \begin{array}{rcll} \langle \mathsf{M}_i, \mathsf{X}\rangle & \le & 0\,, & i\in[1\!:\!m], \\ \langle \mathsf{A}_k, \mathsf{X}\rangle & = & 0\,, & k\in[1\!:\!p], \\ \langle \mathsf{B}_k, \mathsf{X}\rangle & = & 0\,, & k\in[1\!:\!p], \\ \langle \mathsf{J}_0, \mathsf{X}\rangle & = & 1 & \end{array} \right\} \text{ and} \\ z^*_{\mathrm{Burer},D} & := & \sup \left\{ y_0 : \begin{array}{l} \mathsf{y} = (y_0, \mathsf{u}) \in \mathbb{R}\times\mathbb{R}^m_+\,, \\ (\mathsf{w}, \mathsf{z}) \in \mathbb{R}^p \times \mathbb{R}^p\,, \\ \mathsf{Z}_{\mathrm{Burer}}(\mathsf{y}, \mathsf{w}, \mathsf{z}) \in \mathcal{C}^\star \end{array} \right\} \end{array}\right\} \qquad (30)$$

with $\mathsf{Z}_{\mathrm{Burer}}(\mathsf{y}, \mathsf{w}, \mathsf{z}) = \mathsf{Z}_+(\mathsf{y}, \mathsf{w}) + \sum\limits_{k=1}^{p} z_k \mathsf{B}_k$. Since $\mathsf{Z}_{\mathrm{Burer}}(\mathsf{y}, \mathsf{w}, \mathsf{o}) = \mathsf{Z}_+(\mathsf{y}, \mathsf{w})$, we get

$$z^*_{\mathrm{semi}} = z^*_{CD} \le z^*_{\mathrm{Burer},D} \le z^*_{\mathrm{Burer},P} \le z^*_+$$

and similarly $z^*_{CP} \le z^*_{\mathrm{Burer},P} \le z^*_+$.

For linearly constrained quadratic problems with binarity constraints which are formulated as $q_j(\mathsf{x}) = x_j - x_j^2 = 0$ (and relaxed as $\langle \mathsf{M}(q_j), \mathsf{X}\rangle = 0$ with multipliers $u_j \in \mathbb{R}$), the duality gap is zero. Indeed, for $\mathsf{u} = t\mathsf{e}$ and $\mathsf{y} = (y_0, \mathsf{u})$,

$$\mathsf{Z}_{\mathrm{Burer}}(\mathsf{y}, \mathsf{o}, \mathsf{o}) = \begin{bmatrix} y_0 & (t\mathsf{e}^\top - \mathsf{b}_0)^\top \\ (t\mathsf{e} - \mathsf{b}_0) & -2t\mathsf{I}_n + \mathsf{Q}_0 \end{bmatrix}$$

can always be made strictly copositive in light of Lemma 6.1 below, e.g. for $t = \min\{3\lambda_{\min}(\mathsf{Q}_0), -1\}$. Observe that in this case, no sign restrictions to u apply.

Burer showed in [10] that under a mild condition, this relaxation is always tight, $z^*_{\mathrm{Burer},D} = z^*_{\mathrm{Burer},P} = z^*_+$. He also extended this result for problems with additional quadratic *equality* constraints, e.g., complementarity constraints, but did focus on reformulation rather than on relaxation (and the problem (14) with *inequality* constraints was not treated there).

The same strategy as before, replacing $\mathcal{C}$ with $\mathcal{D}_d$ or $\mathcal{C}^\star$ with $\mathcal{D}_d^\star$, would therefore tighten the approximative bounds even beyond the Semi-Lagrangian dual, at the cost of dealing with additional constraints. See Subsection 7.3 for further discussion.

6 Strict feasibility and strong duality

It can easily be shown that strict feasibility of (13) implies strict feasibility of (8). Moreover, if Q_i is (strictly) positive-definite for at least one $i\in[1:m]$, then also (7) is strictly feasible, so that full strong duality holds for the primal-dual SDP pair; see [1, 24]. Under these assumptions, we arrive at

$$z^*_{LD} = z^*_{SD} = z^*_{SP} \leq z^* \,.$$

Now we pass to the problem (14) with linear constraints. By analogous reasons, if one Q_i is positive-definite and if there is a $\widehat{\mathsf{x}} \in P$ with $q_i(\widehat{\mathsf{x}}) < 0$ for all $i\in[1:m]$, then strong duality for the SDP pair (25) and its dual (19) holds: both optimal objective values are attained and equal the dual full Lagrangian bound, $z^*_{LD,+} = z^*_{SD,+} = z^*_{SP,+}$.

We proceed to develop a similar theory for the copositive formulation. This is not as straightforward as it may seem at first sight, as not all relations carry over directly from the (self-dual) psd cone to the pair of dual cones $(\mathcal{C}, \mathcal{C}^\star)$. But before we show that the employed conditions on Q_i also guarantee attainability of (14).

6.1 Sufficient conditions for attainability of original problem

We first need the following auxiliary result:

Lemma 6.1 *Given arbitrary* $\mathsf{d} \in \mathbb{R}^n$ *and a symmetric* $n \times n$ *matrix* H*, consider* $q(\mathsf{x}) = \mathsf{x}^\top \mathsf{H} \mathsf{x} - 2\mathsf{d}^\top \mathsf{x}$*. For any* $\mu \in \mathbb{R}$ *define via (1)*

$$\mathsf{S}_\mu := \mathsf{M}(q) + \mu \mathsf{J}_0 = \mathsf{M}(q + \mu)\,.$$

If H *is strictly copositive, then*

(a) *there is a* $\bar\mu \geq 0$ *such that* S_μ *are strictly copositive for all* $\mu \geq \bar\mu$*;*

(b) q *is bounded from below over* $\mathbb{R}^n_+$*.*

Proof. (a) Since H is strictly copositive, $\sigma := \min\left\{\mathsf{y}^\top \mathsf{H} \mathsf{y} : \mathsf{y} \in \Delta\right\} > 0$. Further define

$$\bar\mu := \frac{2}{\sigma} \max\left\{(\mathsf{d}^\top \mathsf{y})^2 + 1 : \mathsf{y} \in \Delta\right\} > 0\,.$$

Now pick an arbitrary $\mathsf{z} = [x_0, \mathsf{x}^\top]^\top \in \mathbb{R}^{n+1}_+ \setminus \{\mathsf{o}\}$. If $\mathsf{x} = \mathsf{o}$, then $x_0 > 0$ and $\mathsf{z}^\top \mathsf{S}_{\bar\mu} \mathsf{z} = \bar\mu x_0^2 > 0$. If $\mathsf{x} \neq \mathsf{o}$, then $\mathsf{y} := \frac{1}{\mathsf{e}^\top \mathsf{x}} \mathsf{x} \in \Delta$ and $y_0 := \frac{1}{\mathsf{e}^\top \mathsf{x}} x_0 \geq 0$. We conclude

$$\mathsf{z}^\top \mathsf{S}_{\bar\mu} \mathsf{z} = (\mathsf{e}^\top \mathsf{x})^2 [\bar\mu y_0^2 - 2 y_0 \mathsf{d}^\top \mathsf{y} + \mathsf{y}^\top \mathsf{H} \mathsf{y}] \geq (\mathsf{e}^\top \mathsf{x})^2 [\bar\mu y_0^2 - 2 y_0 \mathsf{d}^\top \mathsf{y} + \sigma]\,.$$

Now the strictly convex function $\psi(t) = \bar{\mu}t^2 - 2(\mathsf{d}^\top\mathsf{y})t + \sigma$ attains its minimum over the positive half-ray ($t \geq 0$) either at $t = 0$ with value $\psi(0) = \sigma$, or else at $\bar{t} = \frac{\mathsf{d}^\top\mathsf{y}}{\bar{\mu}}$ with value $\psi(\bar{t}) = \sigma - \frac{(\mathsf{d}^\top\mathsf{y})^2}{\bar{\mu}} \geq \frac{\sigma}{2} > 0$. Hence $\mathsf{z}^\top\mathsf{S}_{\bar{\mu}}\mathsf{z} \geq (\mathsf{e}^\top\mathsf{x})^2\frac{\sigma}{2} > 0$, and claim (a) follows. Assertion (b) then is a consequence of (a) and Corollary 2.1(b). □

One may wonder whether there is a "weak" version of Lemma 6.1(a). However, the example $\mathsf{H} = \mathsf{O}$ and $\mathsf{d} = \mathsf{e}$ shows that S_μ is never copositive, although H is. The corresponding observation and the "strict" result for positive-(semi)definiteness is folklore, but by passing from positive-definite matrices to strictly copositive matrices, we will strengthen these findings, and also derive a stronger version of (29) in case of linear constraints.

So let us next consider primal attainability of the original problem (14). By Lemma 6.1(b), strict copositivity of the objective Hessian matrix $2\mathsf{Q}_0$ implies $z_+^* > -\infty$. Further, if $F \cap P \neq \emptyset$, we therefore have a finite optimal value $z_+^* \in \mathbb{R}$.

Theorem 6.1 *Suppose that for at least one $i \in [0:m]$, the matrix Q_i is strictly copositive. If $F \cap P \neq \emptyset$, then z_+^* is attained: there is an $\mathsf{x}^* \in F \cap P$ such that $q_0(\mathsf{x}^*) = z_+^*$. Further, if above i satisfies $i \geq 1$, then $F \cap P$ is compact.*

Proof. Let Q_i be strictly copositive, and define

$$\rho_i := \max\left\{\frac{\mathsf{b}_i^\top\mathsf{y}}{\mathsf{y}^\top\mathsf{Q}_i\mathsf{y}} : \mathsf{y} \in \Delta\right\}$$

as well as $\tau_i := \max\{2\rho_i - c_i, 2\rho_i, 1\}$. Suppose now $\mathsf{x} = t\mathsf{y}$ with $t := \mathsf{e}^\top\mathsf{x} \geq 1$ and $\mathsf{y} \in \Delta$. We deduce from $q_i(\mathsf{x}) = t^2\mathsf{y}^\top\mathsf{Q}_i\mathsf{y} - 2t\mathsf{b}_i^\top\mathsf{y} + c_i$ that

$$q_i(\mathsf{x}) \leq 0 \quad \text{implies} \quad t \leq \tau_i\,. \tag{31}$$

If $i = 0$, we redefine $c_0 := -z_+^* - 1$ (thus $\tau_0 = \max\{2\rho_0 + z^* + 1, 2\rho_0, 1\}$) and infer that $q_0(\mathsf{x}) > z_+^* + 1$ whenever $\mathsf{e}^\top\mathsf{x} > \tau_0$, by (31). Therefore

$$z_+^* = \inf\{q_0(\mathsf{x}) : \mathsf{x} \in F \cap P\} = \min\left\{q_0(\mathsf{x}) : \mathsf{x} \in F \cap P\,,\, \mathsf{e}^\top\mathsf{x} \leq \tau_0\right\}\,.$$

The latter minimum is attained since $\left\{\mathsf{x} \in \mathbb{R}^n_+ : \mathsf{e}^\top\mathsf{x} \leq \tau_i\right\}$ is compact. For $i \in [1:m]$, we deduce in the same way directly from (31)

$$F \cap P \subseteq \left\{\mathsf{x} \in \mathbb{R}^n_+ : q_i(\mathsf{x}) \leq 0\right\} \subseteq \left\{\mathsf{x} \in \mathbb{R}^n_+ : \mathsf{e}^\top\mathsf{x} \leq \tau_i\right\}$$

and thus z_+^* must be attained as a minimum of the continuous function q_0 over the compact set $F \cap P$. □

6.2 Strong duality in the copositive relaxation

Now we turn to strong duality of the copositive problem.

Theorem 6.2 *Consider the copositive relaxation (27) and (28) of (14).*

(a) *Suppose that* Q_i *is strictly copositive for at least one* $i \in [0:m]$*. Then there is a* $\mathsf{y} = (y_0, \mathsf{u}) \in \mathbb{R} \times \mathbb{R}^m_+$ *such that* $u_j > 0$ *for all* $j \in [1:m]$ *and such that the matrix* $\mathsf{Z}(\mathsf{y}) = \mathsf{Z}_+(\mathsf{y}, \mathsf{o})$ *is strictly copositive.*

(b) *Suppose that there is an* $\widehat{\mathsf{x}} \in \mathbb{R}^n_+$ *such that* $\mathsf{A}\widehat{\mathsf{x}} = \mathsf{a}$ *and* $q_i(\widehat{\mathsf{x}}) < 0$ *for all* $i \in [1:m]$*. Then there is a matrix* X *in the interior of* $\mathcal{C}$ *such that* $\langle \mathsf{J}_0, \mathsf{X} \rangle = 1$ *and* $\langle \mathsf{M}_i, \mathsf{X} \rangle < 0$ *for all* $i \in [1:m]$*.*

(c) *Under the assumptions of (a) and (b), full strong duality for the primal-dual conic pair (27),(28) holds: both optimal values are attained at certain* $\mathsf{X}^* \in \mathcal{C}$ *and* $(\mathsf{y}^*, \mathsf{w}^*) \in \mathbb{R} \times \mathbb{R}^m_+ \times \mathbb{R}^p$*, and there is no duality gap:*

$$z^*_{CD} = y^*_0 = \langle \mathsf{M}_0, \mathsf{X}^* \rangle = z^*_{CP} \qquad \text{and} \qquad \langle \mathsf{X}^*, \mathsf{Z}_+(\mathsf{y}^*, \mathsf{w}^*) \rangle = 0\,.$$

Proof. (a) By assumption on Q_i, the bound $\sigma := \min \left\{ \mathsf{x}^\top \mathsf{Q}_i \mathsf{x} : \mathsf{x} \in \Delta \right\} > 0$. Further define

$$\alpha := \min \left\{ \sum_{j \neq i} \mathsf{x}^\top \mathsf{Q}_j \mathsf{x} : \mathsf{x} \in \Delta \right\} \in \mathbb{R}$$

and put $u_i = \max \left\{ 1, -\frac{2\alpha}{\sigma} \right\} > 0$. Then for all $\mathsf{x} \in \Delta$ we get by construction

$$\mathsf{x}^\top (u_i \mathsf{Q}_i + \sum_{j \neq i} \mathsf{Q}_j) \mathsf{x} \geq u_i \sigma + \alpha = \max \{ -\alpha, \sigma + \alpha \} > 0\,.$$

By positive homogeneity, we arrive at strict copositivity of the matrix $2\mathsf{H}_\mathsf{u} = \mathsf{Q}_0 + \sum_{j=1}^{m} u_i \mathsf{Q}_i$ by setting $u_j := 1 > 0$ for all $j \neq i$ if $i \geq 1$, and else $u_j := \frac{1}{u_0} > 0$ if $i = 0$. By Lemma 6.1(a) and $D^2 L_0(\mathsf{x}; \mathsf{u}) = \mathsf{H}_\mathsf{u}$, we infer that the slack matrix $\mathsf{Z}_+(\mathsf{y}, \mathsf{o}) = \mathsf{Z}(\mathsf{y}) = \mathsf{M}(L_0(\cdot; \mathsf{u})) + \bar{t} \mathsf{J}_0$ as defined in (22) is strictly copositive for $y_0 = \mathsf{c}^\top \mathsf{u} - \bar{t}$ if $\bar{t} > 0$ is large enough.
(b) Given $\widehat{\mathsf{x}}$ as in the assumption, select $\mathsf{x}^{(j)} \in P$ as in (15) for all $j \in [1:n]$, and define

$$\mathsf{x} := (1 - \varepsilon) \widehat{x} + \tfrac{\varepsilon}{n} \sum_{j=1}^{n} \mathsf{x}^{(j)}$$

where $\varepsilon > 0$ is chosen so small that still $q_i(\mathsf{x}) < 0$ holds for all i. This is possible by continuity of all q_i. Then $x_j > 0$ for all $j \in [1:n]$ by construction and also $\mathsf{x} \in F \cap P$. Next put $\mathsf{z} = [1, \mathsf{x}^\top]^\top$ and $\mathsf{X} = (1-\varepsilon)\mathsf{z}\mathsf{z}^\top + \varepsilon \mathsf{I}_{n+1}$. If necessary, decrease $\varepsilon > 0$ further such that still $\langle \mathsf{M}_i, \mathsf{X} \rangle < 0$ holds; again, this is possible by continuity and because

$$\langle \mathsf{M}_i, \mathsf{z}\mathsf{z}^\top \rangle = \mathsf{z}^\top \mathsf{M}_i \mathsf{z} = q_i(\mathsf{x}) < 0 \quad \text{for all } i\,.$$

Hence we can write $\mathsf{X} = [\mathsf{f}|\mathsf{B}][\mathsf{f}|\mathsf{B}]^\top$ where $\mathsf{f} = \sqrt{1-\varepsilon}\,\mathsf{z}$ has all coordinates strictly positive and $\mathsf{B} = \sqrt{\varepsilon}\,\mathsf{I}_{n+1}$ has full rank, and therefore X lies in the interior of $\mathcal{C}$ due to the improved characterization in [14]. Of course, $\langle \mathsf{J}_0, \mathsf{X} \rangle = 1$ by construction.
(c) follows from Slater's theorem for convex optimization. □

7 Tightness and second-order optimality conditions

When is the Semi-Lagrangian/copositive bound tight ?

A first answer is given by Theorem 5.1. But how is this reflected in terms of the original problem (14), i.e., of the (bordered) Hessian of the Lagrangian? Below, we will give an answer which also reveals a second-order condition sufficient for global optimality, which is weaker than the conditions derived from tightness of the Lagrangian relaxation. Note that neither F nor $F \cap P$ are, in general, convex, so strict feasibility would not imply the KKT conditions at a (local) solution, as Slater's theorem does not apply. However, tightness of the relaxations basically enforces the KKT conditions without any further constraint qualifications on (13) or on (14); in the latter case with the Semi-Lagrangian dual in a moderately generalized form though.

7.1 Recap: the full Lagrangian case

Let us shortly go back to the problem (13) without linear constraints. Consider again the conditions guaranteeing strong duality for its SDP relaxation, namely (a) at least one of the Q_i is (strictly) positive-definite; and (b) there is an $\bar{\mathsf{x}} \in \mathbb{R}^n$ such that $q_i(\bar{\mathsf{x}}) < 0$ for all i. Under these conditions, [1] proved that tightness of the semidefinite relaxation for problem (13), i.e. the equality $z^*_{SP} = z^*$, is equivalent to $\mathsf{Z}(q_0(\mathsf{x}^*), \mathsf{u}^*) \succeq \mathsf{O}$ for some $\mathsf{u}^* \in \mathbb{R}^m_+$ which satisfies the KKT conditions at a global solution x^* of (13).

We can say even more: if $(\bar{\mathsf{x}}, \bar{\mathsf{u}})$ is a KKT pair of (13) such that $\mathsf{H}_{\bar{\mathsf{u}}} \succeq \mathsf{O}$, then $\bar{\mathsf{x}}$ is a global solution to (13). In case of the trust region problem where $m = 1$ or a co-centered problem with two constraints where $m = 2$ and all $\mathsf{b}_i = \mathsf{o}$, also the converse is true, so that we have always $z^*_{SP} = z^*$ in these cases, or, equivalently, for any global solution x^* there is a multiplier $\mathsf{u}^* \in \mathbb{R}^m_+$ satisfying the KKT conditions such that $\mathsf{H}_{\mathsf{u}^*} \succeq \mathsf{O}$. However, for the CDT problem (inhomogeneous case of $m = 2$), H_u can be indefinite at the global optimum [3] for all KKT multipliers u at x^* (generically but not always u is unique), and then there is a positive gap, $z^*_{SP} < z^*_Q$. So the converse does not hold in general, not even for problem (13) without linear constraints.

With minimal effort, one can translate above results to the full Lagrangian dual of (14), and arrive at a similar sufficient global optimality condition: if at a KKT pair $(\bar{\mathsf{x}}; \bar{\mathsf{u}}, \bar{\mathsf{v}}, \bar{\mathsf{w}})$, the slack matrix $\mathsf{Z}_+(q_0(\bar{\mathsf{x}}), \bar{\mathsf{u}}, \bar{\mathsf{w}}) \in \mathcal{D}^\star_\diamond$, then $\bar{\mathsf{x}}$ is a global solution to (14). The next subsection will present a much stronger result.

7.2 Second-order optimality condition and Semi-Lagrangian tightness

Here, we go a step further and prove a counterpart of above findings for the Semi-Lagrangian relaxation of problem (14). As this is, again, not a straightforward generalization from positive-semidefiniteness to copositivity, we need to relax the KKT conditions, too: let us say that the pair $(\mathsf{x}; \mathsf{u}, \mathsf{w}) \in (F \cap P) \times \mathbb{R}^m_+ \times \mathbb{R}^p$ is a *generalized KKT pair* for (14) if and only if

$$\left.\begin{array}{rcll} x_j(\mathsf{H}_\mathsf{u}\mathsf{x} - \mathsf{d}_\mathsf{u} - \mathsf{A}^\top \mathsf{w})_j &=& 0 & \text{for all } j\in[1:n]\,, \\ u_i q_i(\mathsf{x}) &=& 0 & \text{for all } j\in[1:n] \text{ and} \\ w_k(a_k - \mathsf{z}_k^\top \mathsf{x}) &=& 0 & \text{for all } k\in[1:p]\,. \end{array}\right\} \quad (32)$$

Let $\mathsf{v} := \mathsf{H}_\mathsf{u}\mathsf{x} - \mathsf{d}_\mathsf{u} - \mathsf{A}^\top \mathsf{w}$; then (32) is equivalent to stipulating equation $\nabla L(\mathsf{x}; \mathsf{u}, \mathsf{v}, \mathsf{w}) = \mathsf{o}$ under the conditions $v_j x_j = 0$, $w_k(a_k - \mathsf{z}_k^\top \mathsf{x}) = 0$ and $u_i q_i(\mathsf{x}) = 0$ for all i, j, k, but without requiring $v_j \geq 0$ now.

Theorem 7.1 *Under the assumptions of Theorem 6.2(c), the following assertions are equivalent:*

(a) *The Semi-Lagrangian relaxation is tight,* $z^*_{\text{semi}} = z^*_+$;

(b) *for all global solutions* x^* *to (14), there is a* $(\mathsf{u}^*, \mathsf{w}^*) \in \mathbb{R}^m_+ \times \mathbb{R}^p$ *such that* $(\mathsf{x}^*; \mathsf{u}^*, \mathsf{w}^*)$ *is a generalized KKT pair and such that*

$$\mathsf{Z}_+(\mathsf{y}^*, \mathsf{w}^*) \in \mathcal{C}^\star \quad \textit{for} \quad \mathsf{y}^* = (q_0(\mathsf{x}^*), \mathsf{u}^*)\,;$$

(c) *there is a global solution* x^* *to (14) and a* $(\mathsf{u}^*, \mathsf{w}^*) \in \mathbb{R}^m_+ \times \mathbb{R}^p$ *such that* $(\mathsf{x}^*; \mathsf{u}^*, \mathsf{w}^*)$ *is a generalized KKT pair and such that*

$$\mathsf{Z}_+(\mathsf{y}^*, \mathsf{w}^*) \in \mathcal{C}^* \quad \textit{for} \quad \mathsf{y}^* = (q_0(\mathsf{x}^*), \mathsf{u}^*)\,;$$

(d) *there is a generalized KKT pair* $(\bar{\mathsf{x}}; \bar{\mathsf{u}}, \bar{\mathsf{w}}) \in (F \cap P) \times \mathbb{R}^m_+ \times \mathbb{R}^p$ *such that*

$$\mathsf{Z}_+(\bar{\mathsf{y}}, \bar{\mathsf{w}}) \in \mathcal{C}^\star \quad \textit{for} \quad \bar{\mathsf{y}} = (q_0(\bar{\mathsf{x}}), \bar{\mathsf{u}})\,.$$

Proof. Under the assumption, there exists an optimal solution x^* to (14) by Theorem 6.2(a). So only (a) $\Rightarrow$ (b) and (d) $\Rightarrow$ (a) need a proof. For the first implication, form again $\mathsf{X}^* = \mathsf{z}\mathsf{z}^\top \in \mathcal{C}$ with $\mathsf{z}^\top = [1, (\mathsf{x}^*)^\top] \in \mathbb{R}^{n+1}_+$. Then $\langle \mathsf{M}_i, \mathsf{X}^* \rangle = q_i(\mathsf{x}^*) \le 0$ for all $i \in [1\!:\!m]$ and $\langle \mathsf{J}_0, \mathsf{X}^* \rangle = 1$, so that X^* is feasible for (28). The (in)equality chain

$$z^*_+ = z^*_{\text{semi}} = z^*_{CD} = z^*_{CP} \le \langle \mathsf{M}_0, \mathsf{X}^* \rangle = q_0(\mathsf{x}^*) = z^*_+$$

establishes optimality of X^*. By strong duality due to Theorem 6.2(c), there is a dually optimal $(\mathsf{y}^*, \mathsf{w}^*) = (\mathsf{y}^*_0, \mathsf{u}^*, \mathsf{w}^*) \in \mathbb{R} \times \mathbb{R}^m_+ \times \mathbb{R}^p$ such that $\mathsf{Z}_+(\mathsf{y}^*, \mathsf{w}^*) \in \mathcal{C}^\star$ and $\langle \mathsf{Z}_+(\mathsf{y}^*, \mathsf{w}^*), \mathsf{X}^* \rangle = 0$. This complementary slackness implies, at first, that

$$\left.\begin{array}{rcccl} u^*_i q_i(\mathsf{x}^*) &=& u^*_i \langle \mathsf{M}_i, \mathsf{X}^* \rangle &=& 0 \quad \text{for all } i \in [1\!:\!m] \quad \text{and} \\ w^*_k (a_k - \mathsf{z}_k^\top \mathsf{x}^*) &=& w^*_k \langle \mathsf{A}_k, \mathsf{X}^* \rangle &=& 0 \quad \text{for all } k \in [1\!:\!p]\,. \end{array}\right\} \tag{33}$$

In particular, we get $(\mathsf{a} - \mathsf{A}\mathsf{x}^*)^\top \mathsf{w}^* = \sum_{k=1}^{p} w^*_k (a_k - \mathsf{z}_k^\top \mathsf{x}^*) = 0$, so that

$$\mathsf{Z}_+(\mathsf{y}^*, \mathsf{w}^*)\mathsf{X}^* = \begin{bmatrix} \mathsf{c}^\top \mathsf{u}^* - y^*_0 - \mathsf{d}_{\mathsf{u}^*}^\top \mathsf{x}^* + \mathsf{a}^\top \mathsf{w}^* & [\mathsf{c}^\top \mathsf{u}^* - y^*_0 - \mathsf{d}_{\mathsf{u}^*}^\top \mathsf{x}^*](\mathsf{x}^*)^\top \\ \mathsf{H}_{\mathsf{u}^*}\mathsf{x}^* - \mathsf{d}_{\mathsf{u}^*} - \mathsf{A}^\top \mathsf{w}^* & [\mathsf{H}_{\mathsf{u}^*}\mathsf{x}^* - \mathsf{d}_{\mathsf{u}^*} - \mathsf{A}^\top \mathsf{w}^*](\mathsf{x}^*)^\top \end{bmatrix}. \tag{34}$$

But by [25, Thm.2.1(a)] we know that $\langle \mathsf{Z}_+(\mathsf{y}^*, \mathsf{w}^*), \mathsf{X}^* \rangle = 0$ also implies $\operatorname{diag}(\mathsf{Z}_+(\mathsf{y}^*, \mathsf{w}^*)\mathsf{X}^*) = \mathsf{o}$, since $\mathsf{X}^* \in \mathcal{C}$ and $\mathsf{Z}_+(\mathsf{y}^*, \mathsf{w}^*) \in \mathcal{C}^\star$, so we infer $y^*_0 = \mathsf{c}^\top \mathsf{u}^* - \mathsf{d}_{\mathsf{u}^*}^\top \mathsf{x}^* + \mathsf{a}^\top \mathsf{w}^*$ and

$$x^*_j (\mathsf{H}_{\mathsf{u}^*}\mathsf{x}^* - \mathsf{d}_{\mathsf{u}^*} - \mathsf{A}^\top \mathsf{w}^*)_j = 0 \quad \text{for all } j \in [1\!:\!n]\,. \tag{35}$$

(note that [25, Thm.2.1(b)] says that the j-th row of $\mathsf{Z}_+(\mathsf{y}^*, \mathsf{w}^*)\mathsf{X}^*$ vanishes if $j = 0$ or else $x^*_j > 0$, which, by (34), exactly amounts to the same). Hence $(\mathsf{x}^*; \mathsf{u}^*, \mathsf{w}^*) \in (F \cap P) \times \mathbb{R}^m_+ \times \mathbb{R}^p$ form a generalized KKT pair for (14). Now (35) also implies $(\mathsf{x}^*)^\top \mathsf{H}_{\mathsf{u}^*}\mathsf{x}^* = (\mathsf{d}_{\mathsf{u}^*} + \mathsf{A}^\top \mathsf{w}^*)^\top \mathsf{x}^*$ and therefore

$$\begin{array}{rcl} y^*_0 &=& \mathsf{c}^\top \mathsf{u}^* - \mathsf{d}_{\mathsf{u}^*}^\top \mathsf{x}^* + \mathsf{a}^\top \mathsf{w}^* \\ &=& \mathsf{c}^\top \mathsf{u}^* + \mathsf{a}^\top \mathsf{w}^* - \mathsf{d}_{\mathsf{u}^*}^\top \mathsf{x}^* + (\mathsf{a} - \mathsf{A}\mathsf{x}^*)^\top \mathsf{w}^* \\ &=& \mathsf{c}^\top \mathsf{u}^* + 2\mathsf{a}^\top \mathsf{w}^* - (\mathsf{d}_{\mathsf{u}^*} + \mathsf{A}^\top \mathsf{w}^*)^\top \mathsf{x}^* \\ &=& \mathsf{c}^\top \mathsf{u}^* + 2\mathsf{a}^\top \mathsf{w}^* - 2(\mathsf{d}_{\mathsf{u}^*} + \mathsf{A}^\top \mathsf{w}^*)^\top \mathsf{x}^* + (\mathsf{x}^*)^\top \mathsf{H}_{\mathsf{u}^*}\mathsf{x}^* \\ &=& L(\mathsf{x}^*; \mathsf{u}^*, \mathsf{o}, \mathsf{w}^*) = q_0(\mathsf{x}^*) \end{array}$$

by (33), and assertion (b) is established.
To show (d) $\Rightarrow$ (a), put $\bar{y}_0 := q_0(\bar{\mathsf{x}})$, $\bar{\mathsf{y}} = [\bar{y}_0, \bar{\mathsf{u}}^\top]$, and $\bar{\mathsf{z}}^\top = [1, \bar{\mathsf{x}}^\top]$ as well as $\bar{\mathsf{X}} = \bar{\mathsf{z}}\bar{\mathsf{z}}^\top \in \mathcal{C}$. By (35), we infer again $(\mathsf{d}_{\bar{\mathsf{u}}} + \mathsf{A}^\top \bar{\mathsf{w}})^\top \bar{\mathsf{x}} = \bar{\mathsf{x}}^\top \mathsf{H}_{\bar{\mathsf{u}}} \bar{\mathsf{x}}$, so that

$$\begin{aligned} \bar{y}_0 &= q_0(\bar{\mathsf{x}}) + \sum_{i=1} \bar{u}_i q_i(\bar{\mathsf{x}}) + 2\sum_{k=1}^p \bar{w}_k (a_k - \mathsf{z}_k^\top \bar{\mathsf{x}}) \\ &= \mathsf{c}^\top \bar{\mathsf{u}} + 2\mathsf{a}^\top \bar{\mathsf{w}} - 2(\mathsf{d}_{\bar{\mathsf{u}}} + \mathsf{A}^\top \bar{\mathsf{w}})^\top \bar{\mathsf{x}} + \bar{\mathsf{x}}^\top \mathsf{H}_{\bar{\mathsf{u}}} \bar{\mathsf{x}}\,, \end{aligned}$$

and therefore

$$0 = (\mathsf{c}^\top \bar{\mathsf{u}} - \bar{y}_0 + 2\mathsf{a}^\top \bar{\mathsf{w}}) - 2(\mathsf{d}_{\bar{\mathsf{u}}} + \mathsf{A}^\top \bar{\mathsf{w}})^\top \bar{\mathsf{x}} + \bar{\mathsf{x}}^\top \mathsf{H}_{\bar{\mathsf{u}}} \bar{\mathsf{x}} = \bar{\mathsf{z}}^\top \mathsf{Z}_+(\bar{\mathsf{y}}, \bar{\mathsf{w}}) \bar{\mathsf{z}}\,.$$

Hence $\langle \bar{\mathsf{X}}, \mathsf{Z}_+(\bar{\mathsf{y}}, \bar{\mathsf{w}}) \rangle = \bar{\mathsf{z}}^\top \mathsf{Z}_+(\bar{\mathsf{y}}, \bar{\mathsf{w}}) \bar{\mathsf{z}} = 0$, so that $(\bar{\mathsf{X}}, \mathsf{Z}_+(\bar{\mathsf{y}}, \bar{\mathsf{w}}))$ form an optimal primal-dual pair for the copositive problem (27) and (28) with zero duality gap. We conclude

$$z_+^* \le q_0(\bar{\mathsf{x}}) = \bar{y}_0 = z_{CD}^* = z_{CP}^* = z_{\text{semi}}^* \le z_+^*$$

which shows tightness of the Semi-Lagrangian relaxation. □

In fact, we have obtained the following sufficient second-order global optimality condition; for the role of copositivity in second-order optimality conditions for general smooth optimization problems, refer to [5].

Corollary 7.1 *Let* $(\bar{\mathsf{x}}; \bar{\mathsf{u}}, \bar{\mathsf{w}}) \in (F \cap P) \times \mathbb{R}_+^m \times \mathbb{R}^p$ *be a generalized KKT pair for (14). If the matrix*

$$\begin{bmatrix} \mathsf{c}^\top \bar{\mathsf{u}} - q_0(\bar{\mathsf{x}}) + 2\mathsf{a}^\top \bar{\mathsf{w}} & -(\mathsf{d}_{\bar{\mathsf{u}}} + \mathsf{A}^\top \bar{\mathsf{w}})^\top \\ -(\mathsf{d}_{\bar{\mathsf{u}}} + \mathsf{A}^\top \bar{\mathsf{w}}) & \mathsf{H}_{\bar{\mathsf{u}}} \end{bmatrix} \tag{36}$$

is copositive, then $\bar{\mathsf{x}}$ *is a global solution to (14).*

Proof. Observe that in the proof of $(d) \Rightarrow (a)$ of Theorem 7.1 above, we never used one of the conditions in Theorem 6.2. So regardless of these, global optimality of $\bar{\mathsf{x}}$ holds, along with tightness and zero duality gap, $z_{\text{semi}}^* = z_+^* = z_{CP}^* = z_{CD}^* = q_0(\bar{\mathsf{x}})$. □

Problem (14) may have many (generalized) KKT points $\bar{\mathsf{x}}$, some of which can be detected with not too much effort by local optimization procedures; cf. [28]. Next, we may solve the linear equations for $(\bar{\mathsf{u}}, \bar{\mathsf{w}})$, and then test a sufficient copositivity criterion for the matrix in (36), to get a certificate for global optimality of $\bar{\mathsf{x}}$. The condition is weaker than that addressed at the end of Subsection 7.1 in two aspects: it deals with *generalized* KKT pairs, and it requires only $Z_+(\bar{\mathsf{y}}, \bar{\mathsf{w}}) \in \mathcal{C}^\star$ rather than $Z_+(\bar{\mathsf{y}}, \bar{\mathsf{w}}) \in \mathcal{D}_\diamond^\star$. Recall that the sub-zero level approximation cone $\mathcal{D}_\diamond^\star$ is much smaller than $\mathcal{C}^\star$.

The difference can also be expressed in properties of the Hessian $\mathsf{H}_{\bar{\mathsf{u}}}$ of the Lagrangian: indeed, the condition $Z_+(\bar{\mathsf{y}}, \bar{\mathsf{w}}) \in \mathcal{D}_\diamond^\star$ implies that its lower right principal submatrix $\mathsf{H}_{\bar{\mathsf{u}}}$ has to be psd, and we know this is too strong in some cases, whereas $Z_+(\bar{\mathsf{y}}, \bar{\mathsf{w}}) \in \mathcal{C}^\star$, by the same argument, only yields copositivity of $\mathsf{H}_{\bar{\mathsf{u}}}$.

Violation of the assumption in Theorem 6.2(b) will play a role in the following subsection.

7.3 Replacing all linear constraints by one quadratic

Finally, let us deal with the approach to replace the p linear constraints $\mathsf{Ax} = \mathsf{a}$ by one quadratic constraint $q_{m+1}(\mathsf{x}) = \|\mathsf{Ax} - \mathsf{a}\|^2$. Of course, we cannot expect full strong duality for the original copositive formulation (27), and neither for the more accurate version, namely the copositive representation of the Semi-Lagrangian dual of this alternative:

$$\left.\begin{array}{l} z^*_{CP,\mathrm{lin}} := \inf\left\{\langle \mathsf{M}_0, \mathsf{X}\rangle : \langle \mathsf{M}_i, \mathsf{X}\rangle \le 0, \langle \mathsf{J}_0, \mathsf{X}\rangle = 1, \langle \mathsf{M}_{m+1}, \mathsf{X}\rangle = 0, \mathsf{X} \in \mathcal{C}\right\} \\ z^*_{CD,\mathrm{lin}} := \sup\left\{y_0 : \mathsf{Z}(\mathsf{y}) \in \mathcal{C}^\star, \mathsf{y} = [y_0, \mathsf{u}^\top, u_{m+1}]^\top \in \mathbb{R} \times \mathbb{R}^m_+ \times \mathbb{R}\right\}\,, \end{array}\right\} \quad (37)$$

Obviously, we have

$$z^*_{LD,+} \le z^*_{CD,\mathrm{lin}} \le z^*_{CP,\mathrm{lin}} \le z^*_+\,.$$

Theorem 7.2 *Consider the case $q_{m+1}(\mathsf{x}) = \|\mathsf{Ax} - \mathsf{a}\|^2$. Suppose that at least one Q_i is strictly copositive for $i \in [0 : m+1]$ (note that $\mathsf{Q}_{m+1} = \mathsf{A}^\top \mathsf{A}$ is so if and only if ker $\mathsf{A} \cap \mathbb{R}^n_+ = \{\mathsf{o}\}$). Then both (28) and (37) have zero duality gap and the primal optimal value is attained if there is an $\bar{\mathsf{x}} \in F \cap P$: for some $\mathsf{X}^* \in \mathcal{C}$ such that $\langle \mathsf{M}_i, \mathsf{X}^*\rangle \le 0$ for all $i \in [1 : m]$ as well as $\langle \mathsf{J}_0, \mathsf{X}^*\rangle = 1$ and $\langle \mathsf{M}_{m+1}, \mathsf{X}^*\rangle = 0$, we have*

$$z^*_{CD,\mathrm{lin}} = z^*_{CP,\mathrm{lin}} = \langle \mathsf{M}_0, \mathsf{X}^*\rangle\,.$$

Proof. We construct a strictly feasible $Z(\mathsf{y})$ as in Theorem 6.2(a) and infer the result from Slater's principle. Indeed, the primal problem in (37) is feasible as $\mathsf{X} = \mathsf{zz}^\top$ with $\mathsf{z}^\top = [1, \bar{\mathsf{x}}^\top]$ satisfies all constraints. □

We have seen in Subsection 5.2 above that Burer's relaxation (30) is tighter, but for large instances this problem may have too many constraints.

So one may replace both constraints $\langle \mathsf{A}_k, \mathsf{X}\rangle = 0$ and $\langle \mathsf{B}_k, \mathsf{X}\rangle = 0$ by a linear combination of them, e.g., by looking at

$$\mathsf{C}_k := a_k \mathsf{A}_k + \mathsf{B}_k = \begin{bmatrix} a_k^2 & -a_k \mathsf{r}_k^\top \\ -a_k \mathsf{r}_k & \mathsf{r}_k \mathsf{r}_k^\top \end{bmatrix} = [a_k, \mathsf{r}_k^\top]^\top [a_k, \mathsf{r}_k^\top] \succeq \mathsf{O}\,.$$

Now all C_k are psd., so for any $\mathsf{X} \in \mathcal{D}_\diamond^\star$, the conditions $\langle \mathsf{C}_k, \mathsf{X}\rangle = 0$ for all $k\in[1:p]$ are equivalent to

$$\sum_{k=1}^{p} \langle \mathsf{C}_k, \mathsf{X}\rangle = 0\,,$$

i.e., to a single homogeneous linear constraint. Now

$$\sum_{k=1}^{p} \mathsf{C}_k = \begin{bmatrix} \mathsf{a}^\top \mathsf{a} & -\mathsf{a}^\top \mathsf{A} \\ -\mathsf{A}^\top \mathsf{a} & \mathsf{A}^\top \mathsf{A} \end{bmatrix} = \mathsf{M}_{m+1}\,,$$

so we exactly arrive again at $z^*_{CD,\text{lin}}$ or $z^*_{CP,\text{lin}}$ this way. Interestingly, the idea to aggregate constraints in copositive optimization formulations recently emerged almost simultaneously and independently by the different approaches in [2, 16].

8 Conclusion and outlook

This paper deals with problems to optimize a quadratic function subject to quadratic and linear constraints, where the linear ones are treated separately. By relaxing everything except the sign constraints, we arrive at a Semi-Lagrangian dual which apparently has not been analyzed before in the literature. Here we have reformulated both the Lagrangian dual and the Semi-Lagrangian dual as conic optimization problems. While the latter is a copositive problem, the former can be seen as a natural relaxation of the latter, namely arising from an approximation of the copositive problem at a sub-zero level. This low level is important in regimes where every additional linear inequality constraint severely slows down algorithmic performance, which is typical in very large problems.

The development lead us to propose us a new approximation hierarchy which may avoid above drawbacks, so that a significant tightening of the bounds becomes tractable. Furthermore, we studied properties of the problem which ensure strong duality of the conic relaxations; specified necessary and sufficient copositivity-based conditions to guarantee that the Semi-Lagrangian relaxation is exact; and proposed a hierarchy of seemingly new, sufficient, second-order global optimality conditions for a KKT point of the original problem which can be tested in polynomial time, demanding much less than the familiar ones which require positive-semidefiniteness of the Hessian of the Lagrangian.

Burer's famous copositive relaxation is shown to be even tighter than the Semi-Lagrangian relaxation; however, this alternative may require too much effort in large instances. We also discuss the alternative to replace all linear equality constraints by a single convex quadratic one.

Building upon this findings, there are several directions of future research, among them:

- to tighten other variants of SDP formulations of the full Lagrangian relaxation [19], and to interpret them in terms of properties of the Lagrangian function of the original problem (in some formulation);
- to define a strategy which balances computational effort identifying and using additional linear constraints (i.e., other than those defining $\mathcal{D}_\diamond$), against efficient strengthening of the resulting bounds;
- to clarify the relation between z^*_{CD} and $z^*_{CD,\mathrm{lin}}$;
- to explore the quality of the relaxation if the A_k constraints are simply replaced by the B_k constraints, and to relate the result with above dual bounds.

Acknowledgement. The author is indebted to Peter J.C. Dickinson and Luis Zuluaga for valuable comments and stimulating discussions on an earlier draft of this paper. Thanks are also due to the Isaac Newton Institute at Cambridge University for providing a stimulating environment when the author participated as a visiting fellow in the Polynomial Optimization Programme 2013, organized by Joerg Fliege, Jean-Bernard Lasserre, Adam Letchford and Markus Schweighofer.

References

[1] Wenbo Ai and Shuzhong Zhang. Strong duality for the CDT subproblem: a necessary and sufficient condition. *SIAM J. Optim.*, 19(4):1735–1756, 2009.

[2] Naohiko Arima, Sunyoung Kim, and Masakazu Kojima. Simplified copositive and Lagrangian relaxations for linearly constrained quadratic optimization problems in continuous and binary variables. Research Report B-469, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, `http://www.optimization-online.org/DB\_HTML/2012/10/3647.html`, 2012.

[3] Amir Beck and Yonina Eldar. Strong duality in nonconvex quadratic optimization with two quadratic constraints. *SIAM J. Optim.*, 17(3):844–860, 2006.

[4] Immanuel M. Bomze. Copositive optimization – recent developments and applications. *European J. Oper. Res.*, 216:509–520, 2012.

[5] Immanuel M. Bomze. Copositivity for second-order optimality conditions in general smooth optimization problems. Preprint NI13033-POP, Isaac Newton Institute, Cambridge UK, 2013.

[6] Immanuel M. Bomze, Mirjam Dür, and Chung-Piaw Teo. Copositive optimization. *Optima – MOS Newsletter*, 89:2–10, 2012.

[7] Immanuel M. Bomze and Florian Jarre. A note on Burer's copositive representation of mixed-binary QPs. *Optim. Lett.*, 4(3):465–472, 2010.

[8] Stefan Bundfuss and Mirjam Dür. Algorithmic copositivity detection by simplicial partition. *Linear Algebra Appl.*, 428(7):1511–1523, 2008.

[9] Stefan Bundfuss and Mirjam Dür. An adaptive linear approximation algorithm for copositive programs. *SIAM J. Optim.*, 20(1):30–53, 2009.

[10] Samuel Burer. On the copositive representation of binary and continuous nonconvex quadratic programs. *Math. Program.*, 120(2, Ser. A):479–495, 2009.

[11] Samuel Burer. Copositive programming. In Miguel F. Anjos and Jean Bernard Lasserre, editors, *Handbook of Semidefinite, Cone and Polynomial Optimization: Theory, Algorithms, Software and Applications*, International Series in Operations Research and Management Science, pages 201–218. Springer, New York, 2012.

[12] Samuel Burer, Sunyoung Kim, and Masakazu Kojima. Faster, but weaker, relaxations for quadratically constrained quadratic programs. *Comput. Optim. Appl.*, to appear. DOI:10.1007/s10589–013–9618–8, 2013.

[13] Jieqiu Chen and Samuel Burer. Globally solving nonconvex quadratic programming problems via completely positive programming. *Math. Program. Comput.*, 4(1):33–52, 2012.

[14] Peter J. C. Dickinson. An improved characterisation of the interior of the completely positive cone. *Electron. J. Linear Algebra*, 20:723–729, 2010.

[15] Peter J. C. Dickinson and Luuk Gijben. On the computational complexity of membership problems for the completely positive cone and its dual. *Comput. Optim. Appl.*, to appear. DOI:10.1007/s10589–013–9594–z, 2013.

[16] Peter J.C. Dickinson. *The copositive cone, the completely positive cone and their generalisations.* Ph.D. dissertation, University of Groningen, The Netherlands, 2013.

[17] Mirjam Dür. Copositive programming — a survey. In Moritz Diehl, Francois Glineur, Elias Jarlebring, and Wim Michiels, editors, *Recent Advances in Optimization and its Applications in Engineering*, pages 3–20. Springer, Berlin Heidelberg New York, 2010.

[18] Alexander Engau, Miguel F. Anjos, and Immanuel M. Bomze. Constraint selection in a build-up interior-point cutting-plane method for solving relaxations of the stable-set problem. *Math. Methods of OR*, 78(1):35–59, 2013.

[19] Alain Faye and Frédéric Roupin. Partial Lagrangian relaxation for general quadratic programming. *4OR*, 5:75–88, 2007.

[20] Sunyoung Kim, Masakazu Kojima, and Kim-Chuan Toh. A Lagrangian-DNN relaxation: a fast method for computing tight lower bounds for a class of quadratic optimization problems. Preprint, `http://www.optimization-online.org/DB\_HTML/2013/10/4073.html`, 2013.

[21] Cheng Lu, Shu-Cherng Fang, Qingwei Jin, Zhengbo Wang, and Wenxun Xing. KKT solution and conic relaxation for solving quadratically constrained quadratic programming problems. *SIAM J. Optim.*, 21(4):1475–1490, 2011.

[22] Jérome Malick and Frédéric Roupin. On the bridge between combinatorial optimization and nonlinear optimization: a family of semidefinite bounds for 01 quadratic problems leading to quasi-Newton methods. *Math. Program.*, 140(Ser. B):99–124, 2013.

[23] Katta G. Murty and Santosh N. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Math. Program.*, 39(2):117–129, 1987.

[24] Yuri Nesterov, Henry Wolkowicz, and Yinyu Ye. Nonconvex quadratic optimization. In Henry Wolkowicz, Romesh Saigal, and Lieven Vandenberghe, editors, *Handbook of Semidefinite Programming*, pages 361–420. Kluwer, New York, 2000.

[25] Naomi Shaked-Monderer, Abraham Berman, Immanuel M. Bomze, Florian Jarre, and Werner Schachinger. New results on the cp rank and related properties of co(mpletely)positive matrices. Preprint, `http://arxiv.org/abs/1305.0737`, 2013.

[26] Naomi Shaked-Monderer, Immanuel M. Bomze, Florian Jarre, and Werner Schachinger. On the cp-rank and minimal cp factorizations of a completely positive matrix. *SIAM J. Matrix Anal. Appl.*, 34(2):355–368, 2013.

[27] Naum Z. Shor. Quadratic optimization problems. *Izv. Akad. Nauk SSSR Tekhn. Kibernet.*, 1:128–139, 222, 1987.

[28] Yinyu Ye. On the complexity of approximating a KKT point of quadratic programming. *Math. Program.*, 80(Ser. A):195–211, 1998.