EURO Journal on Computational Optimization 9 (2021) 100021

Contents lists available at ScienceDirect

EURO Journal on Computational Optimization

journal homepage: www.elsevier.com/locate/ejco

Conic optimization: A survey with special focus on copositive optimization and binary quadratic problems

Mirjam Dür[a,*], Franz Rendl[b]

[a] Department of Mathematics, Augsburg University, Augsburg 86135, Germany
[b] Department of Mathematics, Alpen-Adria-University Klagenfurt, Klagenfurt 9020, Austria

ARTICLE INFO

MSC:
90C22
90C20
90C27

Keywords:
Semidefinite programming
Second order cone programming
Copositive optimization
Binary quadratic optimization

ABSTRACT

A conic optimization problem is a problem involving a constraint that the optimization variable be in some closed convex cone. Prominent examples are linear programs (LP), second order cone programs (SOCP), semidefinite problems (SDP), and copositive problems. We survey recent progress made in this area. In particular, we highlight the connections between nonconvex quadratic problems, binary quadratic problems, and copositive optimization. We review how tight bounds can be obtained by relaxing the copositivity constraint to semidefiniteness, and we discuss the effect that different modelling techniques have on the quality of the bounds. We also provide some new techniques for lifting linear constraints and show how these can be used for stable set and coloring relaxations.

1. Introduction

A conic optimization problem is a problem involving a constraint that the optimization variable be in some closed convex cone. The field of conic optimization is a broad one, as any convex optimization problem can be cast as a conic problem, see Nesterov and Nemirovski (1992). In this paper, we will focus on more specific conic problems which appear naturally when solving quadratic or combinatorial optimization problems. In particular, we will highlight developments in second order cone programming (SOCP), semidefinite programming (SDP), and copositive optimization.

1.1. The general linear conic problem and its dual

Consider a proper cone $\mathcal{K}$, i.e., a closed convex and full dimensional cone which is also pointed, meaning that $\mathcal{K}$ does not contain a straight line, or equivalently, that $\mathcal{K} \cap (-\mathcal{K}) = \{0\}$. For example, when considering linear optimization problems, the cone involved is the nonnegative orthant $\mathcal{K} = \mathbb{R}^n_+ := \{x \in \mathbb{R}^n \mid x \geq 0\}$. In this paper, we will mostly consider optimization problems in matrix variables, in which case $\mathcal{K}$ is a proper cone in the set S_n of symmetric matrices. For this reason, we will use capital letters for the data matrices and variables, and we say that a linear conic optimization problem over $\mathcal{K}$ is a problem of the form

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

where C, X, A_i are matrices (or vectors) of suitable dimension, and $b_i \in \mathbb{R}$ for all $i = 1, \dots, m$. In case of matrices, $\langle \cdot, \cdot \rangle$ denotes the Frobenius inner product $\langle A, B\rangle := \text{trace}(A^T B)$, in case of vectors, it denotes the Euclidean inner product. Problem (P) therefore aims to minimize a linear function over the intersection of a proper cone and an affine subspace.

As in linear programming, a primal problem of the form (P) always comes with a dual problem which involves the dual cone: given an arbitrary cone $\mathcal{K} \subseteq \mathbb{R}^{m\times n}$, the dual cone $\mathcal{K}^*$ is defined as

$$\mathcal{K}^* := \{X \in \mathbb{R}^{m\times n} \mid \langle X, K\rangle \geq 0 \text{ for all } K \in \mathcal{K}\}.$$

As usual, the Lagrangian function $L : \mathcal{K} \times \mathbb{R}^m \to \mathbb{R}$ is defined as

$$L(X, y) := \langle C, X\rangle + \sum_{i=1}^m y_i(b_i - \langle A_i, X\rangle).$$

This gives the dual problem

$$\max_{y\in\mathbb{R}^m} \min_{X\in\mathcal{K}} L(X, y) = \max_{y\in\mathbb{R}^m} \left[\langle b, y\rangle + \min_{X\in\mathcal{K}} \langle C - \sum_{i=1}^m y_i A_i, X\rangle\right].$$

For the inner minimization problem to be finite, we require that $\langle C - \sum_{i=1}^m y_i A_i, X\rangle \geq 0$ for all $X \in \mathcal{K}$, in other words, we require $C - \sum_{i=1}^m y_i A_i \in \mathcal{K}^*$. Therefore, we arrive at the dual problem

$$\begin{array}{lll} d^* = & \max & \langle b, y\rangle \\ & \text{s.t.} & \sum_{i=1}^m y_i A_i + Z = C \\ & & Z \in \mathcal{K}^*, y \in \mathbb{R}^m. \end{array} \tag{D}$$

* Corresponding author.
E-mail addresses: mirjam.duer@math.uni-augsburg.de (M. Dür), franz.rendl@aau.at (F. Rendl).

https://doi.org/10.1016/j.ejco.2021.100021
Received 14 June 2021; Received in revised form 10 November 2021; Accepted 14 November 2021

It is easy to see that the duality gap $\langle C, X\rangle - \langle b, y\rangle$ equals the inner product of the primal and dual variables: $\langle C, X\rangle - \langle b, y\rangle = \langle Z, X\rangle$. Since $\langle Z, X\rangle \geq 0$ for any pair of primal/dual feasible points $X \in \mathcal{K}, Z \in \mathcal{K}^*$, we immediately get weak duality.

Clearly, if the duality gap is zero for a pair of primal/dual feasible points $X \in \mathcal{K}$ and $(y, Z) \in \mathbb{R}^m \times \mathcal{K}^*$, then X is optimal for (P) and (y, Z) is optimal for (D). The converse is, however, not true in general: a positive duality gap may exist, or the optimal value of (P) or (D) may not be attained. Examples for this phenomenon in second order cone programming can be found in Alizadeh and Goldfarb (2003) or in Ben-Tal and Nemirovski (2001, Section 2.4.1). For the SDP case, examples can be found in Helmberg (2002), and a thorough analysis of this behavior can be found in Pataki (2019).

In order to get strong duality, we need constraint qualifications:

Definition 1.1. A point X is called strictly feasible for (P) if X is feasible for (P) and $X \in \text{int}\mathcal{K}$. A pair (y, Z) is called strictly feasible for (D) if (y, Z) is feasible for (D) and $Z \in \text{int}\mathcal{K}^*$. If such points exist, then we say that the problem fulfills the primal (resp. dual) Slater condition.

Note that strict feasibility can always be enforced by considering the so called skew-symmetric embedding of the original problem, see de Klerk et al. (1997). Assuming strict feasibility gives us strong duality:

Theorem 1.2 (Strong Duality Theorem). *Assume that problem (D) has a strictly feasible solution (y, Z). Then the primal and dual optimal values are equal: $p^* = d^*$, and if $p^* < +\infty$, then p^* is attained, i.e., there exists a primal feasible solution X^* with $p^* = \langle C, X^*\rangle$.*

Conversely, assume that problem (P) has a strictly feasible solution X. Then the primal and dual optimal values are equal: $p^* = d^*$, and if $d^* > -\infty$, then d^* is attained, i.e., there exists a dual feasible solution (y^*, Z^*) with $d^* = \langle b, y^*\rangle$.

A proof of this theorem along with a thorough discussion of conic duality can be found for example in Ben-Tal and Nemirovski (2001). It has been shown in Dür et al. (2017) that the Slater condition (and hence strong duality) is a generic property of conic problems which loosely speaking means that Slater's condition is fulfilled (and hence strong duality holds) for almost all feasible conic problems which are parameterized in the space of data $(C, A_1, ..., A_m, b_1, ..., b_m)$. We stress that a constraint qualification is unnecessary if the cone $\mathcal{K}$ is polyhedral like in linear programming, where $\mathcal{K} = \mathbb{R}^n_+$. The reason why a positive duality gap may occur in general conic programming lies in the geometry of the problem and happens if the feasible set is contained in a nontrivial face of the cone.

Note that existence of strictly feasible points is important not only for theoretical purposes to ensure strong duality, but also many optimization algorithms require this property. In the absence of strictly feasible points, a solver may not terminate or may produce a "solution" with no useful meaning. This is a feature that distinguishes general conic optimization from linear programming. Consequently, very careful modelling is needed, since often existence of strictly feasible points can be guaranteed if the problem is modelled in a proper way. We will return to this point in Section 5.

Two approaches have been developed to tackle conic problems that fail to fulfill a constraint qualification: (i) Facial reduction attempts to identify the so called minimal cone $\mathcal{F}_{\min}$ for problem (P), such that problem (P) with $\mathcal{K}$ replaced by $\mathcal{F}_{\min}$ is strictly feasible and has the same optimal solution as (P). This facial reduction technique goes back to Borwein and Wolkowicz (1981a,b). (ii) Other approaches (e.g. Ramana, 1997) work on the dual side and construct an extended dual which achieves strong duality without assuming a constraint qualification. A good exposition of these two approaches can be found in Pataki (2013).

1.2. LP, SOCP, and SDP

Depending on which cone $\mathcal{K}$ is considered, conic optimization includes various classes of problems: If $\mathcal{K} = \mathbb{R}^n_+$, then (P) is a linear problem, a well studied class which appears in numerous applications. LPs are used to model not only straightforward linear constraints, but also constraints involving ℓ_1- or ℓ_∞- norms or absolute values.

If the cone $\mathcal{K}$ in (P) is the second order cone, then (P) is called a second order cone problem (SOCP). The second order cone in $\mathbb{R}^n$ (sometimes also called Lorentz cone or ice cream cone) is defined as $\mathcal{L}_n := \{(x_0, x) \in \mathbb{R} \times \mathbb{R}^{n-1} \mid x_0 \geq \|x\|_2\}$. It appears in optimization problems involving Euclidean norms: for example, the constraint $\|Ax + b\|_2 \leq c^T x + d$ can be written as $(c^T x + d, Ax + b) \in \mathcal{L}_{n+1}$. This is often used in robust optimization when an ellipsoidal uncertainty set is considered (Ben-Tal et al., 2009). Other applications of SOCP can be found in Alizadeh and Goldfarb (2003), Lobo et al. (1998). Certain risk measures in stochastic optimization may also lead to optimization problems over the so called p-order cone $\mathcal{L}^p_n := \{(x_0, x) \in \mathbb{R} \times \mathbb{R}^{n-1} \mid x_0 \geq \|x\|_p\}$ with $p \geq 1$, see e.g. Vinel and Krokhmal (2014). Any p-order cone with p rational can be modeled as SOCP. We refer to Mosek (2021-10-11) for details.

A third prominent setting is semidefinite programming (SDP), where $\mathcal{K}$ is considered to be the cone of symmetric positive semidefinite matrices $S^+_n := \{X \in \mathbb{R}^{n\times n} \mid X = X^T, X \succeq 0\}$. SDPs are used to model problems with linear matrix inequalities. They appear in eigenvalue optimization and control theory, see Vandenberghe and Boyd (1996), Helmberg (2002). Arguably two of the most important areas of application for SDP are robust and combinatorial optimization. For an in depth discussion of SDP in robust optimization, we refer to the book (Ben-Tal et al., 2009) and the recent survey paper (Yanıkoğlu et al., 2019). The role of SDP in relaxations of combinatorial problems will be covered in more detail below.

The cones $\mathbb{R}^n_+$, $\mathcal{L}_n$, and S^+_n are self-dual, whereas the dual of $\mathcal{L}^p_n$ is $\mathcal{L}^q_n$ with q such that $\frac{1}{p} + \frac{1}{q} = 1$. We mention that $\mathbb{R}^n_+$, $\mathcal{L}_n$, and S^+_n are instances of so called symmetric cones that can be studied in the unifying framework of Euclidean Jordan algebras, see Faybusovich (2008) and references therein.

These three problem classes have been studied for decades because of their countless applications and because they can be solved efficiently: it has been shown in the vast literature on interior point methods pioneered by Nesterov and Nemirovskii (1994) (see also Renegar, 2001) that these algorithms are able to solve LPs, SOCPs, and SDPs in polynomial time. A different class of algorithms that solves SDPs is conic bundle methods, see Helmberg and Rendl (2000), Helmberg et al. (2014). Numerous software implementations are available, and we refer the reader to Hans Mittelmann's website (Mittelmann, 2021) for an up-to-date list of the various packages.

1.3. Variants of SDP and SOCP

So far, we discussed linear conic optimization problems. However, the enormous modelling power of semidefinite and second order cone programming only unfolds if we allow for nonlinearities or integer variables:

Mixed integer conic optimization problems are linear conic problems with a constraint that some of the variables are integer valued:

$$\begin{aligned} \min \quad & \langle C, X\rangle \\ \text{s.t.} \quad & \langle A_i, X\rangle = b_i \quad (i = 1, \dots, m) \\ & X \in \mathcal{K}, \\ & X_{ij} \in \mathbb{Z} \ \ ((i, j) \in J). \end{aligned}$$

Sometimes binary constraints $X_{ij} \in \{0, 1\}$ for $(i, j) \in J$ are used instead.

Nonlinear conic problems are nonlinear problems that involve a cone constraint, mostly a semidefiniteness constraint ($\mathcal{K} = S^+_n$) or an SOCP constraint ($\mathcal{K} = \mathcal{L}_n$). Naturally, mixed integer nonlinear conic problems have been studied likewise.

It would be beyond the scope of this paper to discuss the development in mixed integer nonlinear conic optimization here. We mention just a few applications:

Nonlinear SOCPs appear for example in facility location (Chen et al., 2011). Mixed integer SOCPs appear in engineering (e.g. turbine balancing problems), in service system design (Góez and Anjos, 2019), in finance (e.g. cardinality-constrained portfolio optimization), or in combinatorial problems like the Euclidean Steiner Tree Problem, see (Gally et al., 2018) and references therein. Solution approaches for these problems include semismooth Newton methods (Chen et al., 2011), outer approximation algorithms (Drewes and Ulbrich, 2012), cutting plane algorithms (Atamtürk and Narayanan, 2007; Drewes and Pokutta, 2014; Kobayashi and Takano, 2020), and Branch-and-Bound algorithms (Gally et al., 2018).

Mixed integer SDPs have applications in truss topology optimization (Gally et al., 2018), in certain clustering problems (Aloise and Hansen, 2009), or in sparse principal component analysis (Li and Xie, 2020). References to numerous fields of application of nonlinear SDPs in engineering, (robust) control, finance and others can be found in Andreani et al. (2020) and Yamashita and Yabe (2015).

Solution algorithms for these problems include augmented Lagrangian methods, sequential SDP methods, and primal-dual interior point methods, see e.g., Andreani et al. (2020), Kocvara and Stingl (2012), Yamashita and Yabe (2015). For pointers to software implementations, we refer again to Hans Mittelmann's website (Mittelmann, 2021).

2. Conic reformulations of quadratic problems

Conic optimization problems play a particularly fruitful role in the theory of quadratic and binary quadratic optimization problems. This is accomplished by a technique called lifting, which was pioneered by Shor (1987) and Lovász and Schrijver (1991). The main idea can be seen as follows: consider a quadratic expression x^TQx with a symmetric matrix $Q \in \mathbb{R}^{n\times n}$ and $x \in \mathbb{R}^n$. If we introduce a new variable $X \in \mathbb{R}^{n\times n}$ to represent the rank-1 matrix xx^T, then we get

$$x^TQx = \text{trace}(x^TQx) = \text{trace}(Qxx^T) = \langle Q, xx^T\rangle = \langle Q, X\rangle. \tag{1}$$

By this technique, quadratic terms in $x \in \mathbb{R}^n$ become linear terms in $X \in \mathbb{R}^{n\times n}$. Since many optimization problems considered in the sequel contain nonnegativity constraints $x \geq 0$, this leads to the definition of a convex matrix cone that turns out very useful for modelling purposes: The cone of completely positive matrices is defined as

$$\mathcal{CP} := \text{conv}\{xx^T \mid x \geq 0\},$$

and its dual cone, the cone of copositive matrices, is defined as

$$\mathcal{COP} := \{X \in S \mid z^TXz \geq 0 \text{ for all } z \geq 0\}.$$

For the ease of notation, we will omit the index n in notations like $\mathcal{COP}_n$ or $\mathcal{CP}_n$ unless it is necessary to stress the dimension. Both $\mathcal{COP}$ and $\mathcal{CP}$ are proper cones and have been studied for decades in the linear algebra literature, see Berman and Shaked-Monderer (2003) and references therein. They have numerous interesting properties but are still not fully understood, cf. Berman et al. (2015). Note that the two cones are given in different form: $\mathcal{CP}$ is given by its extreme rays which are precisely the rank-1 matrices xx^T with $x \geq 0$, whereas $\mathcal{COP}$ is given as the solution set of (infinitely many) inequalities. This fact plays a role when considering approximations of these cones, see Section 3. A characterization of the extremal rays of $\mathcal{COP}_n$ has only been given for $n \leq 6$, cf. Afonin et al. (2021). Likewise, only limited knowledge is available about the facial structure of $\mathcal{CP}$ and $\mathcal{COP}$, cf. Dickinson (2011).

The earliest use of these cones in optimization was a paper by Preisig (1996) who studied a particular fractional quadratic problem, and by Quist et al. (1998), who were the first to introduce a conic optimization perspective while deriving relaxations for general quadratic optimization problems. Bomze et al. (2000) introduced the term "copositive optimization" and showed for the first time equivalence of a nonconvex quadratic optimization problem and a linear problem over $\mathcal{CP}$ resp. $\mathcal{COP}$: They considered standard quadratic optimization problems, i.e., nonconvex quadratic problems over the standard simplex $\Delta := \{x \in \mathbb{R}^n_+ \mid e^Tx = 1\}$ where $e \in \mathbb{R}^n$ denotes the all-ones vector. Given a symmetric matrix $Q \in \mathbb{R}^{n\times n}$, a standard quadratic problem is of the form

$$\begin{array}{ll} \min & x^TQx \\ \text{s.t.} & e^Tx = 1 \\ & x \geq 0. \end{array} \tag{StQP}$$

In spite of its simple structure, (StQP) is an NP-hard problem if Q has a negative eigenvalue, see Pardalos and Vavasis (1991). Alternatively, NP-hardness of (StQP) can be seen from the fact that the maximum clique problem can be formulated as (StQP): consider a graph with n vertices. Denote its adjacency matrix by A, its clique number by ω, and define $J := ee^T$. It was shown by Motzkin and Straus (1965) that

$$\frac{1}{\omega} = \min\{x^T(J - A)x : x \in \Delta\}. \tag{2}$$

The max clique problem is a particularly difficult NP-hard problem, and even computing an approximation of any reasonable quality is NP-hard (Håstad, 1999). We will see below how copositive optimization can be used to tackle this and other NP-hard problems.

By squaring the constraint in (StQP) and applying the lifting transformation outlined in (1), it is easy to see that the following problem is a relaxation of (StQP):

$$\begin{array}{ll} \min & \langle Q, X\rangle \\ \text{s.t.} & \langle J, X\rangle = 1 \\ & X \in \mathcal{CP}, \end{array} \tag{3}$$

as is its dual problem

$$\max\{\lambda \in \mathbb{R} \mid \lambda J - Q \in \mathcal{COP}\}. \tag{4}$$

It is easy to verify that Slater's condition and hence strong duality holds for (3) and (4). Since the objective function of (3) is linear and the feasible set is convex, it follows that the optimal solution is attained at an extreme point of the feasible set, which can be shown to be the matrices of the form xx^T with $x \in \Delta$, cf. Bomze et al. (2000). This implies that (3) and (4) are not merely relaxations but exact reformulations of (StQP) in the sense that all three problems have the same optimal value. The optimal solutions of (StQP) and (3) fulfill the following relation: if x^* is optimal for (StQP), then $X^* := x^*(x^*)^T$ is optimal for (3). Conversely, if X^* is optimal for (3), then it can be decomposed as $X^* = \sum_{i=1}^p x_i(x_i)^T$ for some $x_i \in \mathbb{R}^n_+$ $(i = 1, \ldots, p)$. Then each $x_i^* := \frac{1}{e^Tx_i}x_i$ is an optimal solution of (StQP).

These reformulations (3) and (4) are interesting because they show that the NP-hard (StQP) can be reformulated equivalently as a linear problem over the convex cones $\mathcal{CP}$ or $\mathcal{COP}$. In these formulations, all local minima vanish, and the complexity of the problem is entirely moved into the cone constraint. This indicates that $\mathcal{CP}$ and $\mathcal{COP}$ must be intractable. Indeed, it was shown in Dickinson and Gijben (2014), that checking membership in $\mathcal{CP}$ is NP-hard. Checking membership in $\mathcal{COP}$ is co-NP-complete, as was shown in Murty and Kabadi (1987). Whether or not checking membership in $\mathcal{CP}$ is also in NP is still one of the many open problems related to these cones, cf. Berman et al. (2015).

Returning to the maximum clique problem for a graph with adjacency matrix A, it follows from Bomze et al. (2000) that the clique number ω equals the optimal value of the following copositive problem:

$$\omega = \min\{\lambda \mid \lambda(J - A) - J \in \mathcal{COP}\}.$$

Many other graph parameters have a representation as a copositive or completely positive problem. We refer to Dür (2010) for references.

Copositive optimization experienced a breakthrough with Burer's 2009 paper Burer (2009). He showed that every quadratic problem with

linear and binary constraints can be rewritten as such a problem. More precisely, he showed that a quadratic binary problem of the form

$$\begin{array}{ll} \min & x^T Q x + 2c^T x \\ \text{s.t.} & a_i^T x = b_i \quad (i = 1, \dots, m) \\ & x \geq 0 \\ & x_j \in \{0, 1\} \quad (j \in B) \end{array} \tag{5}$$

with $Q \in S_n$, $c, a_i \in \mathbb{R}^n$ $(i = 1, \dots, m)$, $b \in \mathbb{R}^m$, and $B \subseteq \{1, \dots, n\}$ can equivalently be reformulated as the following completely positive problem:

$$\begin{array}{ll} \min & \langle Q, X \rangle + 2c^T x \\ \text{s.t.} & a_i^T x = b_i \quad (i = 1, \dots, m) \\ & \langle a_i a_i^T, X \rangle = b_i^2 \quad (i = 1, \dots, m) \\ & x_j = X_{jj} \quad (j \in B) \\ & \begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} \in \mathcal{CP}, \end{array} \tag{6}$$

provided that (5) satisfies the so-called key condition, i.e., $a_i^T x = b_i$ for all i and $x \geq 0$ implies $x_j \leq 1$ for all $j \in B$. As noted by Burer, this condition can be enforced without loss of generality. Doing so may, however, have consequences when relaxing the cone constraint (Bomze et al., 2017; Bomze and Jarre, 2010; Jarre, 2012).

Similar techniques can be used to derive copositive or completely positive formulations for problems involving quadratic constraints or replacing the constraint $x \in \mathbb{R}^n_+$ in (5) by other closed convex cones. This leads to reformulations involving so called *generalized copositive and completely positive cones*: given a closed convex cone $\mathcal{K} \subset \mathbb{R}^n$, one can define

$$\mathcal{CP}_{\mathcal{K}} := \text{conv}\{xx^T \mid x \in \mathcal{K}\},$$

and its dual cone of generalized copositive matrices

$$\mathcal{COP}_{\mathcal{K}} := \{X \in S \mid z^T X z \geq 0 \text{ for all } z \in \mathcal{K}\}.$$

These cones were introduced in Quist et al. (1998) and studied also in Eichfelder and Jahn (2008). As shown by Burer (2012), the problem where the nonnegativity constraints in (5) are replaced by the constraint $x \in \mathcal{K}$ is (under mild conditions) equivalent to a linear conic program over the cone of matrices which are completely positive over $\mathbb{R}_+ \times \mathcal{K}$, i.e., over the cone $\text{conv}\{yy^T \mid y \in \mathbb{R}_+ \times \mathcal{K}\}$. Eichfelder and Povh (2013), Dickinson et al. (2013) generalize this even more to the case where $\mathcal{K}$ is an arbitrary set, and they also give a formulation for problems involving one quadratic constraint.

When considering quadratic constraints, reformulating the problem as a conic problem becomes more involved. Burer (2009) already considered certain special cases, namely binary constraints (which can be viewed as quadratic equations $x_i^2 = x_i$), and complementarity constraints. For more general quadratically constrained quadratic problems, similar reformulations have been obtained: consider a quadratically constrained quadratic problem of the form

$$\begin{array}{ll} \min & x^T Q^0 x + 2(c^0)^T x \\ \text{s.t.} & x^T Q^i x + 2(c^i)^T x \leq b_i \quad (i = 1, \dots, m) \end{array} \tag{QCQP}$$

with $Q^i \in S_n$, $c^i \in \mathbb{R}^n$ $(i = 0, \dots, m)$, and $b \in \mathbb{R}^m$. Burer and Dong (2012) show two different ways of formulating a (QCQP) as such a generalized completely positive problem over $\mathcal{CP}_{\mathcal{K}}$: one where $\mathcal{K}$ is a direct product of $\mathbb{R}^n_+$ and second-order cones, and another where $\mathcal{K}$ is the direct product of $\mathbb{R}^n_+$ and semidefinite cones (viewed as vectors by stacking the columns on top of each other). Bai et al. (2016) and Arima et al. (2013) derive similar formulations under milder assumptions.

We remark that a different generalization of $\mathcal{CP}$ is the cone of completely positive semidefinite matrices, i.e., the cone consisting of all $n \times n$ matrices that admit a Gram representation by positive semidefinite matrices. This cone appears when studying quantum analogues of graph parameters like the stability or chromatic numbers. We refer to Laurent and Piovesan (2015) for an in-depth discussion.

It goes without saying that generalized completely positive and copositive cones are even harder to work with than $\mathcal{CP}$ or $\mathcal{COP}$. The appeal of the formulations discussed above lies in the fact that by this technique difficult, NP-hard problems can be reformulated as linear problems over a convex cone. Hence these reformulations are convex problems which do not possess local minima, and the hardness of the problem is completely captured by the cone constraint. Therefore, any progress made in understanding the cones can be used to help solving a whole range of different problems. As a first step, the approximation schemes for $\mathcal{COP}$ and $\mathcal{CP}$ discussed in Section 3 can be extended to $\mathcal{COP}_{\mathcal{K}}$ and $\mathcal{CP}_{\mathcal{K}}$. However, more research on algorithmic approaches for problems over generalized copositive and completely positive cones is needed to make these approaches work numerically for bigger problems.

2.1. Extensions: polynomial optimization and infinite dimensional conic problems

It is easy to see that any polynomial optimization problem can be rewritten as a quadratic problem by introducing extra variables and constraints. For example, by defining an extra variable and constraint $y = x_j x_k$, the cubic term $x_i x_j x_k$ becomes the quadratic term $x_i y$. So the reformulations discussed above can in principle be applied to polynomial problems, as well. A different line of research has worked with the *cone of completely positive tensors*. This concept was originally introduced by Dong (2013) and consititutes another natural extension of $\mathcal{CP}$. Peña et al. (2015) consider optimization problems involving n-variate polynomials and show that under certain assumptions these can be reformulated as linear problems over the cone of completely positive tensors of order d and dimension $n+1$. They also show that in case of a compact feasible set order $d = 4$ is sufficient. The approach has been extended in Kuang and Zuluaga (2018), Xia and Zuluaga (2017). We also refer to Anjos and Lasserre (2012) for more discussion on the connection between conic and polynomial optimization.

We mentioned above that the maximum clique problem (and, equivalently, the stable set problem) on a graph with vertex set $V = \{1, \dots, n\}$ can be formulated as an (StQP) and consequently as a copositive or completely positive problem (3) and (4). This can be generalized to infinite graphs, i.e., to the setting where the vertex set is not finite but a compact metric space V equipped with a probability measure μ. The problem of determining the stability number of an infinite graph appears e.g. in the kissing number problem (Dobre et al., 2016) and other packing problems, see DeCorte et al. (2021). Dobre et al. (2016) generalized the concept of copositive matrices to the infinite dimensional setting by defining *copositive kernels*: a kernel is a continuous function $K : V \times V \to \mathbb{R}$. Such a kernel K is called a *copositive kernel* if for all continuous nonnegative functions $f : V \to \mathbb{R}_+$ we have

$$\int_V \int_V K(x, y) f(x) f(y) d\mu(x) d\mu(y) \geq 0.$$

It can be shown that the cone of copositive kernels is independent of the choice of μ. This cone as well as its dual (the cone of completely positive measures) can be used to derive exact copositive and completely positive reformulations of the stability number problem for infinite graphs, see Dobre et al. (2016). Since these cones are intractable, approximations have been proposed in Kuryatnikova and Vera Lizcano (2017), Kuryatnikova and Vera (2018). This in turn has been used to derive good bounds for the underlying problems.

3. Approximation hierarchies for $\mathcal{COP}$ and $\mathcal{CP}$

Since the cones $\mathcal{COP}$ and $\mathcal{CP}$ are computationally intractable, approximations have to be used in order to solve an optimization problem over one of these cones. As outlined in (1), the motivation to introduce the cone $\mathcal{CP}$ was by introducing a symmetric matrix X to represent the rank-1 matrix xx^T. So a first straightforward relaxation is to replace the constraint $X = xx^T$ by $X \succeq xx^T$ (meaning that $X - xx^T \in S_n^+$), which by

Schur's complement lemma is equivalent to

$$\begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} \in S^+_{n+1}.$$

This relaxation goes back to Shor (1987) and corresponds to the simple fact that $\mathcal{CP}_n \subseteq S^+_n$ for any n. It is interesting to note that the SDP-relaxation of a quadratic problem corresponds to the Lagrangian dual of that problem, whereas considering partial Lagrangian duals (i.e., dualizing the problem only with respect to a subset of the constraints) leads to various copositive relaxations, cf. Bomze (2015). The Shor relaxation can be improved by adding more constraints to the SDP, or by using some relaxation-linearization techniques, yielding stronger SDP-relaxations. This has been discussed in detail in Bao et al. (2011).

Shor's approximation can be strengthened by using better approximations to $\mathcal{COP}$ and $\mathcal{CP}$: Denote by $\mathcal{N}_n$ the set of symmetric entry-wise nonnegative $n \times n$ matrices. Then it is obvious from the definition that $\mathcal{N}_n \subseteq \mathcal{COP}_n$. We therefore get that $\mathcal{N}_n + S^+_n \subseteq \mathcal{COP}_n$, and by duality $\mathcal{CP}_n \subseteq \mathcal{N}_n \cap S^+_n$. Interestingly, both inclusions are equalities for $n \leq 4$ and are strict for $n \geq 5$, cf. Maxfield and Minc (1962/63). Matrices in $\mathcal{N}_n \cap S^+_n$ are sometimes called doubly nonnegative (both the entries and the eigenvalues are nonnegative). This cone has been frequently used to obtain bounds for certain combinatorial problems, with the most prominent case being the Lovász–Schrijver bound $\vartheta^+(G)$ (sometimes called $\vartheta'(G)$) on the clique number of a graph G, see Schrijver (1979).

In order to get better approximations of $\mathcal{COP}$ and $\mathcal{CP}$, a number of techniques have been developed which often lead to so called approximation hierarchies, i.e., monotonic sequences of inner or outer approximations of $\mathcal{COP}$ or $\mathcal{CP}$ which are, in some sense, exact in the limit. The approximating cones are constructed in such a way that optimizing over them amounts to solving an LP, an SOCP, or an SDP, all of which can be done in polynomial time. Several of these hierarchies have been proposed, and we discuss the most important ones next. Note that these hierarchies were originally designed to approximate either $\mathcal{COP}$ or $\mathcal{CP}$. However, it should be clear that any inner (resp. outer) approximation hierarchy of one cone by duality yields an outer (resp. inner) approximation hierarchy for the dual cone.

3.1. Inner approximation hierarchies for $\mathcal{COP}$

Parrilo (2000) was the first to propose a hierarchy approximating $\mathcal{COP}$ from the interior. The basic idea is to reformulate the copositivity condition as a nonnegativity condition for certain polynomials, and then to use the sufficient condition that a polynomial is nonnegative if it can be represented as a sum of squares (sos) of other polynomials. Suppose we are given a matrix $A \in S_n$ and we would like to determine whether or not $A \in \mathcal{COP}_n$. To this end, consider the polynomial

$$P_A(x) := \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i^2 x_j^2 \tag{7}$$

and observe that $A \in \mathcal{COP}_n$ if and only if $P_A(x) \geq 0$ for all $x \in \mathbb{R}^n$. A sufficient condition for this is that $P_A(x)$ is sos. Parrilo showed that the set of matrices A for which $P_A(x)$ is sos equals $S^+_n + \mathcal{N}_n$.

Moreover, he was able to refine this by using a result by Pólya (1928) and considering higher order polynomials. For any $r \in \mathbb{N}$, define the cone

$$\mathcal{K}^r := \left\{ A \in S \mid P_A(x)\left(\sum_{i=1}^n x_i^2\right)^r \text{ has an sos decomposition} \right\}.$$

Parrilo showed that

$$S^+ + \mathcal{N} = \mathcal{K}^0 \subset \mathcal{K}^1 \subset \ldots \subset \mathcal{COP} \quad \text{and} \quad \operatorname{int}(\mathcal{COP}) \subseteq \bigcup_{r \in \mathbb{N}} \mathcal{K}^r,$$

so the cones $\mathcal{K}^r$ approximate $\mathcal{COP}$ from the interior. The sos condition can be written as a system of linear matrix inequalities (LMIs), and therefore optimizing over $\mathcal{K}^r$ amounts to solving an SDP. However, it should be noted that for increasing values of r, the size of these SDPs increases rapidly, resulting in problems that are beyond the range of current SDP-solvers even for moderate values of r and n.

Ahmadi and Majumdar (2019) developed a more general theory for nonnegativity of polynomials which when applied in our context boils down to relaxing the sos-condition by requiring that $P_A(x)$ resp. $P_A(x)\left(\sum_{i=1}^n x_i^2\right)^r$ has a decomposition as a *sum of squares of binomials*. This is clearly a weaker sufficient condition for nonnegativity of $P_A(x)$, but the advantage is that this condition can be verified by solving an SOCP. They also consider scaled versions and obtain hierarchies that they call $rDSOS_n$ and $rSDSOS_n$, referring to the rth level of the hierarchies corresponding to *diagonally dominant sum of squares (DSOS)* and *scaled diagonally dominant sum of squares* (*SDSOS*).

An alternative sufficient condition for nonnegativity of a polynomial is that all of its coefficients are nonnegative. Exploiting this idea, de Klerk and Pasechnik (2002), cf. also Bomze and De Klerk (2002), define the cones

$$C^r := \left\{ A \in S \mid P_A(x)\left(\sum_{i=1}^n x_i^2\right)^r \text{ has nonnegative coefficients} \right\}.$$

They showed that

$$\mathcal{N} = C^0 \subset C^1 \subset \ldots \subset \mathcal{COP} \quad \text{and} \quad \operatorname{int}(\mathcal{COP}) \subseteq \bigcup_{r \in \mathbb{N}} C^r.$$

Each of the cones C^r is polyhedral, so optimizing over one of them is solving an LP.

Peña et al. (2007) refined the above approaches and derived a hierarchy of cones $\mathcal{Q}^r$ which in a sense sits between C^r and $\mathcal{K}^r$, i.e., it fulfills $C^r \subseteq \mathcal{Q}^r \subseteq \mathcal{K}^r$ for all $r \in \mathbb{N}$. These cones can be described by LMIs as well, so optimizing over $\mathcal{Q}^r$ is again an SDP, however, optimizing over $\mathcal{Q}^r$ provides SDPs of smaller size compared to $\mathcal{K}^r$.

Each of the above hierarchies provides a uniform inner approximation to $\mathcal{COP}$, i.e., the approximation quality is independent of which part of the cone is considered. However, this may not be desirable when considering a specific optimization problem over $\mathcal{COP}$. In this case, one would rather like to obtain a good approximation of $\mathcal{COP}$ in the vicinity of the optimal solutions, whereas in other parts of the feasible set a coarse approximation is sufficient. This idea gave rise to the approach by Bundfuss and Dür (2009): It is easy to see that the definition of $\mathcal{COP}$ is equivalent to $\mathcal{COP} = \{A \in S \mid x^T A x \geq 0 \text{ for all } x \in \Delta\}$. Now Bundfuss and Dür (2009) consider partitions $\mathcal{P} = \{S^1, \ldots, S^m\}$ of Δ into subsimplices and give conditions ensuring nonnegativity of $x^T A x$ over each S^i. Let $S = \operatorname{conv}\{v_1, \ldots, v_n\} \subseteq \Delta$ be such a simplex. Then $x \in S$ can be written as a convex combination $x = \sum_{i=1}^n \lambda_i v_i$ with $\sum_{i=1}^n \lambda_i = 1$ and $\lambda_i \geq 0$ for all i. Copositivity of a matrix A then means that

$$0 \leq x^T A x = \left(\sum_{i=1}^n \lambda_i v_i\right)^T A \left(\sum_{j=1}^n \lambda_j v_j\right) = \sum_{i,j=1}^n (v_i^T A v_j) \lambda_i \lambda_j. \tag{8}$$

Since $\lambda_i \geq 0$ by construction, a sufficient condition for (8) is that $v_i^T A v_j \geq 0$ for all i, j. Note that this constitutes a system of linear inequalities for the entries of A. Therefore,

$$\begin{aligned} \mathcal{I}_{\mathcal{P}} := \{A \in S \mid\ & v^T A v \geq 0 \ \text{ for all vertices } v \text{ of simplices in } \mathcal{P}, \\ & u^T A v \geq 0 \ \text{ for all edges } \{u, v\} \text{ of simplices in } \mathcal{P}\}. \end{aligned}$$

is a polyhedral inner approximation of $\mathcal{COP}$. It is shown in Bundfuss and Dür (2009) how the partition $\mathcal{P}$ can be refined in order to obtain a sequence of inner approximations that can either be tailored to yield a uniform approximation of $\mathcal{COP}$ or an adaptive approximation with good quality in the vicinity of the optimal solution of the underlying copositive optimization problem.

3.2. Outer approximation hierarchies for $\mathcal{COP}$

Since we can write $\mathcal{COP} = \{A \in S \mid x^T A x \geq 0 \text{ for all } x \in \Delta\}$, outer approximations of $\mathcal{COP}$ can be obtained by picking suitable (possibly infinite) subsets $I \subset \Delta$ and considering $\{A \in S \mid x^T A x \geq 0 \text{ for all } x \in I\}$.

One option studied by Yıldırım (2012) is to consider regular grids of rational points on the unit simplex defined as

$$\delta(r) := \bigcup_{k=0}^{r} \{x \in \Delta \mid (k+2)x \in \mathbb{N}_0^n\}. \tag{9}$$

Then the set $\mathcal{O}^r := \{A \in S \mid x^T Ax \geq 0 \text{ for all } x \in \delta(r)\}$ is clearly a polyhedral outer approximation of $\mathcal{COP}$ for any $r \in \mathbb{N}$, and one can show that $\mathcal{O}^0 \supset \mathcal{O}^1 \supset \cdots \supset \mathcal{COP}$ and $\mathcal{COP} = \bigcap_{r\in\mathbb{N}} \mathcal{O}^r$. This approximation scheme gives again uniform approximations of $\mathcal{COP}$ and allows for exact assessment of the quality of the approximation.

Alternatively one can use an approach developed by Lasserre in a series of papers which makes use of the vast body of theory on positive (resp. nonnegative) polynomials and polynomial optimization. Denote by $\mathcal{P}_{n,d}^+$ the cone of n-variate polynomials of total degree $\leq d$ which are nonnegative on $\mathbb{R}^n$ (note that such a polynomial necessarily has even degree). Then $A \in S_n$ is copositive if and only if the polynomial P_A from (7) fulfills $P_A \in \mathcal{P}_{n,4}^+$. The Riesz–Haviland Theorem tells us that the dual of $\mathcal{P}_{n,d}^+$ is the so called moment cone. Exploiting this, one can obtain another hierarchy of cones approximating $\mathcal{COP}$ resp. $\mathcal{CP}$ which is a special case of the Lasserre–hierarchy applied to the setting of copositivity and complete positivity. We refer to the book by Lasserre (2010) and the survey by Laurent (2009) which both give excellent introductions to the general moment approach for polynomial optimization. The paper by Lasserre (2014) explicitly describes how to construct hierarchies of outer approximations of $\mathcal{COP}$ and inner approximations of $\mathcal{CP}$ by using this moment approach. It should be noted that these hierarchies are based on conditions that can be expressed as LMIs, and hence optimizing over these hierarchies amounts to solving SDPs.

A third option is the adaptive approximation approach by Bundfuss and Dür (2009) detailed in Section 3.1 which gives the outer approximation

$$\mathcal{O}_p := \{A \in S \mid v^T Av \geq 0 \text{ for all vertices } v \text{ of simplices in } \mathcal{P}\}.$$

This yields a hierarchy of polyhedral approximations that can again be tailored to either yield a uniform outer approximation of $\mathcal{COP}$ or a finer approximation in the vicinity of the set of optimal solutions but only a coarse approximation in the remaining parts.

3.3. Inner approximation hierarchies for $\mathcal{CP}$

Recall that $\mathcal{CP}_n = \text{conv}\{xx^T \mid x \in \mathbb{R}_+^n\} = \text{cone conv}\{xx^T \mid x \in \Delta\}$. Therefore, inner approximations of $\mathcal{CP}$ can be constructed analogous to outer approximations of $\mathcal{COP}$, namely by chosing suitable subsets $I \subset \Delta$ and considering $C(I) := \text{cone conv}\{xx^T \mid x \in I\}$. A thorough treatment investigating properties of the approximation in dependence of the set I is given in Yıldırım (2017).

If the set I stems from a finite discretization of Δ, then $C(I)$ is polyhedral. A different approach was developed by Gouveia et al. (2020) based on similar work in Ahmadi and Majumdar (2019). They consider the cone

$$\text{SDD}_+^n := \text{conv}\{xx^T \mid x \in \mathbb{R}_+^n,\ |\text{supp}(x)| \leq 2\} \subseteq \mathcal{CP}_n,$$

where the support of a vector x is defined as $\text{supp}(x) := \{i \mid x_i \neq 0\}$. It can be shown (see Ahmadi and Majumdar, 2019; Gouveia et al., 2020 and references therein) that $A \in \text{SDD}_+^n$ if and only if A is scaled diagonally dominant, i.e., if there exists a diagonal matrix D with positive diagonal entries such that DAD is diagonally dominant. From the definition we get that that $A \in \text{SDD}_+^n$ if and only if A can be written as $A = \sum_{i<j} M^{ij}$, where M^{ij} are symmetric nonnegative and positive semidefinite matrices whose entries are zero everywhere except at the positions ii, ij, ji, jj. Observe that positive semidefiniteness of 2×2 symmetric matrices can be characterized by second order conditions. Indeed, we have

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} \in S_2^+ \iff \begin{matrix} a \geq 0,\ c \geq 0, \\ ac - b^2 \geq 0 \end{matrix} \iff \begin{matrix} a \geq 0, \\ c \geq 0, \end{matrix} \left\| \begin{pmatrix} 2b \\ a - c \end{pmatrix} \right\|_2 \leq a + c,$$

and the latter condition is equivalent to the second order cone constraint $(a + c, 2b, a - c)^T \in \mathcal{L}_3$. Therefore, optimizing over SDD_+^n amounts to solving an SOCP. In Gouveia et al. (2020), this approach is further refined by considering scaled variants of SDD_+^n which can be tailored to obtain either uniform or problem-dependent approximation schemes.

4. Examples of binary quadratic problems

In this section, we discuss a few combinatorial problems which can be formulated as binary quadratic optimization problems. These problems are typically NP-hard, so it will be useful to consider reformulations and relaxations which are tractable. It will turn out that relaxations based on conic optimization are particularly useful. We will mostly focus on relaxations in the cone of positive semidefinite matrices. Throughout this section, assume we are given an undirected graph $G = (V, E)$ with $V = \{1, \ldots, n\}$ and adjacency matrix $A \in S_n$.

4.1. Unconstrained binary quadratic optimization and MaxCut

An unconstrained binary quadratic optimization problem takes as input a symmetric $n \times n$ matrix Q and asks to find

$$\begin{matrix} \min & x^T Qx \\ \text{s.t.} & x \in \{0,1\}^n. \end{matrix} \tag{10}$$

Since $x_i = x_i^2$, a possible linear term in the objective function could be integrated in the main diagonal of Q, so it is not necessary to explicitly include a linear term in this model.

The **MaxCut problem** is defined by an edge weighted graph, given through its adjacency matrix A. Hence A is symmetric, but we do not impose any further restrictions to the entries of A. In particular, $a_{ij} < 0$ is possible. If $[i, j]$ is not an edge of the graph, we set $a_{ij} = 0$. The *Laplacian matrix* L associated to A is defined by

$$l_{ij} := -a_{ij} \text{ for } i \neq j, \quad \text{and} \quad l_{ii} := \sum_k a_{ik}.$$

Note that $Le = 0$, and if $A \geq 0$, then L is diagonally dominant and hence $L \succeq 0$. It is a simple exercise to verify that for $y \in \{-1, 1\}^n$ the value of the cut defined by $S := \{i \mid y_i = 1\}$ is given by

$$\sum_{i<j} \tfrac{1}{2} a_{ij}(1 - y_i y_j) = \tfrac{1}{4} y^T Ly,$$

so the MaxCut problem can be formulated as

$$\max \tfrac{1}{4} y^T Ly \text{ such that } y \in \{-1, 1\}^n.$$

Setting $x := \frac{1}{2}(y + e) \in \{0, 1\}^n$ and using $Le = 0$, we get that $\frac{1}{4} y^T Ly = x^T Lx$ which shows that MaxCut and the binary quadratic optimization problem (10) are indeed equivalent optimization problems. Lasserre (2016) showed an even more general result: Considering the linearly constrained binary quadratic problem

$$\min\{c^T x + x^T Fx \mid Ax = b, x \in \{0,1\}^n\}, \tag{11}$$

Lasserre showed that this can by reformulated as a MaxCut problem on a graph with $n + 1$ nodes that can be explicitly constructed from the data of the problem. So in a sense MaxCut is a canonical model for linear and quadratic binary problems. Note that (11) is a special instance of the problem (5) studied by Burer. It can therefore be formulated as a copositive problem, and then approximation hierarchies from Section 3 can be used.

Historically, linear and then semidefinite relaxations were studied before copositivity came into play. In a celebrated paper, Goemans and Williamson (1995) took the following approach: since $y^T Ly = \langle L, yy^T \rangle$ they introduce the matrix Y taking the role of yy^T. Then $Y \succeq 0$ and $\text{diag}(Y) = e$ must hold, and this yields the semidefinite relaxation

$$\max \tfrac{1}{4} \langle L, Y \rangle \text{ such that } \text{diag}(Y) = e,\ Y \succeq 0. \tag{12}$$

For graphs with nonnegative edge weights, they were able to show the celebrated result that the optimal value of (12) is at most 13.83% higher than the optimal value of MaxCut. In other words, the SDP relaxation has a performance guarantee of $\approx 87\%$.

4.2. Partition and clustering problems

We briefly discuss various extensions of unconstrained binary quadratic optimization problems which lead to additional linear constraints on the binary variables.

k-cluster problems

The simplest extension of problem (10) consists in asking that exactly k of the variables in (10) are set to 1. Let $A \geq 0$ be a symmetric $n \times n$ matrix. We may think of a_{ij} as a measure for the interaction between i and j. The problem

$$\max \tfrac{1}{2} x^T A x \text{ such that } e^T x = k \text{ and } x \in \{0,1\}^n$$

asks for a subset of k vertices having maximum total pairwise interaction. Such a set may be viewed as a "cluster" in the sense that it collects a set of k items with maximum mutual interaction. This type of problem has found increased interest from applications in data mining, see for instance (Fortunato, 2010). Various SDP relaxations have been discussed in Malick and Roupin (2012).

A different application of this type of problem is related to the stable set problem. Let A be the adjacency matrix of an unweighted graph G and let k be given. Consider the minimization problem

$$z(k) := \min \tfrac{1}{2} x^T A x \text{ such that } e^T x = k \text{ and } x \in \{0,1\}^n.$$

If the optimal value fulfills $z(k) > 0$, then we have a proof that G has no stable set of size k, so that the stability number $\alpha(G)$ fulfills $\alpha(G) \leq k-1$. This idea will be further exploited in Section 4.3.

MAX-k-CUT

The MAXCUT problem may also be seen as a very simple graph partition problem as it asks to separate the vertices of the graph into two parts such as to maximize the weight of the edges joining the two partition blocks. It is a natural generalization to consider vertex partitions into (at most) k partition blocks for some fixed $k \geq 2$.

We represent k-partitions of $V = \{1, \ldots, n\}$ by 0-1 matrices X of order $n \times k$ satisfying $Xe = e$. This condition simply states that $(Xe)_i = 1$ for all i, meaning that vertex i is in exactly one partition block, namely block j in case $x_{ij} = 1$. The sum of the elements of column j of X equals the number of vertices (possibly zero) in partition block j. It is a simple exercise to verify that the total weight of edges joining vertices in distinct partition blocks is given by

$$\tfrac{1}{2}\text{trace}(LXX^T).$$

Therefore, the MAX-k-CUT problem may be formulated as

$$\max \tfrac{1}{2}\text{trace}(LXX^T) \text{ such that } Xe = e \text{ and } X \in \{0,1\}^{n \times k}.$$

The **k-partition problem** is obtained by further constraining the partitions to have exactly k partition blocks, and to require that partition block j contains exactly $m_j \in \mathbb{N}$ vertices, where $\sum_j m_j = n$. We collect the cardinalities m_j in the vector $m \in \mathbb{N}^k$, so that feasible partitions are represented by matrices $X \in \{0,1\}^{n \times k}$ satisfying

$$Xe = e \text{ and } X^T e = m.$$

The special case where all the m_i are equal is of special interest in certain telecommunication problems, and we refer to Lisser and Rendl (2003) for a discussion of these applications and relaxations based on semidefinite optimization.

4.3. Stable sets and graph coloring

We briefly look at formulations for the stability and the chromatic number of a graph in connection with quadratic binary optimization.

A subset S of vertices of a graph G is called *stable* (or *independent*), if the subgraph of G induced by S is empty. The *stability number* $\alpha(G)$ denotes the cardinality of a largest stable set in G. Determining $\alpha(G)$ is considered an extremely difficult combinatorial optimization problem, cf. Håstad (1999). The following binary quadratic optimization problem determines $\alpha(G)$:

$$\alpha(G) = \max \sum x_i \text{ such that } x_i x_j = 0 \text{ for all } [i,j] \in E(G),\ x \in \{0,1\}^n.$$

The optimal value $z(k)$ of the following optimization problem can be used to check whether G contains a stable set of size k. As usual, A denotes the adjacency matrix of G. Consider the following problem which was already briefly discussed in Section 4.2:

$$z(k) := \min \tfrac{1}{2} x^T A x \text{ such that } e^T x = k,\ x \in \{0,1\}^n. \tag{13}$$

If $z(k) = 0$, then G contains a stable set of size k given by $S = \{i \mid x_i = 1\}$. On the other hand, $z(k) > 0$ shows that G has no stable set of size k and therefore $\alpha(G) < k$.

Let us now turn to vertex colorings of G. A k-coloring of $V(G)$ can be seen as a vertex partition of $V(G)$ into k stable sets (the color classes). The *chromatic number* $\chi(G)$ denotes the smallest number k such that G is k-colorable. A formulation to compute $\chi(G)$ as the solution of a copositive problem was given in Gvozdenović and Laurent (2008).

Expressing $\chi(G)$ as a binary optimization problem is usually done as follows. Let $S = \{s_1, s_2, \ldots\}$ be the collection of characteristic vectors of stable sets in G.

$$\chi(G) = \min \sum_i \lambda_i \text{ such that } \sum_i \lambda_i s_i = e,\ \lambda_i \in \{0,1\}.$$

This is a linear optimization problem in binary variables λ_i. Unfortunately, there may be an exponential number of them. Weakening the condition $\lambda_i \in \{0,1\}$ to $0 \leq \lambda_i \leq 1$ for all i leads us to the *fractional chromatic number* $\chi_f(G)$:

$$\chi_f(G) = \min \sum_i \lambda_i \text{ such that } \sum_i \lambda_i s_i = e,\ 0 \leq \lambda_i \leq 1 \text{ for all } i.$$

Computing the optimal value $\chi_f(G)$ of this linear program is again known to be NP-hard, see for instance (Lund and Yannakakis, 1994).

Let us now consider testing whether G contains a k-coloring for some given k. We introduce the $n \times k$ binary matrix X and require $\sum_{r=1}^k x_{ir} = 1$ for all $1 \leq i \leq n$, which we write in a slight abuse of notation as $Xe = e$. This condition asks that each row of X contains exactly one entry equal to one, so that the columns of X provide a vertex partition of $V(G)$ into (at most) k partition blocks. Since each row of X has exactly one nonzero entry, we see that $\sum_{r=1}^k x_{ir} x_{jr} = 1$ for some $i \neq j$ is only possible if i and j both belong to the same partition block r for some $r \in \{1, \ldots, k\}$. As a consequence

$$\sum_{[i,j] \in E(G)} \sum_r x_{ir} x_{jr} = \tfrac{1}{2}\langle X, AX \rangle$$

counts the number of edges joining vertices in the same partition block. We introduce

$$z(k) := \min \tfrac{1}{2}\langle X, AX \rangle \text{ such that } Xe = e,\ X \in \{0,1\}^{n \times k}. \tag{14}$$

In a slight abuse of notation, we use $z(k)$ again for the optimal value of the relaxation with exactly k columns in X. If $z(k) = 0$, then the optimal X provides a partitioning of $V(G)$ into (at most) k stable sets and therefore $\chi(G) \leq k$. On the other hand $z(k) > 0$ implies that no k-partition exists where all partition blocks are stable sets, and therefore $\chi(G) > k$.

We will come back to SDP relaxations for these problems and investigate connections to the ϑ number in Section 5.2.

4.4. Quadratic set cover

The (linear) set cover problem is defined as follows. We are given a set $C := \{v_1, \ldots, v_n\}$ of n elements and a collection S of m subsets of C such that their union equals C. Each subset S_i in S has cost q_i. The task is to select subsets in S such that their union is C and such that the cost of the selected subsets in minimized. This problem is one of Karp's 21 NP-complete problems.

To state this problem formally, define an $n \times m$ binary matrix A with $a_{ij} = 1$ if $v_i \in S_j$. Row i of A indicates which subsets S_j contain v_i, column j of A is the incidence vector of subset S_j. With this, the linear set cover problem reads:

$$z^* := \min \sum_{j=1}^m q_j x_j \text{ s.t. } \sum_{j=1}^m a_{ij} x_j \geq 1 \text{ for all } i = 1, \ldots, n \text{ and } x \in \{0,1\}^m.$$

Let us denote the largest row sum of A by f and the largest column sum of A by g. The following approximation results go back to the 1980's. Hochbaum (1982) introduces a primal-dual LP-rounding heuristic which gives (in polynomial time) a feasible solution to set cover with value z at most fz^*, i.e. $z \le fz^*$. Chvátal (1979) proposes a greedy rounding heuristic which yields (in polynomial time) a feasible solution to set cover with value z at most $(1+\log(g))z^*$, i.e., $z \le (1+\log(g))z^*$.

The quadratic set cover problem differs from the linear one only in the objective function which now also may contain quadratic terms:

$$\sum_{i=1}^{m}\sum_{j=1}^{m} q_{ij}x_ix_j.$$

It is clear that if $q_{ij}=0$ for all $i \neq j$, then we recover the linear set cover problem (as $x_j = x_j^2$).

A good summary on complexity issues related to quadratic set cover is given by Escoffier and Hammer (2007). They relate quadratic set cover to deciding whether a graph has chromatic number 3: Let G be a (nonbipartite) graph on n vertices and construct a quadratic set cover instance as follows. The ground set is $V=\{1,\dots,n\}$ and we have $3n$ subsets S_{ir} with

$$S_{ir} := \{i\} \quad \text{for } r=1,2,3,$$

so each set consists of only one element, and we have three copies of each set. The n covering conditions ask that

$$x_{i1}+x_{i2}+x_{i3} \ge 1 \quad \text{for all } i=1,\dots,n.$$

We may think of these constraints as asking that each vertex should receive color 1 or 2 or 3. Thus we do not allow that all three of these variables are zero but more than one of them may be set to one. Escoffier and Hammer (2007) show the following theorem.

Theorem 4.1. *Let G be a nonbipartite graph and consider*

$$\begin{aligned} z^* := \min \quad & \sum_{[i,j]\in E(G)} x_{i1}x_{j1}+x_{i2}x_{j2}+x_{i3}x_{j3} \\ \text{s.t.} \quad & x_{i1}+x_{i2}+x_{i3} \ge 1 \quad \text{for all } i=1,\dots,n \\ & x_{ir} \in \{0,1\} \quad \text{for all } i=1,\dots,n,\ r=1,2,3. \end{aligned}$$

Then $z^*=0$ if and only if $\chi(G)=3$.

As a consequence, it is NP-hard to decide whether a quadratic set cover problem has optimal value 0 or greater than 0. The covering problem in this construction is quite simple. Each set consists of only one element ($S_{ir}=\{i\}$), and each cover constraint involves only three elements from the ground set. This problem would therefore be trivial to solve with a linear objective function.

4.5. Quadratic assignment problem

The Quadratic Assignment Problem (QAP) asks to minimize a quadratic objective function over the set of permutation matrices. It contains many prominent NP-hard problems as special cases, see for instance (Pardalos et al., 1994). We define it through three data matrices A, B and C of order $n\times n$ and assume that A and B are symmetric. The set of $n\times n$ permutation matrices is denoted by Π_n or Π for short. The QAP then reads:

$$\min\langle AXB+C, X\rangle \text{ such that } X\in\Pi.$$

Note that Π is contained in the affine space

$$\mathcal{E} := \{X\in\mathbb{R}^{n\times n} \mid Xe = X^Te = e\}$$

of all matrices having row and column sums equal to 1. Let the $(n-1)\times n$ matrix V represent a basis of $e^\perp$, the orthogonal complement to the vector $e\in\mathbb{R}^n$. It is well known that any $X\in\mathcal{E}$ may be written as

$$X = \tfrac{1}{n}J + VMV^T$$

where $J := ee^T$ and $M\in S$ is an arbitrary symmetric matrix of order $n-1$. Note in particular that VMV^T lies in the linear space of matrices having row and column sums equal to 0.

Povh and Rendl (2009) proposed a copositive formulation of QAP and semidefinite relaxations based upon it. To this end, it is useful to rewrite the objective function of QAP in terms of $x := \mathrm{vec}(X)$, where $\mathrm{vec}(X)$ is a vector obtained from the matrix X by stacking the columns of X on top of each other. Using the Kronecker product $B\otimes A$ of the matrices B and A, it is not difficult to see that

$$\langle AXB, X\rangle = x^T(B\otimes A)x$$

We also set $c := \mathrm{vec}(C)$ and derive

$$\langle AXB+C, X\rangle = x^T(B\otimes A)x + c^Tx.$$

We are now interested in the set

$$\mathcal{P} := \mathrm{conv}\{xx^T : x=\mathrm{vec}(X), X\in\Pi\}.$$

We have just seen that

$$x = \mathrm{vec}(X) = \mathrm{vec}(\tfrac{1}{n}J + VMV^T) = \tfrac{1}{n}e\otimes e + (V\otimes V)m$$

using $m := \mathrm{vec}(M)$. Let $z := \binom{1}{m}$ and set $W := (\frac{1}{n}e\otimes e, V\otimes V)$. Then $x = Wz$ and $xx^T = Wzz^TW^T$. The definition of z implies $(zz^T)_{1,1}=1$. The SDP relaxation of QAP is now obtained by allowing any semidefinite matrix R with $(R)_{1,1}=1$ in place of zz^T. For ease of notation we introduce $Y := WRW^T$, and we get the following semidefinite relaxation of QAP:

$$\begin{aligned} &\min\langle B\otimes A + \mathrm{diag}(c), Y\rangle \text{ such that } Y = WRW^T, (R)_{1,1}=1, \\ &y = \mathrm{diag}(Y),\ Y - yy^T \succeq 0. \end{aligned}$$

Since Y takes the role of xx^T we may think of the $n^2\times n^2$ matrix Y as being partitioned into $n\times n$ matrices $Y^{i,j}$ such that $Y^{i,j}$ corresponds to the matrix $X_{\cdot,i}\cdot X_{\cdot,j}^T$. Since $X_{i,k}X_{i,l}=0$ for $k\neq l$, it follows immediately that the submatrix $Y^{i,i}$ is 0 outside its main diagonal, i.e., $(Y^{i,i})_{k,l}=0$ for all $k\neq l$. In a similar way we conclude that $\mathrm{diag}(Y^{i,j})=0$ for $i\neq j$. We refer to Povh and Rendl (2009) for further details.

5. Modelling linear equalities and inequalities in SDP relaxations

In this section we take a closer look at modelling issues related to combinatorial optimization problems. We stress that formulations which are equivalent in the binary setting may give different results when we move to conic relaxations. For instance, suppose we have binary variables $x_i\in\{0,1\}$ and we would like to express the constraint that for a given pair i,j at most one of the associated variables x_i and x_j is allowed to be equal to 1. This could be done either by imposing the linear inequality $x_i+x_j\le 1$ or by requiring the quadratic equation $x_ix_j=0$ to hold. In the binary setting, both conditions are equivalent, but once we move to relaxations, they may yield different results.

Moreover, various ways of constructing SDP relaxations have been proposed in the literature which may also lead to bounds of varying quality. We refer to Anjos et al. (2021) for a recent discussion of various SDP models related to the stable set problem.

Finally, the idea of using approximation hierarchies as detailed in Section 3 has found a lot of interest. Applied to combinatorial optimization problems, these hierarchies typically have the property that the quality of the relaxation gets tighter as one moves up in the hierarchy, yielding the integer optimum as one moves up high enough in the hierarchy. Unfortunately, it is computationally challenging to tackle even the first few levels in these hierarchies. If the initial problem has n binary variables, the SDP in the first level of the hierarchy is formulated in matrices of order $n+1$, but already the second level uses matrices of order $\binom{n}{2}$ which is prohibitive once n is much larger than 100.

5.1. Lifting linear constraints

Here we focus on the modelling issue and investigate how linear constraints may be lifted into the SDP relaxation. The precise method of lifting linear constraints can be a subtle issue and can, if not done carefully, destroy strict feasibility of the lifted problem. For given $a_0 > 0$, we consider the SDP relaxation of

$$\max x^T C x \text{ such that } x \in \{0,1\}^n, a^T x = a_0.$$

This is a binary quadratic optimization problem with a single linear equality constraint $a^T x = a_0$. How should this equation be included in the SDP relaxation?

The semidefiniteness constraint $X - xx^T \succeq 0$ with $\operatorname{diag}(X) = x$ immediately shows that

$$a^T X a \geq (a^T x)^2.$$

Since we should have equality it seems plausible to optimize over the set

$$\mathcal{F}_1 := \{(X,x) \mid \begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} \succeq 0,\ \operatorname{diag}(X) = x,\ a^T x = a_0,\ a^T X a = a_0^2\}.$$

Unfortunately, this construction makes feasible matrices singular, as we show in the next lemma:

Lemma 5.1. *Let $(X,x) \in \mathcal{F}_1$. Then the matrix $\begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix}$ is singular and $Xa = a_0 x$.*

Proof. We first note that

$$\begin{pmatrix} a \\ -a_0 \end{pmatrix}^T \begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} \begin{pmatrix} a \\ -a_0 \end{pmatrix} = a^T X a - 2a_0 a^T x + a_0^2 = 0.$$

Since $\begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} \succeq 0$, this implies $\begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} \begin{pmatrix} a \\ -a_0 \end{pmatrix} = 0$, and therefore $Xa = a_0 x$. □

It is well known that the set of singular positive semidefinite matrices equals the boundary of S_n^+ and the set of positive definite matrices equals the interior of S_n^+. Therefore, Lemma 5.1 implies that if we work with the set $\mathcal{F}_1$ in an SDP-relaxation, then the Slater condition is necessarily violated, and we already saw that this is disadvantageous both from a theoretical perspective (strong duality may not hold) and from a practical perspective (solvers may not be able to handle the problem). As an alternative, we propose the set

$$\mathcal{F}_2 := \{(X,x) \mid X \succeq 0,\ \operatorname{diag}(X) = x,\ a^T x = a_0,\ Xa = a_0 x\}.$$

A formulation similar to $\mathcal{F}_2$ appeared in Burer, 2010, Proposition 1. We next show that the two sets are actually equal:

Lemma 5.2. *We have that $\mathcal{F}_1 = \mathcal{F}_2$.*

Proof. We first take $(X,x) \in \mathcal{F}_1$. Then $X \succeq 0$, $\operatorname{diag}(X) = x$ and $a^T x = a_0$. Lemma 5.1 shows that $Xa = a_0 x$ and therefore $(X,x) \in \mathcal{F}_2$.

Conversely, let $(X,x) \in \mathcal{F}_2$. We immediately get that $a^T x = a_0$ and $a^T X a = a_0^2$. It remains to show that

$$Y := \begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} = \begin{pmatrix} X & \frac{1}{a_0} Xa \\ \left(\frac{1}{a_0} Xa\right)^T & 1 \end{pmatrix} \succeq 0.$$

Let $Xv = 0$. Then $w := \binom{v}{0}$ is in the null-space of Y, as $Yw = \begin{pmatrix} Xv \\ \frac{1}{a_0} a^T Xv \end{pmatrix} = 0$. This shows that the null-space of X (extended with an additional component equal to 0) is contained in the null-space of Y. We also have $Y\binom{a}{-a_0} = 0$, so that X and Y have the same rank. If $X \succeq 0$ then all nonzero eigenvalues of Y are positive by the interlacing property between the eigenvalues of X and Y. □

Remark 5.3. The matrix Y above is sometimes called a *flat extension* of X. It is a well known fact that flat extensions of semidefinite matrices are also semidefinite.

Next we investigate the situation where a linear term $a^T x$ is required to be contained in some interval, say $|a^T x| \leq a_0$, with $a_0 > 0$. The constraints $X - xx^T \succeq 0$ and $\operatorname{diag}(X) = x$ imply

$$a^T X a \geq (a^T x)^2.$$

We conclude that $a_0^2 \geq a^T X a$ is at least as strong as the original inequality $|a^T x| \geq a_0$.

Finally, we observe that the situation is different for one-sided linear inequalities of the form

$$a^T x \geq a_0$$

with $a_0 > 0$. Arguing as before we see that

$$a^T X a \geq (a^T x)^2,$$

hence the original inequality $a^T x \geq a_0$ is at least as strong as $a^T X a \geq a_0^2$.

5.2. Stable set and coloring relaxations

We recall the formulation for the stability number $\alpha(G)$ for a graph G:

$$\alpha(G) := \max\{e^T x \mid x_i x_j = 0 \text{ for all } [i,j] \in E(G),\ x \in \{0,1\}^n\}.$$

One of its first relaxations using SDP was introduced in a seminal paper by Lovász (1979). It may be derived from nonzero binary vectors x by introducing $X := \frac{1}{x^T x} xx^T$. Note that $e^T x = e^T X e$ holds for any feasible x. Let

$$\vartheta(G) := \max\{e^T X e \mid X \succeq 0,\ x_{ij} = 0 \text{ for all } [i,j] \in E(G)\}.$$

As any characteristic vector x of a stable set leads to a feasible matrix X for this problem, it is clear that $\alpha(G) \leq \vartheta(G)$. Lovász and Schrijver (1991) showed that $\vartheta(G)$ can also be obtained as the optimal value of the following SDP:

$$\vartheta(G) = \max\{e^T x \mid x = \operatorname{diag}(X),\ X_{ij} = 0 \,\forall [i,j] \in E(G),\ \begin{pmatrix} X & x \\ x^T & 1 \end{pmatrix} \succeq 0\}.$$

We now return to the parameter $z(k)$ from (13) and derive upper bounds on $\alpha(G)$ based on it. We denote by A the adjacency matrix of our graph G. For given $k \in \mathbb{N}$ we have

$$z(k) := \min \tfrac{1}{2} x^T A x \text{ such that } e^T x = k \text{ and } x \in \{0,1\}^n.$$

If $z(k) > 0$, then clearly G has no stable set of size k and therefore $\alpha(G) < k$. Computing $z(k)$ is an NP-complete problem, see for instance (Billionnet and Roupin, 2008), so we consider the following tractable relaxation denoted by $P(t)$ for some $t > 1$:

$$P(t) \quad \min \tfrac{1}{2}\langle A, Y\rangle \text{ such that } Y \succeq 0,\ Y \geq 0,\ \operatorname{trace}(Y) = t,\ Ye = t\operatorname{diag}(Y)$$

Note that this is a doubly nonnegative relaxation, since the constraints $Y \succeq 0$, $Y \geq 0$ mean that Y should be in the cone of doubly nonnegative matrices. We denote the optimal value of $P(t)$ by $\operatorname{val}(P(t))$. Note that $\operatorname{val}(P(t)) > 0$ implies $\alpha(G) \leq \lfloor t \rfloor$, so we are interested in finding the smallest t such that $\operatorname{val}(P(t)) > 0$. It turns out that the answer to this question is closely related to Schrijver's refinement $\vartheta^+ = \vartheta^+(G)$ (sometimes denoted by $\vartheta'(G)$) of the original theta function $\vartheta(G)$:

$$\begin{aligned} \vartheta^+ := \max\ & \operatorname{trace}(Y) \\ \text{s.t.}\ & Y - yy^T \succeq 0,\ \operatorname{diag}(Y) = y, \\ & y_{ij} = 0 \text{ for all } [i,j] \in E(G),\ Y \geq 0. \end{aligned} \tag{S_1}$$

Before we establish this connection we recall the following alternative formulation of ϑ^+:

$$\begin{aligned} \vartheta^+ = \max\ & \langle J, X\rangle \\ \text{s.t.}\ & X \succeq 0,\ \operatorname{trace}(X) = 1, x_{ij} = 0 \text{ for all } [i,j] \in E(G), X \geq 0. \end{aligned} \tag{S_2}$$

We set $n := |V(G)|$ and $m := |E(G)|$. The next two theorems can be used to get bounds for $\alpha(G)$ and $\chi(G)$. Contrary to the computation of $\vartheta^+(G)$ which requires the solution of an SDP with more than m equality constraints, the SDP relaxation $P(t)$ contains $n+1$ equality constraints,

independent of m. As a drawback, one has to guess the proper value t which might require solving several SDPs for different values of t.

The following property of optimal solutions to (S_1) will be used later on.

Lemma 5.4. *Let (Y^*, y^*) be optimal for (S_1). Then $Y^*e = \vartheta^+ y^*$.*

Proof. We first note that $X^* := \frac{1}{\vartheta^+} Y^*$ is feasible for (S_2) and therefore $\langle J, X^*\rangle \le \vartheta^+$. This means that $\langle J, Y^*\rangle \le (\vartheta^+)^2$. On the other hand, $Y^* - y^*(y^*)^T \succeq 0$ so that $\langle J, Y^*\rangle \ge (e^T y^*)^2 = (\vartheta^+)^2$. Thus we have shown that $\langle J, Y^*\rangle = (\vartheta^+)^2$. This implies that $e^T(Y^* - y^*(y^*)^T)e = 0$ and together with the semidefiniteness of the matrix $Y^* - y^*(y^*)^T$ we get $(Y^* - y^*(y^*)^T)e = 0$, showing that $Y^*e = \vartheta^+ y^*$. □

We are now ready to show the following result.

Theorem 5.5. *Let A be the adjacency matrix of a graph G. Then* $\mathrm{val}(P(t)) > 0$ *if and only if* $t > \vartheta^+(G)$.

Proof. We first consider the problem $P(t)$ for the value $t = \vartheta^+ := \vartheta^+(G)$:

$$P(\vartheta^+) \quad \min \tfrac{1}{2}\langle A, Y\rangle \text{ s.t. } Y \succeq 0,\ Y \ge 0,\ \mathrm{trace}(Y) = \vartheta^+,\ Ye = \vartheta^+ \mathrm{diag}(Y).$$

Now take an optimal solution (Y, y) for problem (S_1), so $\mathrm{trace}(Y) = \vartheta^+$, $Y - yy^T \succeq 0$, $y_{ij} = 0$ for all $[i,j] \in E(G)$, $y_{ij} \ge 0$ for all $[i,j] \notin E(G)$. Lemma 5.4 shows us that $Ye = \vartheta^+ y$. We conclude that (Y, y) is feasible for problem $P(\vartheta^+)$. Since $y_{ij} = 0$ on $E(G)$ we conclude that $\mathrm{val}(P(\vartheta^+)) = 0$.

Next, suppose that $\mathrm{val}(P(t)) = 0$ and consider t' with $1 < t' < t$. Suppose that (Y, y) is optimal for $P(t)$. It is a simple exercise to verify that

$$Y' := \frac{t'}{t(t-1)}\left[(t'-1)Y + (t-t')\mathrm{diag}(Y)\right]$$

is feasible for $P(t')$ with objective value 0, hence $\mathrm{val}(P(t)) = 0$ for all $1 < t \le \vartheta^+$.

Finally, the definition of problem (S_1) shows that ϑ^+ is the largest possible value for the trace of a matrix Y which satisfies $Y \succeq 0$, $Y \ge 0$, $y_{ij} = 0$ for all $[i,j] \in E(G)$, and therefore $\mathrm{val}(P(t)) > 0$ for any $t > \vartheta^+$. □

A similar approach can also be used to get lower bounds for the chromatic number χ. We recall the binary quadratic problem from (14):

$$z(k) := \min \tfrac{1}{2}\langle X, AX\rangle \text{ such that } Xe = e,\ X \in \{0,1\}^{n\times k}.$$

If $z(k) > 0$ for some given k, then $\chi(G) > k$. As before, we need a tractable relaxation for this problem. It is obtained by first extending X with an additional row of all ones, so we introduce

$$\tilde{X} := \begin{pmatrix} X \\ e^T \end{pmatrix}$$

and observe that

$$\tilde{X}\tilde{X}^T = \begin{pmatrix} XX^T & e \\ e^T & k \end{pmatrix}.$$

The main diagonal of the matrix XX^T clearly equals the all-ones vector, because each row of the 0-1 matrix X has exactly one entry equal to 1. A relaxation is obtained by allowing arbitrary matrices Y instead of XX^T. We get

$$(P(t)) \quad \min \tfrac{1}{2}\langle A, Y\rangle \text{ such that } \begin{pmatrix} Y & e \\ e^T & t \end{pmatrix} \succeq 0,\ \mathrm{diag}(Y) = e,\ Y \ge 0.$$

From $\mathrm{val}(P(t)) > 0$ we may conclude that $\chi(G) \ge \lceil t \rceil$. Thus we would like to find the largest value t such that $\mathrm{val}(P(t)) > 0$. It turns out that the strengthening of the ϑ number towards the chromatic number $\chi(G)$ proposed by Szegedy (1994) provides the answer to this question. The parameter $\vartheta^-(\overline{G})$ defined for the complement $\overline{G}$ of G was introduced by Szegedy as a lower bound on the chromatic number of G:

$$\vartheta^-(\overline{G}) := \min\ t \quad \text{such that } \begin{pmatrix} Y & e \\ e^T & t \end{pmatrix} \succeq 0,\ \mathrm{diag}(Y) = e,$$
$$Y \ge 0,\ y_{ij} = 0 \text{ for all } [i,j] \in E(\overline{G}).$$

We have $\vartheta(\overline{G}) \le \vartheta^-(\overline{G}) \le \chi(G)$.

Theorem 5.6. *For a given graph G we have that*

$\mathrm{val}(P(t)) > 0$ if and only if $t < \vartheta^-(\overline{G})$.

Proof. We first consider problem $P(t)$ for $t = \vartheta^-(\overline{G})$. Let Y be an optimal solution for $\vartheta^-(\overline{G})$. Then Y is also feasible for $P(\vartheta^-(\overline{G}))$ with value 0, so $P(\vartheta^-(\overline{G})) = 0$. The solution Y remains feasible for any $t' > t = \vartheta^-(\overline{G})$ so that $\mathrm{val}(P(t')) = 0$ also in this case. Finally, the definition of $\vartheta^-(\overline{G})$ shows that for any $t < \vartheta^-(\overline{G})$, the system

$$\begin{pmatrix} Y & e \\ e^T & t \end{pmatrix} \succeq 0,\ \mathrm{diag}(Y) = e,\ Y \ge 0,\ y_{ij} = 0 \text{ for all } [i,j] \in E(\overline{G})$$

is infeasible. Therefore, any Y satisfying $\begin{pmatrix} Y & e \\ e^T & t \end{pmatrix} \succeq 0$, $\mathrm{diag}(Y) = e$, $Y \ge 0$ will have an entry $y_{ij} > 0$ for some $[i,j] \in E(\overline{G})$, and hence $\mathrm{val}(P(t)) > 0$. □

6. Conclusions

We have seen that conic optimization is an extremely versatile tool with an abundance of applications. Depending on the cone in question, different complexities may occur: while linear programming, second order cone programming as well as semidefinite programming are solvable in polynomial time, optimizing over the cones of copositive or completely positive matrices is NP-hard.

The cones $\mathcal{COP}$ and $\mathcal{CP}$ are highly useful modelling tools for nonconvex quadratic or combinatorial optimization. In many cases, it is possible to reformulate such problems equivalently as linear problems over $\mathcal{COP}$ or $\mathcal{CP}$. Relaxing the cone constraint to a semidefiniteness or double nonnegativity constraints yields very good bounds which are often provably tighter than LP-based bounds. When using approximation hierarchies, one can often show that some finite level of the hierarchy gives the exact solution of the underlying combinatorial problem.

In this paper, we have discussed various conic approaches to *binary* quadratic problems. It is noteworthy that not much literature is available for conic approaches to nonconvex quadratic problems involving general integer or mixed integer variables, i.e., problems with a variable $x \in \mathbb{R}^p_+ \times \mathbb{Z}^r_+$. This type of problems is studied in Burer and Letchford (2014) and Buchheim and Traversi (2015), however many open questions remain in this area.

In contrast to linear programming, the existence of strictly feasible solutions plays a crucial role in conic optimization. In the absence of strictly feasible points, strong duality may not hold and algorithms may fail to solve the problem. This is therefore a point that should be carefully considered when modelling the problem in question as a conic optimization problem.

At the moment, the main bottleneck for using SDP relaxations or conic optimization in a broader context of applications is the lack of algorithms that can solve large scale problems in reasonable time. Semidefinite relaxations of a problem in $\mathbb{R}^n$ clearly involve matrices of order at least $n \times n$, so the number of variables is roughly squared. If one works with approximation hierarchies, the the SDPs get larger at each level of the hierarchy. Very quickly, these SDPs are out of reach for current computational algorithms.

A possible remedy when the underlying combinatorial problem is highly structured is to exploit the symmetry by using the theory of matrix $*$-algebras. Roughly speaking, the idea is to pre-process the SDP by applying a suitable unitary transformation in such a way that the resulting matrices in the SDP exhibit block diagonal structure. This

structure can then be exploited by interior point methods. We refer to de Klerk (2010) for a survey on this approach, and to Dobre and Vera (2015) for a discussion on how it can be used in approximation hierarchies.

Unfortunately however, not all SDPs exhibit symmetries, and so there is a need for faster algorithms. Maybe in the future other algorithms than interior point methods will turn out to be efficient. First attempts are a semismooth Newton-CG augmented Lagrangian method (Yang et al., 2015) and an augmented Lagrangian method combined with a suitable randomization technique (Yurtsever et al., 2021). The ADMM method has already proved successful when applied to SDP relaxations of binary quadratic problems (Wen et al., 2010), the quadratic assignment problem (Oliveira et al., 2018), or the quadratic shortest path problem (Hu and Sotirov, 2020).

As we have outlined, the past decades have seen an enormous progress in understanding conic problems and using them for modelling purposes. The next decades should be particularly devoted to the numerical side.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We wish to thank the referee for his/her extremely careful reading and highly detailed comments which have greatly improved the presentation, as well as the editors for the efficient handling of our paper.

References

Afonin, A., Hildebrand, R., Dickinson, P.J.C., 2021. The extreme rays of the 6×6 copositive cone. J. Glob. Optim. 79 (1), 153–190. doi:10.1007/s10898-020-00930-y.

Ahmadi, A.A., Majumdar, A., 2019. DSOS and SDSOS optimization: more tractable alternatives to sum of squares and semidefinite optimization. SIAM J. Appl. Algebra Geom. 3 (2), 193–230. doi:10.1137/18M118935X.

Alizadeh, F., Goldfarb, D., 2003. Second-order cone programming. Math. Program. 95 (1, Series B), 3–51. doi:10.1007/s10107-002-0339-5. ISMP 2000, Part 3 (Atlanta, GA)

Aloise, D., Hansen, P., 2009. A branch-and-cut SDP-based algorithm for minimum sum-of-squares clustering. Pesqui. Operacional 29 (3), 503–516. doi:10.1590/S0101-74382009000300002.

Andreani, R., Haeser, G., Viana, D.S., 2020. Optimality conditions and global convergence for nonlinear semidefinite programming. Math. Program. 180 (1-2, Series A), 203–235. doi:10.1007/s10107-018-1354-5.

Anjos, M.F., Emine, Y., Lodi, A., Sun, Z., 2021. A note on the Lasserre hierarchy for different formulations of the maximum independent set problem. Oper. Res. Lett. 49 (1), 30–34. doi:10.1016/j.orl.2020.10.009.

Anjos, M.F., Lasserre, J.B. (Eds.). Handbook on Semidefinite, Conic and Polynomial Optimization, International Series in Operations Research & Management Science, (Vol. 166), 2012. Springer, New York doi:10.1007/978-1-4614-0769-0.

Arima, N., Kim, S., Kojima, M., 2013. A quadratically constrained quadratic optimization model for completely positive cone programming. SIAM J. Optim. 23 (4), 2320–2340. doi:10.1137/120890636.

Atamtürk, A., Narayanan, V., 2007. Cuts for conic mixed-integer programming. In: Integer Programming and Combinatorial Optimization. In: Lecture Notes in Computing Science, 4513. Springer, Berlin, pp. 16–29. doi:10.1007/978-3-540-72792-7_2.

Bai, L., Mitchell, J.E., Pang, J.-S., 2016. On conic QPCCs, conic QCQPs and completely positive programs. Math. Program. 159 (1-2, Series A), 109–136. doi:10.1007/s10107-015-0951-9.

Bao, X., Sahinidis, N.V., Tawarmalani, M., 2011. Semidefinite relaxations for quadratically constrained quadratic programming: a review and comparisons. Math. Program. 129 (1, Series B), 129–157. doi:10.1007/s10107-011-0462-2.

Ben-Tal, A., El Ghaoui, L., Nemirovski, A., 2009. Robust Optimization. Princeton series in applied mathematics. Princeton University Press, Princeton, NJ doi:10.1515/9781400831050.

Ben-Tal, A., Nemirovski, A., 2001. Lectures on Modern Convex Optimization. MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Programming Society (MPS), Philadelphia, PA doi:10.1137/1.9780898718829.

Berman, A., Dür, M., Shaked-Monderer, N., 2015. Open problems in the theory of completely positive and copositive matrices. Electron. J. Linear Algebra 29, 46–58. doi:10.13001/1081-3810.2943.

Berman, A., Shaked-Monderer, N., 2003. Completely Positive Matrices. World Scientific Publishing Co., Inc., River Edge, NJ doi:10.1142/9789812795212.

Billionnet, A., Roupin, F., 2008. A deterministic approximation algorithm for the densest k-subgraph problem. Int. J. Oper. Res. 3 (3), 301–314. doi:10.1504/IJOR.2008.017534.

Bomze, I.M., 2015. Copositive relaxation beats Lagrangian dual bounds in quadratically and linearly constrained quadratic optimization problems. SIAM J. Optim. 25 (3), 1249–1275. doi:10.1137/140987997.

Bomze, I.M., Cheng, J., Dickinson, P.J.C., Lisser, A., 2017. A fresh CP look at mixed-binary QPs: new formulations and relaxations. Math. Program. 166 (1-2, Series A), 159–184. doi:10.1007/s10107-017-1109-8.

Bomze, I.M., De Klerk, E., 2002. Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. J. Glob. Optim. 24 (2), 163–185. doi:10.1023/A:1020209017701.

Bomze, I.M., Dür, M., de Klerk, E., Roos, C., Quist, A.J., Terlaky, T., 2000. On copositive programming and standard quadratic optimization problems. J. Glob. Optim. 18 (4), 301–320. doi:10.1023/A:1026583532263.

Bomze, I.M., Jarre, F., 2010. A note on Burer's copositive representation of mixed-binary QPs. Optim. Lett. 4 (3), 465–472. doi:10.1007/s11590-010-0174-1.

Borwein, J., Wolkowicz, H., 1981a. Facial reduction for a cone-convex programming problem. J. Aust. Math. Soc. Ser. A 30 (3), 369–380.

Borwein, J., Wolkowicz, H., 1981b. Regularizing the abstract convex program. J. Math. Anal. Appl. 83 (2), 495–530. doi:10.1016/0022-247X(81)90138-4.

Buchheim, C., Traversi, E., 2015. On the separation of split inequalities for non-convex quadratic integer programming. Discret. Optim. 15, 1–14. doi:10.1016/j.disopt.2014.08.002.

Bundfuss, S., Dür, M., 2009. An adaptive linear approximation algorithm for copositive programs. SIAM J. Optim. 20 (1), 30–53. doi:10.1137/070711815.

Burer, S., 2009. On the copositive representation of binary and continuous nonconvex quadratic programs. Math. Program. 120 (2, Series A), 479–495. doi:10.1007/s10107-008-0223-z.

Burer, S., 2010. Optimizing a polyhedral-semidefinite relaxation of completely positive programs. Math. Program. Comput. 2 (1), 1–19.

Burer, S., 2012. Copositive programming. In: Handbook on Semidefinite, Conic and Polynomial Optimization. In: International Series in Operations Research and Management Science, 166. Springer, New York, pp. 201–218. doi:10.1007/978-1-4614-0769-0_8.

Burer, S., Dong, H., 2012. Representing quadratically constrained quadratic programs as generalized copositive programs. Oper. Res. Lett. 40 (3), 203–206. doi:10.1016/j.orl.2012.02.001.

Burer, S., Letchford, A.N., 2014. Unbounded convex sets for non-convex mixed-integer quadratic programming. Math. Program. 143 (1-2, Series A), 231–256. doi:10.1007/s10107-012-0609-9.

Chen, J.-S., Pan, S., Ko, C.-H., 2011. A continuation approach for the capacitated multifacility Weber problem based on nonlinear SOCP reformulation. J. Glob. Optim. 50 (4), 713–728. doi:10.1007/s10898-010-9632-7.

Chvátal, V., 1979. A greedy heuristic for the set-covering problem. Math. Oper. Res. 4 (3), 233–235. doi:10.1287/moor.4.3.233.

DeCorte, E., de Oliveira Filho, F.M., Vallentin, F., 2021. Complete positivity and distance-avoiding sets. Math. Program. doi:10.1007/s10107-020-01562-6, OnlineFirst.

Dickinson, P.J.C., 2011. Geometry of the copositive and completely positive cones. J. Math. Anal. Appl. 380 (1), 377–395. doi:10.1016/j.jmaa.2011.03.005.

Dickinson, P.J.C., Eichfelder, G., Povh, J., 2013. Erratum to: On the set-semidefinite representation of nonconvex quadratic programs over arbitrary feasible sets. Optim. Lett. 7 (6), 1387–1397. doi:10.1007/s11590-013-0645-2.

Dickinson, P.J.C., Gijben, L., 2014. On the computational complexity of membership problems for the completely positive cone and its dual. Comput. Optim. Appl. 57 (2), 403–415. doi:10.1007/s10589-013-9594-z.

Dobre, C., Dür, M., Frerick, L., Vallentin, F., 2016. A copositive formulation for the stability number of infinite graphs. Math. Program. 160 (1-2, Series A), 65–83. doi:10.1007/s10107-015-0974-2.

Dobre, C., Vera, J., 2015. Exploiting symmetry in copositive programs via semidefinite hierarchies. Math. Program. 151 (2, Series B), 659–680. doi:10.1007/s10107-015-0879-0.

Dong, H., 2013. Symmetric tensor approximation hierarchies for the completely positive cone. SIAM J. Optim. 23 (3), 1850–1866. doi:10.1137/100813816.

Drewes, S., Pokutta, S., 2014. Symmetry-exploiting cuts for a class of mixed-0/1 second-order cone programs. Discret. Optim. 13, 23–35. doi:10.1016/j.disopt.2014.04.002.

Drewes, S., Ulbrich, S., 2012. Subgradient based outer approximation for mixed integer second order cone programming. In: Mixed Integer Nonlinear Programming. In: The IMA Volumes in Mathematics and its Applications, 154. Springer, New York, pp. 41–59.

Dür M., 2010. Copositive programming – a survey. In: Diehl, M., Glineur, F., Jarlebring, E., Michiels, W. (Eds.), Recent Advances in Optimization and its Applications in Engineering. Springer.

Dür, M., Jargalsaikhan, B., Still, G., 2017. Genericity results in linear conic programming—a tour d'horizon. Math. Oper. Res. 42 (1), 77–94. doi:10.1287/moor.2016.0793.

Eichfelder, G., Jahn, J., 2008. Set-semidefinite optimization. J. Convex Anal. 15 (4), 767–801.

Eichfelder, G., Povh, J., 2013. On the set-semidefinite representation of nonconvex quadratic programs over arbitrary feasible sets. Optim. Lett. 7 (6), 1373–1386. doi:10.1007/s11590-012-0450-3.

Escoffier, B., Hammer, P.L., 2007. Approximation of the quadratic set cover problem. Discret. Optim. 4 (3), 378–386. doi:10.1016/j.disopt.2007.10.001.

Faybusovich, L., 2008. Several Jordan-algebraic aspects of optimization. Optimization 57 (3), 379–393. doi:10.1080/02331930701523510.

Fortunato, S., 2010. Community detection in graphs. Phys. Rep. 486 (3-5), 75–174. doi:10.1016/j.physrep.2009.11.002.

Gally, T., Pfetsch, M.E., Ulbrich, S., 2018. A framework for solving mixed-

integer semidefinite programs. Optim. Methods Softw. 33 (3), 594–632. doi:10.1080/10556788.2017.1322081.
Goemans, M.X., Williamson, D.P., 1995. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. J. Assoc. Comput. Mach. 42 (6), 1115–1145. doi:10.1145/227683.227684.
Góez, J.C., Anjos, M.F., 2019. Second-order cone optimization formulations for service system design problems with congestion. In: Modeling and Optimization: Theory and Applications. In: Springer Proceedings in Mathematics & Statistics, 279. Springer, Cham, pp. 97–120. doi:10.1007/978-3-030-12119-8_5.
Gouveia, J., Pong, T.K., Saee, M., 2020. Inner approximating the completely positive cone via the cone of scaled diagonally dominant matrices. J. Glob. Optim. 76 (2), 383–405. doi:10.1007/s10898-019-00861-3.
Gvozdenović, N., Laurent, M., 2008. The operator Ψ for the chromatic number of a graph. SIAM J. Optim. 19 (2), 572–591. doi:10.1137/050648237.
Håstad, J., 1999. Clique is hard to approximate within $n^{1-\epsilon}$. Acta Math. 182 (1), 105–142. doi:10.1007/BF02392825.
Helmberg, C., 2002. Semidefinite programming. Eur. J. Oper. Res. 137 (3), 461–482. doi:10.1016/S0377-2217(01)00143-6.
Helmberg, C., Overton, M.L., Rendl, F., 2014. The spectral bundle method with second-order information. Optim. Methods Softw. 29 (4), 855–876. doi:10.1080/10556788.2013.858155.
Helmberg, C., Rendl, F., 2000. A spectral bundle method for semidefinite programming. SIAM J. Optim. 10 (3), 673–696. doi:10.1137/S1052623497328987.
Hochbaum, D.S., 1982. Approximation algorithms for the set covering and vertex cover problems. SIAM J. Comput. 11 (3), 555–556. doi:10.1137/0211045.
Hu, H., Sotirov, R., 2020. On solving the quadratic shortest path problem. INFORMS J. Comput. 32 (2), 219–233.
Jarre, F., 2012. Burer's key assumption for semidefinite and doubly nonnegative relaxations. Optim. Lett. 6 (3), 593–599. doi:10.1007/s11590-010-0269-8.
de Klerk, E., 2010. Exploiting special structure in semidefinite programming: a survey of theory and applications. Eur. J. Oper. Res. 201 (1), 1–10. doi:10.1016/j.ejor.2009.01.025.
de Klerk, E., Pasechnik, D.V., 2002. Approximation of the stability number of a graph via copositive programming. SIAM J. Optim. 12 (4), 875–892. doi:10.1137/S1052623401383248.
de Klerk, E., Roos, C., Terlaky, T., 1997. Initialization in semidefinite programming via a self-dual skew-symmetric embedding. Oper. Res. Lett. 20 (5), 213–221. doi:10.1016/S0167-6377(97)00011-4.
Kobayashi, K., Takano, Y., 2020. A branch-and-cut algorithm for solving mixed-integer semidefinite optimization problems. Comput. Optim. Appl. 75 (2), 493–513. doi:10.1007/s10589-019-00153-2.
Kocvara, M., Stingl, M., 2012. PENNON: software for linear and nonlinear matrix inequalities. In: Handbook on Semidefinite, Conic and Polynomial Optimization. In: International Series in Operations Research and Management Science, 166. Springer, New York, pp. 755–791. doi:10.1007/978-1-4614-0769-0_26.
Kuang, X., Zuluaga, L.F., 2018. Completely positive and completely positive semidefinite tensor relaxations for polynomial optimization. J. Glob. Optim. 70 (3), 551–577. doi:10.1007/s10898-017-0558-1.
Kuryatnikova, O., Vera, J. C., 2018. Positive semidefinite approximations to the cone of copositive kernels. ArXiv Preprint No. 1812.00274.
Kuryatnikova, O., Vera Lizcano, J.C., 2017. Approximating the cone of copositive kernels to estimate the stability number of infinite graphs. In: LAGOS'17—IX Latin and American Algorithms, Graphs and Optimization. In: Electronic Notes in Discrete Mathematics, 62. Elsevier Science B.V., Amsterdam, pp. 303–308. doi:10.1016/j.endm.2017.10.052.
Lasserre, J.B., 2010. Moments, Positive Polynomials and Their Applications. Imperial College Press Optimization Series, 1. Imperial College Press, London.
Lasserre, J.B., 2014. New approximations for the cone of copositive matrices and its dual. Math. Program. 144 (1-2, Series A), 265–276. doi:10.1007/s10107-013-0632-5.
Lasserre, J.B., 2016. A MAX-CUT formulation of 0/1 programs. Oper. Res. Lett. 44 (2), 158–164. doi:10.1016/j.orl.2015.12.014.
Laurent, M., 2009. Sums of squares, moment matrices and optimization over polynomials. In: Emerging Applications of Algebraic Geometry. In: The IMA Volumes in Mathematics and its Applications, 149. Springer, New York, pp. 157–270. doi:10.1007/978-0-387-09686-5_7.
Laurent, M., Piovesan, T., 2015. Conic approach to quantum graph parameters using linear optimization over the completely positive semidefinite cone. SIAM J. Optim. 25 (4), 2461–2493. doi:10.1137/14097865X.
Li, Y., Xie, W., 2020. Exact and approximation algorithms for sparse PCA. ArXiv Preprint No. 2008.12438.
Lisser, A., Rendl, F., 2003. Graph partitioning using linear and semidefinite programming. Math. Program. 95 (1, Series B), 91–101. doi:10.1007/s10107-002-0342-x.
Lobo, M.S., Vandenberghe, L., Boyd, S., Lebret, H., 1998. Applications of second-order cone programming. Linear Algebra Appl. 284 (1-3), 193–228. doi:10.1016/S0024-3795(98)10032-0.
Lovász, L., 1979. On the Shannon capacity of a graph. IEEE Trans. Inf. Theory 25 (1), 1–7. doi:10.1109/TIT.1979.1055985.
Lovász, L., Schrijver, A., 1991. Cones of matrices and set-functions and 0-1 optimization. SIAM J. Optim. 1 (2), 166–190. doi:10.1137/0801013.
Lund, C., Yannakakis, M., 1994. On the hardness of approximating minimization problems. J. Assoc. Comput. Mach. 41 (5), 960–981. doi:10.1145/185675.306789.
Malick, J., Roupin, F., 2012. Solving k-cluster problems to optimality with semidefinite programming. Math. Program. 136 (2, Series B), 279–300. doi:10.1007/s10107-012-0604-1.
Maxfield, J.E., Minc, H., 1962/63. On the matrix equation $X'X = A$. Proc. Edinburgh Math. Soc. (2) 13, 125–129. doi:10.1017/S0013091500014681.
Mittelmann, H.D., 2021. Decision tree for optimization software . http://plato.asu.edu/guide.html.
Mosek, 2021-10-11. The Mosek modeling cookbook . http://docs.mosek.com/MOSEKModelingCookbook-a4paper.pdf.
Motzkin, T.S., Straus, E.G., 1965. Maxima for graphs and a new proof of a theorem of Turán. Can. J. Math. 17, 533–540. doi:10.4153/CJM-1965-053-6.
Murty, K.G., Kabadi, S.N., 1987. Some NP-complete problems in quadratic and nonlinear programming. Math. Program. 39 (2), 117–129. doi:10.1007/BF02592948.
Nesterov, Y., Nemirovski, A., 1992. Conic formulation of a convex programming problem and duality. Optim. Methods Softw. 1 (2), 95–115. doi:10.1080/10556789208805510.
Nesterov, Y., Nemirovskii, A., 1994. Interior-Point Polynomial Algorithms in Convex Programming. SIAM Studies in Applied Mathematics, 13. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA doi:10.1137/1.9781611970791.
Oliveira, D.E., Wolkowicz, H., Xu, Y., 2018. ADMM for the SDP relaxation of the QAP. Math. Program. Comput. 10 (4), 631–658. doi:10.1007/s12532-018-0148-3.
Pardalos, P.M., Rendl, F., Wolkowicz, H., 1994. The quadratic assignment problem: a survey and recent developments. In: Quadratic Assignment and Related Problems (New Brunswick, NJ, 1993). In: DIMACS Series Discrete Mathematics and Theoretical Computer Science, 16. American Mathematical Society, Providence, RI, pp. 1–42. doi:10.1090/dimacs/016/01.
Pardalos, P.M., Vavasis, S.A., 1991. Quadratic programming with one negative eigenvalue is NP-hard. J. Glob. Optim. 1 (1), 15–22. doi:10.1007/BF00120662.
Parrilo, P., 2000. Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization. California Institute of Technology Ph.D. thesis.
Pataki, G., 2013. Strong duality in conic linear programming: facial reduction and extended duals. In: Computational and Analytical Mathematics. In: Springer Proceedings in Mathematics & Statistics, 50. Springer, New York, pp. 613–634. doi:10.1007/978-1-4614-7621-4_28.
Pataki, G., 2019. Characterizing bad semidefinite programs: normal forms and short proofs. SIAM Rev. 61 (4), 839–859. doi:10.1137/17M1140844.
Peña, J., Vera, J., Zuluaga, L.F., 2007. Computing the stability number of a graph via linear and semidefinite programming. SIAM J. Optim. 18 (1), 87–105. doi:10.1137/05064401X.
Peña, J., Vera, J.C., Zuluaga, L.F., 2015. Completely positive reformulations for polynomial optimization. Math. Program. 151 (2, Series B), 405–431. doi:10.1007/s10107-014-0822-9.
Pólya, G., 1928. Über positive Darstellung von Polynomen. Vierteljahresschr. Naturforschenden Ges. Zürich 73, 141–145.
Povh, J., Rendl, F., 2009. Copositive and semidefinite relaxations of the quadratic assignment problem. Discret. Optim. 6 (3), 231–241. doi:10.1016/j.disopt.2009.01.002.
Preisig, J.C., 1996. Copositivity and the minimization of quadratic functions with nonnegativity and quadratic equality constraints. SIAM J. Control Optim. 34 (4), 1135–1150. doi:10.1137/S0363012993251894.
Quist, A.J., De Klerk, E., Roos, C., Terlaky, T., 1998. Copositive relaxation for general quadratic programming. Optim. Methods Softw. 9 (1-3), 185–208. doi:10.1080/10556789808805692.
Ramana, M.V., 1997. An exact duality theory for semidefinite programming and its complexity implications. Math. Program. 77 (2, Series B), 129–162. doi:10.1016/S0025-5610(96)00082-2.
Renegar, J., 2001. A Mathematical View of Interior-Point Methods in Convex Optimization. MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Programming Society (MPS), Philadelphia, PA doi:10.1137/1.9780898718812.
Schrijver, A., 1979. A comparison of the Delsarte and Lovász bounds. IEEE Trans. Inf. Theory 25 (4), 425–429. doi:10.1109/TIT.1979.1056072.
Shor, N.Z., 1987. Quadratic optimization problems. Soviet J. Comput. Syst. Sci. 25, 111.
Szegedy, M., 1994. A note on the θ number of Lovász and the generalized Delsarte bound. In: Proceedings of the 35th Annual Symposium on Foundations of Computer Science, 20-22 November 1994, Santa Fe, New Mexico, USA. IEEE, pp. 36–39. doi:10.1109/SFCS.1994.365707.
Vandenberghe, L., Boyd, S., 1996. Semidefinite programming. SIAM Rev. 38 (1), 49–95. doi:10.1137/1038003.
Vinel, A., Krokhmal, P.A., 2014. Polyhedral approximations in p-order cone programming. Optim. Methods Softw. 29 (6), 1210–1237. doi:10.1080/10556788.2013.877905.
Wen, Z., Goldfarb, D., Yin, W., 2010. Alternating direction augmented Lagrangian methods for semidefinite programming. Math. Program. Comput. 2 (3-4), 203–230. doi:10.1007/s12532-010-0017-1.
Xia, W., Zuluaga, L.F., 2017. Completely positive reformulations of polynomial optimization problems with linear constraints. Optim. Lett. 11 (7), 1229–1241. doi:10.1007/s11590-017-1123-z.
Yamashita, H., Yabe, H., 2015. A survey of numerical methods for nonlinear semidefinite programming. J. Oper. Res. Soc. Jpn. 58 (1), 24–60. doi:10.15807/jorsj.58.24.
Yang, L., Sun, D., Toh, K.-C., 2015. SDPNAL+: a majorized semismooth Newton-CG augmented Lagrangian method for semidefinite programming with nonnegative constraints. Math. Program. Comput. 7 (3), 331–366. doi:10.1007/s12532-015-0082-6.
Yanıkoğlu, I., Gorissen, B.L., den Hertog, D., 2019. A survey of adjustable robust optimization. Eur. J. Oper. Res. 277 (3), 799–813. doi:10.1016/j.ejor.2018.08.031.
Yıldırım, E.A., 2012. On the accuracy of uniform polyhedral approximations of the copositive cone. Optim. Methods Softw. 27 (1), 155–173. doi:10.1080/10556788.2010.540014.
Yıldırım, E.A., 2017. Inner approximations of completely positive reformulations of mixed binary quadratic programs: a unified analysis. Optim. Methods Softw. 32 (6), 1163–1186. doi:10.1080/10556788.2016.1245732.
Yurtsever, A., Tropp, J.A., Fercoq, O., Udell, M., Cevher, V., 2021. Scalable semidefinite programming. SIAM J. Math. Data Sci. 3 (1), 171–200. doi:10.1137/19M1305045.