THE MILP ROAD TO MIQCP

SAMUEL BURER* AND ANUREET SAXENA†

Abstract. This paper surveys results on the NP-hard mixed-integer quadratically constrained programming problem. The focus is strong convex relaxations and valid inequalities, which can become the basis of efficient global techniques. In particular, we discuss relaxations and inequalities arising from the algebraic description of the problem as well as from dynamic procedures based on disjunctive programming. These methods can be viewed as generalizations of techiniques for mixed-integer linear programming. We also present brief computational results to indicate the strength and computational requirements of these methods.

1. Introduction. More than fifty years have passed since Dantzig et. al. [25] solved the 50-city travelling salesman problem. An achievement in itself at the time, their seminal paper gave birth to one of the most succesful disciplines in computational optimization, Mixed Integer Linear Programming (MILP). Five decades of wonderful research, both theoretical and computational, have brought mixed integer programming to a stage where it can solve many if not all MILPs arising in practice (see [43]). The ideas discovered during the course of this development have naturally influenced other disciplines. Constraint programming, for instance, has adopted and refined many of the ideas from MILP to solve more general classes of problems [2].

Our focus in this paper is to track the influence of MILP in solving mixed integer quadratically constrained problems (MIQCP). In particular, we survey some of the recent research on MIQCP and establish their connections to well known ideas in MILP. The purpose of this is two-fold. First, it helps to catalog some of the recent results in a form that is accessible to a researcher with reasonable background in MILP. Second, it defines a roadmap for further research in MIQCP; although significant progress has been made in the field of MIQCP, the "breakthrough" results are yet to come and we believe that the past of MILP holds the clues to the future of MIQCP.

Specifically, we focus on the following mixed integer quadratically constrained problem

$$\begin{array}{ll} \min & x^T C x + c^T x \\ \text{s.t.} & x \in \mathcal{F} \end{array} \qquad (\mathbf{MIQCP})$$

*Department of Management Sciences, University of Iowa, Iowa City, IA 52242-1994, USA (samuel-burer@uiowa.edu). Author supported in part by NSF Grant CCF-0545514.

† Axioma Inc., 8800 Roswell Road, Building B. Suite 295, Atlanta GA, 30350, USA (asaxena@axiomainc.com).

J. Lee and S. Leyffer (eds.), *Mixed Integer Nonlinear Programming*, The IMA Volumes in Mathematics and its Applications 154, DOI 10.1007/978-1-4614-1927-3_13,

where

$$\mathcal{F} := \left\{ x \in \mathbb{R}^n \; : \; \begin{array}{rl} x^T A_k x + a_k^T x \le b_k & \forall\, k = 1, \ldots, m \\ l \le x \le u & \\ x_i \in \mathbb{Z} & \forall\, i \in I \end{array} \right\}.$$

The data of (**MIQCP**) is

- $(C, c) \in \mathcal{S}^n \times \mathbb{R}^n$
- $(A_k, a_k, b_k) \in \mathcal{S}^n \times \mathbb{R}^n \times \mathbb{R}$ for all $k = 1, \ldots, m$
- $(l, u) \in (\mathbb{R} \cup \{-\infty\})^n \times (\mathbb{R} \cup \{+\infty\})^n$
- $I \subseteq [n]$

where, in particular, S^n is the set of all $n \times n$ symmetric matrices and $[n] := \{1, \ldots, n\}$. Without loss of generality, we assume $l < u$, and for all $i \in I$, $(l_i, u_i) \in (\mathbb{Z} \cup \{-\infty\}) \times (\mathbb{Z} \cup \{+\infty\})$. Note that any specific lower or upper bound may be infinite.

If all $A_k = 0$, then (**MIQCP**) reduces to MILP. So (**MIQCP**) is NP-hard. In fact, the continuous variant of MIQCP, namely a non-convex QCP, is already NP-hard and a well-known problem in global optimization [45, 46]. The computational intractability of MIQCP is quite notorious and can be traced to the result of Jeroslow [30] from the seventies that shows that the variant of MIQCP without explicit non-infinite lower/upper bounds on some of the varibles is undecidable. (**MIQCP**) is itself a special case of mixed integer nonlinear programming (MINLP); we refer the reader to the website *MINLP World* [41] for surveys, software, and test instances for MINLP. We also note that any polynomial optimization problem may be reduced to (**MIQCP**) by the introduction of auxiliary variables and constraints to reduce all polynomial degrees to 2, e.g., a cubic term $x_1x_2x_3$ could be modeled as x_1X_{23} with $X_{23} = x_2x_3$.

Note that if the objective function and constraints in MIQCP are convex, then the resulting optimization problem can be solved using standard techniques for solving convex MINLP (see [16] for more details). Most of the ideas and methods discussed in this paper specifically exploit the non-convex quadratic nature of the objective and constraints of (**MIQCP**). In fact, our viewpoint is that many ideas from the solution of MILPs can be adapted in interesting ways for the study of (**MIQCP**). In this sense, we view (**MIQCP**) as a natural progression from MILP rather than, say, a special case of MINLP.

We are also not specifically concerned with the global optimization of (**MIQCP**). Rather, we focus on generating strong convex relaxations and valid inequalities, which could become the basis of efficient global techniques.

In Section 2, we review the idea of *lifting*, which is commonly used to convexify (**MIQCP**) and specifically the feasible set $\mathcal{F}$. We then discuss the generation of various types of linear, second-order-cone, and semidefinite valid inequalities which strengthen the convexification. These inequalities have the property that they arise directly from the algebraic form of $\mathcal{F}$.

In this sense, they generalize the basic LP relaxation often used in MILP. We also catalog several known and new results establishing the strength of these inequalities for certain specifications of $\mathcal{F}$. Then, in Section 3, we describe several related approaches that shed further light on convex relaxations of (**MIQCP**).

In Section 4, we discuss methods for dynamically generating valid inequalities, which can further improve the relaxations. One of the fundamental tools is that of disjunctive programming, which has been used in the MILP community for five decades. However, the disjunctions employed herein are new in the sense that they truly exploit the quadratic form of (**MIQCP**). Recently, Belotti [11] studies disjunctive cuts for general MINLP.

Finally, in Section 5, we consider a short computational study to give some sense of the computational effort and effect of the methods surveyed in this paper.

1.1. Notation and terminology. Most of the notation used in this paper is standard. We define here just a few perhaps atypical notations. For symmetric matrices A and B of conformable dimensions, we define $\langle A, B\rangle = tr(AB)$; a standard fact is that the quadratic form $x^T Ax$ can be represented as $\left\langle A, xx^T\right\rangle$. For a set P in the space of variables (x, y), $\text{proj}_x(P)$ denotes the coordinate projection of P onto the space x. $\text{clconv}\, P$ is the closed convex hull of P. For a square matrix A, $\text{diag}(A)$ denotes the vector of diagonal entries of A. The vector e is the all-ones vector, and e_i is the vector having all zeros except a 1 in position i. The notation $X \succeq 0$ means that X is symmetric positive semidefinite; $X \preceq 0$ means symmetric negative semidefinite.

2. Convex relaxations and valid inequalities. In this section, we describe strong convex relaxations of (**MIQCP**) and $\mathcal{F}$, which arise from the algebraic description of $\mathcal{F}$. For the purposes of presentation, we partition the indices $[m]$ of the quadratic constraints into three groups:

$$\begin{aligned} \text{“linear”} \quad & LQ := \{k : A_k = 0\} \\ \text{“convex quadratic”} \quad & CQ := \{k : A_k \neq 0,\ A_k \succeq 0\} \\ \text{“nonconvex quadratic”} \quad & NQ := \{k : A_k \neq 0,\ A_k \not\succeq 0\}. \end{aligned}$$

For those $k \in CQ$, there exists a rectangular matrix B_k (not necessarily unique) such that $A_k = B_k^T B_k$. Using the B_k's, it is well known that each convex quadratic constraint can be represented as a second-order-cone constraint.

PROPOSITION 2.1 (Alizadeh and Goldfarb [5]). *Let $k \in CQ$ with $A_k = B_k^T B_k$. A point $x \in \mathbb{R}^n$ satisfies $x^T A_k x + a_k^T x \leq b_k$ if and only if*

$$\left\| \begin{pmatrix} B_k x \\ \frac{1}{2}(1 + a_k^T x - b_k) \end{pmatrix} \right\| \leq \frac{1}{2}(1 - a_k^T x + b_k).$$

So $\mathcal{F} \subseteq \Re^n$ may be rewritten as

$$\mathcal{F} = \left\{ x \ : \ \begin{array}{rl} a_k^T x \le b_k & \forall\, k \in LQ \\ \left\| \begin{pmatrix} B_k^T x \\ \frac{1}{2}(a_k^T x - b_k + 1) \end{pmatrix} \right\| \le \frac{1}{2}(1 - a_k^T x + b_k) & \forall\, k \in CQ \\ x^T A_k x + a_k^T x \le b_k & \forall\, k \in NQ \\ l \le x \le u & \\ x_i \in \mathbb{Z} & \forall\, i \in I \end{array} \right\}.$$

If so desired, we can model the bounds $l \le x \le u$ within the linear constraints. However, since bounds often play a special role in approaches for (**MIQCP**), we leave them separate.

2.1. Lifting, convexification, and relaxation. An idea fundamental to many methods for (**MIQCP**) is *lifting* to a higher dimensional space. The simplest lifting idea is to introduce auxiliary variables X_{ij}, which model the quadratic terms $x_i x_j$ via equations $X_{ij} = x_i x_j$ for all $1 \le i, j \le n$. The single symmetric matrix equation $X = xx^T$ also captures this lifting.

As an immediate consequence of lifting, the quadratic objective and constraints may be expressed linearly in (x, X), e.g.,

$$x^T C x + c^T x \quad \overset{X = xx^T}{=} \quad \langle C, X \rangle + c^T x.$$

So (**MIQCP**) becomes

$$\begin{array}{rl} \min & \langle C, X \rangle + c^T x \\ \text{s.t.} & (x, X) \in \widehat{\mathcal{F}} \end{array} \tag{2.1}$$

where

$$\widehat{\mathcal{F}} := \left\{ (x, X) \in \mathbb{R}^n \times \mathcal{S}^n \ : \ \begin{array}{rl} \langle A_k, X \rangle + a_k^T x \le b_k & \forall\, k = 1, \ldots, m \\ l \le x \le u & \\ x_i \in \mathbb{Z} & \forall\, i \in I \\ X = xx^T & \end{array} \right\}.$$

This provides an interesting perspective: the "hard" quadratic objective and constraints of (**MIQCP**) are represented as "easy" linear ones in the space (x, X). The trade-off is the nonconvex equation $X = xx^T$, and of course the non-convex integrality conditions remain.

The linear objective in (2.1) allows convexification of the feasible region without change to the optimal value. From standard convex optimization:

PROPOSITION 2.2. *The problem (2.1), and hence also (*__MIQCP__*), is equivalent to*

$$\min \left\{ \langle C, X \rangle + c^T x : (x, X) \in \operatorname{clconv} \widehat{\mathcal{F}} \right\}.$$

Thus, many lifting approaches may be interpreted as attempting to better understand $\operatorname{clconv}\widehat{\mathcal{F}}$. We adopt this point of view.

A straightforward linear relaxation of $\operatorname{clconv}\widehat{\mathcal{F}}$, which is analogous to the basic linear relaxation of a MILP, is gotten by simply dropping $X = xx^T$ and $x_i \in \mathbb{Z}$:

$$\widehat{\mathcal{L}} := \left\{ (x, X) \in \mathbb{R}^n \times \mathcal{S}^n \ : \ \begin{array}{c} \langle A_k, X\rangle + a_k^T x \le b_k \quad \forall\, k = 1, \ldots, m \\ l \le x \le u \end{array} \right\}.$$

There are many ways to strengthen $\widehat{\mathcal{L}}$ as discussed in the following three subsections.

2.2. Valid linear inequalities. The most common idea for constructing valid linear inequalities for $\operatorname{clconv}\widehat{\mathcal{F}}$ is the following. Let $\alpha^T x \le \alpha_0$ and $\beta^T x \le \beta_0$ be any two valid linear inequalites for $\mathcal{F}$ (possibly the same). Then the quadratic inequality

$$0 \le (\alpha_0 - \alpha^T x)(\beta_0 - \beta^T x) = \alpha_0\beta_0 - \beta_0\, \alpha^T x - \alpha_0\, \beta^T x + x^T \alpha\beta^T x$$

is also valid for $\mathcal{F}$, and so the linear inequality

$$\alpha_0\beta_0 - \beta_0\, \alpha^T x - \alpha_0\, \beta^T x + \left\langle \beta\alpha^T, X \right\rangle \ge 0$$

is valid for $\operatorname{clconv}\widehat{\mathcal{F}}$.

The above idea works with *any* pair of valid inequalities, e.g., ones given explicitly in the description of $\mathcal{F}$ or derived ones. For those explicitly given (the bounds $l \le x \le u$ and the constraints corresponding to $k \in LQ$), the complete list of derived valid quadratic inequalities is

$$\left.\begin{array}{r} (x_i - l_i)(x_j - l_j) \ge 0 \\ (x_i - l_i)(u_j - x_j) \ge 0 \\ (u_i - x_i)(x_j - l_j) \ge 0 \\ (u_i - x_i)(u_j - x_j) \ge 0 \end{array}\right\} \quad \forall \quad (i, j) \in [n] \times [n],\ i \le j \tag{2.2a}$$

$$\left.\begin{array}{r} (x_i - l_i)(b_k - a_k^T x) \ge 0 \\ (u_i - x_i)(b_k - a_k^T x) \ge 0 \end{array}\right\} \quad \forall \quad (i, k) \in [n] \times LQ \tag{2.2b}$$

$$(b_\ell - a_\ell^T x)(b_k - a_k^T x) \ge 0 \ \} \quad \forall \quad (\ell, k) \in LQ \times LQ,\ \ell \le k. \tag{2.2c}$$

The linearizations of (2.2) were considered in [37] and are sometimes referred to as *rank-2 linear inequalities* [33]. So we denote the collection of all (x, X), which satisfy these linearizations, as R_2, i.e.,

$$R_2 := \{\ (x, X)\ :\ \text{linearized versions of (2.2) hold}\ \}.$$

In particular, the linearized versions of (2.2a) are called the RLT inequalities after the "reformulation-linearization technique" of [55], though they first appeared in [3, 4, 40]. These inequalities have been studied extensively because of their wide applicability and simple structure. Specifically, the RLT constraints provide simple bounds on the entries of X, which otherwise may be weakly constrained in $\widehat{\mathcal{L}}$, especially when m is small. After linearization via $X_{ij} = x_i x_j$, the four inequalities (2.2a) for a specific (i,j) are

$$\left.\begin{array}{r} l_i x_j + x_i l_j - l_i l_j \\ u_i x_j + x_i u_j - u_i u_j \end{array}\right\} \le X_{ij} \le \left\{\begin{array}{l} x_i u_j + l_i x_j - l_i u_j \\ x_i l_j + u_i x_j - u_i l_j. \end{array}\right. \tag{2.3}$$

Using the symmetry of X, the entire collection of RLT inequalities in matrix form is

$$\left.\begin{array}{r} lx^T + xl^T - ll^T \\ ux^T + xu^T - uu^T \end{array}\right\} \le X \le xu^T + lx^T - lu^T. \tag{2.4}$$

It can be shown easily that the original inequalities $l \le x \le u$ are implied by (2.4) if both l and u are finite in all components.

Since the RLT inequalities are just a portion of the inequalities defining R_2, it might be reasonable to consider R_2 generally and not the RLT constraints specifically. However, it is sometimes convenient to study the RLT constraints on their own. So we write $(x,X) \in \text{RLT}$ when (x,X) satisfies (2.4) but not necessarily all rank-2 linear inequalities, i.e.,

$$\text{RLT} := \{\ (x,X)\ :\ (2.4) \text{ holds}\ \}.$$

2.3. Valid second-order-cone inequalities. Similar to the derivation of the inequalities defining R_2, the linearizations of the following quadratic inequalities are valid for $\operatorname{clconv}\widehat{\mathcal{F}}$: for all $(k,\ell) \in LQ \times CQ$,

$$(b_k - a_k^T x)\left(\frac{1}{2}(1 - a_\ell^T x + b_\ell) - \left\|\begin{pmatrix} B_\ell^T x \\ \frac{1}{2}(a_\ell^T x - b_\ell + 1)\end{pmatrix}\right\|\right) \ge 0. \tag{2.5}$$

We call the linearizations *rank-2 second-order inequalities* and denote by S_2 the set of all satisfying (x,X), i.e.,

$$S_2 := \{\ (x,X)\ :\ \text{linearized versions of (2.5) hold}\ \}.$$

As an example, suppose we have the two valid inequalities $x_1 + x_2 \le 1$ and $x_1^2 + x_2^2 \le 2/3$, the second of which, via Proposition 2.1, is equivalent to the second-order cone constraint $\|x\| \le \sqrt{2/3}$. Then we multiply $1 - x_1 - x_2 \ge 0$ with $\sqrt{2/3} - \|x\| \ge 0$ to obtain the valid inequality

$$\begin{aligned} 0 &\le (1 - x_1 - x_2)(\sqrt{2/3} - \|x\|) \\ &= \sqrt{2/3}(1 - x_1 - x_2) - \|(1 - x_1 - x_2)x\|, \end{aligned}$$

which is linearized as

$$\left\|\begin{pmatrix} x_1 - X_{11} - X_{12} \\ x_2 - X_{21} - X_{22}\end{pmatrix}\right\| \le \sqrt{2/3}(1 - x_1 - x_2).$$

2.4. Valid semidefinite inequalities. The application of SDP to (**MIQCP**) arises from the following fundamental observation:

LEMMA 2.1 (Shor [56]). *Given $x \in \mathbb{R}^n$, it holds that*

$$\begin{pmatrix} 1 & x^T \\ x & xx^T \end{pmatrix} = \begin{pmatrix} 1 \\ x \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix}^T \succeq 0.$$

Thus, the linearized semidefinite inequality

$$Y := \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0 \tag{2.6}$$

is valid for $\operatorname{clconv} \widehat{\mathcal{F}}$. We define

$$\text{PSD} := \{\ (x, X)\ :\ (2.6) \text{ holds }\}.$$

Instead of enforcing $(x, X) \in \text{PSD}$, i.e., the full PSD condition (2.6), one can enforce relaxations of it. For example, since all principal submatrices of $Y \succeq 0$ are semidefinite, one could enforce just that all or some of the 2×2 principal submatrices of Y are positive semidefinite. This has been done in [32], for example.

2.5. The strength of valid inequalities. From the preceding subsections, we have the following result by construction:

PROPOSITION 2.3. $\operatorname{clconv} \widehat{\mathcal{F}} \subseteq \widehat{\mathcal{L}} \cap \text{RLT} \cap R_2 \cap S_2 \cap \text{PSD}$.

Even though $R_2 \subseteq \text{RLT}$, we retain RLT in the above expression for emphasis. We next catalog and discuss various special cases in which equality is known to hold in Proposition 2.3.

2.5.1. Simple bounds. We first consider the case when $\mathcal{F}$ is defined by simple, finite bounds, i.e., $\mathcal{F} = \{x \in \mathbb{R}^n : l \le x \le u\}$ with (l, u) finite in all components. In this case, $R_2 = \text{RLT} \subseteq \widehat{\mathcal{L}}$ and S_2 is vacuous. So Proposition 2.3 can be stated more simply as $\operatorname{clconv} \widehat{\mathcal{F}} \subseteq \text{RLT} \cap \text{PSD}$. Equality holds if and only if $n \le 2$:

THEOREM 2.1 (Anstreicher and Burer [6]). *Let $\mathcal{F} = \{x \in \mathbb{R}^n : l \le x \le u\}$ with (l, u) finite in all components. Then* $\operatorname{clconv} \widehat{\mathcal{F}} \subseteq \text{RLT} \cap \text{PSD}$ *with equality if and only if $n \le 2$.*

For $n > 2$, [6] and [19] derive additional valid inequalities but are still unable to determine an exact representation by valid inequalities even for $n = 3$. ([6] does give an exact *disjunctive* representation for $n = 3$.)

We also mention a classical result, which is in some sense subsumed by Theorem 2.1. Even still, this result indicates the strength of the RLT inequalities and can be useful when one-variable quadratics $X_{ii} = x_i^2$ are not of interest. The result does not fully classify $\operatorname{clconv} \widehat{\mathcal{F}}$ but rather coordinate projections of it.

THEOREM 2.2 (Al-Khayyal and Falk [4]). *Let $\mathcal{F} = \{x \in \mathbb{R}^n : l \le x \le u\}$ with (l, u) finite in all components. Then, for all $1 \le i < j \le n$,*

$\text{proj}_{(x_i,x_j,X_{ij})}(\text{clconv}\,\widehat{\mathcal{F}}) = \text{RLT}_{ij}$, *where* $\text{RLT}_{ij} := \{(x_i, x_j, X_{ij}) \in \mathbb{R}^3 :$ (2.3) *holds*$\}$.

2.5.2. Binary integer grid. We next consider the case when $\mathcal{F}$ is a binary integer grid: that is, $\mathcal{F} = \{x \in \mathbb{Z}^n : l \le x \le u\}$ with $u = l + e$ and l finite in all components. Note that this is simply a shift of the standard 0-1 binary grid and that $\text{clconv}\,\widehat{\mathcal{F}}$ is a polytope. In this case, $R_2 = \text{RLT} \subseteq \widehat{\mathcal{L}}$ and S_2 is vacuous. So Proposition 2.3 states that $\text{clconv}\,\widehat{\mathcal{F}} \subseteq \text{RLT} \cap \text{PSD}$. Also, some additional, simple linear equations are valid for $\text{clconv}\,\widehat{\mathcal{F}}$.

PROPOSITION 2.4. *Suppose that* $i \in I$ *has* $u_i = l_i + 1$ *with* l_i *finite. Then the equation* $X_{ii} = (1 + 2\,l_i)x_i - l_i - l_i^2$ *is valid for* $\text{clconv}\,\widehat{\mathcal{F}}$.

Proof. The shift $x_i - l_i$ is either 0 or 1. Hence, $(x_i - l_i)^2 = x_i - l_i$. After linearization with $X_{ii} = x_i^2$, this quadratic equation becomes the claimed linear one. ∎

When $I = [n]$, the individual equations $X_{ii} = (1 + 2\,l_i)x_i - l_i - l_i^2$ can be collected as $\text{diag}(X) = (e + 2\,l) \circ x - l - l^2$. We remark that the next result does not make use of PSD.

THEOREM 2.3 (Padberg [44]). *Let* $\mathcal{F} = \{x \in \mathbb{Z}^n : l \le x \le u\}$ *with* $u = l + e$ *and* l *finite in all components. Then*

$$\text{clconv}\,\widehat{\mathcal{F}} \subseteq \text{RLT} \cap \big\{(x, X) : \text{diag}(X) = (e - 2\,l) \circ x - l - l^2\big\}$$

with equality if and only if $n \le 2$.

In this case, $\text{clconv}\,\widehat{\mathcal{F}}$ is closely related to the *boolean quadric polytope* [44]. For $n > 2$, additional valid inequalities, such as the triangle inequalities described by Padberg [44], provide an even better approximation of $\text{clconv}\,\widehat{\mathcal{F}}$. For general n, a full description should not be easily available unless $P = NP$.

2.5.3. The nonnegative orthant and standard simplex. We now consider a case arising in the study of completely positive matrices in optimization and linear algebra [13, 23, 24]. A matrix Y is *completely positive* if $Y = NN^T$ for some nonnegative, rectangular matrix N, and the set of all completely positive matrices is a closed, convex cone. Although this cone is apparently intractable [42], it can be approximated from the outside to any precision using a seqence of polyhedral-semidefinite relaxations [26, 47]. The simplest approximation is by the so-called *doubly nonnegative matrices*, which are matrices Y satisfying $Y \succeq 0$ and $Y \ge 0$. Clearly, every completely positive matrix Y is doubly nonnegative. The converse holds if and only if $n \le 4$ [39].

Let $\mathcal{F} = \{x \in \mathbb{R}^n : x \ge 0\}$. Then, since $\text{RLT} = R_2$ and S_2 is vacuous, Proposition 2.3 states $\text{clconv}\,\widehat{\mathcal{F}} \subseteq \widehat{\mathcal{L}} \cap \text{RLT} \cap \text{PSD}$. Note that, because u is infinite in this case, $(x, X) \in \text{RLT}$ does not imply $x \ge 0$, and so we must include $\widehat{\mathcal{L}}$ explicitly to enforce $x \ge 0$. It is easy to see that

$$\widehat{\mathcal{L}} \cap \text{RLT} \cap \text{PSD} = \left\{(x, X) \ge 0 : \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0\right\} \tag{2.7}$$

which is the set of all doubly nonnegative matrices of size $(n+1) \times (n+1)$ having a 1 in the top-left corner.

For general n, it appears that equality in Proposition 2.3 does not hold in this case. However, it does hold for $n \leq 3$, which we prove next. As far as we are aware, this result has not appeared in the literature.

We first characterize the difference between $\operatorname{conv} \widehat{\mathcal{F}}$ and $\widehat{\mathcal{L}} \cap \text{RLT} \cap \text{PSD}$ for $n \leq 3$. The following lemma shows that this difference is precisely the recession cone of $\widehat{\mathcal{L}} \cap \text{RLT} \cap \text{PSD}$, which equals

$$\operatorname{rcone}(\widehat{\mathcal{L}} \cap \text{RLT} \cap \text{PSD}) = \{(0, D) \geq 0 : D \succeq 0\}. \tag{2.8}$$

LEMMA 2.2. *Let* $\mathcal{F} = \{x \in \mathbb{R}^n : x \geq 0\}$ *with* $n \leq 3$. *Then*

$$\widehat{\mathcal{L}} \cap \text{RLT} \cap \text{PSD} = \operatorname{conv} \widehat{\mathcal{F}} + \operatorname{rcone}(\widehat{\mathcal{L}} \cap \text{RLT} \cap \text{PSD}).$$

Proof. To prove the statement, we show containment in both directions; the containment $\supseteq$ is easy. To show $\subseteq$, let $(x, X) \in \widehat{\mathcal{L}} \cap \text{RLT} \cap \text{PSD}$ be arbitary. By (2.7),

$$Y := \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix}$$

is doubly nonnegative of size $(n+1) \times (n+1)$. Since $n \leq 3$, the "$n \leq 4$" result of [39] implies that Y is completely positive. Hence, there exists a rectangular $N \geq 0$ such that $Y = NN^T$. By decomposing each column $N_{\cdot j}$ of N as

$$N_{\cdot j} = \begin{pmatrix} \zeta_j \\ z_j \end{pmatrix}, \qquad (\zeta_j, z_j) \in \mathbb{R}_+ \times \mathbb{R}^n_+$$

we can write

$$\begin{aligned} Y = \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} &= \sum_j \begin{pmatrix} \zeta_j \\ z_j \end{pmatrix} \begin{pmatrix} \zeta_j \\ z_j \end{pmatrix}^T \\ &= \sum_{j:\zeta_j > 0} \zeta_j^2 \begin{pmatrix} 1 \\ \zeta_j^{-1} z_j \end{pmatrix} \begin{pmatrix} 1 \\ \zeta_j^{-1} z_j \end{pmatrix}^T + \sum_{j:\zeta_j = 0} \begin{pmatrix} 0 \\ z_j \end{pmatrix} \begin{pmatrix} 0 \\ z_j \end{pmatrix}^T \\ &= \sum_{j:\zeta_j > 0} \zeta_j^2 \begin{pmatrix} 1 & \zeta_j^{-1} z_j^T \\ \zeta_j^{-1} z_j & \zeta_j^{-2} z_j z_j^T \end{pmatrix} + \sum_{j:\zeta_j = 0} \begin{pmatrix} 0 & 0^T \\ 0 & z_j z_j^T \end{pmatrix}. \end{aligned}$$

This shows that $\sum_{j:\zeta_j > 0} \zeta_j^2 = 1$ and, from (2.8), that (x, X) is expressed as the convex combination of points in $\widehat{\mathcal{F}}$ plus the sum of points in $\operatorname{rcone}(\widehat{\mathcal{L}} \cap \text{RLT} \cap \text{PSD})$, as desired. ∎

Using the lemma, we can prove equality in Proposition 2.3 for $n \leq 3$.

THEOREM 2.4. *Let* $\mathcal{F} = \{x \in \mathbb{R}^n : x \geq 0\}$. *Then* $\operatorname{clconv} \widehat{\mathcal{F}} \subseteq \widehat{\mathcal{L}} \cap \text{RLT} \cap \text{PSD}$ *with equality if* $n \leq 3$.

Proof. The first statement of the theorem is just Proposition 2.3. Next, for contradiction, suppose there exists $(\bar{x}, \bar{X}) \in \widehat{\mathcal{L}} \cap \mathrm{RLT} \cap \mathrm{PSD} \setminus \mathrm{clconv}\, \widehat{\mathcal{F}}$. By the separating hyperplane theorem, there exists (c, C) such that

$$\min\{\langle C, X\rangle + c^T x : x \in \mathrm{clconv}\, \widehat{\mathcal{F}}\} \geq 0 > \langle C, \bar{X}\rangle + c^T \bar{x}.$$

Since $(\bar{x}, \bar{X}) \in \widehat{\mathcal{L}} \cap \mathrm{RLT} \cap \mathrm{PSD}$, by the lemma there exists $(z, Z) \in \mathrm{conv}\, \widehat{\mathcal{F}}$ and $(0, D) \in \mathrm{rcone}(\widehat{\mathcal{L}} \cap \mathrm{RLT} \cap \mathrm{PSD})$ such that $(\bar{x}, \bar{X}) = (z, Z + D)$. Thus, $\langle C, D\rangle < 0$.

Since $D \geq 0$, $D \succeq 0$, and $n \leq 3$, D is completely positive, i.e., there exists rectangular $N \geq 0$ such that $D = NN^T$. We have $\langle C, NN^T\rangle < 0$, which implies $d^T C d < 0$ for some nonzero column $d \geq 0$ of N. It follows that d is a negative direction of recession for the function $x^T C x + c^T x$. In other words,

$$\min\{\langle C, X\rangle + c^T x : x \in \mathrm{clconv}\, \widehat{\mathcal{F}}\} = -\infty,$$

a contradiction. ∎

A related result occurs for a bounded slice of the nonnegative orthant, e.g., the standard simplex $\{x \geq 0 : e^T x = 1\}$. In this case, however, the boundedness, the linear constraint, and R_2 ensure that equality holds in Proposition 2.3 for $n \leq 4$.

THEOREM 2.5 (Anstreicher and Burer [6]). *Let* $\mathcal{F} := \{x \geq 0 : e^T x = 1\}$. *Then* $\mathrm{clconv}\, \widehat{\mathcal{F}} \subseteq \widehat{\mathcal{L}} \cap \mathrm{RLT} \cap R_2 \cap \mathrm{PSD}$ *with equality if and only if* $n \leq 4$.

[36] and [6] also give related results where $\mathcal{F}$ is an affine transformation of the standard simplex.

2.5.4. Half ellipsoid. Let $\mathcal{F}$ be a half ellipsoid, that is, the intersection of a linear half-space and a possibly degenerate ellipsoid. In contrast to the previous cases considered, [57] proved that this case achieves equality in Proposition 2.3 regardless of the dimension n. On the other hand, the number of constraints is fixed. In particular, all simple bounds are infinite, $|LQ| = 1$, $|CQ| = 1$, and $NQ = \emptyset$ in which case Proposition 2.3 states simply $\mathrm{clconv}\, \widehat{\mathcal{F}} \subseteq \widehat{\mathcal{L}} \cap S_2 \cap \mathrm{PSD}$.

THEOREM 2.6 (Sturm and Zhang [57]). *Suppose*

$$\mathcal{F} = \left\{ x \in \mathbb{R}^n : \begin{array}{r} a_1^T x \leq b_1 \\ x^T A_2 x + a_2^T x \leq b_2 \end{array} \right\}$$

with $A_2 \succeq 0$ *is nonempty. Then* $\mathrm{clconv}\, \widehat{\mathcal{F}} = \widehat{\mathcal{L}} \cap S_2 \cap \mathrm{PSD}$.

As far as we are aware, this is the only case where Proposition 2.3 is provably strengthened via use of the rank-2 second-order inequalities enforced by S_2.

2.5.5. Bounded quadratic form. The final case we consider is that of a bounded quadratic form. Specifically, for a given quadratic form $x^T Ax + a^T x$ and bounds $-\infty \le b_l \le b_u \le +\infty$, let $\mathcal{F}$ be the set of points such that the form falls within the bounds, i.e., $\mathcal{F} = \{x : b_l \le x^T Ax + a^T x \le b_u\}$. No assumptions are made on A, e.g., we do not assume that A is positive semidefinite. As far as we are aware, the result proved below is new, but closely related results can be found in [57, 61].

Since there are no explicit bounds and no linear or convex quadratic inequalities, Proposition 2.3 states simply that $\operatorname{clconv} \widehat{\mathcal{F}} \subseteq \widehat{\mathcal{L}} \cap \mathrm{PSD}$, where

$$\widehat{\mathcal{L}} \cap \mathrm{PSD} = \left\{ (x, X) : \begin{array}{c} b_l \le \langle A, X\rangle + a^T x \le b_u \\ \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \succeq 0 \end{array} \right\}. \tag{2.9}$$

In general, it appears that equality in Proposition 2.3 does not hold in this case. However, we can still characterize the difference between $\operatorname{clconv} \widehat{\mathcal{F}}$ and $\widehat{\mathcal{L}} \cap \mathrm{PSD}$. As it turns out in the theorem below, this difference is precisely the recession cone of $\widehat{\mathcal{L}} \cap \mathrm{PSD}$, which equals

$$\operatorname{rcone}(\widehat{\mathcal{L}} \cap \mathrm{PSD}) = \left\{ (0, D) : \begin{array}{cl} 0 \le \langle A, D\rangle & \text{if } -\infty < b_l \\ \langle A, D\rangle \le 0 & \text{if } b_u < +\infty \\ D \succeq 0. & \end{array} \right\}.$$

In particular, if $\operatorname{rcone}(\widehat{\mathcal{L}} \cap \mathrm{PSD})$ is trivial, e.g., when $A \succ 0$ and $b_l = b_u$ are finite, then we will have equality in Proposition 2.3. The proof makes use of an important external lemma but otherwise is self-contained. Related proof techniques have been used in [10, 18].

LEMMA 2.3 (Pataki [48]). *Consider a consistent feasibility system in the symmetric matrix variable Y, which enforces $Y \succeq 0$ as well as p linear equalities and q linear inequalities. Suppose $\bar{Y}$ is an extreme point of the feasible set, and let $\bar{r} := \operatorname{rank}(Y)$ and $\bar{s}$ be the number of inactive linear inequalities at $\bar{Y}$. Then $\bar{r}(\bar{r}+1)/2 + \bar{s} \le p + q$.*

THEOREM 2.7. *Let $\mathcal{F} = \{x \in \mathbb{R}^n : b_l \le x^T Ax + a^T x \le b_u\}$ with $-\infty \le b_l \le b_u \le +\infty$ be nonempty. Then*

$$\operatorname{clconv} \widehat{\mathcal{F}} \subseteq \widehat{\mathcal{L}} \cap \mathrm{PSD} = \operatorname{conv} \widehat{\mathcal{F}} + \operatorname{rcone}(\widehat{\mathcal{L}} \cap \mathrm{PSD}).$$

Proof. Proposition 2.3 gives the inclusion $\operatorname{clconv} \widehat{\mathcal{F}} \subseteq \widehat{\mathcal{L}} \cap \mathrm{PSD}$. So we need to prove $\widehat{\mathcal{L}} \cap \mathrm{PSD} = \operatorname{clconv} \widehat{\mathcal{F}} + \operatorname{rcone}(\widehat{\mathcal{L}} \cap \mathrm{PSD})$. The containment $\supseteq$ is straightforward by construction.

For the containment $\subseteq$, recall that any point in a convex set may be written as a convex combination of finitely many extreme points plus a finite number of extreme rays. Hence, to complete the proof, it suffices to show that every extreme point of $\widehat{\mathcal{L}} \cap \mathrm{PSD}$ is in $\widehat{\mathcal{F}}$.

So let $(\bar{x}, \bar{X})$ be any extreme point of $\widehat{\mathcal{L}} \cap \mathrm{PSD}$. Examining (2.9) in the context of Lemma 2.3, $\widehat{\mathcal{L}} \cap \mathrm{PSD}$ can be represented as a feasibility system in

$$Y := \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix}$$

under four scenarios based on the values of (b_l, b_u):

(i) $(p, q) = (1, 0)$ if both b_l and b_u are infinite;

(ii) $(p, q) = (1, 1)$ if exactly one is finite;

(iii) $(p, q) = (1, 2)$ if both are finite and $b_l < b_u$; or

(iv) $(p, q) = (2, 0)$ if both are finite and $b_l = b_u$.

Define $\bar{Y}$ according to $(\bar{x}, \bar{X})$ and $(\bar{r}, \bar{s})$ as in the lemma. In the three cases (i), (ii), and (iv), $p + q \le 2$, and since $\bar{s} \ge 0$, we have

$$\bar{r}(\bar{r} + 1)/2 \le \bar{r}(\bar{r} + 1)/2 + \bar{s} \le p + q \le 2,$$

in which case $\bar{r} \le 1$. In the case (iii), $p + q = 3$ but $\bar{s} \ge 1$ since $b_l < b_u$, and so

$$\bar{r}(\bar{r} + 1)/2 \le p + q - \bar{s} \le 2,$$

which implies $\bar{r} \le 1$ as well. Overall, $\bar{r} \le 1$ means that $(\bar{x}, \bar{X})$ satisfies $\bar{X} = \bar{x}\bar{x}^T$, in which case $(\bar{x}, \bar{X}) \in \widehat{\mathcal{F}}$, as desired. ∎

We remark that, if one were able to prove $\text{rcone}(\text{clconv}\, \widehat{\mathcal{F}}) = \text{rcone}(\widehat{\mathcal{L}} \cap \text{PSD})$, then one would have equality in general. Using Lemma 2.3, it is possible to show that $\text{rcone}(\widehat{\mathcal{L}} \cap \text{PSD}) = \widehat{\mathcal{D}}$, where

$$\mathcal{D} := \left\{ d \in \mathbb{R}^n : \begin{array}{ll} 0 \le d^T A d & \text{if } -\infty < b_l \\ d^T A d \le 0 & \text{if } b_u < +\infty \end{array} \right\}$$

$$\widehat{\mathcal{D}} := \left\{ (0, dd^T) \in \mathbb{R}^n \times \mathcal{S}^n : d \in \mathcal{D} \right\},$$

but the relationship between $\widehat{\mathcal{D}}$ and $\text{rcone}(\text{clconv}\, \widehat{\mathcal{F}})$ is unclear.

3. Convex relaxations and valid inequalities: Related topics.

3.1. Another approach to convexification. We have presented Section 2 in terms of the set $\text{clconv}\, \widehat{\mathcal{F}}$. Another common approach [29, 58] is to study the so-called *convex* and *concave envelopes* of nonconvex functions. For example, in the formulation (**MIQCP**), suppose $f(x) := x^T C x + c^T x$ is nonconvex, and let $f_-(x)$ be any convex function that underestimates $f(x)$ in $\mathcal{F}$, i.e., $f_-(x) \le f(x)$ holds for all $x \in \mathcal{F}$. Then one can relax $f(x)$ by $f_-(x)$. Considering all such $f_-(x)$, since the point-wise supremum of convex functions is convex, there is a single convex $\hat{f}_-(x)$ which most closely underestimates the objective. By definition, $\hat{f}_-(x)$ is the convex envelope for $f(x)$ over $\mathcal{F}$. The convex envelope is closely related to the closed convex hull of the graph or epigraph of $f(x)$, i.e., $\text{clconv}\{(x, f(x)) : x \in \mathcal{F}\}$ or $\text{clconv}\{(x, w) : x \in \mathcal{F}, f(x) \le w\}$. Concave envelopes apply similar ideas to overestimation.

One can obtain a convex relaxation of (**MIQCP**) by relaxing the objective and all nonconvex constraints using convex envelopes. This can be seen as an alternative to lifting via $X = xx^T$. Either approach is generically hard. Another alternative is to perform some mixture of lifting and convex envelopes as is done in [36].

Like the various results in Section 2 with clconv $\widehat{\mathcal{F}}$, there are relatively few cases (typically low-dimensional) for which exact convex envelopes are known. For example, the following gives another perspective on Theorem 2.2 and equation (2.3) above:

THEOREM 3.1 (Al-Khayyal and Falk [4]). *Let $\mathcal{F} = \{x \in \mathbb{R}^n : l \le x \le u\}$ with (l,u) finite in all components. For all $1 \le i < j \le n$, the convex and concave envelopes of $f(x_i, x_j) = x_i x_j$ over $\mathcal{F}$ are, respectively,*

$$\max\{l_i x_j + x_i l_j - l_i l_j, u_i x_j + x_i u_j - u_i u_j\}$$
$$\min\{x_i u_j + l_i x_j - l_i u_j, x_i l_j + u_i x_j - u_i l_j\}.$$

These basic formulas can be used to construct convex underestimators (not necessarily envelopes) for any quadratic function by separating that quadratic function into pieces based on all $x_i x_j$. Such techniques are used in the software BARON [51]. Also, [21] generalizes the above theorem to the case where f is the product of two linear functions having disjoint support.

3.2. A SOCP relaxation. [31] proposes an SOCP relaxation for (**MIQCP**), which does not explicitly require the lifting $X = xx^T$ and is related to ideas in difference-of-convex programming [28].

First, the authors assume without loss of generality that the objective of (**MIQCP**) is linear. This can be achieved, for example, by introducing a new variable $t \in \mathbb{R}$ as well as a new quadratic constraint $x^T C x + c^T x \le t$ and then minimizing t. Next, for each $k \in NQ$, A_k is written as the difference of two (carefully chosen) positive semidefinite $A_k^+, A_k^- \succeq 0$, i.e., $A_k = A_k^+ - A_k^-$, so that k-th constraint may be expressed as

$$x^T A_k^+ x + a_k^T x \le b_k + x^T A_k^- x.$$

Then, an auxiliary variable $z_k \in \mathbb{R}$ is introduced to represent $x^T A_k^- x$ but also immediately relaxed as $x^T A_k^- x \le z_k$ resulting in the convex system

$$x^T A_k^+ x + a_k^T x \le b_k + z_k$$
$$x^T A_k^- x \le z_k.$$

Finally, z_k must be bounded in some fashion, say as $z_k \le \mu_k \in \mathbb{R}$, or else the relaxation will in fact be useless. Bounding z_k depends very much on the problem and the choice of A_k^+, A_k^-. [31] provides strategies for bounding z_k.

The relaxation thus obtained can be modeled as a problem having only linear and convex quadratic inequalities, which can in turn be represented as an SOCP. In total, the relaxation obtained by the authors is

$$\left\{ x \in \mathbb{R}^n \ : \ \begin{array}{rl} a_k^T x \le b_k & \forall\ k \in LQ \\ x^T A_k x + a_k^T x \le b_k & \forall\ k \in CQ \\ x^T A_k^+ x + a_k^T x \le b_k + z_k & \forall\ k \in NQ \\ x^T A_k^- x \le z_k & \forall\ k \in NQ \\ l \le x \le u,\ z \le \mu & \end{array} \right\}.$$

This SOCP model is shown to be dominated by the SDP relaxation $\widehat{\mathcal{L}} \cap \mathrm{PSD}$, while it is not directly comparable to the basic LP relaxation $\widehat{\mathcal{L}}$.

The above relaxation was recently revisited in [53]. The authors studied the relaxation obtained by the following splitting of the A_k matrices,

$$A_k = \sum_{\lambda_{kj} > 0} \lambda_{kj} v_{kj} v_{kj}^T - \sum_{\lambda_{kj} < 0} \lambda_{kj} v_{kj} v_{kj}^T \ ,$$

where $\{\lambda_{k1}, \ldots, \lambda_{kn}\}$ and $\{v_{k1}, \ldots, v_{kn}\}$ are the set of eigenvalues and eigenvectors of A_k, respectively. The constraint $x^T A_k x + a_k^T x \le b_k$ can thus be reformulated as, $\sum_{\lambda_{kj}>0} \lambda_{kj} \left(v_{kj}^T x\right)^2 + a_k^T x \le b_k + \sum_{\lambda_{kj}<0} \lambda_{kj} \left(v_{kj}^T x\right)^2$. The non-convex terms $\left(v_{kj}^T x\right)^2$ $(\lambda_{kj} < 0)$ can be relaxed by using their secant approximation to derive a convex relaxation of the above constraint. Instances of (**MIQCP**) tend to have geometric correlations along those v_{kj} with $\lambda_{kj} < 0$, which can be captured by projection techniques, and embedded within the polarity framework to derive strong cutting planes. We refer the reader to [53] for further details.

3.3. Results relating simple bounds and the binary integer grid. Motivated by Theorems such as 2.1 and 2.3 and the prior work of Padberg [44] and Yajima and Fujie [60], Burer and Letchford [19] studied the relationship between the two convex hulls

$$\mathrm{clconv}\left\{(x, xx^T) : \ x \in [0,1]^n \ \right\} \tag{3.1a}$$

$$\mathrm{clconv}\left\{(x, xx^T) : x \in \{0,1\}^n\right\}. \tag{3.1b}$$

The convex hull (3.1a) has been named QPB_n by the authors because of its relationship to "quadratic programming over the box." The convex hull (3.1b) is essentially the well known boolean quadric polytope BQP_n [44]. In fact, the authors show that BQP_n is simply the coordinate projection of (3.1b) on the variables x_i $(1 \le i \le n)$ and X_{ij} $(1 \le i < j \le n)$. Note that nothing is lost in the projection because $X_{ii} = x_i$ and $X_{ji} = X_{ij}$ are valid for (3.1b).

We let π represent the coordinate projection just mentioned, i.e., onto the variables x_i and X_{ij} $(i < j)$. The authors' result can be stated as $\pi(QPB_n) = BQP_n$, which immediately implies the following:

THEOREM 3.2 (Burer and Letchford [19]). *Any inequality in the variables x_i $(1 \le i \le n)$ and X_{ij} $(1 \le i < j \le n)$, which is valid for BQP_n, is also valid for QPB_n.*

For proper interpretation of the theorem, it is important to keep in mind that QPB_n still involves the variables X_{ii}, while those same variables have been projected out to obtain BQP_n. So another way to phrase the theorem is as follows: a valid inequality for BQP_n becomes valid for QPB_n when the variables X_{ii} are introduced into the inequality with zero coefficients.

This result shows that, in a certain sense, describing QPB_n is at least as hard as describing BQP_n, and since many classes of valid inequalities are already known for BQP_n, it also gives many classes of valid inequalities for QPB_n. Indeed, the authors prove many classes of facets for BQP_n are in fact facets for QPB_n. The authors also demonstrate that PSD and the RLT inequalities for pairs (i, i) help describe QPB_n beyond BQP_n.

3.4. Completely positive programming. Burer [18] has recently studied a special case of (**MIQCP**) having

$$\mathcal{F} = \left\{ x \geq 0 : \begin{array}{c} Ax = b \\ x_i x_j = 0 \;\; \forall\, (i,j) \in E \\ x_i \in \{0,1\} \;\; \forall\, i \in I \end{array} \right\}, \tag{3.2}$$

where, in particular, A is a rectangular matrix and $E \subseteq [n] \times [n]$ is symmetric. The author considers this specific form because it is amenable to analysis and yet is still fairly general. The results in [18] do not appear to hold, for example, under the general quadratic constraints of (**MIQCP**).

Proposition 2.3 and similar logic as in Theorem 2.3 show that

$$\text{clconv}\, \widehat{\mathcal{F}} \subseteq \widehat{\mathcal{H}}_{\text{PSD}} := \widehat{\mathcal{L}} \cap \text{RLT} \cap R_2 \cap \text{PSD} \cap \{(x, X) : X_{ii} = x_i \;\forall\, i \in I\}.$$

The following simplifying proposition is fairly easy to show:

PROPOSITION 3.1 (Burer [17]). *It holds that*

$$\widehat{\mathcal{H}}_{\text{PSD}} = \left\{ (x, X) \in \text{PSD} : \begin{array}{c} (x, X) \geq 0 \\ Ax = b, \;\text{diag}(AXA^T) = b^2 \\ X_{ij} = 0 \;\; \forall\, (i,j) \in E \\ X_{ii} = x_i \;\; \forall\, i \in I \end{array} \right\}.$$

Actually, $(x, X) \in \text{clconv}\, \widehat{\mathcal{F}}$ satisfies a stronger convex condition than $(x, X) \in \text{PSD}$. Recall that $(x, X) \in \text{PSD}$ is derived from

$$\begin{pmatrix} 1 & x^T \\ x & xx^T \end{pmatrix} = \begin{pmatrix} 1 \\ x \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix}^T \succeq 0.$$

Using that $x \in \mathcal{F}$ has $x \geq 0$, the above matrix is completely positive, not just positive semidefinite; see Section 2.5.3. We write

$$(x, X) \in \text{CPP} \quad \Longleftrightarrow \quad \begin{pmatrix} 1 & x^T \\ x & X \end{pmatrix} \text{ is completely positive}$$

and define $\widehat{\mathcal{H}}_{\text{CPP}} := \widehat{\mathcal{H}}_{\text{PSD}} \cap \text{CPP}$. The result below establishes that $\text{clconv}\,\widehat{\mathcal{F}} = \widehat{\mathcal{H}}_{\text{CPP}}$.

THEOREM 3.3 (Burer [18], Bomze and Jarre [15]). *Let $\mathcal{F}$ be defined as in (3.2). Define $J := \{j : \exists\, k \text{ s.t. } (j,k) \in E$ or $(k,j) \in E\}$, and suppose x_i is bounded in $\{x \geq 0 : Ax = b\}$ for all $i \in I \cup J$. Then* $\text{clconv}\,\widehat{\mathcal{F}} = \widehat{\mathcal{H}}_{\text{CPP}}$.

We emphasize that the result holds regardless of the boundedness of $\mathcal{F}$ as a whole; it is only important that certain variables are bounded. Completely positive representations of $\text{clconv}\,\widehat{\mathcal{F}}$ for different $\mathcal{F}$, which are not already covered by the above theorem, can also be found in [49, 50].

Starting from the above theorem, Burer [17] has implemented a specialized algorithm for optimizing the relaxation $\widehat{\mathcal{H}}_{\text{PSD}}$. We briefly discuss this implementation in Section 5.

3.5. Higher-order liftings and projections. Whenever it is not possible to capture $\text{clconv}\,\widehat{\mathcal{F}}$ exactly in the lifted space (x, X), it is still possible to lift into ever higher dimensional spaces and to linearize, say, cubic, quartic, or higher-degree valid inequalities there. This is quite a deep and powerful technique for capturing $\text{clconv}\,\widehat{\mathcal{F}}$. We refer the reader to the following papers: [9, 14, 33, 34, 35, 37, 54].

One of the most famous results in this area is the *sequential convexification* result for mixed 0-1 linear programs (M01LPs). Balas [8] showed that M01LPs are special cases of facial disjunctive programs which possess the *sequential convexifiability property*. Simply put, this means that the closed convex hull of all feasible solutions to a M01LP can be obtained by imposing the 0-1 condition on the binary variables *sequentially*, i.e., by imposing the 0-1 condition on the first binary variable and convexifying the resulting set, followed by imposing the 0-1 condition on the second variable, and so on. This is stated as the following theorem.

THEOREM 3.4 (Balas [8]). *Let $\mathcal{F}$ be the feasible set of a M01LP, i.e.,*

$$\mathcal{F} = \left\{x \in \{0,1\}^n : a_k^T x \leq b_k \;\forall\; k = 1, \ldots, m\right\}$$

and define $\mathcal{L}$ to be its basic linear relaxation in x. For each $i = 1, \ldots, n$, define $T_i := \{x : x_i \in \{0,1\}\}$. and

$$\begin{aligned} S_0 &:= \mathcal{L} \\ S_i &:= \text{clconv}\,(S_{i-1} \cap T_i) \qquad \forall \;\; i = 1, \ldots, n. \end{aligned}$$

Then $S_n = \text{clconv}\,\mathcal{F}$.

There exists an analogous sequential convexficiation for the continuous case of (**MIQCP**) for general quadratic constraints.

THEOREM 3.5 (Saxena et al. [52]). *Suppose that the feasible region $\mathcal{F}$ of (***MIQCP***) is bounded with $I = \emptyset$, i.e., no integer variables. For each $i = 1, \ldots, n$, define $\widehat{T}_i := \{(x, X) : X_{ii} \leq x_i^2\}$. Also define*

$$\begin{aligned} \widehat{S}_0 &:= \widehat{\mathcal{L}} \cap \text{PSD} \\ \widehat{S}_i &:= \text{clconv}\left(\widehat{S}_{i-1} \cap \widehat{T}_i\right) \qquad \forall \;\; i = 1, \ldots, n. \end{aligned}$$

Then $\widehat{S}_n = \text{clconv}\,\widehat{\mathcal{F}}$.

Part of the motivation for this theorem comes from the fact that

$$\text{PSD} \cap \bigcap_{i=1}^{n} \widehat{T}_i = \{(x, X) : X = xx^T\}$$

i.e., enforcing all $\widehat{T}_i$ along with positive semidefiniteness recovers the non-convex condition $X = xx^T$. This is analogous to the fact that $\bigcap_{i=1}^{n} T_i$ recovers the integer condition in Theorem 3.4.

There is one crucial difference between Theorems 3.4 and 3.5. Note that a M01LP with a single binary variable is polynomial-time solvable; Balas [8] gave a polynomial-sized LP for this problem. On the other hand, the analogous problem in the context of (**MIQCP**) involves minimizing a linear function over a nonconvex set of the form

$$\widehat{\mathcal{L}} \cap \text{PSD} \cap \{(x, X) : X_{ii} \leq x_i^2\}.$$

It is not immediately clear if this is a polynomial-time solvable problem. Indeed, it is likely to be NP-hard [38].

An immediate consequence of any sequential convexification theorem is that it decomposes the non-convexity of the original problem (M01LP or MIQCP) into a set of simple *atomic* non-convex conditions, such as $x_j \in \{0, 1\}$ or $X_{ii} \leq x_i^2$ that can be handled separately. For instance, Balas, Ceria and Cornuéjols [9] studied a lifted LP formulation of M01LP with a single binary variable and combined it with projection techniques to derive a family of cutting planes for M01LP, widely known as lift-and-project cuts. In order to apply the same idea to (**MIQCP**) we need systematic techniques for deriving valid cutting planes for the set $\widehat{\mathcal{L}} \cap \text{PSD} \cap \{(x, X) : X_{ii} \leq x_i^2\}$; a disjunctive programming based approach is described in the following section.

Theorems 3.4 and 3.5 can actually be combined to convexify any bounded $\mathcal{F}$ having a mix of binary and continuous variables. Also, Theorem 3.5 holds if the sets $\widehat{T}_i$ are defined with respect to any orthonormal basis $\{v_1, \ldots, v_n\}$, i.e., $\widehat{T}_i = \{(x, X) : \langle v_i v_i^T, X\rangle \leq (v_i^T x)^2\}$, not just the standard basis $\{e_1, \ldots, e_n\}$. We refer the reader to [52] for proofs of these results.

4. Dynamic approaches for generating valid inequalities.

Our starting point in this section is the lifted version $\widehat{\mathcal{F}}$ of the feasible set $\mathcal{F}$, whose convex hull can be relaxed, for example, as $\text{clconv}\,\widehat{\mathcal{F}} \subseteq \widehat{\mathcal{L}} \cap \text{RLT} \cap R_2 \cap S_2 \cap \text{PSD}$ (see Proposition 2.3). We are particularly interested in improving this relaxation through valid inequalities coming from a certain disjunctive programming approach.

Besides the presence of the integrality constraints $x_i \in \mathbb{Z}$, the only nonconvex constraint in $\widehat{\mathcal{F}}$ is the nonlinear equation $X = xx^T$ which can be represented exactly by the pair of SDP inequalities $X - xx^T \succeq 0$ and

$X - xx^T \preceq 0$. In fact, by the Schur complement theorem, the former is equivalent to the inequality (2.6), which is enforced by PSD. However, the latter is nonconvex. So relaxing $\widehat{\mathcal{F}}$ to $\widehat{\mathcal{L}} \cap \text{PSD}$ can be viewed as simply dropping $X - xx^T \preceq 0$. Said differently, $\widehat{\mathcal{F}} = \widehat{\mathcal{L}} \cap \text{PSD} \cap \{(x, X) : X - xx^T \preceq 0\}$. Harnessing the power of the inequality $X - xx^T \preceq 0$ constitutes the emphasis of this section.

As an aside, we would like to mention that all the results presented in this section exploit the continuous non-convexities in (**MIQCP**) to generate cutting planes. Non-convexities arising from presence of integer variables can be handled in a manner that is usally done in MILP; we refer the reader to [52] for computational results on disjunctive cuts for (**MIQCP**) derived from elementary 0-1 disjunctions, and to [7] for mixed integer rounding cuts for conic programs with 0-1 variables.

4.1. A procedure for generating disjunctive cuts. For any orthonormal basis $\{v_1, \ldots, v_n\}$ of $\mathbb{R}^n$,

$$\widehat{\mathcal{F}} = \widehat{\mathcal{L}} \cap \text{PSD} \cap \{(x, X) : X - xx^T \preceq 0\} \tag{4.1}$$

$$= \widehat{\mathcal{L}} \cap \text{PSD} \cap \left\{(x, X) : \langle X, v_i v_i^T \rangle \le (v_i^T x)^2 \ \forall\, i = 1, \ldots, n\right\}. \tag{4.2}$$

Given an arbitrary incumbent solution $(\hat{x}, \hat{X})$, say, from optimizing over $\widehat{\mathcal{L}}$ or $\widehat{\mathcal{L}} \cap \text{PSD}$, we would like to choose a basis $\{v_1, \ldots, v_n\}$ whose corresponding reformulation most effectively elucidates the infeasibility of $(\hat{x}, \hat{X})$ with respect to (4.1). The problem of choosing such a basis can be formulated as the following optimization problem that focuses on maximizing the violation of $(\hat{x}, \hat{X})$ with respect to the set of nonconvex constraints $\langle X, v_i v_i^T \rangle \le \left(v_i^T x\right)^2$:

$$\begin{array}{ll} \max & \max_{i=1\ldots n} \langle \hat{X}, v_i v_i^T \rangle - \left(v_i^T \hat{x}\right)^2 \\ \text{s.t.} & \{v_1, \ldots, v_n\} \text{ is an orthonormal basis.} \end{array}$$

Clearly, a set of orthonormal eigenvectors of $\hat{X} - \hat{x}\hat{x}^T$ is an optimal solution to the above problem. This exercise of choosing a reformulation that hinges on certain characteristics of the incumbent solution (in this case the spectral decomposition of $(\hat{x}, \hat{X})$) can be viewed as a *dynamic* reformulation technique that rotates the coordinate axes so as to most effectively highlight the infeasibility of the incumbent solution.

Having chosen an orthonormal basis, we need a systematic technique to derive cutting planes for $\text{clconv}\,\widehat{\mathcal{F}}$ using (4.2). We use the framework of disjunctive programming to accomplish this goal. Classical disjunctive programming of Balas [8] requires a linear relaxation $\widehat{\mathcal{P}}$ of $\widehat{\mathcal{F}}$ and a disjunction that is satisfied by all $(x, X) \in \widehat{\mathcal{F}}$. The linear relaxation $\widehat{\mathcal{P}}$ could be taken equal to $\widehat{\mathcal{L}}$ but could also incorporate cutting planes generated from previous incumbent solutions. As for the choice of disjunctions, we seek the sources of nonconvexities in (4.2). Evidently, (4.2) has two of these, namely, the integrality conditions on the variables x_i for $i \in I$ and the non-convex constraints $\langle X, v_i v_i^T\rangle \le \left(v_i^T x\right)^2$. Integrality constraints have been used to derive disjunctions in MILP for the past five decades; examples of such disjunctions include elementary 0-1 disjunctions, split disjunctions, GUB disjunctions, etc. We do not detail these disjunctions here. For constraints of the type $\langle Y, vv^T\rangle \le \left(v^T x\right)^2$ for fixed $v \in \mathbb{R}^n$, Saxena et. al. [52] proposed a technique to derive a valid disjunction, which we detail next. Following [52], we refer to $\langle X, vv^T\rangle \le \left(v^T x\right)^2$ as a *univariate expression*.

Let

$$\begin{aligned} \eta_L(v) &:= \min\left\{v^T x \mid (x, X) \in \widehat{\mathcal{P}}\right\} \\ \eta_U(v) &:= \max\left\{v^T x \mid (x, X) \in \widehat{\mathcal{P}}\right\} \\ \theta &\in (\eta_L(v), \eta_U(v))\,. \end{aligned}$$

In their computational experiments, Saxena et. al. [52] chose $\theta = v^T \hat{x}$, where $(\hat{x}, \hat{X})$ is the current incumbent. Every $(x, X) \in \widehat{\mathcal{F}}$ satisfies the following disjunction:

$$\left[\begin{array}{c} \eta_L(v) \le v^T x \le \theta \\ -(v^T x)(\eta_L(v) + \theta) + \theta\eta_L(v) \le -\langle X, vv^T\rangle \end{array}\right] \bigvee \left[\begin{array}{c} \theta \le v^T x \le \eta_U(v) \\ -(v^T x)(\eta_U(v) + \theta) + \theta\eta_U(v) \le -\langle X, vv^T\rangle \end{array}\right]. \tag{4.3}$$

This disjunction can be derived by splitting the range $[\eta_L(v), \eta_U(v)]$ of the function $v^T x$ over $\widehat{\mathcal{P}}$ into the two intervals $[\eta_L(v), \theta]$ and $[\theta, \eta_U(v)]$ and constructing a secant approximation of the function $-(v^T x)^2$ in each of the intervals, respectively (see Figure 1 for an illustration).

The disjunction (4.3) can then be embedded within the framework of Cut Generation Linear Programs (CGLPs) to derive disjunctive cuts as discussed in the following theorem.

THEOREM 4.1 ([8]). [1] *Let a polyhedral set $P = \{x : Ax \ge b\}$, a disjunction $D = \bigvee_{k=1}^{q} (D_k x \ge d_k)$ and a point $\hat{x} \in P$ be given. Then $\hat{x} \in Q := \operatorname{clconv} \cup_{k=1}^{q} \{x \in P \mid D_k x \ge d_k\}$ if and only if the optimal value of the following Cut Generation Linear Program (CGLP) is non-negative:*

[1] We caution the reader that the notation used in this theorem is not specifically tied to the notation for $\mathcal{F}$ and related sets.

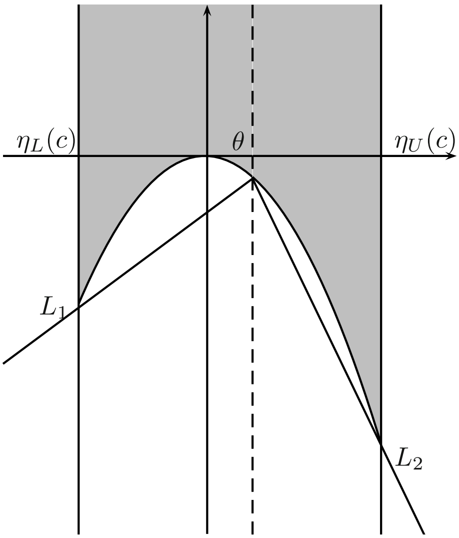

FIG. 1. *The constraint $-(v^T x)^2 \le -\langle X, vv^T\rangle$ and the disjunction (1) represented in the space spanned by $v^T x$ (horizontal axis) and $-\langle X, vv^T\rangle$ (vertical axis). The feasible region is the grey area above the parabola between $\eta_L(v)$ and $\eta_U(v)$. Disjunction (4.3) is obtained by taking the piecewise-linear approximation of the parabola, using a breakpoint at θ, and given by the two lines L_1 and L_2. Clearly, if $\eta_L(v) \le v^T x \le \theta$ then (x, X) must be above L_1 to be in the grey area; if $\theta \le v^T x \le \eta_U(v)$ then (x, X) must be above L_2.*

$$\begin{aligned} \min \quad & \alpha^T \hat{x} - \beta & & \qquad \text{(CGLP)} \\ \text{s.t.} \quad & A^T u^k + D_k^T v^k = \alpha & & k = 1, \ldots, q \\ & b^T u^k + d_k^T v^k \ge \beta & & k = 1, \ldots, q \\ & u^k, v^k \ge 0 & & k = 1, \ldots, q \\ & \sum_{k=1}^{q} \left(\xi^T u^k + \xi_k^T v^k\right) = 1 & & \end{aligned}$$

where ξ and ξ_k $(k = 1, \ldots, q)$ are any non-negative vectors of conformable dimensions that satisfy $\xi_k > 0$ $(k = 1, \ldots, q)$. If the optimal value of (CGLP) is negative, and (α, β) are part of an optimal solution, then $\alpha^T x \ge \beta$ is a valid inequality for Q which cuts off $\hat{x}$.

Next, we illustrate the above procedure for deriving disjunctive cuts on a small example. Consider the following instance of (**MIQCP**) derived from the st_ph11 instance from the GLOBALLib repository [27]:

$$\begin{aligned} \min \quad & x_1 + x_2 + x_3 - \frac{1}{2}\left(x_1^2 + x_2^2 + x_3^2\right) \\ \text{s.t.} \quad & 2x_1 + 3x_2 + 4x_3 \le 35 \\ & 0 \le x_1, x_2, x_3 \le 4. \end{aligned}$$

An optimal solution to the linear-semidefinite relaxation

$$\min\left\{x_1 + x_2 + x_3 - \frac{1}{2}(X_{11} + X_{22} + X_{33}) : (x, X) \in \widehat{\mathcal{L}} \cap \mathrm{PSD}\right\}$$

is

$$\hat{x} = \begin{pmatrix} 4 \\ 4 \\ 3.75 \end{pmatrix}, \quad \hat{X} = \begin{pmatrix} 16 & 16 & 15 \\ 16 & 16 & 15 \\ 15 & 15 & 15 \end{pmatrix}$$

and so

$$\hat{X} - \hat{x}\hat{x}^T = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0.9375 \end{pmatrix}$$

has exactly one non-zero eigenvalue. The associated eigenvector and univariate expression are given by $c^T = (0, 0, 1)$ and $X_{33} \leq x_3^2$, respectively. Note that $(\hat{x}, \hat{X})$ satisfies the secant approximation $X_{33} \leq 4x_3$ of $X_{33} \leq x_3^2$ at equality; hence *the secant inequality does not cut off this point.* Choosing $\theta = 2$ in (4.3), we get the following disjunction which is satisfied by every feasible solution $(x, X) \in \widehat{\mathcal{F}}$ for this example:

$$\begin{bmatrix} 0 \leq x_3 \leq 2 \\ 2x_3 - X_{33} \geq 0 \end{bmatrix} \bigvee \begin{bmatrix} 2 \leq x_3 \leq 4 \\ 6x_3 - X_{33} \geq 8 \end{bmatrix}.$$

In order to derive a disjunctive cut, for each term in the disjunction we sum a non-negative weighted combination of its constraints together with the original linear constraints of $\widehat{\mathcal{F}}$ to construct a new constraint valid for that term in the disjunction. If the separate weights in each term can be chosen in such a way that the resulting constraints for both terms are the same, then that constraint is a disjunctive cut. In particular, using the weighting scheme

$$\begin{bmatrix} 2x_3 - y_{33} \geq 0 & (14.70588) \\ -x_1 \geq -4 & (15.68627) \\ -x_2 \geq -4 & (23.52941) \\ x_3 \geq 0 & (27.45098) \end{bmatrix} \bigvee$$
$$\begin{bmatrix} 6x_3 - y_{33} \geq 8 & (14.70588) \\ -2x_1 - 3x_2 - 4x_3 \geq -35 & (7.84314) \end{bmatrix},$$

we arrive at the disjunctive cut

$$-15.68627x_1 - 23.52941x_2 + 56.86275x_3 - 14.70588y_{33} \geq -156.86274.$$

This disjunctive cut is violated by $(\hat{x}, \hat{X})$.

This example highlights a very important aspect of disjunctive programming: its ability to involve additional problem constraints in deriving strong cuts for $\operatorname{clconv}\widehat{\mathcal{F}}$. For illustration, consider the same example and the relaxation $\widehat{\mathcal{L}} \cap \mathrm{RLT} \cap \mathrm{PSD}$. Note that this relaxation only incorporates the effect of the general linear constraint $2x_1 + 3x_2 + 4x_3 \le 35$ via the set $\widehat{\mathcal{L}}$. Defining

$$x^1 = \begin{pmatrix} 4 \\ 4 \\ 4 \end{pmatrix}, \ X^1 = x^1(x^1)^T \qquad x^2 = \begin{pmatrix} 4 \\ 4 \\ 0 \end{pmatrix}, \ X^2 = x^2(x^2)^T$$

and $(\hat{x}, \hat{X}) := \frac{15}{16}(x^1, X^1) + \frac{1}{16}(x^2, X^2)$, i.e., the same $(\hat{x}, \hat{X})$ is in the example, it holds that $(\hat{x}, \hat{X}) \in \widehat{\mathcal{L}} \cap \mathrm{RLT} \cap \mathrm{PSD}$. Note, however, that the endpoint (x^1, X^1) is not in $\operatorname{clconv}\widehat{\mathcal{F}}$ since $x^1 \notin \mathcal{F}$, which implies $(x^1, X^1) \notin \widehat{\mathcal{L}}$ in this case. So it remains possible that $(\hat{x}, \hat{X})$ is not in $\operatorname{clconv}\widehat{\mathcal{F}}$. Indeed, by explicitly involving the linear constraint in a more powerful way during the convexification process, disjunctive programming cuts off $(\hat{x}, \hat{X})$ from $\operatorname{clconv}\widehat{\mathcal{F}}$.

In fact, something stronger can be said. For this same example, define $\mathcal{F}_{lu} := \{x : 0 \le x_1, x_2, x_3 \le 4\}$ so that $\mathcal{F} = \mathcal{F}_{lu} \cap \{x : 2x_1 + 3x_2 + 4x_3 \le 35\}$; also define $\widehat{\mathcal{F}}_{lu}$ accordingly. Now consider the stronger relaxation $\widehat{\mathcal{L}} \cap \operatorname{clconv}\widehat{\mathcal{F}}_{lu}$ of $\operatorname{clconv}\widehat{\mathcal{F}}$, which completely convexifies with respect to the bounds $l \le x \le u$. Still, even this stronger relaxation contains $(\hat{x}, \hat{X})$, and so we see that convexifying with respect to the bounds is simply not enough to cut off $(\hat{x}, \hat{X})$. One must incorporate the general linear inequality in a more aggressive fashion such as disjunctive programming does.

4.2. Computational insights. In [52], the authors report computational results with a cutting plane procedure based on these ideas. For instances of (**MIQCP**) coming from GLOBALLib, the authors solved five separate relaxations of

$$v_* := \min \left\{ \langle C, X \rangle + c^T x : x \in \operatorname{clconv}\widehat{\mathcal{F}} \right\}.$$

These relaxations were (with accompanying "version numbers" and optimal values for reference)

$$\begin{array}{llll} \text{(V0)} & v_{\mathrm{RLT}} := \min & \langle C, X \rangle + c^T x \\ & \qquad\quad \text{s.t.} & x \in \widehat{\mathcal{L}} \cap \mathrm{RLT}\,, \\ \text{(V1)} & v_{\mathrm{PSD}} := \min & \langle C, X \rangle + c^T x \\ & \qquad\quad \text{s.t.} & x \in \widehat{\mathcal{L}} \cap \mathrm{RLT} \cap \mathrm{PSD}\,, \end{array}$$

$$\begin{array}{lll}
\text{(V2)} & v_{\text{dsj}} := \min & \langle C, X\rangle + c^T x \\
& \quad\text{s.t.} & x \in \widehat{\mathcal{L}} \cap \text{RLT} \cap \text{PSD} \cap \text{“disjunctive cuts”}, \\
\text{(V2-SI)} & v_{\text{sec}} := \min & \langle C, X\rangle + c^T x \\
& \quad\text{s.t.} & x \in \widehat{\mathcal{L}} \cap \text{RLT} \cap \text{PSD} \cap \text{“secant cuts”}, \\
\text{(V2-Dsj)} & v'_{\text{dsj}} := \min & \langle C, X\rangle + c^T x \\
& \quad\text{s.t.} & x \in \widehat{\mathcal{L}} \cap \text{RLT} \cap \text{“disjunctive cuts”}.
\end{array}$$

The secant-cut referred to in the description of V2-SI is obtained by constructing the convex envelope of the non-convex inequality $\langle Y, vv^T\rangle \leq \left(v^T x\right)^2$; using the notation introduced above, the corresponding secant inequality is given by, $\langle Y, vv^T\rangle \leq (\eta_L(v) + \eta_U(v))v^T x - \eta_L(v)\eta_U(v)$. Since the secant inequality can be obtained cheaply once $\eta_L(v)$ and $\eta_U(v)$ have been computed, variant V2-SI helps us assess the marginal importance of using the computationally expensive disjunctive cut as compared to readily available secant inequality.

Note that $v_* \leq v_{\text{dsj}} \leq v_{\text{sec}} \leq v_{\text{PSD}} \leq v_{\text{RLT}}$ and $v_* \leq v_{\text{dsj}} \leq v'_{\text{dsj}} \leq v_{\text{RLT}}$. V0 was used as a base relaxation by which others were judged. In particular, for each of the four remaining relaxations, the metric

$$\text{percent duality gap closed} := \frac{v_{\text{RLT}} - v}{v_{\text{RLT}} - v_*} \times 100$$

was recorded on each instance using the optimal value v for that relaxation. Only instances having $v_{\text{RLT}} > v_*$ were selected for testing (129 instances). We remark that, when present, constraint PSD was enforced with a cutting-plane approach based on convex quadratic cuts rather than a black-box SDP solver.

Table 1
Summary of computational results.

	V1	V2	V2-SI	V2-Dsj
>99.99 %	16	23	24	1
98-99.99 %	1	44	4	29
75-98 %	10	23	17	10
25-75 %	11	22	26	29
0-25 %	91	17	58	60
Average Gap Closed	24.80%	76.49%	44.40%	41.54%

Table 1 summarizes the authors' key results on the 129 instances. Each of the main columns gives, for that version, the number of instances in several bins of the metric "percentage gap closed." Some comments are in order.

First, the variant V2 code that uses disjunctive cuts closes 50% more duality gap than the SDP relaxation V1. In fact, relaxations obtained by adding disjunctive cuts close more than 98% of the duality gap on 67 out of 129 instances; the same figure for SDP relaxations is 17 out of 129 instances. Second, the authors were able to close 99% of the duality gap on some of the instances such as st_qpc-m3a, st_ph13, st_ph11, ex3_1_4, st_jcbpaf2, ex2_1_9, etc., on which the SDP relaxation closes 0% of the duality gap.

Third, the variant V2-SI of the code that uses the secant inequality instead of disjunctive cuts does close a significant proportion (44.40%) of the duality gap. However, using disjunctive cuts improves this statistic to 76.49% thereby demonstrating the marginal benefits of disjunctive programming. Fourth, it is worth observing that both variants V2 and V2-SI have access to the same kinds of nonconvexities, namely, univariate expressions $\langle X, vv^T\rangle \le (v^T x)^2$ derived from eigenvectors v of $\hat{X} - \hat{x}\hat{x}^T$. Despite this commonality, why does V2, which has access to the CGLP apparatus, outperform V2-SI? The answer to this question lies in the way the individual frameworks process the nonconvex expression $\langle X, vv^T\rangle \le (v^T x)^2$. While V2-SI takes a local view of the problem and convexifies $\langle X, vv^T\rangle \le (v^T x)^2$ in the 2-dimensional space spanned by $v^T x$ and $\langle X, vv^T\rangle$, V2 takes a global view of the problem and combines disjunctive terms with other problem constraints. It is precisely this ability to derive stronger inferences by combining disjunctive information with other problem constraints that allows V2 to outperform its local counterpart V2-SI.

Fifth, it is worth observing that removing PSD has a debilitating effect on the cutting plane algorithm presented in [52] as demonstrated by the performance of V2-Dsj relative to V2. While the CGLP apparatus allows us to take a global view of the problem, its ability to derive strong disjunctive cuts is limited by the strength of the initial relaxation. By removing PSD, the relaxation is significantly weakened, and this subsequently has a deteriorating effect on the strength of disjunctive cuts later derived.

TABLE 2
Selection criteria.

% Duality Gap closed by		
V1	V2	Instance Chosen
$<$ 10 %	$>$ 90 %	st_jcbpaf2
$>$ 40%	$<$ 60%	ex9_2_7
$<$ 10%	$<$ 10%	ex7_3_1

The basic premise of the work in [52] lies in generating valid cutting planes for $\operatorname{clconv}\widehat{\mathcal{F}}$ from the spectrum of $\hat{X} - \hat{x}\hat{x}^T$, where $(\hat{x}, \hat{X})$ is the incumbent solution. In order to highlight the impact of these cuts on

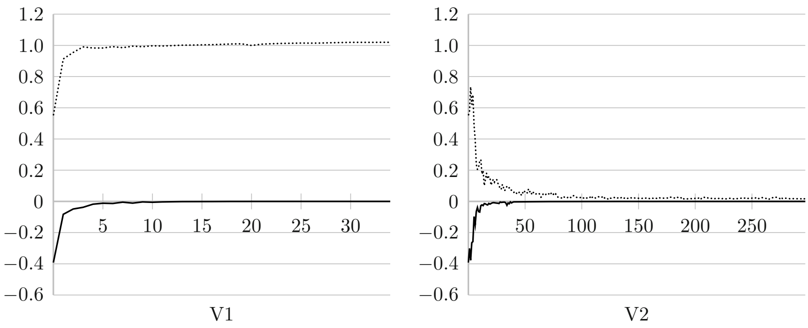

FIG. 2. *Plot of the sum of positive and negative eigenvalues for st_jcbpaf2 with V1–V2.*

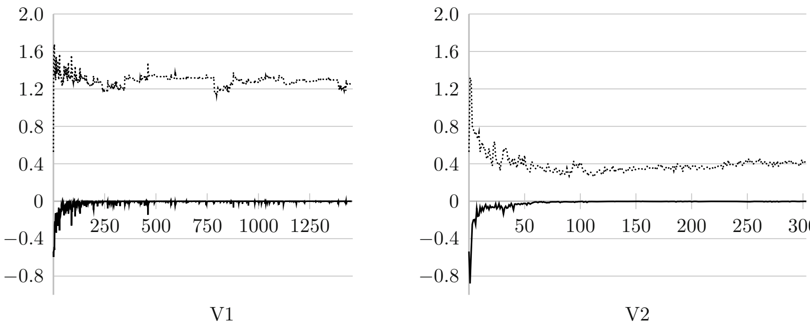

FIG. 3. *Plot of the sum of positive and negative eigenvalues for ex_9_2_7 with V1–V2.*

the spectrum itself, the authors presented details on three instances listed in Table 2 which we reproduce here for the sake of illustration. Figures 2–4 report the key results. The horizontal axis represents the number of iterations while the vertical axis reports the sum of the positive eigenvalues of $\hat{X} - \hat{x}\hat{x}^T$ (broken line) and the sum of the negative eigenvalues of $\hat{X} - \hat{x}\hat{x}^T$ (solid line) . Some remarks are in order.

First, the graph of the sum of negative eigenvalues converges to zero much faster than the corresponding graph for positive eigenvalues. This is not surprising because the problem of eliminating the negative eigenvalues is a convex programming problem, namely an SDP; the approach of adding convex-quadratic cuts is just an iterative cutting-plane based technique to impose the $X - xx^T \succeq 0$ condition. Second, V1 has a widely varying effect on the sum of positive eigenvalues of $X - xx^T$. This is to be expected because the $X - xx^T \succeq 0$ condition imposes no constraint on the positive eigenvalues of $X - xx^T$. Furthermore, the sum of positive eigenvalues represents the part of the nonconvexity of $\widehat{\mathcal{F}}$ that is not captured by PSD. Third, it is interesting to note that the variant that uses disjunctive cuts,

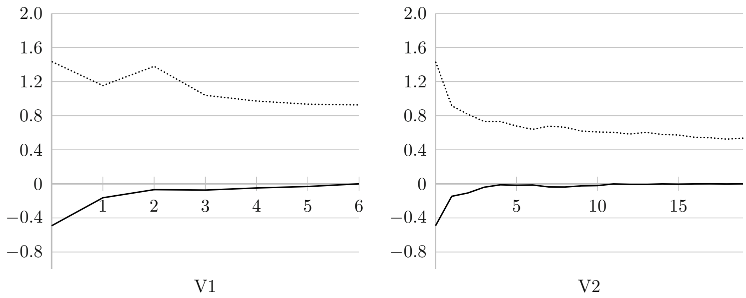

FIG. 4. *Plot of the sum of the positive and negative eigenvalues for ex_7_3_1 with V1–V2.*

namely V2, is able to force both positive and negative eigenvalues to converge to 0 for the st_jcbpaf2 thereby generating an almost feasible solution to this problem.

4.3. Working with only the original variables. Finally, we would like to emphasize that all of the relaxations of clconv $\widehat{\mathcal{F}}$ discussed until now are defined in the lifted space of (x, X). While the additional variable X enhances the expressive power of the formulation, it also increases the size of the formulation drastically, resulting in an enormous computational overhead which would be, for example, incurred at every node of a branch-and-bound tree. Ideally, we would like to extract the strength of these extended reformulations in the form of cutting planes that are defined only in the space of the original x variable. Systematic approaches for constructing such convex relaxations of (**MIQCP**) are described in a recent paper by Saxena et. al. [53]. We briefly reproduce some of these results to expose the reader to this line of research.

Consider the relaxation $\widehat{\mathcal{L}} \cap \text{RLT} \cap \text{PSD}$ of clconv $\widehat{\mathcal{F}}$, and define $\mathcal{Q} := \text{proj}_x(\widehat{\mathcal{L}} \cap \text{RLT} \cap \text{PSD})$, which is a relaxation of clconv $\mathcal{F}$ (not clconv $\widehat{\mathcal{F}}$!) in the space of the original variable x—but one that retains the power of $\widehat{\mathcal{L}} \cap \text{RLT} \cap \text{PSD}$. Can we separate from $\mathcal{Q}$, hence enabling us to work solely in the x-space? Specifically, given a point $\hat{x}$ that satisfies at least the simple bounds $l \le x \le u$, we desire an algorithmic framework that either shows $\hat{x} \in \mathcal{Q}$ or finds an inequality valid for $\mathcal{Q}$ which cuts off $\hat{x}$. Note that $\hat{x} \in \mathcal{Q}$ if and only if the following system is feasible in X with $\hat{x}$ fixed:

$$\begin{aligned} &\langle A_k, X\rangle + a_k^T \hat{x} \le b_k \quad \forall \;\; k = 1, \ldots, m \\ &\left.\begin{array}{r} l\hat{x}^T + \hat{x}l^T - ll^T \\ u\hat{x}^T + \hat{x}u^T - uu^T \end{array}\right\} \le X \le \hat{x}u^T + l\hat{x}^T - lu^T \\ &\begin{pmatrix} 1 & \hat{x}^T \\ \hat{x} & X \end{pmatrix} \succeq 0. \end{aligned}$$

As is typical, if this system is infeasible, then duality theory provides a cut (in this case, a convex quadratic cut) cutting off $\hat{x}$ from Q. Further, one can optimize to obtain a deep cut. We refer the reader to [53] for further details where the authors report computational results to demonstrate the computational dividends of working in the space of original variables possibly augmented by a *few* additional variables. We reproduce a slice of their computational results in Section 5.

5. Computational case study. To give the reader an impression of the computational requirements of the relaxations and techniques proposed in this paper, we compare three implementations for solving the relaxation

$$\min\{\langle C, X\rangle + c^T x : (x, X) \in \widehat{\mathcal{L}} \cap \text{RLT} \cap \text{PSD}\} \tag{5.1}$$

of the following particular case of (**MIQCP**), which is called *quadratic programming over the box*:

$$\min\{x^T Cx + c^T x : x \in [0, 1]^n\}. \tag{5.2}$$

We compare a black-box interior-point-method SDP solver (called IPM for "interior-point method"), the specialized completely-positive solver of [17] mentioned in Section 3.4 (called CP for "completely positive"), and the projection cutting plane method of [53] discussed in Section 4.3 (called PROJ for "projection"). We refer the reader to the original papers for full details of the implementations.

Methods IPM and CP work with the formulation (5.1). On the other hand, PROJ first reformulates (5.2) as

$$\min\left\{t : \begin{array}{c} x \in [0,1]^n \\ x^T Qx + c^T x \le t \end{array}\right\} \tag{5.3}$$

and then, in a pre-processing step, calculates several convex quadratic constraints as cutting planes for the relaxation $\text{proj}_{(t,x)}(\widehat{\mathcal{L}} \cap \text{RLT} \cap \text{PSD})$ of the reformulation. The procedure for calculating the cutting planes is outlined briefly in Section 4.3. Theoretically, if all possible cuts are generated, then the power of (5.1) is recovered. In practice, however, it is hoped that a few deep cuts will recover most of the power of (5.1) but save a significant amount of computation time. Finally, letting $\alpha_k t^2 + \beta_k t + x^T A_k x + a_k^T x \le b_k$ represent the derived convex cuts, PROJ solves the relaxation

$$\min\left\{t : \begin{array}{c} x \in [0,1]^n \\ \alpha_k t^2 + \beta_k t + x^T A_k x + a_k^T x \le b_k \quad \forall\ k \end{array}\right\}. \tag{5.4}$$

Nine instances from [20] are tested. Their relevant characteristics under relaxations (5.1) and (5.4) are given in Table 3, and the timings (in seconds) are give in Table 4. Also, in Table 5, we give the percentage gaps closed by the three methods relative to the pure linear relaxation $\widehat{\mathcal{L}} \cap \text{RLT}$ (see Section 4.2 for a careful definition of the percentage gap closed). A few comments are in order.

TABLE 3
Sizes of tested instances.

	# Variables		# Constraints			
			Linear		Convex	
Instance	IPM/CP	PROJ	IPM/CP	PROJ	IPM/CP (SDP)	PROJ (quad)
spar100-025-1	5151	203	20201	156	1	119
spar100-025-2	5151	201	20201	151	1	95
spar100-025-3	5151	201	20201	150	1	114
spar100-050-1	5151	201	20201	150	1	98
spar100-050-2	5151	201	20201	150	1	113
spar100-050-3	5151	201	20201	150	1	97
spar100-075-1	5151	201	20201	150	1	131
spar100-075-2	5151	201	20201	150	1	109
spar100-075-3	5151	199	20201	147	1	90

TABLE 4
Computational utility of projected relaxations.

	Time (sec)		
Instances	IPM	CP	PROJ (pre-process + solve)
spar100-025-1	5719.42	59	670.15 + 1.14
spar100-025-2	10185.65	54	538.03 + 1.52
spar100-025-3	5407.09	58	656.59 + 1.24
spar100-050-1	10139.57	76	757.14 + 1.07
spar100-050-2	5355.20	92	929.91 + 1.26
spar100-050-3	7281.26	76	747.46 + 0.82
spar100-075-1	9660.79	101	1509.96 + 2.00
spar100-075-2	6576.10	100	936.61 + 1.23
spar100-075-3	10295.88	81	657.84 + 0.87

First, on each instance, IPM is not competitive with either CP or PROJ. This illustrates a recognized trend in solving relaxations of this sort, namely that, at this point in time, specialized solvers perform better than black-box ones. Perhaps this will change as black-box solvers become more robust. Second, CP performs best in terms of overall time on each instance, but PROJ, discounting its pre-processing phase, solves its relaxation the quickest while still closing most of the gap that IPM and CP do. Within the context of using PROJ within branch-and-bound, this accrues significance due to two observations: (i) most contemporary branch-and-bound procedures generate cutting planes primarily at the root node and only sparingly at other nodes; and (ii) such a relaxation would be solved hundreds or thousands of times within the tree. So the pre-processing time of PROJ can be effectively amortized over the entire branch-and-bound tree.

TABLE 5
Computational utility of projected relaxations.

	% Gap Closed	
Instances	IPM/CP	PROJ
spar100-025-1	98.93%	92.36%
spar100-025-2	99.09%	92.16%
spar100-025-3	99.33%	93.26%
spar100-050-1	98.17%	93.62%
spar100-050-2	98.57%	94.13%
spar100-050-3	99.39%	95.81%
spar100-075-1	99.19%	95.84%
spar100-075-2	99.18%	96.47%
spar100-075-3	99.19%	96.06%

We also mention that, while PROJ is currently being solved by a nonlinear programming algorithm, the convex quadratic constraints of PROJ could actually be approximated by polyhedral relaxations introduced by Ben-Tal and Nemirovski [12] (also see [59]) yielding LP relaxations of these problems. Such LP relaxations are extremely desirable for branch-and-bound algorithms for two reasons. One, they can be efficiently re-optimized using warm-starting capabilities of LP solvers thereby reducing the computational overhead at nodes of the enumeration tree. Two, these LP relaxations can easily avail techniques, such as branching strategies, cutting planes, heuristics, etc., which have been developed by the MILP community in the past five decades (see [1] for application of these techniques in the context of convex MINLPs).

6. Conclusion. Table 6 catalogues the results covered in this paper. The first column lists the main concepts while the following two columns list their manifestations for M01LP and MIQCP, respectively. Some remarks are in order.

First, while linear programming based relaxations are almost universally used in M01LP, the same does not hold for MIQCP. There is a wide variety of relaxations for MIQCP that can be used, starting from the extended RLT+SDP relaxations to the compact eigen-reformulations (see [53]) defined in the original space of variables. It must be noted that all of these relaxations are currently solved by interior point methods that lack efficient re-optimization capabilities making them bottlenecks in a branch-and-bound procedure.

Second, there is a well established theory of exact formulations in M01LP (see [22]). Many of these results were obtained as byproducts of the tremendous amount of research that went into proving the perfect graph conjecture. Unfortunately, the progress in this direction in MIQCP

TABLE 6
M01LP vs. MIQCP.

Concept	**M01LP**	**MIQCP**
Relaxation	LP relaxation	$\widehat{\mathcal{L}} \cap \text{RLT} \cap R_2 \cap S_2 \cap \text{PSD}$ projected SDP eigen-reformulation
Exact Description	total unimodularity; perfect, ideal, and balanced matrices	theorems in Section 2.5
Elementary Non-Convexity	$x_j \in \{0, 1\}$	$X_{ii} \leq x_i^2$
Linear Transformed Non-Convexity	$(\pi x \leq \pi_0) \vee (\pi x \geq \pi_0 + 1)$	$\langle X, vv^T \rangle \leq (v^T x)^2$
Sequential Convexification	Balas [8]	Saxena et al. [52]

has been rather slow, and exact descriptions are unknown for most classes of problems except for some very small problem instances.

Third, there is an interesting connection between cuts derived from the univariate expression $\langle X, vv^T \rangle \leq (v^T x)^2$ for **MIQCP** and split cuts derived from split disjunctions $(\pi x \leq \pi_0) \vee (\pi x \geq \pi_0 + 1)$ $(\pi \in \mathbb{Z}^n)$ in M01LP. To see this, note that $\langle X, vv^T \rangle \leq (v^T x)^2$ can be obtained from the elementary non-convex constraint $X_{ii} \leq x_i^2$ by the linear transformation $(x, X) \longrightarrow (v^T x, \langle X, vv^T \rangle)$ where the linear transformation is chosen depending on the incumbent solution; for example, Saxena et al. [52] derive the v vector from the spectral decomposition of $\hat{X} - \hat{x}\hat{x}^T$. Similarly, the split disjunction $(\pi x \leq \pi_0) \vee (\pi x \geq \pi_0 + 1)$ can be obtained from elementary 0-1 disjunction $(x_j \leq 0) \vee (x_j \geq 1)$ by the linear transformation $x \longrightarrow \pi x$ where the linear transformation is chosen depending on the incumbent solution; for instance, the well known mixed integer Gomory cuts can be obtained from split disjunctions derived by monoidal strengthening of elementary 0-1 disjunctions, wherein the monoid that is chosen to strengthen the cut depends on the incumbent solution (see [9]).

Acknowledgements. The authors are in debt to two anonymous referees for many helpful suggestions that have improved the paper significantly.

REFERENCES

[1] K. Abhishek, S. Leyffer, and J.T. Linderoth, *FilMINT: An outer-approximation-based solver for convex mixed-integer nonlinear programs*, INFORMS Journal on Computing, **22** (2010), pp. 555–567.

[2] T.K. Achterberg, T. Berthold and K.T. Wolter, *Constraint integer programming: A new approach to integrate cp and mip*, Lecture Notes in Computer Science, **5015** (2008), pp. 6–20.
[3] F.A. Al-Khayyal, *Generalized bilinear programming, Part i: Models, applications, and linear programming relaxations*, European Journal of Operational Research, **60** (1992), pp. 306–314.
[4] F.A. Al-Khayyal and J.E. Falk, *Jointly constrained biconvex programming*, Math. Oper. Res., **8** (1983), pp. 273–286.
[5] F. Alizadeh and D. Goldfarb, *Second-order cone programming*, Math. Program., **95** (2003), pp. 3–51. ISMP 2000, Part 3 (Atlanta, GA).
[6] K.M. Anstreicher and S. Burer, *Computable representations for convex hulls of low-dimensional quadratic forms*, with K. Anstreicher, Mathematical Programming (Series B), **124**(1-2), pp. 33-43 (2010).
[7] A. Atamtürk and V. Narayanan, *Conic mixed-integer rounding cuts*, Math. Program., **122** (2010), pp. 1–20.
[8] E. Balas, *Disjunctive programming: properties of the convex hull of feasible points*, Discrete Appl. Math., **89** (1998), pp. 3–44.
[9] E. Balas, S. Ceria, and G. Cornuéjols, *A lift-and-project cutting plane algorithm for mixed 0-1 programs*, Mathematical Programming, **58** (1993), pp. 295–324.
[10] A. Beck, *Quadratic matrix programming*, SIAM J. Optim., **17** (2006), pp. 1224–1238 (electronic).
[11] P. Belotti, *Disjunctive cuts for non-convex MINLP*, IMA Volume Series, Springer 2010, accepted.
http://myweb.clemson.edu/~pbelott/papers/belotti-disj-MINLP.pdf.
[12] A. Ben-Tal and A. Nemirovski, *On polyhedral approximations of the second-order cone*, Math. Oper. Res., **26** (2001), pp. 193–205.
[13] A. Berman and N. Shaked-Monderer, *Completely Positive Matrices*, World Scientific, 2003.
[14] D. Bienstock and M. Zuckerberg, *Subset algebra lift operators for 0-1 integer programming*, SIAM J. Optim., **15** (2004), pp. 63–95 (electronic).
[15] I. Bomze and F. Jarre, *A note on Burer's copositive representation of mixed-binary QPs*, Optimization Letters, **4** (2010), pp. 465–472.
[16] P. Bonami, L. Biegler, A. Conn, G. Cornuéjols, I. Grossmann, C. Laird, J. Lee, A. Lodi, F. Margot, N. Sawaya, and A. Wächter, *An algorithmic framework for convex mixed-integer nonlinear programs.*, Discrete Optimization, **5** (2008), pp. 186–204.
[17] S. Burer, *Optimizing a polyhedral-semidefinite relaxation of completely positive programs*, Mathematical Programming Computation, **2**(1), pp 1–19 (2010).
[18] ———, *On the copositive representation of binary and continuous nonconvex quadratic programs*, Mathematical Programming, **120** (2009), pp. 479–495.
[19] S. Burer and A.N. Letchford, *On nonconvex quadratic programming with box constraints*, SIAM J. Optim., **20** (2009), pp. 1073–1089.
[20] S. Burer and D. Vandenbussche, *Globally solving box-constrained nonconvex quadratic programs with semidefinite-based finite branch-and-bound*, Comput. Optim. Appl., **43** (2009), pp. 181–195.
[21] D. Coppersmith, O. Günlük, J. Lee, and J. Leung, *A polytope for a product of a real linear functions in 0/1 variables*, manuscript, IBM, Yorktown Heights, NY, December 2003.
[22] G. Cornuéjols, *Combinatorial optimization: packing and covering*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2001.
[23] R.W. Cottle, G.J. Habetler, and C.E. Lemke, *Quadratic forms semi-definite over convex cones*, in Proceedings of the Princeton Symposium on Mathematical Programming (Princeton Univ., 1967), Princeton, N.J., 1970, Princeton Univ. Press, pp. 551–565.

[24] G. Danninger and I.M. Bomze, *Using copositivity for global optimality criteria in concave quadratic programming problems*, Math. Programming, **62** (1993), pp. 575–580.
[25] G. Dantzig, R. Fulkerson, and S. Johnson, *Solution of a large-scale traveling-salesman problem*, J. Operations Res. Soc. Amer., **2** (1954), pp. 393–410.
[26] E. de Klerk and D.V. Pasechnik, *Approximation of the stability number of a graph via copositive programming*, SIAM J. Optim., **12** (2002), pp. 875–892.
[27] See the website: www.gamsworld.org/global/globallib/globalstat.htm.
[28] R. Horst and N.V. Thoai, *DC programming: overview*, J. Optim. Theory Appl., **103** (1999), pp. 1–43.
[29] M. Jach, D. Michaels, and R. Weismantel, *The convex envelope of* $(N-1)$*-convex fucntions*, SIAM J. Optim., **19** (2008), pp. 1451–1466.
[30] R. Jeroslow, *There cannot be any algorithm for integer programming with quadratic constraints.*, Operations Research, **21** (1973), pp. 221–224.
[31] S. Kim and M. Kojima, *Second order cone programming relaxation of nonconvex quadratic optimization problems*, Optim. Methods Softw., **15** (2001), pp. 201–224.
[32] ———, *Exact solutions of some nonconvex quadratic optimization problems via SDP and SOCP relaxations*, Comput. Optim. Appl., **26** (2003), pp. 143–154.
[33] M. Kojima and L. Tunçel, *Cones of matrices and successive convex relaxations of nonconvex sets*, SIAM J. Optim., **10** (2000), pp. 750–778.
[34] J.B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim., **11** (2001), pp. 796–817.
[35] M. Laurent, *A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for 0-1 programming*, Math. Oper. Res., **28** (2003), pp. 470–496.
[36] J. Linderoth, *A simplicial branch-and-bound algorithm for solving quadratically constrained quadratic programs*, Math. Program., **103** (2005), pp. 251–282.
[37] L. Lovász and A. Schrijver, *Cones of matrices and set-functions and* 0-1 *optimization*, SIAM Journal on Optimization, **1** (1991), pp. 166–190.
[38] T. Matsui, *NP-hardness of linear multiplicative programming and related problems*, J. Global Optim., **9** (1996), pp. 113–119.
[39] J.E. Maxfield and H. Minc, *On the matrix equation* $X'X = A$, Proc. Edinburgh Math. Soc. (2), **13** (1962/1963), pp. 125–129.
[40] G.P. McCormick, *Computability of global solutions to factorable nonconvex programs. I. Convex underestimating problems*, Math. Programming, **10** (1976), pp. 147–175.
[41] See the website: http://www.gamsworld.org/minlp/.
[42] K.G. Murty and S.N. Kabadi, *Some NP-complete problems in quadratic and nonlinear programming*, Math. Programming, **39** (1987), pp. 117–129.
[43] G. Nemhauser and L. Wolsey, *Integer and Combinatorial Optimization*, Wiley-Interscience, 1999.
[44] M. Padberg, *The Boolean quadric polytope: some characteristics, facets and relatives*, Math. Programming, **45** (1989), pp. 139–172.
[45] P. Pardalos, *Global optimization algorithms for linearly constrained indefinite quadratic problems*, Computers and Mathematics with Applications, **21** (1991), pp. 87–97.
[46] P.M. Pardalos and S.A. Vavasis, *Quadratic programming with one negative eigenvalue is NP-hard*, J. Global Optim., **1** (1991), pp. 15–22.
[47] P. Parrilo, *Structured Semidefinite Programs and Semi-algebraic Geometry Methods in Robustness and Optimization*, PhD thesis, California Institute of Technology, 2000.
[48] G. Pataki, *On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues*, Mathematics of Operations Research, **23** (1998), pp. 339–358.
[49] J. Povh and F. Rendl, *A copositive programming approach to graph partitioning*, SIAM J. Optim., **18** (2007), pp. 223–241.

[50] ———, *Copositive and semidefinite relaxations of the quadratic assignment problem*, Discrete Optim., **6** (2009), pp. 231–241.

[51] N.V. Sahinidis, *BARON: a general purpose global optimization software package*, J. Glob. Optim., **8** (1996), pp. 201–205.

[52] A. Saxena, P. Bonami, and J. Lee, *Convex relaxations of non-convex mixed integer quadratically constrained programs: Extended formulations*, Mathematical Programming (Series B), **124**(1-2), pp. 383–411 (2010). http://dx.doi.org/10.1007/s10107-010-0371-9.

[53] ———, *Convex relaxations of non-convex mixed integer quadratically constrained programs: Projected formulations*, 2010. To appear in *Mathematical Programming*. http://dx.doi.org/10.1007/s10107-010-0340-3.

[54] H.D. Sherali and W.P. Adams, *A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems*, SIAM J. Discrete Math., **3** (1990), pp. 411–430.

[55] H.D. Sherali and W.P. Adams, *A Reformulation-Linearization Technique (RLT) for Solving Discrete and Continuous Nonconvex Problems*, Kluwer, 1997.

[56] N. Shor, *Quadratic optimization problems*, Soviet Journal of Computer and Systems Science, **25** (1987), pp. 1–11. Originally published in *Tekhnicheskaya Kibernetika*, **1**:128–139, 1987.

[57] J.F. Sturm and S. Zhang, *On cones of nonnegative quadratic functions*, Math. Oper. Res., **28** (2003), pp. 246–267.

[58] M. Tawarmalani and N.V. Sahinidis, *Convexification and global optimization in continuous and mixed-integer nonlinear programming*, Vol. **65** of Nonconvex Optimization and its Applications, Kluwer Academic Publishers, Dordrecht, 2002. Theory, algorithms, software, and applications.

[59] J.P. Vielma, S. Ahmed, and G.L. Nemhauser, *A lifted linear programming branch-and-bound algorithm for mixed-integer conic quadratic programs*, INFORMS J. Comput., **20** (2008), pp. 438–450.

[60] Y. Yajima and T. Fujie, *A polyhedral approach for nonconvex quadratic programming problems with box constraints*, J. Global Optim., **13** (1998), pp. 151–170.

[61] Y. Ye and S. Zhang, *New results on quadratic minimization*, SIAM J. Optim., **14** (2003), pp. 245–267 (electronic).