Affine FR : an effective facial reduction algorithm for semidefinite relaxations of combinatorial problems

Hao Hu*[1] and Boshi Yang†[2]

[1,2]School of Mathematical and Statistical Sciences, Clemson University, Clemson, SC, USA.
[1]hhu2@clemson.edu, [2]boshiy@clemson.edu

2:40pm, September 14, 2023

Abstract

We develop a new method called *affine FR* for recovering Slater's condition for semidefinite programming (SDP) relaxations of combinatorial optimization (CO) problems. Affine FR is a user-friendly method, as it is fully automatic and only requires a description of the problem. We provide a rigorous analysis of differences between affine FR and the existing methods. We also present numerical results to demonstrate the effectiveness of affine FR in reducing the size of SDP relaxations for CO problems.

Key Words: facial reduction, exposing vector, semidefinite programming, Slater's condition, combinatorial optimization

*Research supported by the Air Force Office of Scientific Research under award number FA9550-23-1-0508.
†Research supported by the Air Force Office of Scientific Research under award number FA9550-23-1-0508.

1 Introduction

We have developed a novel facial reduction algorithm for semidefinite programming (SDP) relaxations of combinatorial optimization (CO) problems. Our algorithm specifically targets CO problems that involve binary variables, which naturally represent yes/no decisions in various applications. For instance, these problems arise in radio frequency assignment, time-tabling, scheduling, cargo transportation, large network communication, satellite network design, and resource allocation.

CO problems are mathematically challenging, and the state-of-the-art solution methods struggle to efficiently solve large-scale problems that are crucial for business and scientific needs. SDP, on the other hand, is a convex optimization problem with a linear objective function subject to affine and positive semidefinite constraints, see [2, 36]. SDP approaches have gained attention in solving CO problems for over three decades due to their ability to provide tight bounds for the optimal value, accelerating the solving process. Notable results have been achieved for SDP relaxations of challenging problems such as for the max-cut problem [13, 28] and the quadratic assignment problem [38]. Despite recent advances in first-order methods making SDP approaches more practical, there are still significant challenges when it comes to the large-scale problems and handling numerical instability caused by the lack of regularity conditions in semidefinite programs, see [11, 17, 23, 39]. To be more precise, the dimension of the SDP relaxation grows quickly, and it is well-known that SDP solvers do not scale as well as linear programming (LP) solvers. And semidefinite programs often lack regularity conditions due to the problem structure, which makes solvers numerically unstable.

Facial reduction is a preprocessing technique for the regularization of semidefinite programs proposed by Borwein and Wolkowicz, see [5, 6]. It has been proven effective in restoring Slater's condition for SDP relaxations, as demonstrated in the quadratic assignment problem [38]. If facial reduction restores Slater's condition for the SDPs, then it often yields a significant acceleration in the computation. While facial reduction can always restore Slater's condition in theory, its implementation is highly non-trivial and thus special methods are developed to implement facial reduction in practice, see [26, 40].

In our paper, we introduce *affine FR*, a new implementation of facial reduction algorithm for SDP relaxations of CO problems. Our approach leverages the geometric structure in the original CO problem and its connection to the SDP relaxation. Through theoretical analysis and experimental results, we establish several advantages of the affine FR method:

- Effectiveness: For CO problems, our method outperforms existing general approaches by exploiting the underlying problem structure.
- Applicability: Affine FR can be applied to a wider range of CO problems than existing approaches for CO problems.
- User-friendly: Affine FR is fully automatic and only requires a problem description, minimizing user efforts in the preprocessing step.

Affine FR is effective and can be applied to all mixed binary integer programming problems. It keeps the users' efforts in the preprocessing step at a minimum level. This paves the way for the community to make semidefinite programs a more practical tool in operations research.

2 Preliminaries

2.1 SDP relaxation for CO problems

A combinatorial optimization (CO) problem seeks an optimal solution containing some discrete variables. To simplify the presentation, we discuss CO problems whose feasible region F is the intersection of the set of binary vectors and a polyhedron P, i.e.,

$$F = P \cap \{0,1\}^n. \tag{1}$$

Here, we assume that a linear system $Ax \leq b$ is provided for defining the polyhedron $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$. Our method also works for the more general mixed-binary linear programming problems which contain both continuous and binary variables, see Section 5 and Remark 4.2. Throughout we assume that P is not empty and each binary variable x_i is between zero and one for any $x \in P$.

Assumption 2.1. *$P \neq \emptyset$. If $x \in P$, then $0 \leq x \leq 1$.*

We are interested in maximizing or minimizing a given objective function over F. For most applications, the objective function is a linear or quadratic function. In this case, SDP relaxations are often considered, as they usually provide tight bounds for the optimal value. SDP is a convex optimization problem with a linear objective function subject to affine and positive semidefinite constraints. To construct an SDP relaxation for the feasible set F, we consider the lifted feasible set

$$F_1 := \left\{ \begin{bmatrix} 1 \\ x \end{bmatrix} \begin{bmatrix} 1 \\ x \end{bmatrix}^T \;\middle|\; x \in F \right\} \subset \mathbb{S}^{n+1}_+. \tag{2}$$

Denote by $\mathbb{S}^{n+1}_+$ and $\mathbb{S}^{n+1}_{++}$ the set of positive semidefinite matrices and the set of positive definite matrices of size $n+1$, respectively. We introduce a matrix variable $Y \in \mathbb{S}^{n+1}$ to represent a relaxation for the nonlinear expression $\left[\begin{smallmatrix} 1 \\ x \end{smallmatrix}\right]\left[\begin{smallmatrix} 1 \\ x \end{smallmatrix}\right]^T$ in F_1. Clearly, the matrix variable Y must be positive semidefinite. Along with the positive semidefinite constraint, valid linear equality and inequality constraints on the matrix variable Y can be added to the SDP relaxation. For example, an SDP relaxation K for F_1 with equality constraints can be represented as

$$K := \left\{ Y \in \mathbb{S}^{n+1}_+ \mid \mathcal{A}(Y) = d \right\}, \tag{3}$$

where $\mathcal{A} : \mathbb{S}^{n+1}_+ \to \mathbb{R}^m$ is a linear operator and $d \in \mathbb{R}^m$. Since there is a one-to-one correspondence between F_1 and F, we also call $K \subseteq \mathbb{S}^{n+1}$ a relaxation for $F \subseteq \mathbb{R}^n$.

In practice, there are many different SDP relaxations and we describe one of the most popular SDP relaxation called *the Shor's SDP relaxation* in Example 2.1. The key is to find relaxations with a good tradeoff between its accuracy and computational costs. In [27,35], the authors discuss different constructions of SDP relaxations for the CO problems. The Shor's SDP relaxation is used for a more general class of problems called *quadratically constrained quadratic programming (QCQP).* And the mixed-binary linear programming problem studied in this paper is a special case of QCQP. In [3], the authors present a comprehensive comparison of different types of SDP relaxations for QCQP. The exactness of different SDP relaxations including Shor's relaxation is studied in [34].

While SDP relaxations provide tight bounds for CO problems, it has issues with scaling or degeneracy. SDP scales poorly with the size of the data, and they are also expensive to solve. Moreover, due to the special structure of CO problems, SDP relaxations often suffer from degeneracy. A theoretical framework for addressing these issues is discussed in the next section.

Example 2.1. *The Shor's SDP relaxation for the binary set* $F := \{x \in \{0,1\}^n \mid Ax \le b\}$ *for some* $A \in \mathbb{R}^{m \times n}$ *and* $b \in \mathbb{R}^m$ *is*

$$S := \left\{Y \in \mathbb{S}_+^{n+1} \mid \operatorname{Arrow}(Y) = e_0,\ \mathcal{A}(Y) \le 0\right\}. \tag{4}$$

Here, $e_0 \in \mathbb{R}^{n+1}$ *is the first standard unit vector. The operator* $\mathcal{A} : \mathbb{S}_+^{n+1} \to \mathbb{R}^m$ *is given by*

$$(\mathcal{A}(Y))_i := \left\langle \begin{bmatrix} -b_i & \frac{1}{2}a_i^T \\ \frac{1}{2}a_i & 0 \end{bmatrix}, Y \right\rangle \text{ for } i = 1, \dots, m,$$

where $a_i^T x = b_i$ *is the ith constraint in the system* $Ax \le b$*. And the arrow operator* Arrow : $\mathbb{S}_+^{n+1} \to \mathbb{R}^{n+1}$ *is*

$$\operatorname{Arrow}(Y) := \begin{bmatrix} Y_{00} \\ Y_{11} - \frac{1}{2}(Y_{01} + Y_{10}) \\ \vdots \\ Y_{nn} - \frac{1}{2}(Y_{0n} + Y_{n0}) \end{bmatrix} \in \mathbb{R}^{n+1}. \tag{5}$$

The arrow constraint requires $Y_{00} = 1$ *and* $Y_{0,i} = Y_{i,i}$ $(i = 1, \dots, n)$*; and it is a relaxation for the non-linear binary constraint* $x_i = x_i^2$ *in (1). Thus the matrix variable* Y *is in the form of*

$$Y = \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \text{ for some } x, X \text{ satisfying } x_i = X_{ii}.$$

2.2 Facial reduction

Facial reduction, proposed by Borwein and Wolkowicz in [5, 6], is a regularization technique for solving SDP problems. We say that *Slater's condition* holds for the set K in (3) if there exists a positive definite feasible solution $Y \in K$. When Slater's condition is satisfied, we can optimize over K with improved numerical stability; otherwise, solvers may run into serious numerical issues and fail to find the correct optimal solutions. Facial reduction is a framework for restoring Slater's condition, and its implementation is referred to as a *facial reduction algorithm* (FRA). While the theory of facial reduction is well-established, its implementation in practice is still in its infancy.

The theory of facial reduction relies on the fact that if the set K does not satisfy Slater's condition, then it must be contained in a proper face of the positive semidefinite cone. To identify such a face, we can identify a so-called *exposing vector*[1] as every face of the positive semidefinite cone is exposed. A nonzero matrix $W \in \mathbb{S}_+^{n+1}$ is called an exposing vector for the set K in (3) if

$$\langle W, Y \rangle = 0 \quad \forall\, Y \in K, \tag{6}$$

where $\langle \cdot, \cdot \rangle$ is the Frobenius inner product of two matrices. An exposing vector allows us to regularize K. For example, for the first standard unit vector $e_0 = (1, 0, \dots, 0)^T$, if $W = e_0 e_0^T$ is an exposing vector for K, then (6) implies that the first row and the first column of any $Y \in K$ must be zero. In this case, we replace $Y_+ \in \mathbb{S}^{n+1}$ by a smaller positive semidefinite constraint $R \in \mathbb{S}_+^n$, and reformulate the set K correspondingly. Given any non-trivial exposing vector W, we can reformulate K equivalently as

$$K = V\mathcal{R}V^T, \quad \text{where } \mathcal{R} = \left\{R \in \mathbb{S}_+^r \mid \mathcal{A}(VRV^T) = d\right\}, \tag{7}$$

[1] Although the variables in SDP are matrices, we still call them exposing vectors as a convention.

$r := \dim(\text{null}(W))$, and $V \in \mathbb{R}^{(n+1)\times r}$ is the so-called *facial range vector* which is any matrix with linearly independent columns such that $\text{range}(V) = \text{null}(W)$.

If W is of maximum rank, then we restore Slater's condition, and this situation is referred to as *complete facial reduction*; otherwise, we call it *partial facial reduction*. Although partial facial reduction does not restore Slater's condition, it reduces the problem size of the SDP and enhances the numerical stability. In practice, facial reduction, whether complete or partial, can reduce the solving time of SDPs.

To find an exposing vector for K in (3), Borwein and Wolkowicz in [5,6] consider the so-called *FR auxiliary problem*

$$\{\mathcal{A}^*(y) \in \mathbb{S}^{n+1}_+ \mid d^T y = 0\}, \tag{8}$$

where $\mathcal{A}^*$ is the adjoint operator of $\mathcal{A}$. Their fundamental result shows that if K does not satisfy Slater's condition, then the FR auxiliary problem (8) must have a non-trivial solution $\mathcal{A}^*(y)$ which is an exposing vector for K. This allows us to reformulate the original problem as in (7) and thus reduce the problem size. This process called a *facial reduction step*. In general, we need to apply this procedure to the reformulated problem iteratively until we can't find any non-trivial solution from the auxiliary problem (8). In this case, we can conclude that Slater's condition is restored.

Since the size of the reformulated problem is reduced strictly after each FR step, the total number of facial reduction steps required is at most $n+1$. Jos Sturm in [31] defines the *singularity degree* of a spectrahedron K to be the smallest number of facial reduction steps to restore Slater's condition, and it is used to derive error bounds for linear matrix inequalities. Since then, singularity degree becomes a very important parameter in conic optimization, see [9,16,19,21,24,25,30,32].

The main issue in practice is that the auxiliary problem (8) can be equally challenging to solve as the original optimization problem. Therefore, a more practical approach is to target exposing vectors that may not have the maximum rank. Over time, a number of facial reduction algorithms have been developed for generating such exposing vectors. In Section 3, we introduce a new FRA which exploits the structures in the underlying CO problems, and it allows us to find exposing vectors more effectively.

2.3 Polyhedron theory and linear algebra

Let $P = \{x \in \mathbb{R}^n \mid Ax \le b\}$ be a polyhedron. We say $c^T x \le \delta$ is *valid* for P if $c^T x \le \delta$ for every $x \in P$. The *recession cone* of P is $\text{rec}\, P := \{x \in \mathbb{R}^n \mid Ax \le 0\}$. It is well-known that the set of valid inequalities for P can be characterized by the set of non-negative combinations of the linear inequalities in $Ax \le b$, see [7].

Lemma 2.1. *Assume that* $\dim P > \dim \text{rec}\, P$. *An inequality* $c^T x \le \delta$ *is valid for* P *if and only if there exists* $u \ge 0$ *such that* $u^T A = c^T$ *and* $u^T b = \delta$.

We note that under Assumption 2.1, the assumption $\dim P > \dim \text{rec}\, P$ in Lemma 2.1 holds. The i-th inequality $a_i^T x \le b_i$ in the system $Ax \le b$ is called an *implicit equality* if $a_i^T x = b_i$ holds for every $x \in P$. The *affine hull* of P is the smallest affine set containing P. Denote by $A^= x = b^=$ the implicit inequalities in $Ax \le b$. Then the affine hull of the polyhedron is given by

$$\text{aff}\, P = \{x \in \mathbb{R}^n \mid A^= x = b^=\}. \tag{9}$$

Complementary slackness is a fundamental relation between the primal optimal solution and the dual optimal solution. For LP problems with a finite optimal value, Goldman and Tucker [14] show that strict complementary slackness condition holds as well.

Theorem 2.1 (Goldman-Tucker theorem)**.** *Consider the primal and dual LP problems with a finite optimal value,*

$$\min\left\{c^T x \mid Ax = b, x \geq 0\right\} \text{ and } \max\left\{b^T y \mid A^T y + s = c, s \geq 0\right\}.$$

There exists primal and dual optimal solutions x^* and (y^*, s^*) such that $x^* + s^* > 0$.

We need the following result about the eigenvalues in Lemma 4.4 in Section 4.

Theorem 2.2 (Cauchy interlacing theorem)**.** *Let $A \in \mathbb{S}^{n+1}$ with eigenvalues $\lambda_1, \ldots, \lambda_{n+1}$. If $B \in \mathbb{S}^n$ is a principal submatrix of A with eigenvalues $\beta_1, \ldots, \beta_n$, then*

$$\alpha_j \leq \beta_j \leq \alpha_{j+1} \text{ for all } j \leq n.$$

3 Affine FR

In this section, we introduce *affine FR* as a means of regularizing SDP relaxations for CO problems. Affine FR serves as a pre-processing algorithm for SDPs, consistent with the principle of *simple and quickly* advocated by Andersen and Andersen in [1]. As the effectiveness of pre-processing algorithms is greatly influenced by the specific problem instances, it is essential for a pre-processing algorithm to strike a balance between time consumption and simplification achieved. If a pre-processing algorithm takes significant time without yielding substantial simplifications, it can have a detrimental impact on the overall computational time.

Affine FR exploits the fact that any affine set containing the feasible region $F \subseteq \{0,1\}^n$ leads to a partial facial reduction. To see this, assume $F \subseteq L$ for some r-dimensional affine subspace L given by

$$L = \left\{x \in \mathbb{R}^n \mid U^T \left[\begin{smallmatrix}1\\x\end{smallmatrix}\right] = 0\right\} \text{ where } U \in \mathbb{R}^{(n+1)\times(n-r)}. \tag{10}$$

Define $W := UU^T \in \mathbb{S}^{n+1}_+$. Then rank $(W) = n - r$. We claim that W is an exposing vector for the lifted set F_1 defined in (2). For any $x \in F$, we define $\tilde{x} = \left[\begin{smallmatrix}1\\x\end{smallmatrix}\right]$. Then it holds that

$$\langle W, Y\rangle = \tilde{x}^T UU^T \tilde{x} = \left\|U^T \tilde{x}\right\|^2 = 0 \text{ for every } Y = \tilde{x}\tilde{x}^T \in F_1.$$

This demonstrates that an r-dimensional affine set containing F generates an exposing vector W of rank $n-r$. As a result, we can conduct a partial facial reduction, decreasing the matrix variable order in any SDP relaxation for F from $n+1$ to $r+1$. Specifically, when $r = \dim(F)$, we have L representing the affine hull of F, maximizing the rank of W. Consequently, Slater's condition is restored, leading to a complete facial reduction. In both scenarios, partial or complete facial reduction is accomplished without the need to solve the challenging auxiliary problem (8).

To determine an affine set L in (10), we suggest using the linear system that defines F. Since we assume that $F = P \cap \{0,1\}^n$, where P represents a given polyhedron defined as $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$, it follows that $F \subseteq P$. Therefore, we have $F \subset \operatorname{aff} F \subseteq \operatorname{aff} P$. An advantage of using aff P is that affine FR does not require a description of the SDP relaxation, and it only depends on the linear system $Ax \leq b$. We will provide an efficient subroutine for computing aff P in Section 3.1.

Now we are able to provide a complete description of affine FR. When provided with a linear system $Ax \leq b$ such that $F = \{x \in \{0,1\}^n \mid Ax \leq b\}$, affine FR executes the following steps.

- Step 1: Compute the affine hull $L := \operatorname{aff}\{x \in \mathbb{R}^n \mid Ax \leq b\}$ using (11).

- Step 2: Form the exposing vector W.

Some comments are in order.

- Tunçel, in [33], shows that the affine hull of F allows one to restore Slater's condition for any SDP relaxations. However, obtaining an explicit description of aff F is often a challenging computational task. In fact, determining the affine hull in general is NP-hard.[2] The main idea behind affine FR is to use an arbitrary affine set that contains F to establish a computationally feasible approach. The trade-off is that affine FR results in a partial facial reduction rather than a complete one.

- Affine FR does not restore Slater's condition in general. This is always the case for any special FRAs, unless the problem has very special structure. While Slater's condition is not restored, the efforts are not in vein. If at least a non-trivial exposing vector is identified, then we still benefit from a strict reduction in the problem size. Moreover, the reduced problem may have a smaller singularity degree. And this implies improved numerical stability.

- It is worth to note that while we usually need exponentially many linear inequalities to describe *conv* F, at most n equalities are needed to define aff F. Thus, it is reasonable to expect that aff P can capture some valid equalities for aff F. And this is indeed the case in our experiments.

We provide a concrete example to clarify affine FR next.

Example 3.1. *Consider the binary feasible set*

$$F := P \cap \{0,1\}^3 = \{(1,0,0),(0,1,0)\},$$

where $P = \{x \in [0,1]^3 \mid 2x_1 + x_2 \le 2, x_1 + 2x_2 \le 2, x_3 \le 0\}$*. The affine hull of* P *is*

$$\operatorname{aff} P := \{x \in \mathbb{R}^3 \mid x_3 = 0\}.$$

The set F *and* aff P *are depicted in Figures 1 and 2. Applying affine FR, we obtain an exposing vector* W *of rank* 1*, and a facial range vector* V *of size* 4 *by* 3*, i.e.,*

$$W = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}^T \quad \text{and } V = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Thus, we can reduce the size of the matrix variable in any SDP relaxation for F *from* 4 *to* 3*.*

[2] For example, given a set of positive integers $a_1, \dots, a_n$, the subset sum problem asks if there exists a subset $T \subseteq \{1, \dots, n\}$ such that $\sum_{i \in T} a_i = \sum_{i \notin T} a_i$. The subset sum problem is NP-hard and its answer is NO if and only if the dimension of the associated affine hull is -1.

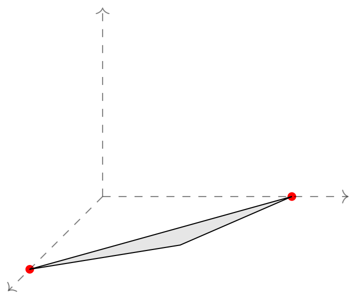

Figure 1: The red dots are F, and the polyhedron P is the grey area.

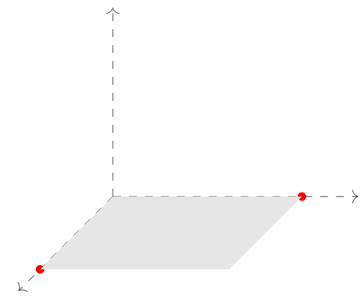

Figure 2: The affine hull of P.

It should be noted that affine FR employed in the previous example does not restore Slater's condition. This fact becomes evident in the continued example provided below.

Example 3.2. *The affine hull of* $F := \{(1,0,0),(0,1,0)\}$ *is*

$$\operatorname{aff} F := \{x \in \mathbb{R}^3 \mid x_1 + x_2 = 1, x_3 = 0\}.$$

The affine hull of F *is* 1*-dimensional, see Figures 3 and 4. This yields an exposing vector* W *of rank* 2*, and a facial range vector* V *of size* 4 *by* 2 *below.*

$$W = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix}^T + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}^T = \begin{bmatrix} 1 & -1 & -1 & 0 \\ -1 & 1 & 1 & 0 \\ -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Thus, if we know the affine hull of F*, then facial reduction can be applied to reduce the size of the matrix variable in any SDP relaxation for* F *from* 4 *to* 2*, and restore Slater's condition.*

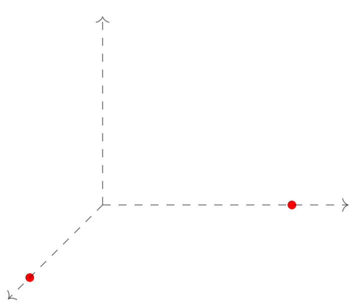

Figure 3: The red dots are F.

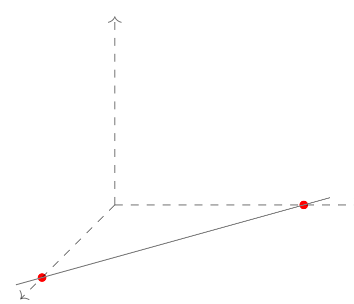

Figure 4: The affine hull of F

3.1 Computing the Affine Hull

In this subsection, we discuss implementation details concerning the computation of the affine hull of a polyhedron P. To determine $\operatorname{aff} P$, it is necessary to identify all implicit equalities in the system $Ax \leq b$ defining P, as indicated in (9). In [10], Fukuda proposes solving at most m distinct LP problems associated with P to find all implicit equalities. However, this approach can be computationally expensive, particularly when the number of inequalities m is significant. Here, we present an alternative approach.

As $P \neq \emptyset$, the strong duality theorem implies the equality between the following primal and dual LP problems

$$\max\{0 \mid Ax \leq b\} = \min\{b^T y \mid y^T A = 0, y \geq 0\}.$$

Note that any primal feasible solution x is optimal. Let y^* be any optimal dual solution. By complementary slackness condition, it holds that $(b - Ax)^T y^* = 0$. Let $I := \{i \mid y_i^* > 0\}$ be the set of positive entries in y^*. The primal constraints $Ax \leq b$ associated with I are always active for any $x \in P$. By definition, this implies that $a_i^T x = b_i$ for $i \in I$ are implicit equality constraints. If the optimal solution y^* contains the maximum number of non-zeros, then I contains all implicit equality constraints by Goldman-Tucker theorem 2.1; and in this case, we have

$$\operatorname{aff} P = \{x \in \mathbb{R}^n \mid a_i^T x = b_i \text{ for every } i \in I\}.$$

Finding an optimal solution y^* with the maximum number of non-zeros can be handled by solving a single LP problem (11), see [4, 22]. We provide a self-contained proof here to show that (11) yields an optimal solution y^* with the maximum number of non-zeros.

$$\begin{array}{rrcl} \max & e^T u & & \\ \text{subject to} & (u+v)^T \begin{bmatrix} A & b \end{bmatrix} & = & 0 \\ & u, v & \geq & 0 \\ & u & \leq & 1. \end{array} \tag{11}$$

Let y^* be an optimal dual solution with the maximum number of non-zeros. Since the dual optimal set is a cone, we can scale y^* by a positive constant such that each positive entry in y^* is at least one; and this does not change the number of non-zeros. Now we can set $u^* = \min\{y^*, e\}$ and $v^* = y^* - u^*$ to obtain a feasible solution for (11) such that $e^T u^*$ is exactly the number of non-zeros in y^*. Conversely, assume u^* and v^* are optimal solutions for (11). By optimality, we can easily see that $u^* \in \{0,1\}^m$, and $u_i^* = 0$ implies $v_i^* = 0$. Thus, $y^* = u^* + v^*$ is a dual optimal solution whose number of non-zeros is exactly $e^T u^*$.

Thus, we can compute the affine hull of a polyhedron with n variables and m inequalities by solving the LP problem (11) with $2m$ variables and $n+1$ inequalities. As the variable u is bounded from above, the bounded-variable simplex method can be applied to solve (11). While the number of variables in the LP problem (11) is doubled, its computational cost is still negligible relative to the cost of solving its SDP relaxation.

Remark 3.1. *We can speed up the costs for solving* (11) *as follows. Let x be a given feasible solution. Then we can verify if an inequality $a_i^T x < b_i$ holds strictly easily, and if so, it is not an implicit equality. Thus, there is no need to introduce variables u_i and v_i in* (11) *for x_i. For certain applications, it is not difficult to generate some feasible solutions and they can be used to reduce the size of* (11) *substantially.*

4 Theoretical comparisons

In this section, we conduct a theoretical comparison for different types of FRAs in the literature. Facial reduction involves the task of finding an exposing vector as outlined in (6). The challenge lies in generating an exposing vector efficiently in practical applications. Existing methods can be categorized into two main classes:

1. The first class of methods is specifically designed to address certain problems using analytical approaches. These methods derive analytical formulas for the exposing vector, making them highly efficient as they require no additional computational work. However, these methods are limited to problems with very specific structures and may necessitate a higher level of expertise for implementation, making them less user-friendly.

2. The second class of methods is suitable for general SDP problems. Here, we consider partial FR in [26] and Sieve-SDP in [40]. They are fully automatic and the performance depends on the formulation of the SDP problem instance. However, these methods do not exploit the structures of the underlying CO problems for the SDP relaxation. As a result, they may fail to achieve sufficient reduction for CO problems, as observed in our analysis and experiments.

The key features and characteristics of these methods are summarized in Table 1.

Method	Applicability	Costs	Effectiveness	User friendly/Automatic
Analytical	Very Low	Zero	High	No
Sieve-SDP	High	Very Low	Low	Yes
Partial FR	High	Low	Low	Yes
Affine FR	High	Low	Medium	Yes

Table 1: Comparison among different FRAs for CO problems.

It is important to emphasize that each of the four FRAs offers distinct advantages depending on the nature of the problem at hand. Existing methods continue to be an excellent choice for special structured instances or general instances. The purpose of this comparison is to provide clarity regarding the differences between these FRAs, enabling users to select the most suitable method for their specific problems. For instance, if a problem exhibits a substantial amount of structure, it would be worth exploring whether an analytical formula can be derived for its affine hull. Conversely, when confronted with an SDP problem lacking explicit structural information, employing Sieve-SDP or Partial FR would be advisable. If the SDP problem serves as a relaxation for a combinatorial optimization problem, affine FR proves to be a favorable option.

The main theoretical result in this section is as follows.

Theorem 4.1. *Let r_A be the size of the matrix variable after applying affine FR to the linear system $Ax \leq b$. Denote by r_P and r_P^+ the size of the matrix variable after applying partial FR to the Shor's relaxation using inner approximations $\mathcal{D}$ and $\mathcal{DD}$, respectively, see details in Section 4.2. Let r_S be the size of the matrix variable after applying Sieve-SDP to the Shor's relaxation. Then*

$$r_A \leq r_P^+ \leq r_P \leq r_S = n + 1.$$

Proof. See Corollaries 4.1 and 4.2 and Lemma 4.5. □

4.1 Analytical approach

In the analytical approach, we need to find an analytical expression for $\operatorname{aff} F$ or an affine set containing F. This yields an exposing vector for the SDP relaxation of F. In [38], Zhao et al. apply the analytical approach to derive several SDP relaxations satisfying Slater's condition for the quadratic assignment problem. Their analysis relies on the fact that the set of permutation matrices has a compact description

$$conv\, F = \{X \in [0,1]^{n\times n} \mid X\mathbf{1} = \mathbf{1}, X^T\mathbf{1} = \mathbf{1}, X \geq 0\}, \tag{12}$$

where $\mathbf{1}$ is the all-ones vector of length n. As $X \geq 0$ are not implicit equalities, the affine hull of F can be defined by the equality constraints in $conv\, F$ above. This allows them to derive $\operatorname{aff} F$ analytically and thus achieve a complete facial reduction.

Assuming the affine hull can be derived, the analytical approach has essentially zero computational cost and thus it is the best choice. However, the analytical approach is very limited. In fact, most of the subsequent studies using the analytical approach are just another variant of the quadratic assignment problem, e.g., the graph partitioning problem [20, 37], the min-cut problem [18], the vertex separator problem [29]. The quadratic cycle covering [8] and the quadratic shortest path problem [15] apply similar arguments to restore Slater's condition for their SDP relaxations. In general, the binary feasible set F does not admit a compact description as in (12), and it is NP-hard to derive an analytical formula. Thus, it is more practical to develop a numerical algorithm for generating exposing vectors for general problems.

If affine FR is applied to the above mentioned CO problems, then it always restores Slater's condition. This is a direct consequence of the following simple result in Lemma 4.1. Hence, when it is uncertain whether a CO problem exhibits any structures that can be used to derive an analytical formula, Lemma 4.1 suggests that using affine FR is a reliable choice.

Lemma 4.1. *Let $F = P \cap \{0,1\}^n$ for some polyhedron P. If $\operatorname{aff} F = \operatorname{aff} P$, then affine FR restores Slater's condition for any SDP relaxations for F.*

4.2 Partial FR

Permenter and Parrilo in [26] propose a special FRA called *partial facial reduction* (partial FR).[3] It replaces the positive semidefinite constraint $X \in \mathbb{S}_+^{n+1}$ in the SDP relaxation by a more tractable convex cone $\mathcal{K} \supseteq \mathbb{S}_+^{n+1}$. Consequently, the auxiliary problem (8) becomes an easier conic optimization problem over the dual cone $\mathcal{K}^*$, i.e.,

$$\{\mathcal{A}^*(y) \in \mathcal{K}^* \mid d^T y = 0\}. \tag{13}$$

If we find a non-trivial solution $W \in \mathcal{A}^*(y)$ in (13), then it is an exposing vector for the original problem as $\mathcal{K}^* \subseteq \mathbb{S}_+^{n+1}$. As in (7), partial FR reformulates the problem using W. We call this *a FR step with respect to* $\mathcal{K}$. This procedure is then repeated for the reformulated problem until we can't identify any non-trivial exposing vector. In general, we have to implement more than one FR step with respect to $\mathcal{K}$, and this can be computationally expensive. For any spectrahedron K, we define $\operatorname{sd}(K, \mathcal{K})$ as the smallest number of iterations needed for partial FR with $\mathcal{K}$ to terminate minus one, and we call $\operatorname{sd}(K, \mathcal{K})$ *the singularity degree of K with respect to $\mathcal{K}$.*

Similar to the singularity degree of K, we can compute $\operatorname{sd}(K, \mathcal{K})$ by picking an exposing vector of maximum rank in (13) at each FR step with respect to $\mathcal{K}$. This yields a lower bound on the iteration complexity for partial FR algorithm. A maximum rank exposing vector for some $\mathcal{K}$ can be obtained by solving an SDP problem which is expensive, see [26].

[3] The same name "partial facial reduction" is also used to describe the situation when facial reduction does not restore Slater's condition in the literature. The intended meaning of "partial facial reduction" is typically clear from the context.

Partial FR makes an inner approximation for the auxiliary problem (8). Thus, it is possible that (8) contains a non-trivial exposing vector W, but (13) is infeasible. In this case, partial FR does not detect any exposing vectors, and Slater's condition is not restored. Clearly, if $\mathcal{K}^*$ is an accurate inner approximation for $\mathbb{S}_+^{n+1}$, then it is more likely to obtain an exposing vector. As suggested in [26], we consider the following two choices for $\mathcal{K}$. And (13) becomes an LP problem that is computationally inexpensive for both choices.

- The cone $\mathcal{D}$ of non-negative diagonal matrices defined by

$$\mathcal{D} := \{W \mid W_{ii} \geq 0 \text{ for every } i\}.$$

- The cone $\mathcal{DD}$ of diagonally dominant matrices given by

$$\mathcal{DD} := \left\{ W \;\middle|\; W_{ii} \geq \sum_{j \neq i} |W_{i,j}| \text{ for every } i \right\}.$$

To compare these methods, we still need to specify the settings. For partial FR, we consider the Shor's SDP relaxation associated with the same linear system $Ax \leq b$ defining F, see its construction in Example 2.1. In this way, these methods have roughly the same computational costs and thus we have a fair comparison. The settings are summarized in the following table.

The matrix size after reduction	Method	Setting
r_A	Affine FR	Apply it to $Ax \leq b$
r_p	Partial FR	(a) The Shor's SDP relaxation (b) Non-negative diagonal matrices $\mathcal{D}$
r_p^+	Partial FR	(a) The Shor's SDP relaxation (b) Diagonally dominant matrices $\mathcal{DD}$

We first state the FR auxiliary problem for the Shor's SDP relaxation more explicitly. An important observation is that the semidefinite constraint for the auxiliary problem has an arrowhead structure, and this is the key in our analysis later.

Lemma 4.2. *The FR auxiliary problem for the Shor's SDP relaxation* (4) *is*

$$W := \begin{bmatrix} -b^T y & \frac{1}{2}(y^T A - z^T) \\ \frac{1}{2}(A^T y - z) & \mathrm{Diag}(z) \end{bmatrix} \in \mathbb{S}_+^{n+1} \text{ and } y \geq 0, \tag{14}$$

where $y \in \mathbb{R}^m$ and $z \in \mathbb{R}^n$. Any feasible W is an exposing vector of a face of $\mathbb{S}_+^n$ containing F.

Proof. We introduce a new variable $s \in \mathbb{R}^m$ to convert the Shor's SDP relaxation (4) into standard form, i.e.,

$$\left\{(Y, s) \in \mathbb{S}_+^n \times \mathbb{R}_+^m \mid \mathcal{A}(Y) + s = 0, \mathrm{Arrow}(Y) = e_0\right\}.$$

Then the FR auxiliary problem is

$$(\mathcal{A}^*(y) + \mathrm{Arrow}^*(\tilde{z}), y) \in \mathbb{S}_+^n \times \mathbb{R}_+^m \text{ and } e_0^T z = 0. \tag{15}$$

where $y \in \mathbb{R}^m$ and $\tilde{z} \in \mathbb{R}^{n+1}$. The equality constraint $e_0^T \tilde{z} = 0$ implies that $\tilde{z}_0 = 0$. Thus, if we define $z \in \mathbb{R}^n$ to be the vector formed by the last n entries in $\tilde{z} \in \mathbb{R}^{n+1}$, then

$$\text{Arrow}^*(\tilde{z}) = \begin{bmatrix} 0 & -\frac{1}{2}z \\ -\frac{1}{2}z & \text{Diag}(z) \end{bmatrix}.$$

The adjoint of $\mathcal{A}$ is given by

$$\mathcal{A}^*(y) = \begin{bmatrix} -b^T y & \frac{1}{2}y^T A \\ \frac{1}{2}A^T y & 0 \end{bmatrix}.$$

The expression in (14) follows. □

4.2.1 Non-negative diagonal matrices

We first show that partial FR with non-negative diagonal matrices removes the rows and columns in the matrix variable that correspond to variables that are fixed at zero in P.

Lemma 4.3. *Let $\mathcal{W}$ be the set of exposing vector that can be obtained from applying partial FR with non-negative diagonal matrices $\mathcal{D}$ to the Shor's SDP relaxation S in* (4)*. Define the diagonal matrix W by*

$$\begin{cases} W_{ii} = 1 & \text{if } x_i = 0 \text{ for every } x \in P, \\ W_{ij} = 0 & \text{otherwise.} \end{cases} \tag{16}$$

It holds that

1. *$W \in \mathcal{W}$ has the maximum rank in $\mathcal{W}$.*
2. *$sd(S, \mathcal{D}) = 1$.*

Proof. If we use $\mathcal{D}$ to inner approximate $\mathbb{S}_+^{n+1}$ in the FR auxiliary problem (14), then (14) is equivalent to the following LP problem

$$A^T y \geq 0, b^T y \leq 0 \text{ and } y \geq 0. \tag{17}$$

By Assumption 2.1, we have P is non-empty and $x \geq 0$ for every $x \in P$. Thus, we can pick a feasible $x^* \in P$. For any y satisfying (17), we have that $0 \leq y^T A x^* \leq y^T b \leq 0$. This means

$$y^T A x^* = b^T y = 0. \tag{18}$$

Thus, if W is an exposing vector of maximum rank, then it corresponds to a solution y in (17) such that $A^T y$ has the maximum number of non-zeros.

We first show that W given in (16) is in $\mathcal{W}$. If $x_i > 0$ for some $x \in P$, then $(A^T y)_i = 0$ by (18). Assume $x_i = 0$ for every $x \in P$. Then strong duality implies that

$$0 = \max\{x_i \mid Ax \leq b\} = \min\{b^T y \mid y^T A = e_i^T, y \geq 0\}. \tag{19}$$

Let y^* be optimal for the dual problem above. Then y^* is also feasible for the auxiliary problem (17), and more importantly, $(A^T y^*)_i = 1 > 0$. This yields an exposing vector with the ith diagonal entry equals one, and zero otherwise. As the feasible set of the auxiliary problem (17) is a polyhedral cone, we can repeat this argument for every index i satisfying $x_i = 0$ for every $x \in P$; this yields W in (16) and thus $W \in \mathcal{W}$.

It remains to show that W has the maximum rank in $\mathcal{W}$. As partial FR removes any variables x_i such that $x_i = 0$ for every $x \in P$ after the first iteration, there exists no variables that are fixed at zero. Thus, partial FR terminates after the second iteration and $\text{sd}(S, \mathcal{D}) = 1$. □

Based on Lemma 4.3, partial FR with $\mathcal{D}$ removes the ith row and column from the matrix variable Y for each index i with $(A^T y)_i > 0$. And this corresponds to removing all variables x_i such that $x_i = 0$ for every $x \in P$. Note that the variables x_i being removed in Lemma 4.3 are fixed at zero with respect to P. It is possible that $x_i > 0$ for some $x \in P$, and $x_i = 0$ for every $x \in F$. Thus, not all variables fixed at zero in the original feasible set F can be detected in this way.

We can pick y^* satisfying (17) such that $A^T y$ has the maximum number of non-zeros, and then partial FR terminates after two iterations. Note that this vector y^* can be computed via linear programming using the same trick in (3.1). Thus, there is no need to solve an SDP for getting a maximum rank exposing vector in [26].

Corollary 4.1. $r_A \le r_P$.

Proof. Applying Lemma 4.3, the inequality is a direct consequence of (9). We include a detailed proof here. Define $w = A^T y$ for any $y \ge 0$. Then $w^T x \le b^T y$ is a valid inequality for P, see (2.1). If y satisfies (17), then $w \ge 0$ and $b^T y = 0$. Since $x \ge 0$ for every $x \in P$, the valid inequality $w^T x \le 0$ implies that if $w_i > 0$ then $x_i = 0$ for every $x \in P$. Therefore, we have that $\operatorname{aff} P \subseteq \{x \in \mathbb{R}^n_+ \mid x_i = 0 \text{ if } w_i > 0\}$, and thus $r_A \le r_P$. □

Next, we provide some examples to clarify the differences between r_A and r_P in handling the Shor's SDP relaxation.

Example 4.1. *Let P be a polyhedron defined by the linear system $Ax \le b$ with*

$$A = \begin{bmatrix} 1 & 1 \\ -1 & -1 \\ 1 & 0 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} \text{ and } b = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{20}$$

Define $F := \{0,1\}^2 \cap P = \{(0,1)\}$. The auxiliary problem from parital FR with $\mathcal{D}$ for the Shor's SDP relaxation of F is

$$\begin{bmatrix} -y_1 + y_2 & \frac{1}{2}(y_1 - y_2 + y_3 - y_4 - z_1) & \frac{1}{2}(y_1 - y_2 - y_5 - z_2) \\ \frac{1}{2}(y_1 - y_2 + y_3 - y_4 - z_1) & z_1 & 0 \\ \frac{1}{2}(y_1 - y_2 - y_5 - z_2) & 0 & z_2 \end{bmatrix} \in \mathcal{D} \text{ and } y \ge 0.$$

As $b^T y = 0$, we have $y_1 = y_2$ and thus

$$A^T y = \begin{bmatrix} y_3 - y_4 \\ -y_5 \end{bmatrix} \ge 0.$$

As $y \ge 0$, this means $y_5 = 0$ and thus an exposing vector $W \in \mathcal{D}$ of maximum rank has rank one. For example, it can be achieved for any $y \ge 0$ such that $y_5 = 0$ and $y_3 - y_4 > 0$, i.e.,

$$W = \begin{bmatrix} 0 & 0 & 0 \\ 0 & y_3 - y_4 & 0 \\ 0 & 0 & 0 \end{bmatrix} \in \mathcal{D} \subset \mathbb{S}^3_+.$$

Thus only the row and column corresponding to x_1 can be removed. Indeed, $x_1 = 0$ for every $x \in P$. The corresponding facial range vector is

$$V = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 2}.$$

This yields a facially reduced problem for the Shor's SDP relaxation. And applying partial FR to the reduced problem we obtain a new auxiliary problem

$$\begin{bmatrix} -y_1 + y_2 & \frac{1}{2}(y_1 - y_2 - y_5 - z_2) \\ \frac{1}{2}(y_1 - y_2 - y_5 - z_2) & z_2 \end{bmatrix} \in \mathcal{D} \text{ and } y \geq 0$$

As $-y_1 + y_2 = 0$, we have $0 \geq -y_5 = z_2 \geq 0$. Thus, the only solution is zero. This shows that there is no reduction in the second iteration, and partial FR terminates.

Example 4.2. *Let us apply affine FR to the same F in the previous example. The affine hull of P is*

$$\operatorname{aff} P = \{x \in \mathbb{R}^2 \mid x_1 + x_2 = 1, x_1 = 0\}.$$

This yields a rank 2 exposing vector W given by

$$W = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ -1 & 1 \end{bmatrix}^T = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 2 & 2 \\ -1 & 2 & 2 \end{bmatrix}.$$

The associated facial range vector is

$$V = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \in \mathbb{R}^{3\times 1}.$$

As $\dim(F) = 0$, the facial range vector V is the smallest possible and thus Slater's condition is restored by affine FR.

4.2.2 Diagonally dominant matrices

We investigate partial FR with diagonally dominant matrices $\mathcal{DD}$. It turns out that partial FR with $\mathcal{DD}$ removes exactly the variables that are fixed at zero or one in the polyhedron P. We prove this by constructing an exposing vector of maximum rank.

Lemma 4.4. *Let $\mathcal{W}$ be the set of exposing vectors that can be obtained from applying partial FR with diagonally dominant matrices $\mathcal{DD}$ to the Shor's SDP relaxation S in (4) for F. It holds that*

1. *There exists a maximum rank exposing vector $W \in \mathcal{W}$ of the form*

$$W = \begin{bmatrix} -\delta & w^T & 0 \\ w & Diag(w) & 0 \\ 0 & 0 & I \end{bmatrix} \in \mathcal{W}, \tag{21}$$

where $\delta \in \mathbb{R}$, w is a vector satisfying $e^T w = \delta$, and I is an identity matrix of proper size.

2. *The rank of W equals to the number of variables fixed at zero or one in P.*

3. *The singularity degree of S with respect to $\mathcal{DD}$ is one, i.e., $sd(S, \mathcal{DD}) = 1$.*

Proof. We first construct an exposing vector in the form of (21). When we replace $\mathbb{S}^{n+1}_+$ with $\mathcal{DD}$, the auxiliary problem (15) becomes

$$\left\{(y,z) \in \mathbb{R}^m \times \mathbb{R}^{n+1} \mid -b^T y \geq \frac{1}{2} e^T \left|A^T y - z\right|,\ z \geq \frac{1}{2}|A^T y - z|,\ y \geq 0\right\}. \tag{22}$$

As valid inequalities for P can be characterized by $w^T x \leq \delta$ with $w := A^T y$ and $\delta := b^T y$ for some $y \geq 0$, see Lemma 2.1. To simplify the analysis, we write (22) in an equivalent form as the set of vectors $(w, \delta, z) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^{n+1}$ satisfying

$$\begin{array}{rcl} -\delta & \geq & \frac{1}{2} e^T \left|w - z\right|, \\ z & \geq & \frac{1}{2}|w - z|, \\ w^T x & \leq & \delta \text{ is a valid inequality for } P. \end{array} \tag{23}$$

If $x_k = 0$ for every $x \in P$, then we know from Lemma 4.3 and the $\mathcal{D} \subset \mathcal{DD}$ that partial FR can detect and remove them from the problem. This corresponds to identity matrix in the third diagonal block of W in (21). In the subsequent analysis, we can assume that $x > 0$ for some $x \in P$. Let (w, δ, z) be feasible for (23).

- We claim that $w \leq 0$. Assume $w_k > 0$ for some k. Define

  $$z_i^* = \begin{cases} z_i & \text{if } i \neq k \\ w_k & \text{if } i = k. \end{cases}$$

  We show that (w, δ, z^*) is also feasible for (23). Since $|w_k - z_k| \geq |w_k - z_k^*| = 0$, it holds that

  $$\begin{array}{rcl} -\delta & \geq & \frac{1}{2} e^T \left|w - z\right| \geq \frac{1}{2} e^T \left|w - z^*\right|, \\ z_k^* & = & w_k > 0 = \frac{1}{2}\left|w_k - z_k^*\right|. \end{array}$$

  The remaining constraints hold trivially. Thus, (w, δ, z^*) is feasible for (23), and it corresponds to an exposing vector such that the k-th row and column are zeros except the diagonal entry $z_k^* = w_k > 0$. This implies that $x_i = 0$ for every $x \in P$, which is not possible.

- Assume $w = 0$. As $\delta \leq 0$ and P is non-empty, this implies that $\delta = 0$ and thus $z = 0$. In this case, the only exposing vector is zero matrix and there is no reduction.

- Assume $w \leq 0$ is non-zero. Define $z^* := -w$. The constraint $z_i \geq \frac{1}{2}\left|w_i - z_i\right| = \frac{1}{2}(z_i - w_i)$ implies that $z_i \geq -w_i$. Note that the absolute value $|w_i - z_i|$ over $z_i \geq -w_i$ is minimized when we set z_i to be $-w_i$. This yields

  $$-\delta \geq \sum \frac{1}{2}\left|w_i - z_i\right| \geq \sum \frac{1}{2}\left|w_i - z_i^*\right| = -\sum w_i.$$

  Recall that $x \in P$ implies $x \in [0,1]^n$ by Assumption 2.1. Since $w \leq 0$, we obtain that

  $$\delta \leq \sum w_i \leq w^T x \text{ for every } x \in P. \tag{24}$$

But $w^T x \leq \delta$ is a valid inequality for P, see (23). This shows that $w^T x = \delta$ for $x \in P$, and we have equalities everywhere in (24). This yields $\sum w_i = \delta$. Let $\tilde{w} \in \mathbb{R}^k$ be the subvector associated with the non-zero elements in w, and the corresponding non-zero submatrix of W is of the form

$$\tilde{W} = \begin{bmatrix} -\delta & \tilde{w}^T \\ \tilde{w} & \mathrm{Diag}(-\tilde{w}) \end{bmatrix} \in \mathbb{R}^{(k+1)\times(k+1)}. \tag{25}$$

Let $\lambda_1 \leq \cdots \leq \lambda_{k+1}$ be the eigenvalues of $\tilde{W}$. Applying Cauchy interlacing theorem 2.2 to $\tilde{W}$ with respect to the k by k principal submatrix $\mathrm{Diag}(-\tilde{w})$, we obtain that $\lambda_1 \leq -\tilde{w}_1 \leq \lambda_2 \leq \cdots \leq -\tilde{w}_k \leq \lambda_{k+1}$. As $-\tilde{w}_1 > 0$, the eigenvalues $\lambda_2, \cdots, \lambda_{k+1} > 0$ are positive. Since $\sum w_i = \delta$, all ones vector is an eigenvector for $\tilde{W}$ associated with eigenvalue zero, and this means $\lambda_1 = 0$. Thus $\mathrm{rank}\, W = \mathrm{rank}\, \tilde{W} = k$.

As we have equalities in (24), if $w_i < 0$, then $x_i = 1$ for every $x \in P$. From Lemma 2.1, if W is an exposing vector of maximum rank, then we must have $w_i < 0$ for every x_i fixed at in P. Putting together, we have constructed an exposing vector W in the form of (21) whose rank is the same as the number of variables fixed at zero or one in P.

Finally, we show that $\mathrm{sd}(S, \mathcal{DD}) = 1$. To this end, we show that partial FR terminates after reformulating the Shor's SDP relaxation using W. Assume the negative entries in w are exactly the first k variables. The facial range vector associated with W is

$$V = \begin{bmatrix} -1 & 0 \\ e_k & 0 \\ 0 & I_{n-k} \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n-k+1)}.$$

The reformulated problem using V is equivalent to the Shor's SDP relaxation applied to a restricted F where we fix the first k variables $x_1, \ldots, x_k$ to be one. To be more precisely, if $A = \begin{bmatrix} \bar{A} & \tilde{A} \end{bmatrix}$ where $\bar{A} \in \mathbb{R}^{m\times k}$ and $\tilde{A} \in \mathbb{R}^{m\times(n-k)}$, then

$$\tilde{F} := \tilde{P} \in \{0,1\}^{n-k} \text{ where } \tilde{P} := \left\{ \tilde{x} \in \mathbb{R}^{n-k} \mid \tilde{A}\tilde{x} \leq b - \bar{A}e \right\}.$$

It is not hard to see that there exist some $\tilde{x} \in \tilde{P}$ such that $\tilde{x} > 0$. If we apply partial FR to $\tilde{F}$ again, then we must have $w = 0$ and thus $\delta = 0$ in (23). Thus, partial FR terminates in the second iteration.

□

Corollary 4.2. $r_A \leq r_P^+$.

Proof. As $\mathrm{aff}\, P$ is contained in the affine subspace defined by equations $x_i = 0$ or $x_i = 1$ for these fixed variables, we have $r_A \leq r_P^+$. □

Note that the proof of Lemma 4.4 also shows that maximum rank exposing vectors can be computed through solving a single LP problem. Thus, it has about the same computational complexity as affine FR.

Corollaries 4.1 and 4.2 show that, for any linear system $Ax \leq b$ defining F, affine FR yields more reduction than partial FR applied to the Shor's SDP relaxation with $\mathcal{D}$ or $\mathcal{DD}$. This suggests that the knowledge of the binary set F underlying its SDP relaxation can be helpful to find exposing vectors. If we use a more accurate inner approximation for $\mathbb{S}_+^n$ in the partial FR, e.g., Factor-width-k matrices, then we may obtain more reductions from partial FR. However, the associated auxiliary problem is also harder to solve. In fact, even the cone of diagonally dominant matrices $\mathcal{DD}$ already involves $\mathcal{O}(n^2)$ variables in general and thus impractical for problems with thousands of variables.[4] Since affine FR uses only n variables, we do not compare it with more accurate inner approximation in partial FR here.

Remark 4.1. *The inequality $r_A \leq r_P$ in Corollary 4.1 can also be derived as a direct corollary of $r_A \leq r_P^+$ in Corollary 4.2. The proof in Corollary 4.1 is of independent interests as it shows the hidden structure behind partial FR with non-negative diagonal matrices, i.e., it finds precisely variables that are fixed at zero.*

Remark 4.2. *All the results in this section can be extended to mixed binary integer programming problems. For example, if F is a mixed binary feasible set defined as the feasible region of* (29)*, then partial FR with $\mathcal{D}$ for the Shor's SDP relaxation only detect binary variables fixed at zero or one in P, while leaving the continuous variables unaffected. Similarly, when partial FR uses $\mathcal{DD}$, then only binary variables fixed at zero or one are removed. The proof of this extension uses a similar argument as* (4.3) *and* (4.4)*.*

4.2.3 Sieve-SDP

In [40], Zhu et al. propose a special FRA called *Sieve-SDP*. Their special FRA attempts to identify a certain pattern in the constraints so that an exposing vector can be obtained at very little cost. We can state the FR component in Sieve-SDP as follows: for each constraint of the form $\langle A_i, Y\rangle = b_i$ or $\langle A_i, Y\rangle \leq b_i$, we check if it is possible to permute the rows and columns of matrix A_i so that it can be written as

$$A_i = \begin{bmatrix} D_i & 0 \\ 0 & 0 \end{bmatrix} \text{ with } D_i \in \mathbb{S}_{++}^{n+1}. \tag{26}$$

If $b_i = 0$, then A_i is an exposing vector, see (6). Then we can implement FR by simply removing the rows and columns in the matrix variable Y associated with D_i. The algorithm is terminated when it can't find any such constraints. Sieve-SDP avoids solving the auxiliary problem (8) completely, and it only requires an incomplete Cholesky factorization for checking positive definiteness of D_i. This makes Sieve-SDP an extremely fast algorithm.

Next, we consider Sieve-SDP applied to the Shor's SDP relaxation.

Lemma 4.5. *Let r_S be the size of the matrix variable after applying Sieve-SDP to the Shor's relaxation. Then $r_S = n + 1$.*

Proof. We simply list all the data matrices in the Shor's relaxation (4). And then we can easily verify that either they do not contain a positive definite principal submatrix or the right hand side is non-zero, and thus Sieve-SDP does not yield any reduction, i.e., $r_S = n+1$. The data matrices in the constraint $\mathcal{A}(Y) \leq 0$ are

$$\begin{bmatrix} -b_i & \frac{1}{2}a_i^T \\ \frac{1}{2}a_i & 0 \end{bmatrix} \text{ for } i = 1, \ldots, m \tag{27}$$

[4]The complexity of computing r_p^+ for the Shor's relaxation is only $\mathcal{O}(n)$ due to the arrowhead structure in the matrix variable.

If any matrix in (27) contains a positive definite principal submatrix, then we must have $a_i = 0$. Since P is not empty, $a_i = 0$ implies $b_i \geq 0$. Thus there is no such a submatrix.

The data matrices from $\text{arrow}(Y) = e_0$ are

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & -\frac{1}{2}e_i^T \\ -\frac{1}{2}e_i & e_i e_i^T \end{bmatrix} \text{ for } i = 1, \ldots, n. \tag{28}$$

The right hand side for the first matrix (28) is 1, and the remaining matrices in (28) do not contain any positive definite principal submatrices. Thus, there is no reduction. □

We can extend Sieve-SDP as follows to capture more exposing vectors. In practice, we often strengthen the Shor's SDP relaxation as follows. Let $a_i^T x = b_i$ be the ith equality constraint. We add a valid equality constraint $\langle W, Y \rangle = 0$, where

$$W_i = \begin{bmatrix} -b_i \\ a_i \end{bmatrix} \begin{bmatrix} -b_i \\ a_i \end{bmatrix}^T.$$

As W_i is positive semidefinite, it is an exposing vector. Thus, if we extend Sieve-SDP by relaxing the condition in (26) to D is positive semidefinite, then the exposing vector W_i can be detected by the extended Sieve-SDP. To implement FR, we can add up all exposing vectors obtained in this way to get a facial range vector. It is interesting to investigate the performance of this extension in the future research.

5 Numerical experiments

In this section, we test the performance of affine FR. All the experiment was conducted on a MacBook Pro with the following machine specifications: Apple M1 Max, Memory 32 GB, macOS. We demonstrate the effectiveness of affine FR by evaluating its performance on a collection of mixed-binary linear programming instances sourced from the official benchmark set of MIPLIB 2017, see [12]. These test instances are available for download at `https://miplib.zib.de/`.

The mixed-binary linear programming instances from [12] are in the form of

$$\begin{array}{rl} \min & c^T x \\ \text{subject to} & Ax = b \\ & Bx \leq d \\ & l \leq x \leq u \\ & x_i \in \{0, 1\} \quad \text{for } i \in \mathcal{B}. \end{array} \tag{29}$$

The problem has n variables denoted by the vector $x \in \mathbb{R}^n$. The linear system $Ax = b$ and $Bx \leq d$ are the equality and inequality constraints, respectively. The vectors l and u specify the lower and upper bounds for x, respectively. If $l_i = -\infty$, then x_i has no lower bound. Similarly $u_i = \infty$ means x_i has no upper bound. The subset $\mathcal{B} \subseteq \{1, \ldots, n\}$ indicates the binary variables.

The Shor's SDP relaxation for (29) is given by

$$\begin{array}{rl} \min & c^T x \\ \text{subject to} & Ax = b \\ & Bx \leq d \\ & l \leq x \leq u \\ & x_i = X_{ii} \text{ for } i \in \mathcal{B} \\ & Y := \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \in \mathbb{S}^{n+1}_+ . \end{array} \tag{30}$$

The size of the matrix variable Y in the Shor's SDP relaxation is $n+1$. Applying affine FR, we identify certain facial range vector $V \in \mathbb{R}^{(n+1)\times r}$ with $r \leq n+1$. Then we can use the relation $Y = VRV^T$ to replace Y by a smaller positive semidefinite matrix variable $R \in \mathbb{S}^r_+$ as in (7). In practice, problems with thousands variables are very common, and the Shor's SDP relaxation is computationally intractable due to memory requirement. From the computational experiments below, we see that the size of the matrix variable can be reduced substantially.

We emphasize that affine FR is independent of the SDP relaxation used for approximating the original BIP problem. It can reduce the size of the matrix variable for any SDP relaxation. For example, it is possible to strengthen the Shor's SDP relaxation by adding non-negativity constraints $Y \geq 0$. This results in the so-called *doubly non-negative relaxation* which often provide the best known lower bounds for many problems. The doubly non-negative relaxation is more expensive to solve than the Shor's SDP relaxation. Affine FR can also result in an equivalent reduction in the size of the matrix variable in the doubly non-negative relaxation.

The MIPLIP dataset consists of 314 instances of mixed binary linear programming, each with a maximum of ten thousand variables. This range effectively encompasses all relevant applications in SDP approaches. Any problem larger than this size typically becomes intractable without assuming additional structures and cannot be solved within a reasonable amount of time.

We list instances with at least one implicit equality from the inequality constraints $Bx \leq d$. In the first column, we specify the name of each instance. The second/third/fourth columns show the number of variables, the number of equality constraints, as well as the number of inequality constraints present in the problem. The size of the matrix variable before and after the reduction is listed in the fifth and sixth columns, respectively. The ratio represents the reduced size divided by the original size, with a smaller ratio indicating a larger reduction in the problem size. Finally, the last column specifies the time in seconds required to compute the reduction using affine FR.

Instance	Variables	Equalities	Inequalities	The size of the matrix variable			Time
				Before Reduction	After Reduction	Ratio	
pb-market-split8-70-4	71	0	17	72	63	0.88	0.00
misc05inf	136	29	272	137	101	0.74	0.02
p0201	201	0	133	202	146	0.72	0.02
misc07	260	35	177	261	208	0.80	0.03
supportcase14	304	107	127	305	179	0.59	0.03
supportcase16	319	111	19	320	193	0.60	0.03
rlp1	461	10	58	462	364	0.79	0.03
nexp-50-20-1-1	490	18	522	491	462	0.94	0.02
bg512142	792	240	1067	793	543	0.68	0.11
ponderthis0517-inf	975	26	52	976	899	0.92	0.18
neos-4333596-skien	1005	225	587	1006	753	0.75	0.14
berlin_5_8_0	1083	1	1531	1084	989	0.91	0.11
nexp-50-20-4-2	1225	17	523	1226	1205	0.98	0.03
acc-tight5	1339	277	2775	1340	442	0.33	0.92
seymour	1372	0	4944	1373	1256	0.91	0.16
seymour1	1372	0	4944	1373	1256	0.91	0.17
acc-tight4	1620	297	2988	1621	661	0.41	0.80
ns1830653	1629	565	2367	1630	562	0.34	0.42
railway_8_1_0	1796	1	2526	1797	1592	0.89	0.17

Figure 5: Instances with 2000 or less variables

We provide additional details about Figure 5. By applying affine FR, the average reduction in the size of the matrix variable is approximately 75%. Notably, the instance "acc-tight5" achieves the most significant reduction, with the matrix variable's size reduced to only 33% of its original size. This reduction in the matrix variable size significantly reduces the solving time for their SDP relaxations. Moreover, the computational time required for executing the proposed FRA is less than 1 second for all instances, which is essentially negligible compared to the cost of solving the facially reduced SDP relaxation. This observation remains valid even for larger instances.

Instance	Variables	Equalities	Inequalities	The size of the matrix variable			Time
				Before Reduction	After Reduction	Ratio	
10teams	2025	135	95	2026	1459	0.72	4.21
dg012142	2080	640	5670	2081	1401	0.67	0.84
graph20-20-1rand	2183	74	5513	2184	1925	0.88	2.76
uct-subprob	2256	901	1072	2257	1356	0.60	1.16
usAbbrv-8-25_70	2312	1	3290	2313	2104	0.91	0.22
app1-1	2480	1226	3700	2481	1253	0.51	0.34
beasleyC3	2500	500	1250	2501	2000	0.80	0.25
mod010	2655	145	1	2656	2430	0.91	3.37
ns2071214	2720	501	5749	2721	1560	0.57	2.32
neos-503737	2850	150	350	2851	2461	0.86	1.49
ns1208400	2883	339	3950	2884	2261	0.78	17.16
s1234	2945	1	8417	2946	2865	0.97	0.16
neos-691058	3006	462	2205	3007	2505	0.83	1.78
app3	3080	462	304	3081	1553	0.50	1.69
supportcase4	3162	1026	8466	3163	1113	0.35	0.39
neos18	3312	2394	9008	3313	964	0.29	0.60
neos-3216931-puriri	3555	579	5410	3556	2992	0.84	4.39
bnatt400	3600	1586	4028	3601	2014	0.56	0.06
hanoi5	3862	0	16399	3863	2976	0.77	0.37
fhnw-binpack4-77	3924	398	4524	3925	3527	0.90	0.21

Figure 6: Instances with 2000 to 4000 variables

Instance	Variables	Equalities	Inequalities	The size of the matrix variable			Time
				Before Reduction	After Reduction	Ratio	
neos-4387871-tavua	4004	554	4000	4005	3723	0.93	0.89
neos-4393408-tinui	4004	554	4000	4005	3723	0.93	0.98
neos-1605061	4111	666	2808	4112	2681	0.65	12.21
supportcase3	4191	1027	11675	4192	2141	0.51	0.85
neos-1601936	4446	681	2450	4447	3001	0.67	18.69
bnatt500	4500	1971	5058	4501	2530	0.56	0.08
peg-solitaire-a3	4552	1367	3220	4553	3186	0.70	4.18
misc04inf	4897	311	1415	4898	4315	0.88	81.01
fhnw-schedule-paira100	5150	27	9973	5151	5124	0.99	0.12
momentum1	5174	558	42122	5175	3088	0.60	90.98
istanbul-no-cutoff	5282	221	20125	5283	5062	0.96	0.40

Figure 7: Instances with 4000 to 6000 variables

Instance	Variables	Equalities	Inequalities	The size of the matrix variable			Time
				Before Reduction	After Reduction	Ratio	
blp-ir98	6097	66	420	6098	6029	0.99	9.58
swath	6805	504	380	6806	6303	0.93	44.50
swath1	6805	504	380	6806	6303	0.93	41.13
swath2	6805	504	380	6806	6303	0.93	36.42
swath3	6805	504	380	6806	6303	0.93	39.55
air05	7195	426	0	7196	5885	0.82	150.83
lrn	7253	1009	7482	7254	5613	0.77	8.78
fhnw-binpack4-58	7550	600	9300	7551	6951	0.92	0.60
cdma	7891	60	9035	7892	6583	0.83	46.71

Figure 8: Instances with 6000 to 8000 variables

Instance	Variables	Equalities	Inequalities	The size of the matrix variable			Time
				Before Reduction	After Reduction	Ratio	
neos-3581454-haast	8112	48	17172	8113	7477	0.92	9.94
air04	8904	823	0	8905	6545	0.73	359.54
neos-738098	9093	497	25352	9094	8452	0.93	20.71
ta2-UUE	9241	2471	216	9242	5622	0.61	1.21
neos-4343293-stony	9400	1254	9396	9401	8769	0.93	3.42
neos-820879	9522	210	151	9523	7971	0.84	1146.66
rentacar	9559	6294	509	9560	1941	0.20	7.17
blp-ic97	9845	92	831	9846	8358	0.85	131.93
neos-498623	9861	147	1900	9862	9715	0.99	18.66
disctom	10000	399	0	10001	9601	0.96	58.41

Figure 9: Instances with 8000 to 10000 variables

6 Conclusion

In this paper, we develop a new facial reduction algorithm called affine FR for SDP relaxations of CO problems. The SDP relaxation is an important tool for solving challenging combinatorial problems, as it provides extremely tight bounds. A crucial bottleneck in solving SDP relaxations is the lack of regularity conditions. This bottleneck obstructs the application of promising SDP-based approaches to broader classes and larger scales of CO problems. Therefore, research advances to overcome this issue can have a significant impact on the optimization society.

Affine FR is a fully automated pre-processing algorithm that exploits the inherent structures present in CO problems, leading to enhanced performance. Affine FR serves as a valuable addition to the existing approaches for CO problems. We provide a theoretical analysis to demonstrate differences between affine FR and the existing methods. Furthermore, based on our experiments conducted in Section 5, we have observed that affine FR demonstrates remarkable potential in preprocessing SDP relaxations of CO problems. Our findings underscore the practical effectiveness and suitability of affine FR.

7 Acknowledgement

The authors would like to thank Professor Gábor Pataki for his valuable insights and comments on Sieve-SDP, which have contributed to the improved presentation of this work.

References

[1] E. D. Andersen and K. D. Andersen, *Presolving in linear programming*, Mathematical Programming, 71 (1995), pp. 221–245.

[2] M. F. Anjos and J. B. Lasserre, *Handbook on semidefinite, conic and polynomial optimization*, vol. 166, Springer Science & Business Media, 2011.

[3] X. Bao, N. V. Sahinidis, and M. Tawarmalani, *Semidefinite relaxations for quadratically constrained quadratic programming: A review and comparisons*, Mathematical programming, 129 (2011), pp. 129–157.

[4] D. Bertsimas and J. N. Tsitsiklis, *Introduction to linear optimization*, vol. 6, Athena scientific Belmont, MA, 1997.

[5] J. Borwein and H. Wolkowicz, *Regularizing the abstract convex program*, Journal of Mathematical Analysis and Applications, 83 (1981), pp. 495–530.

[6] J. M. Borwein and H. Wolkowicz, *Facial reduction for a cone-convex programming problem*, Journal of the Australian Mathematical Society, 30 (1981), pp. 369–380.

[7] M. Conforti, G. Cornuéjols, and G. Zambelli, *Integer programming*, vol. 271, Springer, 2014.

[8] F. De Meijer and R. Sotirov, *SDP-based bounds for the quadratic cycle cover problem via cutting-plane augmented lagrangian methods and reinforcement learning: Informs journal on computing meritorious paper awardee*, INFORMS Journal on Computing, 33 (2021), pp. 1262–1276.

[9] D. Drusvyatskiy, G. Pataki, and H. Wolkowicz, *Coordinate shadows of semidefinite and euclidean distance matrices*, SIAM Journal on Optimization, 25 (2015), pp. 1160–1178.

[10] K. Fukuda, *Lecture: Polyhedral computation, spring 2016*, Institute for Operations Research and Institute of Theoretical Computer Science, ETH Zurich. https://inf. ethz. ch/personal/fukudak/lect/pclect/notes2015/PolyComp2015. pdf, (2016).

[11] E. Gaar, D. Puges, and A. Wiegele, *Strong SDP based bounds on the cutwidth of a graph*, arXiv preprint arXiv:2301.03900, (2023).

[12] A. Gleixner, G. Hendel, G. Gamrath, T. Achterberg, M. Bastubbe, T. Berthold, P. M. Christophel, K. Jarck, T. Koch, J. Linderoth, M. Lübbecke, H. D. Mittelmann, D. Ozyurt, T. K. Ralphs, D. Salvagnin, and Y. Shinano, *MIPLIB 2017: Data-Driven Compilation of the 6th Mixed-Integer Programming Library*, Mathematical Programming Computation, (2021).

[13] M. X. Goemans and D. P. Williamson, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, Journal of the ACM (JACM), 42 (1995), pp. 1115–1145.

[14] A. J. Goldman and A. W. Tucker, *4. theory of linear programming*, in Linear Inequalities and Related Systems.(AM-38), Volume 38, Princeton University Press, 2016, pp. 53–98.

[15] H. Hu and R. Sotirov, *On solving the quadratic shortest path problem*, INFORMS Journal on Computing, 32 (2020), pp. 219–233.

[16] H. Im and H. Wolkowicz, *Revisiting degeneracy, strict feasibility, stability, in linear programming*, European Journal of Operational Research, (2023).

[17] N. Krislock, J. Malick, and F. Roupin, *Biqcrunch: A semidefinite branch-and-bound method for solving binary quadratic problems*, ACM Transactions on Mathematical Software (TOMS), 43 (2017), pp. 1–23.

[18] X. Li, T. K. Pong, H. Sun, and H. Wolkowicz, *A strictly contractive peaceman-rachford splitting method for the doubly nonnegative relaxation of the minimum cut problem*, Computational Optimization and Applications, 78 (2021), pp. 853–891.

[19] Y. Lin, S. B. Lindstrom, B. F. Lourenço, and T. K. Pong, *Optimal error bounds in the absence of constraint qualifications for the generalized power cones, and further applications*, arXiv preprint arXiv:2211.16142, (2022).

[20] A. Lisser and F. Rendl, *Graph partitioning using linear and semidefinite programming*, Mathematical Programming, 95 (2003), pp. 91–101.

[21] B. F. Lourenço, *Amenable cones: error bounds without constraint qualifications*, Mathematical Programming, 186 (2021), pp. 1–48.

[22] M. Mehdiloo, *On finding a relative interior point of a polyhedral set*, Mathematics Interdisciplinary Research, 6 (2021), pp. 121–138.

[23] D. E. Oliveira, H. Wolkowicz, and Y. Xu, *ADMM for the SDP relaxation of the QAP*, Mathematical Programming Computation, 10 (2018), pp. 631–658.

[24] G. Pataki, *On positive duality gaps in semidefinite programming*, arXiv preprint arXiv:1812.11796, (2018).

[25] G. Pataki and A. Touzov, *How do exponential size solutions arise in semidefinite programming?*, arXiv preprint arXiv:2103.00041, (2021).

[26] F. Permenter and P. Parrilo, *Partial facial reduction: simplified, equivalent sdps via approximations of the psd cone*, Mathematical Programming, 171 (2018), pp. 1–54.

[27] S. Poljak, F. Rendl, and H. Wolkowicz, *A recipe for semidefinite relaxation for (0, 1)-quadratic programming: In memory of svata poljak*, Journal of Global Optimization, 7 (1995), pp. 51–73.

[28] F. Rendl, G. Rinaldi, and A. Wiegele, *Solving max-cut to optimality by intersecting semidefinite and polyhedral relaxations*, Mathematical Programming, 121 (2010), pp. 307–335.

[29] F. Rendl and R. Sotirov, *The min-cut and vertex separator problem*, Computational Optimization and Applications, 69 (2018), pp. 159–187.

[30] S. Sremac, H. J. Woerdeman, and H. Wolkowicz, *Error bounds and singularity degree in semidefinite programming*, SIAM Journal on Optimization, 31 (2021), pp. 812–836.

[31] J. F. Sturm, *Error bounds for linear matrix inequalities*, SIAM Journal on Optimization, 10 (2000), pp. 1228–1248.

[32] S.-i. Tanigawa, *Singularity degree of the positive semidefinite matrix completion problem*, SIAM Journal on Optimization, 27 (2017), pp. 986–1009.

[33] L. Tunçel, *On the Slater condition for the SDP relaxations of nonconvex sets*, Operations Research Letters, 29 (2001), pp. 181–186.

[34] A. L. Wang, *On Quadratically Constrained Quadratic Programs and their Semidefinite Program Relaxations*, PhD thesis, Carnegie Mellon University, 2022.

[35] H. Wolkowicz, *Semidefinite and lagrangian relaxations for hard combinatorial problems*, in System Modelling and Optimization: Methods, Theory and Applications. 19 th IFIP TC7 Conference on System Modelling and Optimization July 12–16, 1999, Cambridge, UK 19, Springer, 2000, pp. 269–309.

[36] H. Wolkowicz, R. Saigal, and L. Vandenberghe, *Handbook of semidefinite programming: theory, algorithms, and applications*, vol. 27, Springer Science & Business Media, 2012.

[37] H. Wolkowicz and Q. Zhao, *Semidefinite programming relaxations for the graph partitioning problem*, Discrete Applied Mathematics, 96 (1999), pp. 461–479.

[38] Q. Zhao, S. E. Karisch, F. Rendl, and H. Wolkowicz, *Semidefinite programming relaxations for the quadratic assignment problem*, Journal of Combinatorial Optimization, 2 (1998), pp. 71–109.

[39] Y. Zheng, G. Fantuzzi, A. Papachristodoulou, P. Goulart, and A. Wynn, *Fast ADMM for semidefinite programs with chordal sparsity*, in 2017 American Control Conference (ACC), IEEE, 2017, pp. 3335–3340.

[40] Y. Zhu, G. Pataki, and Q. Tran-Dinh, *Sieve-SDP: a simple facial reduction algorithm to preprocess semidefinite programs*, Mathematical Programming Computation, 11 (2019), pp. 503–586.