Intersection Cuts for Polynomial Optimization

Daniel Bienstock[1], Chen Chen[2], and Gonzalo Muñoz[3(✉)]

[1] IEOR, Columbia University, New York, USA
dano@columbia.edu
[2] ISE, The Ohio State University, Columbus, USA
chen.8018@osu.edu
[3] IVADO Fellow, Polytechnique Montréal, Montreal, Canada
gonzalo.munoz@polymtl.ca

Abstract. We consider dynamically generating linear constraints (cutting planes) to tighten relaxations for polynomial optimization problems. Many optimization problems have feasible set of the form $S \cap P$, where S is a closed set and P is a polyhedron. Integer programs are in this class and one can construct intersection cuts using convex "forbidden" regions, or S-free sets. Here, we observe that polynomial optimization problems can also be represented as a problem with linear objective function over such a feasible set, where S is the set of real, symmetric matrices representable as outer-products of the form xx^T. Accordingly, we study outer-product-free sets and develop a thorough characterization of several (inclusion-wise) maximal intersection cut families. In addition, we present a cutting plane approach that guarantees polynomial-time separation of an extreme point in $P \setminus S$ using our outer-product-free sets. Computational experiments demonstrate the promise of our approach from the point of view of strength and speed.

1 Introduction

In this work we focus on polynomial optimization:

$$\min \ p_0(x)$$
$$(\textbf{PO}) \ \text{ s.t. } p_i(x) \leq 0 \quad i = 1, ..., m,$$

where each p_i is a polynomial function with respect to $x \in \mathbb{R}^n$. We consider the dynamic generation of linear valid inequalities, i.e. cutting planes, to tighten relaxations of **PO**. Cuts for polynomial optimization are typically generated for a single nonlinear term or function (e.g. [7,35,37,41,45,51–53]) over a simple subset of linear constraints such as box constraints. In contrast, we develop general-purpose cuts that have the potential to involve all variables simultaneously. To the best of our knowledge there are two papers (applicable to polynomial optimization) that are similar to our work in this regard. The disjunctive cuts of Saxena, Bonami, and Lee [46,47], and the work of Ghaddar, Vera, and

Anjos [26] who propose a lift-and-project method using moment relaxations. Polynomial-time separation from these procedures is not guaranteed in general.

We adopt the geometric perspective for generating cuts, in which cuts for a region of the form $S \cap P$, with P a polyhedron and S a closed set, are derived from convex forbidden zones, or S-free sets. The S-free approach developed in the context of mixed-integer programming, with S typically considered to be the integer lattice. Practical applications of this technique have so far focused on natural extensions such as conic integer programming (e.g. [4,30,42]) and bilevel mixed-integer linear programming [23]. In contrast, **PO** represents an essentially different domain of application since variables here are continuous.

We work with a representation of **PO** that uses a symmetric matrix of decision variables, and yields an *equivalent* formulation with a linear objective function and a feasible region of the form $S \cap P$, with S the (closed) set of symmetric matrices that can be represented as outer products of the form xx^T— accordingly, we study *outer-product-free* sets. Several families of full-dimensional (inclusion-wise) maximal outer-product-free sets are identified in Theorems 3 and 4 of Sect. 5. Furthermore, we derive an oracle-based outer-product-free set in Sect. 4. With the aforementioned results we develop a cut generation procedure (see Sect. 6) that has (to our knowledge) the following unique properties: any infeasible extreme point of a (lifted) polyhedral relaxation of **PO** can be separated in polynomial time; and variable bounds are not required. In Sect. 7 we demonstrate the practical effectiveness of our approach over a variety of instances using a straightforward pure cutting-plane setup. The speed of our separation routines and the quality of the resulting cuts strongly suggest the viability of our cut families within a full-fledged branch-and-cut solver.

1.1   Notation

Denote the interior of a set int($\cdot$) and its boundary bd($\cdot$). The convex hull of a set is denoted conv($\cdot$), and its closure is clconv($\cdot$); likewise, the conic hull of a set is cone($\cdot$) and its closure clcone($\cdot$). For a point x and nonempty set S in $\mathbb{R}^n$, we define $d(x,S) := \inf_{s \in S}\{\|x - s\|_2\}$; note that for S closed we can replace the infimum with minimum. Denote the ball with center x and radius r to be $\mathcal{B}(x,r)$. $\langle \cdot, \cdot \rangle$ denotes the matrix inner product and $\| \cdot \|_F$ the Frobenius norm. A positive semidefinite matrix may be referred to as a PSD matrix for short, and likewise NSD refers to negative semidefinite.

2   *S*-free Sets and the Intersection Cut

Definition 1. *A set $C \subset \mathbb{R}^n$ is S-free if $int(C) \cap S = \emptyset$ and C is convex.*

For any S-free set C we have $S \cap P \subseteq \text{clconv}(P \setminus \text{int}(C))$, and so any valid inequalities for clconv($P \setminus \text{int}(C)$) are valid for $S \cap P$. Hillestad and Jacobsen [29], and later on Sen and Sherali [48], provide results regarding polyhedrality of clconv($P \setminus \text{int}(C)$). Averkov [5] provides theoretical consideration on how

one can derive cuts from C. In specific instances, $\mathrm{conv}(P \setminus \mathrm{int}(C))$ can be fully described (see [10,11,30,42]), however, separating over $P \setminus \mathrm{int}(C)$ is NP-hard [25]. A standard workaround is to find a simplicial cone P' containing P and apply Balas' intersection cut [6] for $P' \setminus \mathrm{int}(C)$ (also see Tuy [54]). Larger S-free sets can be useful for generating deeper cuts [17].

Definition 2. *An S-free set C is* maximal *if $V \not\supset C$ for all S-free V.*

Under certain conditions (see [9,17,18,30]), maximal S-free sets are sufficient to generate all nontrivial cuts for a problem. When $S = \mathbb{Z}^n$, C is called a lattice-free set. Maximal lattice-free sets are well-studied in integer programming theory [1,2,8,14,20,27,30,36], and the notion of S-free sets was introduced as a generalization [21].

2.1   The Intersection Cut

Let $P' \supseteq P$ be a simplicial conic relaxation of P: a displaced polyhedral cone with apex $\bar{x}$ and defined by the intersection of n linearly independent halfspaces. P' may be written as follows:

$$P' = \{\bar{x} + \sum_{j=1}^{n} \lambda_j r^j : \lambda \geq 0\}. \tag{1}$$

Each extreme ray of P' is of the form $\{\bar{x} + \lambda_j r^j | \lambda_j \geq 0\}$. Alternatively, the simplicial conic relaxation can be given in inequality form

$$P' = \{x | Ax \leq b\}, \tag{2}$$

where A is an invertible matrix. Note that any basis of P would be suitable to derive P'. The apex $\bar{x} = A^{-1}b$, and the rays r^j in (1) can be obtained directly from A: for each j, one can identify $-r^j$ as the jth column of the inverse of A.

We shall assume $\bar{x} \notin S$, so that $\bar{x}$ is to be separated from S via separation from $P' \setminus \mathrm{int}(C)$, with C an S-free set with $\bar{x}$ in its interior. Since $\bar{x} \in \mathrm{int}(C)$, there must exist $\lambda > 0$ such that $\bar{x} + \lambda_j r^j \in \mathrm{int}(C) \; \forall j$. Also, each extreme ray is either entirely contained in C, i.e. $\bar{x} + \lambda_j r^j \in \mathrm{int}(C) \, \forall \lambda_j \geq 0$, or else there is an intersection point with the boundary: $\exists \lambda_j^* : \bar{x} + \lambda_j^* r^j \in \mathrm{bd}(C)$. We refer to λ_j^* as the *step length* in the latter case, and for convenience, we define the step length $\lambda_j^* = \infty$ in the former case. The *intersection cut* is the halfspace whose boundary contains each intersection point (given by $\lambda_j^* < \infty$) and that is parallel to all extreme rays contained in C.

Given $\lambda_j^* \in (0, \infty] \, \forall j = 1, \ldots, n$, Balas [6, Theorem 2] provides a closed-form expression for the intersection cut $\pi x \leq \pi_0$:

$$\pi_0 = \sum_{i=1}^{n} (1/\lambda_i^*) b_i - 1, \;\; \pi_j = \sum_{i=1}^{n} (1/\lambda_i^*) a_{ij}, \tag{3}$$

where a_{ij} are the entries of A in (2) and $1/\infty := 0$ [6, p. 34].

The validity of the intersection cut and a condition in which the cut gives the convex hull of $P' \setminus \text{int}(C)$ is established by Balas [6, Theorem 1]. A more detailed analysis, as well as a strengthening procedure for infinite step lengths is provided in our full-length paper [12].

3   Moment-Based Reformulation of Polynomial Optimization

Our approach to **PO** leverages the moment/sum-of-squares approach (see [33, 34]) from which a definition of the feasible set as $S \cap P$ is naturally obtained.

Let $m_r = [1, x_1, \ldots, x_n, x_1 x_2, \ldots, x_n^2, \ldots, x_n^r]$ be a vector of all monomials up to degree r. Any polynomial may be written as $p_i(x) = m_r^T A_i m_r$ (provided r is sufficiently large), where A_i is a symmetric matrix derived from p_i. We can apply this transformation to **PO** to obtain a lifted representation **LPO**:

$$\min \langle A_0, X \rangle$$
$$\textbf{(LPO)} \text{ s.t. } \langle A_i, X \rangle \leq b_i, \ i = 1, ..., m, \tag{4a}$$
$$X = m_r m_r^T. \tag{4b}$$

Denote $n_r := \binom{n+r}{r}$, i.e. the length of m_r. Here $A_i \in \mathbb{S}^{n_r \times n_r}$ are symmetric matrices of data, and $X \in \mathbb{S}^{n_r \times n_r}$ is a symmetric matrix of decision variables. The problem has linear objective function, linear constraints (4a), and nonlinear constraints (4b). One can replace the moment matrix condition $X = m_r m_r^T$ with the equivalent conditions of $X \succeq 0, \text{rank}(X) \leq 1$ and linear consistency constraints among entries from X representing the same monomial. Dropping the nonconvex rank one constraint yields the standard semidefinite relaxation [50].

On the other hand, the feasible region of **LPO** has a natural description as an intersection of a polyhedron P_{OP}, that corresponds to linear constraints (4a) together with consistency constraints, and the following closed set,

$$S_{OP} := \{X \in \mathbb{S}^{n_r \times n_r} : X = xx^T, x \in \mathbb{R}^{n_r}\}.$$

Accordingly, we shall study sets that are *outer-product-free* (OPF): closed, convex sets in $\mathbb{S}^{n_r \times n_r}$ with interiors that do not intersect with S_{OP}. In what follows, suppose we have an extreme point $\bar{X} \in P_{OP} \setminus S_{OP}$ with spectral decomposition $\bar{X} := \sum_{i=1}^{n_r} \lambda_i d_i d_i^T$ and ordering $\lambda_1 \geq \ldots \geq \lambda_{n_r}$. We seek to separate $\bar{X}$.

4   Oracle-Based Outer-Product-Free Sets

If one has access to a distance oracle to the set S, one can easily construct an OPF set, namely, an OPF *ball*. In the case of **LPO** this corresponds to the distance to the nearest symmetric outer product. This distance is a special case of the following PSD matrix approximation problem, given an integer $q > 0$:

$$\textbf{(PMA)} \quad \min_Y \left\{ \|\bar{X} - Y\| : \text{rank}(Y) \leq q, \ Y \succeq 0 \right\}.$$

Here $\| \cdot \|$ is a unitarily invariant matrix norm such as the Frobenius norm, $\| \cdot \|_F$. Dax [19] proves the following:

Theorem 1 (Dax's Theorem). *Let k be the number of nonnegative eigenvalues of $\bar{X}$. For $q = 1, \ldots, n - 1$, an optimal solution to $\boldsymbol{PMA}$ is given by $Y = \sum_{i=1}^{\min\{k,q\}} \lambda_i d_i d_i^T$.*

When $\bar{X}$ is not NSD, the solution from Dax's theorem coincides with Eckart-Young-Mirsky [22,40] solution to **PMA** without the PSD constraint. The optimal PSD approximant allows us to construct an OPF ball:

$$\mathcal{B}_{\text{oracle}}(\bar{X}) := \begin{cases} \mathcal{B}(\bar{X}, \|\bar{X}\|_F), & \text{if } \bar{X} \text{ is NSD,} \\ \mathcal{B}(\bar{X}, \|\sum_{i=2}^n \lambda_i d_i d_i^T\|_F), & \text{otherwise.} \end{cases}$$

Corollary 1. $\mathcal{B}_{oracle}(\bar{X})$ *is OPF.*

Proof. Setting $q = 1$ in Dax's Theorem, we see that the nearest symmetric outer product is either $\lambda_1 d_1 d_1^T$ if $\lambda_1 > 0$, or else the zeros matrix. □

For **LPO** we can use a simple geometric construction involving Theorem 2 to obtain an OPF *cone* from the oracle ball. This extension is detailed in [12].

5   Maximal Outer-Product-Free Sets

5.1   General Properties of Maximal Outer-Product-Free Sets

We now turn to characterizing and finding maximal OPF sets. Our first Theorem is a building block towards maximality.

Theorem 2. *Let $C \subset \mathbb{S}^{n_r \times n_r}$ be a full-dimensional OPF set. Then clcone(C) is OPF. In particular, every full-dimensional maximal OPF set is a convex cone.*

Proof. Suppose clcone(C) is not OPF; since it is closed and convex, then by definition of OPF sets there must exist $d \in \mathbb{R}^{n_r}$ such that dd^T is in its interior. If d is the zeros vector, then int(C) also contains the origin, which contradicts the condition of C being OPF. Otherwise the ray r^0 emanating from the origin with nonzero direction dd^T is entirely contained in and hence is an interior ray of clcone(C). By convexity, the interior of cone(C) is the same as the interior of its closure, so r^0 is also an interior ray of cone(C). From this, it can be proved that r^0 must pass through the interior of C (see [12]). But every point along r^0 is a symmetric outer-product, which again implies that C is not OPF. □

We can also obtain the following properties regarding the geometry of maximal OPF sets via their supporting halfspaces.

Definition 3. *A* supporting halfspace *of a closed, convex set S contains S and its boundary is a supporting hyperplane of S.*

Lemma 1. *Let C be a full-dimensional maximal OPF set. Every supporting halfspace of C is of the form $\langle A, X \rangle \geq 0$ for some $A \in \mathbb{S}^{n_r \times n_r}$.*

Proof. From Theorem 2 we have that C is a convex cone. From this it follows that a supporting halfspace $\langle A, X \rangle \geq b$ must have $b = 0$. □

From Lemma 1 we may characterize a maximal OPF set as $C = \{X \in \mathbb{S}^{n_r \times n_r} | \langle A_i, X \rangle \geq 0 \ \forall i \in I\}$, with I a potentially infinite index set.

Theorem 3. *The halfspace $\langle A, X \rangle \geq 0$ is maximal OPF iff A is NSD.*

Proof. If A has a strictly positive eigenvalue, then $\langle A, dd^T \rangle > 0$ for some d, and so the halfspace is not OPF. If A is NSD then $\langle A, dd^T \rangle = d^T A d \leq 0 \ \forall d \in \mathbb{R}^{n_r}$, so the halfspace is OPF. For maximality, suppose the halfspace is strictly contained in another OPF set $\bar{C}$. Then there exists some $\bar{X} \in \text{int}(\bar{C})$ such that $\langle A, \bar{X} \rangle < 0$. Thus, $-\bar{X} \in \text{int}(\bar{C})$ and so is the zeros matrix. Thus $\bar{C}$ cannot be OPF. □

5.2 Maximal Outer-Product-Free Sets Derived from 2×2 Submatrices

Theorem 3 provides our first explicit family of maximal OPF sets. Another family is suggested by the following result by Kocuk, Dey, and Sun [31]:

Proposition 1 (KDS Proposition). *A nonzero, Hermitian matrix X is PSD and has rank one iff all the 2×2 minors of X are zero and the diagonal elements of X are nonnegative.*

In what follows, denote the entries of a 2×2 submatrix of X from some rows $i_1 < i_2$ and columns $j_1 < j_2$ as $X_{[[i_1, i_2], [j_1, j_2]]} := \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

Lemma 2. *Let $\lambda \in \mathbb{R}^2$ with $\|\lambda\|_2 = 1$. (5a) and (5b) describe an OPF set:*

$$\lambda_1(a + d)/2 + \lambda_2(b - c)/2 \geq \|(b + c)/2, (a - d)/2\|_2, \tag{5a}$$
$$\lambda_1(b + c)/2 + \lambda_2(a - d)/2 \geq \|(a + d)/2, (b - c)/2\|_2. \tag{5b}$$

Proof. The set defined by $ad \geq bc$ is OPF (Proposition 1). The proof follows from checking that (5a) defines a subset of it. Similarly, (5b) defines a subset of $ad \leq bc$. □

The following Theorem provides an extensive list of *maximal* OPF sets that can be obtained from Lemma 2. We leave the proof in the Appendix.

Theorem 4. *(5a) describes a* maximal *OPF set if*

(i) $\lambda_1 = 1, \lambda_2 = 0$, *and neither b nor c are diagonal entries;*
(ii) $\lambda_1 = 0, \lambda_2 = 1$, *and b is a diagonal entry;*
(iii) $\lambda_1 = 0, \lambda_2 = -1$, *and c is a diagonal entry;*
(iv) $\lambda_1^2 + \lambda_2^2 = 1$, *and none of a, b, c, d are diagonal entries.*

Similarly, (5b) describes a maximal *OPF set if*

(v) $\lambda_1 = 1, \lambda_2 = 0$, *and either b or c is a diagonal entry;*

(vi) $\lambda_1 = 0, \lambda_2 = 1$, *and a but not d is a diagonal entry;*

(vii) $\lambda_1 = 0, \lambda_2 = -1$, *and d but not a is a diagonal entry;*

(viii) $\lambda_1^2 + \lambda_2^2 = 1$, *and none of a, b, c, d are diagonal entries.*

The following theorem shows that the maximal OPF sets we have identified characterizes all such sets in the special case where $n_r = 2$.

Theorem 5. *In $\mathbb{S}^{2\times 2}$ every full-dimensional maximal OPF set is either the cone of PSD matrices or a halfspace of the form $\langle A, X \rangle \geq 0$, where A is NSD.*

Proof. See [12] for details.

6   Implementation of Intersection Cuts

Suppose that we have a simplicial conic relaxation of P_{OP} with apex $\bar{X}$. This section provides a brief overview on how to generate a cutting plane to separate $\bar{X}$ from $P_{OP} \cap S_{OP}$ using results from Sects. 4 and 5.

6.1   Step 1: Selecting an Outer-Product-Free Set

Separation Using the Distance Oracle. As outlined in Sect. 4, $\mathcal{B}_{\text{oracle}}(\bar{X})$, or its conic extension/strengthening can always be used to separate $\bar{X}$.

Separation Using Halfspaces. Theorem 3 shows that certain halfspaces are OPF sets. Moreover, it is not hard to see that the halfspaces of Theorem 3 imply and provide no more than the family of cuts equivalent to the PSD condition:

$$d^T X d \geq 0 \quad \forall d \in \mathbb{R}^n \iff X \succeq 0.$$

Choosing d equal to the eigenvectors of $\bar{X}$ provides polynomial-time separation (given fixed numerical tolerances); this is a well-studied linear outer-approximation procedure for semidefinite programming problems [32, 44, 47, 49], and here we provide a new interpretation of them via the maximal OPF property.

Separation with all 2×2 Submatrices of $\bar{X}$. From Proposition 1 we have that $\bar{X} \notin S_{OP}$ implies a nonzero 2×2 minor or a negative diagonal term. Supposing the nonnegative diagonal constraints are included in P_{OP}, then at least one of the $\mathcal{O}(n^4)$ 2×2 minors will be nonzero. We can show that for any such minor that is nonzero at least one of the sets described in Theorem 4 will strictly contain $\bar{X}$. There is an additional choice of the λ parameters for sets of the form (iv) and (viii), which in our experiments we set to extreme values $\lambda_1 = 1, \lambda_2 = 0$ or $\lambda_1 = 0, \lambda_2 = 1$. Intermediate values for λ are the subject of ongoing research.

Separation with Only Principal 2×2 Submatrices of $\bar{X}$. An alternative characterization to Proposition 1 is given by Chen, Atamtürk and Oren [16]:

Proposition 2 (CAO Proposition). *For $n > 1$ a nonzero Hermitian PSD $n \times n$ matrix X has rank one iff all of its 2×2 principal minors are zero.*

Hence if $\bar{X}$ is not PSD then it is contained in at least one halfspace described in Theorem 3. Otherwise, $\bar{X}$ has at least one of its $\mathcal{O}(n^2)$ principal 2×2 minors strictly positive, and so (5a) is strictly satisfied for case (i) of Theorem 4.

6.2   Step 2: Generating an Intersection Cut

The halfspace OPF sets can generate a cut directly as mentioned above. We only need the eigenvectors of $\bar{X}$. The remaining OPF sets are a ball, for which the step-lengths for the intersection cuts are simply the ball's radius, and second-order cones. We can thus derive (see [12]) computationally trivial closed-form expressions for step lengths from the interior of a second-order cone to its boundary. This is one of the most crucial features of our proposed cutting planes, as they can be generated with little computational effort and thus making them suitable for their incorporation in a branch-and-cut procedure.

7   Numerical Experiments

We present experiments using a pure cutting-plane algorithm using the cuts described in Sect. 6. The experiments are designed to investigate the stand-alone performance of our cuts, particularly separation speed and effectiveness. The cutting plane algorithm solves an LP relaxation and obtains an (extreme point) optimal solution $\bar{X}$, adds cuts separating $\bar{X}$, and repeats until either:

– A time limit of 600 seconds is reached, or
– The objective value does not improve for 10 iterations, or
– The violation of all cuts is not more than 10^{-6}. Here, if $\pi^T x \leq \pi_0$ is the cut and x^* is the candidate solution, we define the violation as $(\pi^T x^* - \pi_0)/\|\pi\|_1$.

For improving stability, we add a maximum of 10 cuts per iteration (selected using violations) and remove non-active cuts every 15 iterations. Computations are run on a 32-core server with an Intel Xeon Gold 6142 2.60 GHz CPU and 512 GB of RAM. Although the machine is powerful, we run the algorithm single-threaded and the experiments do not require a significant amount of memory; we confirmed that similar performance can be obtained with a laptop. The code is written in C++ using the Eigen library for linear algebra [28]. The LP solver is Gurobi 8.0.0 and, for comparisons, we solve SDP relaxations using the C++ Fusion API of Mosek 8 [43]. Our code is available at https://github.com/g-munoz/poly_cuts_cpp.

Test instances are taken from two sources. First, we consider all 27 problem instances from Floudas et al. [24] (available via GLOBALLib [39]) that have quadratic objective and constraints. Our cuts can accommodate arbitrary polynomial terms, however for implementation purposes reading QCQP problems is

more convenient. Second, we consider all 99 instances of BoxQP developed by several authors [15,55]. These problems have simple box constraints $x \in [0, 1]^n$ and a nonconvex quadratic objective function. In a recent paper by Bonami et al. [13], the authors show how to obtain cutting planes for this particular case.

We choose the initial LP relaxation to be the standard RLT relaxation of QCQP: setting $r = 1$ in **LPO** and including McCormick estimators for bilinear terms (see [3,38]). Problem sizes vary from 21×21 to 126×126 symmetric matrices of decision variables for BoxQP instances and from 6×6 to 63×63 for GLOBALLib instances. To obtain variable bounds for some of the GLOBALLib instances we apply a simple bound tightening procedure: minimize/maximize a given variable subject to the RLT relaxation. Lastly, we use *Gap Closed* as a measure of quality of the bounds generated by each approach. This is defined as follows: let OPT denote the optimal value of an instance, RLT the optimal value of the standard RLT relaxation, and GLB the objective value obtained after applying the cutting plane procedure. Then Gap Closed $= \frac{GLB-RLT}{OPT-RLT}$.

Results. In Table 1, we show a performance comparison in the selected GLOB-ALLib instances between our cutting plane algorithm versus the relaxation obtained from adding a PSD requirement for the variable X in the RLT relaxation (SDP). We do not show results for 2 of the instances, as the RLT relaxation is tight for these. The results in Table 1 are very encouraging: in only 4 instances we are not able to reach the same gap closed as the SDP. Moreover, our simple cutting plane approach (almost) always runs in just a few seconds.

In Table 2, we compare our results with the V2 setting used by Saxena, Bonami and Lee [46] in the selected GLOBALLib instances. We chose [46] as a comparison as we find it the most similar to our approach. V2 uses an RLT relaxation for QCQPs and applies two types of cuts: an outer-approximation of the PSD cone and disjunctive cuts for which the separation involves a MIP. We emphasize that these families of cuts are complementary and not competitive. It is also important to mention that the running times in Table 2 for V2 correspond to the reports in [46], published in 2010. While new hardware may improve these times, we believe the conclusions we draw from Table 2 would not change.

For comparison purposes, we turned off our simple bound tightening routine in order to obtain the same initial relaxation value as V2 (and thus the gaps are different than the ones in Table 1). Even doing so, for certain instances of GLOBALLib we did not obtain the same initial bound and thus excluded these from comparison. On the comparable GLOBALLib instances our algorithm terminates with smaller gap closed on average, but it does produce higher gap closed on some instances. The advantage of our cuts is that times are substantially shorter. This is expected, as V2 solves a MIP in the cut generation, while our cuts only require finding eigenvalues and roots of single-variable quadratics. Overall we believe these results are promising, as the cutting planes are able to close a significant amount of gap in many cases, in a very short time.

The results on the BoxQP instances are interesting as well. In the interest of space, we limit ourselves to summarizing them here. The interested reader can find a complete log in https://goo.gl/8wPeY6. We compare with V2 as before,

Table 1. Comparison of intersection cuts versus SDP relaxation in non-convex quadratic GLOBALLib instances.

Instance	SDP		Intersection Cuts		Instance	SDP		Intersection Cuts	
Name	Gap Closed	Time	Gap Closed	Time	Name	Gap Closed	Time	Gap Closed	Time
Ex2_1_1	0.00%	0.01	57.70%	0.02	Ex5_2_5	0.00%	1.84	0.00%	7.83
Ex2_1_5	0.00%	0.03	99.57%	0.01	Ex5_3_2	0.10%	0.74	0.00%	0.94
Ex2_1_6	0.00%	0.02	79.56%	0.09	Ex5_3_3	3.75%	105.01	0.50%	8.12
Ex2_1_7	0.00%	0.46	22.59%	0.96	Ex5_4_2	0.00%	0.02	0.40%	0.08
Ex2_1_8	0.00%	1.11	51.89%	1.91	Ex8_4_1	98.43%	0.84	59.42%	23.03
Ex2_1_9	0.00%	0.02	31.92%	0.98	Ex9_1_4	0.00%	0.06	97.65%	0.08
Ex3_1_1	0.00%	0.05	0.71%	0.32	Ex9_2_1	6.25%	0.07	26.55%	4.45
Ex3_1_2	22.41%	0.01	100.00%	0.00	Ex9_2_2	16.67%	0.04	62.14%	1.55
Ex3_1_4	0.00%	0.01	32.61%	0.02	Ex9_2_3	0.00%	0.24	0.00%	0.30
Ex5_2_2_case1	0.00%	0.04	12.84%	0.79	Ex9_2_4	99.83%	0.03	33.33%	0.10
Ex5_2_2_case2	0.00%	0.04	30.25%	0.56	Ex9_2_6	99.76%	0.27	56.64%	0.20
Ex5_2_2_case3	0.00%	0.04	19.15%	0.50	Ex9_2_7	6.25%	0.05	26.55%	4.58
Ex5_2_4	0.00%	0.02	27.55%	0.20					

Table 2. Comparison of Intersection Cuts versus V2 approach of [46] in Non-Convex Quadratic GLOBALLib Instances. Entries labelled NR were not reported in [46].

Instance	V2		Intersection Cuts		Instance	V2		Intersection Cuts	
Name	Gap Closed	Time	Gap Closed	Time	Name	Gap Closed	Time	Gap Closed	Time
Ex2_1_1	72.62%	704.40	57.70%	0.02	Ex5_2_4	79.31%	68.93	27.14%	0.20
Ex2_1_5	99.98%	0.17	99.68%	0.00	Ex5_2_5	6.27%	3793.17	0.00%	7.59
Ex2_1_6	99.95%	3397.65	88.82%	0.09	Ex5_3_2	7.27%	245.82	0.00%	0.91
Ex2_1_8	84.70%	3632.28	3.08%	1.71	Ex5_3_3	0.21%	3693.76	0.50%	16.73
Ex2_1_9	98.79%	1587.94	32.01%	0.89	Ex5_4_2	27.57%	3614.38	0.56%	0.10
Ex3_1_1	15.94%	3600.27	0.89%	0.11	Ex9_1_4	0.00%	0.60	0.00%	0.06
Ex3_1_2	99.99%	0.08	100.00%	0.00	Ex9_2_1	60.04%	2372.64	27.27%	0.10
Ex3_1_4	86.31%	21.26	32.61%	0.02	Ex9_2_2	88.29%	3606.36	69.88%	8.30
Ex5_2_2_case1	0.00%	0.02	0.00%	0.02	Ex9_2_6	87.93%	2619.02	99.70%	1.76
Ex5_2_2_case2	0.00%	0.05	0.00%	0.08	Ex9_2_8	NR	NR	99.68%	0.03
Ex5_2_2_case3	0.36%	0.36	18.89%	0.11					

and we replicate the same initial relaxation used by Saxena, Bonami and Lee [46], namely the weak RLT relaxation (wRLT)[1]. We also compare against the wRLT relaxation with a PSD constraint (wRLT+SDP). On the 42 BoxQP instances reported in [46], our cuts *always* perform better than both V2 and wRLT+SDP. The latter reaches optimality in seconds, but the relaxation is not strong, as there are missing McCormick inequalities. Intersection Cuts, with a time limit of 600 seconds, is able to close 90.49% gap on average in these instances, while V2 closes 65.28% and wRLT+SDP 51.87%. Even though wRLT is a relaxation that is not used in practice, these experiments still evidence the potential of our proposed cuts, as they close a large amount of gap in a short amount of time, even surpassing the impact of including an explicit SDP constraint.

---

[1] This BoxQP relaxation only adds the "diagonal" McCormick estimates $X_{ii} \leq x_i$.

8   Conclusions

We have introduced intersection cuts in the context of polynomial optimization. Accordingly, we have developed an S-free approach for polynomial optimization, where S is the set of real, symmetric outer products. Our results on full-dimensional maximal OPF sets include a full characterization of such sets when $n_r = 2$ as well as extensive families of maximal OPF sets. Computational experiments have demonstrated the potential of our cuts as a fast way to reduce optimality gaps on a variety of problems using computationally simple closed-form procedures. A full implementation is being considered for future empirical work, incorporating the cuts into a branch-and-cut solver and developing a more sophisticated implementation, e.g. stronger initial relaxations with problem-specific valid inequalities, warm-starting the outer-approximation with an SDP, sparsification of the cuts, advanced cut management, improved scalability, among others.

Appendix

Proof (Theorem 4). The OPF property is given by Lemma 2, so maximality remains. Let C be a set described by (5a) or (5b). It suffices to construct, for *every* symmetric matrix $\bar{X} \notin C$, $Z := zz^T$ such that $Z - \bar{X} \in \text{int}(C)$. This implies $Z \in \text{int}(\text{conv}(C \cup \bar{X}))$. Denote the submatrices of $\bar{X}, Z$:

$$\bar{X}_{[[i_1,i_2],[j_1,j_2]]} := \begin{bmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{bmatrix}, \ Z_{[[i_1,i_2],[j_1,j_2]]} := \begin{bmatrix} a_Z & b_Z \\ c_Z & d_Z \end{bmatrix}.$$

Furthermore, for convenience let us define the following:

$$\bar{p} := (\bar{a} + \bar{d})/2, \ \bar{q} := (\bar{a} - \bar{d})/2, \ \bar{r} := (\bar{b} + \bar{c})/2, \ \bar{s} := (\bar{b} - \bar{c})/2.$$

Construction for (5a): Suppose $\bar{X}$ violates (5a). We propose the following:

$$a_Z = \bar{q} + \lambda_1\|\bar{q},\bar{r}\|_2, \ b_Z = \bar{r} + \lambda_2\|\bar{q},\bar{r}\|_2, c_Z = \bar{r} - \lambda_2\|\bar{q},\bar{r}\|_2, \ d_Z = -\bar{q} + \lambda_1\|\bar{q},\bar{r}\|_2 \tag{6}$$

$$\implies \lambda_1(\bar{a} + \bar{d})/2 + \lambda_2(\bar{b} - \bar{c})/2 \ < \ \|(\bar{b} + \bar{c})/2, (\bar{a} - \bar{d})/2\|_2$$
$$= \lambda_1(a_Z + d_Z)/2 + \lambda_2(b_Z - c_Z)/2$$

where the last equality follows from $\lambda_1^2 + \lambda_2^2 = 1$. This implies

$$\lambda_1((a_Z - \bar{a}) + (d_Z - \bar{d}))/2 + \lambda_2((b_Z - \bar{b}) - (c_Z - \bar{c}))/2 > 0$$

and since $\|((b_Z - \bar{b}) + (c_Z - \bar{c}))/2, ((a_Z - \bar{a}) - (d_Z - \bar{d}))/2\|_2 = 0$, we conclude $Z - \bar{X} \in \text{int}(C)$.

Construction for (5b): If $\bar{X}$ violates (5b), we use the following construction:

$$a_Z = \bar{p} + \lambda_2\|\bar{p}, \bar{s}\|_2, b_Z = \bar{s} + \lambda_1\|\bar{p}, \bar{s}\|_2, c_Z = -\bar{s} + \lambda_1\|\bar{p}, \bar{s}\|_2, d_Z = \bar{p} - \lambda_2\|\bar{p}, \bar{s}\|_2.$$
$$\implies \lambda_1(\bar{b} + \bar{c})/2 + \lambda_2(\bar{a} - \bar{d})/2 \; < \; \|(\bar{a} + \bar{d})/2, (\bar{b} - \bar{c})/2\|_2$$
$$= \lambda_1(b_Z + c_Z)/2 + \lambda_2(a_Z - d_Z)/2,$$
$$\implies \lambda_1((b_Z - \bar{b}) + (c_Z - \bar{c}))/2 + \lambda_2((a_Z - \bar{a}) - (d_Z - \bar{d}))/2 > 0.$$

We conclude $Z - \bar{X} \in \text{int}(C)$ as before, since $\|((a_Z - \bar{a}) + (d_Z - \bar{d}))/2, ((b_Z - \bar{b}) - (c_Z - \bar{c}))/2\|_2 = 0$. It remains to set the other entries of Z and to show it is an outer product.

Claim. For each condition (i)–(viii), $a_Z d_Z = b_Z c_Z$ and all diagonal elements among a_Z, b_Z, c_Z, d_Z are nonnegative.

<u>Proof:</u> First consider conditions (i)–(iv). By construction of (6):

$$a_Z d_Z \; = \; -\bar{q}^2 + \lambda_1^2\|\bar{q}, \bar{r}\|_2^2 \; = \; \bar{r}^2 - \lambda_2^2\|\bar{q}, \bar{r}\|_2^2 \; = \; b_Z c_Z.$$

The second equality is derived from the following identity:

$$\|\bar{q}, \bar{r}\|_2^2 = \bar{q}^2 + \bar{r}^2 \iff -\bar{q}^2 + \lambda_1^2\|\bar{q}, \bar{r}\|_2^2 = \bar{r}^2 - \lambda_2^2\|\bar{q}, \bar{r}\|_2^2.$$

Nonnegativity of diagonal elements follows from $\|\bar{q}, \bar{r}\|_2 \geq \max\{|\bar{q}|, |\bar{r}|\}$. In case (i) only a_Z or d_Z can be diagonal elements, and they are both nonnegative. The other cases can be directly verified. Similarly, for conditions (v)–(viii):

$$a_Z d_Z = \bar{p}^2 - \lambda_2^2\|\bar{p}, \bar{s}\|_2^2 \; = \; -\bar{s}^2 + \lambda_1^2\|\bar{p}, \bar{s}\|_2^2 \; = \; b_Z c_Z.$$

The second equality is derived from the following identity:

$$\|\bar{p}, \bar{s}\|_2^2 = \bar{p}^2 + \bar{s}^2 \iff -\bar{s}^2 + \lambda_1^2\|\bar{p}, \bar{s}\|_2^2 = \bar{p}^2 - \lambda_2^2\|\bar{p}, \bar{s}\|_2^2.$$

Nonnegativity of diagonal elements follows from the same argument as before, by using the fact that $\|\bar{p}, \bar{s}\|_2 \geq \max\{|\bar{p}|, |\bar{s}|\}$. ∎

To maintain symmetry we set $Z_{i_1, j_1} = Z_{j_1, i_1}, Z_{i_1, j_2} = Z_{j_2, i_1}, Z_{i_2, j_1} = Z_{j_1, i_2}, Z_{i_2, j_2} = Z_{j_2, i_2}$. Now denote $\ell = [i_1, i_2, j_1, j_2]$. If $a_Z = b_Z = c_Z = d_Z = 0$, then we simply set all other entries of Z equal to zero and so Z is the outer-product of the vector of zeroes. Otherwise, consider the following cases.

Case 1: ℓ has 4 unique entries. Suppose w.l.o.g we have an upper-triangular entry ($i_1 < i_2 < j_1 < j_2$) and furthermore suppose that b_Z is nonzero. Then set

$$Z_\ell := \begin{bmatrix} 1 & d_Z/b_Z & a_Z & b_Z \\ d_Z/b_Z & d_Z^2/b_Z^2 & c_Z & d_Z \\ a_Z & c_Z & a_Z^2 & a_Z b_Z \\ b_Z & d_Z & a_Z b_Z & b_Z^2 \end{bmatrix}$$

and set all remaining entries of Z to zero. Other orderings of indices or the use of a different nonzero entry is handled by relabeling/rearranging column/row order.

Case 2: ℓ has three unique entries. Then, exactly one of a_Z, b_Z, c_Z, d_Z is a diagonal entry, and so cases (i)–(iii), (v)–(vii) apply. If in any of these cases a_Z or d_Z is on the diagonal, by construction $|b_Z| = |c_Z|$. As $a_Z d_Z = b_Z c_Z$, we have $b_Z = c_Z = 0$ iff exactly one of a_Z or d_Z is zero. Likewise, if b_Z or c_Z is a diagonal element, then $|a_Z| = |d_Z|$ and so $a_Z = d_Z = 0$ iff exactly one of b_Z or c_Z are zero.

Suppose a_Z is a nonzero diagonal entry. We propose:

$$Z_{\ell'} = \begin{bmatrix} a_Z & b_Z & c_Z \\ b_Z & b_Z^2/a_Z & d_Z \\ c_Z & d_Z & c_Z^2/a_Z \end{bmatrix}$$

where ℓ' are the unique entries of ℓ. If $a_Z = 0$ and on the diagonal, then we replace b_Z^2/a_Z and c_Z^2/a_Z with $|d_Z|$. If b_Z, c_Z or d_Z is on the diagonal, we use the same construction but with relabeling/rearranging column/row order.

Case 3: ℓ has two unique entries. All remaining entries of Z are set to zero.

For all cases, our construction ensures that all diagonal entries of Z are nonnegative, and all 2×2 minors are zero; by Proposition 1, Z is an outer-product. $\qquad\square$

References

1. Andersen, K., Louveaux, Q., Weismantel, R.: An analysis of mixed integer linear sets based on lattice point free convex sets. Math. Oper. Res. **35**(1), 233–256 (2010)
2. Andersen, K., Louveaux, Q., Weismantel, R., Wolsey, L.A.: Inequalities from two rows of a simplex tableau. In: Fischetti, M., Williamson, D.P. (eds.) IPCO 2007. LNCS, vol. 4513, pp. 1–15. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-72792-7_1
3. Anstreicher, K.M.: Semidefinite programming versus the reformulation lineariza-tion technique for nonconvex quadratically constrained quadratic programming. J. Glob. Optim. **43**, 471–484 (2009)
4. Atamtürk, A., Narayanan, V.: Conic mixed-integer rounding cuts. Math. Program. **122**, 1–20 (2010)
5. Averkov, G.: On finite generation and infinite convergence of generalized closures from the theory of cutting planes. arXiv preprint arXiv:1106.1526 (2011)
6. Balas, E.: Intersection cuts—a new type of cutting planes for integer programming. Oper. Res. **19**(1), 19–39 (1971)
7. Bao, X., Sahinidis, N.V., Tawarmalani, M.: Multiterm polyhedral relaxations for nonconvex, quadratically constrained quadratic programs. Optim. Methods Softw. **24**(4–5), 485–504 (2009)
8. Basu, A., Conforti, M., Cornuéjols, G., Zambelli, G.: Maximal lattice-free convex sets in linear subspaces. Math. Oper. Res. **35**(3), 704–720 (2010)

9. Basu, A., Conforti, M., Cornuéjols, G., Zambelli, G.: Minimal inequalities for an infinite relaxation of integer programs. SIAM J. Discrete Math. **24**(1), 158–168 (2010)

10. Belotti, P., Góez, J.C., Pólik, I., Ralphs, T.K., Terlaky, T.: On families of quadratic surfaces having fixed intersections with two hyperplanes. Discrete Appl. Math. **161**(16–17), 2778–2793 (2013)

11. Bienstock, D., Michalka, A.: Cutting-planes for optimization of convex functions over nonconvex sets. SIAM J. Optim. **24**, 643–677 (2014)

12. Bienstock, D., Chen, C., Muñoz, G.: Outer-product-free sets for polynomial optimization and oracle-based cuts. arXiv preprint arXiv:1610.04604 (2016)

13. Bonami, P., Günlük, O., Linderoth, J.: Globally solving nonconvex quadratic programming problems with box constraints via integer programming methods. Math. Program. Comput. **10**(3), 333–382 (2018)

14. Borozan, V., Cornuéjols, G.: Minimal valid inequalities for integer constraints. Math. Oper. Res. **34**(3), 538–546 (2009)

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

16. Chen, C., Atamtürk, A., Oren, S.S.: A spatial branch-and-cut method for nonconvex QCQP with bounded complex variables. Math. Program. **165**(2), 549–577 (2017)

17. Conforti, M., Cornuéjols, G., Daniilidis, A., Lemaréchal, C., Malick, J.: Cut-generating functions and S-free sets. Math. Oper. Res. **40**(2), 276–391 (2014)

18. Cornuéjols, G., Wolsey, L., Yıldız, S.: Sufficiency of cut-generating functions. Math. Program. **152**, 1–9 (2013)

19. Dax, A.: Low-rank positive approximants of symmetric matrices. Adv. Linear Algebra Matrix Theory **4**(3), 172–185 (2014)

20. Dey, S.S., Wolsey, L.A.: Lifting integer variables in minimal inequalities corresponding to lattice-free triangles. In: Lodi, A., Panconesi, A., Rinaldi, G. (eds.) IPCO 2008. LNCS, vol. 5035, pp. 463–475. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-68891-4_32

21. Dey, S.S., Wolsey, L.A.: Constrained infinite group relaxations of MIPs. SIAM J. Optim. **20**(6), 2890–2912 (2010)

22. Eckart, C., Young, G.: The approximation of one matrix by another of lower rank. Psychometrika **1**(3), 211–218 (1936)

23. Fischetti, M., Ljubić, I., Monaci, M., Sinnl, M.: A new general-purpose algorithm for mixed-integer bilevel linear programs. Oper. Res. **65**(6), 1615–1637 (2017)

24. Floudas, C.A., et al.: Handbook of Test Problems in Local and Global Optimization, vol. 33. Springer, Boston (2013). https://doi.org/10.1007/978-1-4757-3040-1

25. Freund, R.M., Orlin, J.B.: On the complexity of four polyhedral set containment problems. Math. Program. **33**(2), 139–145 (1985)

26. Ghaddar, B., Vera, J.C., Anjos, M.F.: A dynamic inequality generation scheme for polynomial programming. Math. Program. **156**(1–2), 21–57 (2016)

27. Gomory, R.E., Johnson, E.L.: Some continuous functions related to corner polyhedra. Math. Program. **3**(1), 23–85 (1972)

28. Guennebaud, G., Jacob, B., et al.: Eigen v3 (2010). http://eigen.tuxfamily.org

29. Hillestad, R.J., Jacobsen, S.E.: Reverse convex programming. Appl. Math. Optim. **6**(1), 63–78 (1980)

30. Kılınç-Karzan, F.: On minimal valid inequalities for mixed integer conic programs. Math. Oper. Res. **41**(2), 477–510 (2015)

31. Kocuk, B., Dey, S.S., Sun, X.A.: Matrix minor reformulation and SOCP-based spatial branch-and-cut method for the AC optimal power flow problem. Math. Program. Comput. **10**(4), 557–596 (2018)
32. Krishnan, K., Mitchell, J.E.: A unifying framework for several cutting plane methods for semidefinite programming. Optim. Methods Softw. **21**, 57–74 (2006)
33. Lasserre, J.B.: Global optimization with polynomials and the problem of moments. SIAM J. Optim. **11**, 796–817 (2001)
34. Laurent, M.: Sums of squares, moment matrices and optimization over polynomials. In: Putinar, M., Sullivant, S. (eds.) Emerging Applications of Algebraic Geometry, pp. 157–270. Springer, New York (2009). https://doi.org/10.1007/978-0-387-09686-5_7
35. Locatelli, M., Schoen, F.: On convex envelopes for bivariate functions over polytopes. Math. Program. **144**, 1–27 (2013)
36. Lovász, L.: Geometry of numbers and integer programming. In: Mathematical Programming: Recent Developments and Applications, pp. 177–210 (1989)
37. Luedtke, J., Namazifar, M., Linderoth, J.: Some results on the strength of relaxations of multilinear functions. Math. Program. **136**(2), 325–351 (2012)
38. McCormick, G.P.: Computability of global solutions to factorable nonconvex programs: part I - convex underestimating problems. Math. Program. **10**(1), 147–175 (1976)
39. Meeraus, A.: GLOBALLib. http://www.gamsworld.org/global/globallib.htm
40. Mirsky, L.: Symmetric gauge functions and unitarily invariant norms. Q. J. Math. **11**(1), 50–59 (1960)
41. Misener, R., Floudas, C.A.: Global optimization of mixed-integer quadratically-constrained quadratic programs (MIQCQP) through piecewise-linear and edge-concave relaxations. Math. Program. **136**(1), 155–182 (2012)
42. Modaresi, S., Kılınç, M.R., Vielma, J.P.: Intersection cuts for nonlinear integer programming: convexification techniques for structured sets. Math. Program. **155**, 1–37 (2015)
43. MOSEK ApS: The MOSEK Fusion API for C++ 8.1.0.63 (2018). https://docs.mosek.com/8.1/cxxfusion/index.html
44. Qualizza, A., Belotti, P., Margot, F.: Linear programming relaxations of quadratically constrained quadratic programs. In: Lee, J., Leyffer, S. (eds.) Mixed Integer Nonlinear Programming, pp. 407–426. Springer, New York (2012). https://doi.org/10.1007/978-1-4614-1927-3_14
45. Rikun, A.D.: A convex envelope formula for multilinear functions. J. Glob. Optim. **10**(4), 425–437 (1997)
46. Saxena, A., Bonami, P., Lee, J.: Convex relaxations of non-convex mixed integer quadratically constrained programs: extended formulations. Math. Program. **124**, 383–411 (2010)
47. Saxena, A., Bonami, P., Lee, J.: Convex relaxations of non-convex mixed integer quadratically constrained programs: projected formulations. Math. Program. **130**, 359–413 (2011)
48. Sen, S., Sherali, H.D.: Nondifferentiable reverse convex programs and facetial convexity cuts via a disjunctive characterization. Math. Program. **37**(2), 169–183 (1987)
49. Sherali, H.D., Fraticelli, B.M.P.: Enhancing RLT relaxations via a new class of semidefinite cuts. J. Glob. Optim. **22**, 233–261 (2002)
50. Shor, N.Z.: Quadratic optimization problems. Sov. J. Circ. Syst. Sci. **25**, 6 (1987)
51. Tardella, F.: Existence and sum decomposition of vertex polyhedral convex envelopes. Optim. Lett. **2**, 363–375 (2008)

52. Tawarmalani, M., Richard, J.P.P., Xiong, C.: Explicit convex and concave envelopes through polyhedral subdivisions. Math. Program. **138**, 1–47 (2013)
53. Tawarmalani, M., Sahinidis, N.V.: Convex extensions and envelopes of lower semi-continuous functions. Math. Program. **93**, 247–263 (2002)
54. Tuy, H.: Concave programming under linear constraints. Sov. Math. **5**, 1437–1440 (1964)
55. Vandenbussche, D., Nemhauser, G.: A branch-and-cut algorithm for nonconvex quadratic programs with box constraints. Math. Program. **102**(3), 559–575 (2005)