Abstract. The class of POPs (Polynomial Optimization Problems) over cones covers a wide range of optimization problems such as 0-1 integer linear and quadratic programs, nonconvex quadratic programs and bilinear matrix inequalities. This paper presents a new framework for convex relaxation of POPs over cones in terms of linear optimization problems over cones. It provides a unified treatment of many existing convex relaxation methods based on the lift-and-project linear programming procedure, the reformulation-linearization technique and the semidefinite programming relaxation for a variety of problems. It also extends the theory of convex relaxation methods, and thereby brings flexibility and richness in practical use of the theory.

Key words. Global optimization, Convex relaxation, Nonconvex program, Quadratic program, Semidefinite program, Second-order cone program, Lift-and-project linear programming procedure, Polynomial optimization problem.
1 Introduction

Various convex relaxation methods have been studied intensively and extensively in recent years. For 0-1 integer LPs, a lift-and-project LP (Linear Programming) procedure by Balas-Ceria-Cornuéjols [1], the RLT (Reformulation-Linearization Technique) by Sherali-Adams [17] and an SDP (Semidefinite Programming) relaxation method by Lovász-Schrijver [12] were regarded as their pioneering works. They had been modified, generalized and extended to various problems and methods; the RLT [18, 19] for 0-1 mixed integer polynomial programs, the SCRM (Successive Convex Relaxation Method) [7, 8] for QOPs (Quadratic Optimization Problems), the SDP relaxation [9, 10][4] of polynomial programs, and SOCP (Second Order Cone Programming) relaxations [5, 6] for QOPs. These methods share the following basic idea.

(i) Add (redundant) valid inequality constraints to a target optimization problem in the $n$ dimensional Euclidean space $\mathbb{R}^n$.

(ii) Lift the problem with the additional inequality constraints in $\mathbb{R}^n$ to an equivalent optimization problem in a symmetric matrix space; the resulting problem is an LP with additional rank-1 and positive semidefinite constraints on its matrix variables.

(iii) Relax the rank-1 constraint (and positive semidefinite constraint in cases of the RLT and the lift-and-project LP procedure) so that the resulting feasible region is convex.

(iv) Project the relaxed lifted problem in the matrix space back to the original Euclidean space $\mathbb{R}^n$.

In some special cases such as the max cut problem [3] or some classes of 0-1 QOPs [6, 15, 23, 24, 26], (i) is not included, but it is necessary in general. In fact, (i) is a key issue in the papers [1, 5, 7, 8, 9, 10, 12, 17, 18, 19] mentioned above; they employ various techniques of constructing effective valid inequality constraints to strengthen their convex relaxations. Lasserre’s SDP relaxation method [9] is very powerful as a theoretical result in the sense that optimal values of fairly general polynomial programs having a compact feasible region can be approximated as closely as desired by solving a finite sequence of SDP relaxations. However, as we require higher accuracy for approximate optimal values, the size of SDP relaxations to be solved increases rapidly. This creates a major difficulty in applying the SDP relaxation method even to medium scale problems. Practically, computing better bounds for optimal values efficiently (inexpensively) is a critical and important issue, specifically when convex relaxation is utilized in the branch and bound method. Related to this issue, Kim and Kojima [5] recently showed that their SOCP relaxation is a reasonable compromise between the effectiveness of the SDP relaxation and the low computational cost of the lift-and-project LP relaxation. Some comparison among the techniques employed in [9], [12] and [17] was reported in [11].

The purpose of this article is to present a general and flexible framework for convex relaxation methods for POPs (polynomial optimization problems) over cones. This new framework is a slight extension of the existing framework for polynomial programs [9, 10, 18,

---

4Lasserre called his relaxation presented in these papers the LMI (Linear Matrix Inequality) relaxation, but we call it the SDP relaxation because the latter name is more popular than the former.
However, it not only provides a unified treatment of various existing convex relaxation methods mentioned above, but also leads to their further extensions using LOPs (Linear Optimization Problems) over cones [16]; in particular, it highlights SOCP relaxations [5, 6, 14] of POPs over cones, and bring richness in the theory and practice of convex relaxations of nonconvex programs.

In Section 2, we begin with a POP over cones in the Euclidean space; the problem can be either a target problem itself to be solved or the one that we have derived from a 0-1 LP, a QOP or a polynomial program by adding some valid polynomial constraints over cones as in (i). (See the problem (1) below for an example of a POP over cones.) Then, instead of (ii) and (iii), we just relax the problem by an LOP over cones in a higher dimensional Euclidean space. This is described in the latter part of Section 2. The linearization technique used there is the same as in the lift-and-project LP procedure [1] and the RLT [17, 18, 19].

We should emphasize that adding valid polynomial constraints over cones to an optimization problem to be solved is essential to have tighter convex relaxations. An important feature of the new framework is flexibility and variety in constructing valid constraints, which can be polynomial SDP constraints, polynomial SOCP constraints or even more general polynomial constraints over cones. In Sections 3 and 4, we discuss how we construct such constraints in detail.

In Section 5, we show how our convex relaxation by LOPs over cones works on the problem (1) below.

\[
\begin{align*}
\text{maximize} \quad & -2x_1 + x_2 \\
\text{subject to} \quad & x_1 \geq 0, \ x_2 \geq 0, \ x_1^2 + (x_2 - 1)^2 - 1 \geq 0, \ \left\| \begin{pmatrix} x_1 + 1 \\
 x_2 \end{pmatrix} \right\| \leq 2.
\end{align*}
\] (1)

Here \(x_1\) and \(x_2\) denote real variables. See Figure 1 on page 12. This problem is used as an illustrative example of POPs over cones throughout the paper.

Section 6 is devoted to the theory of the convex relaxation of POPs over cones. Based on the duality theorem of LOPs over cones, we present a sufficient condition for the convex relaxation to attain a given bound for the objective values of a POP over cones. This part is regarded as a generalization of Theorem 4.2 (b) of [9] where Lasserre presented a characterization of bounds attained by his SDP relaxation for the objective values of a polynomial program. We also generalize a characterization of the SDP relaxation given by Fujie-Kojima [2] (see also [7]) for QOPs to POPs over cones.

## 2 Polynomial optimization problems over cones and their linearization

Let \(\mathcal{K}\) be a closed convex cone in the \(m\)-dimensional Euclidean space \(\mathbb{R}^m\), \(f_0(x), f_1(x), \ldots, f_m(x)\) polynomials in real variables \(x_1, x_2, \ldots, x_n, x = (x_1, x_2, \ldots, x_n)^T\) a variable vector in \(\mathbb{R}^n\), and \(f(x) = (f_1(x), f_2(x), \ldots, f_m(x))^T \in \mathbb{R}^m\) a system of polynomials. We are concerned with the POP (Polynomial Optimization Problem) over the cone \(\mathcal{K}\):

\[
\begin{align*}
\text{maximize} \quad & f_0(x) \text{ subject to } x \in \mathcal{F} \equiv \{ x \in \mathbb{R}^n : f(x) \in \mathcal{K} \}.
\end{align*}
\] (2)

In our succeeding discussions, we often deal with the cases where \(\mathcal{K}\) is represented as the Cartesian product of closed convex cones \(\mathcal{K}_i \subset \mathbb{R}^{m_i} \ (i = 1, 2, \ldots, k)\) such that \(\mathcal{K} =\)
\[ \mathcal{K}_1 \times \mathcal{K}_2 \times \cdots \times \mathcal{K}_k, \text{ where we assume that } \sum_{i=1}^{k} m_i = m. \] The POP in such cases is written as

\[
\text{maximize } f_0(x) \text{ subject to } x \in \mathcal{F} \equiv \{ x \in \mathbb{R}^n : f_i(x) \in \mathcal{K}_i (i = 1, 2, \ldots, k) \}. \tag{3}
\]

Here each \( f_i(x) \in \mathbb{R}^{m_i} \) denotes a system of polynomials in real variables \( x_1, x_2, \ldots, x_n \).

The nonnegative orthant \( \mathbb{R}_+^\ell \) and the cone \( \mathcal{S}^\ell_+ \) of \( \ell \times \ell \) symmetric positive semidefinite matrices are popular closed convex cones which have been used in LPs (Linear Programs) and SDPs (Semidefinite Programs), respectively. We identify the space \( \mathcal{S}^\ell \) of \( \ell \times \ell \) symmetric matrices with the \( \ell \times \ell \)-dimensional Euclidean space in our succeeding discussion; hence we regard the positive semidefinite matrix cone \( \mathcal{S}^\ell_+ \) in \( \mathcal{S}^\ell \) as a closed convex cone in \( \mathbb{R}^{\ell \times \ell} \). We also consider the \( p \)-order cone

\[
\mathcal{N}_{p}^{1+\ell} \equiv \{ v = (v_0, v_1, v_2, \ldots, v_\ell)^T \in \mathbb{R}^{1+\ell} : \sum_{i=1}^{\ell} |v_i|^p \leq v_0^p \}. 
\]

Here \( p \geq 1 \). Following the convention, we use the notation

\[
\mathcal{N}_{\infty}^{1+\ell} \equiv \{ v = (v_0, v_1, v_2, \ldots, v_\ell)^T \in \mathbb{R}^{1+\ell} : |v_i| \leq v_0 (i = 1, 2, \ldots, \ell) \}. 
\]

Among the \((1 + \ell)\)-dimensional \( p \)-order cones \( \mathcal{N}_{p}^{1+\ell} \) \(( p \geq 1 \)\), we are particularly interested in the case where \( p = 2 \) and the case where \( p = +\infty \). The former is corresponding to the second order cone, and LOPs over second order cones are known as SOCPs (Second Order Cone Programs). In the latter, \( v \in \mathcal{N}_{\infty}^{1+\ell} \) is characterized as a system of linear inequalities \(-v_0 \leq v_i \leq v_0 \ (i = 1, 2, \ldots, \ell)\). For example, we can rewrite lower and upper bound constraints \( \bar{b}_i \leq v_i \leq \bar{b}_i \) \((i = 1, 2, \ldots, \ell)\) as

\[
\begin{pmatrix}
1 \\
(2v_1 - (\bar{b}_1 + \bar{b}_1)) / (\bar{b}_1 - \bar{b}_1) \\
(2v_2 - (\bar{b}_2 + \bar{b}_2)) / (\bar{b}_2 - \bar{b}_2) \\
\vdots \\
(2v_m - (\bar{b}_m + \bar{b}_m)) / (\bar{b}_m - \bar{b}_m)
\end{pmatrix} \in \mathcal{N}_{\infty}^{1+\ell}.
\]

Here \(-\infty < \bar{b}_1 < \bar{b}_i < +\infty \ (i = 1, 2, \ldots, \ell)\). More general \( p \)-order cones were used in [22] where Xue and Ye proposed interior-point methods for minimization of a sum of \( p \)-norms. It should be noted that \( \mathcal{N}_{p}^{1+\ell} \subset \mathcal{N}_{q}^{1+\ell} \) if \( 1 \leq p \leq q \).

A good example which shows a difference between polynomial programs [9, 10, 19, 21] and POPs over cones is a BMI (Bilinear Matrix Inequality): Find \( x = (1, x_1, x_2, \ldots, x_p)^T \in \mathbb{R}^p \) and \( y = (1, y_1, y_2, \ldots, y_q)^T \in \mathbb{R}^q \) satisfying \( \sum_{i=0}^{p} \sum_{j=0}^{q} A_{ij}x_iy_j \in \mathcal{S}_{+}^{\ell}, \) where \( A_{ij} \in \mathcal{S}^{\ell} \) \((i = 0, 1, 2, \ldots, p, \ j = 0, 1, 2, \ldots, q)\). We can rewrite the BMI as a POP over cones:

\[
\text{maximize } \lambda \text{ subject to } \sum_{i=0}^{p} \sum_{j=0}^{q} A_{ij}x_iy_j - \lambda I \in \mathcal{S}_{+}^{\ell} \text{ and } 1 - \lambda \geq 0, 
\]

where \( I \) denotes the \( \ell \times \ell \) identity matrix; the BMI has a solution if and only if the maximal objective value of the POP over cones above is nonnegative. In their paper [7], Kojima-Tuncel dealt with this problem as a special case of the conic quadratic inequality representation.
If we take \( n = 2, k = 4 \),

\[
\begin{align*}
g_0(x) &= -2x_1 + x_2, \quad g_1(x) = x_1, \quad g_2(x) = x_2, \\
g_3(x) &= x_1^2 + x_2^2 - 2x_2, \quad g_4(x) = \begin{pmatrix} 2 \\ x_1 + 1 \end{pmatrix}, \\
K_1 &= K_2 = K_3 = \mathbb{R}_+ \quad \text{(the cone of nonnegative numbers)}, \\
K_4 &= \mathbb{N}_2^3 \quad \text{(the 3-dimensional second order cone)},
\end{align*}
\]

we can rewrite the problem (1) as a POP

\[
\begin{align*}
\text{maximize} & \quad g_0(x) \\
\text{subject to} & \quad g_j(x) \in \mathbb{R}_+ \quad (j = 1, 2, 3), \quad g_4(x) \in \mathbb{N}_2^3.
\end{align*}
\] (4)

Let us illustrate how we create valid polynomial constraints over cones in the example (4) above. First we observe that

\[
g_5(x) \equiv x_1x_1 \in \mathbb{R}_+, \quad g_6(x) \equiv x_1x_2 \in \mathbb{R}_+, \quad g_7(x) \equiv x_2x_2 \in \mathbb{R}_+ \quad \text{(5)}
\]

are valid polynomial constraints over the cone \( \mathbb{R}_+ \) of nonnegative numbers for the problem (4). We consider the pair of constraints \( g_1(x) \in \mathbb{R}_+ \) and \( g_4(x) \in \mathbb{N}_2^3 \) over cones in the problem (4). Another constraint is obtained by “multiplying” them as

\[
g_8(x) \equiv g_1(x)g_4(x) \in \mathbb{N}_2^3. \quad \text{(6)}
\]

Similarly we can derive valid constraint

\[
g_9(x) \equiv g_2(x)g_4(x) \in \mathbb{N}_2^3. \quad \text{(7)}
\]

By adding the valid constraints in (5), (6) and (7) to the problem (4), we obtain a POP over cones

\[
\begin{align*}
\text{maximize} & \quad g_0(x) \\
\text{subject to} & \quad g_j(x) \in \mathbb{R}_+ \quad (j = 1, 2, 3, 5, 6, 7), \quad g_k(x) \in \mathbb{N}_2^3 \quad (k = 4, 8, 9),
\end{align*}
\] (8)

which is equivalent to the original problem (1).

Now we return to the general POP (2), and show how we linearize it. Let \( \mathbb{Z}_+ \) denote the set of nonnegative integers. For every variable vector \( x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n \) and every \( \mathbf{a} = (a_1, a_2, \ldots, a_n) \in \mathbb{Z}_+^n \), we use the notation \( x^\mathbf{a} \) for the term \( x_1^{a_1}x_2^{a_2}\cdots x_n^{a_n} \). Then we can write any polynomial \( f(x) \) in the variables \( x_1, x_2, \ldots, x_n \) in the form \( f(x) = \sum_{\mathbf{a} \in \mathcal{A}} c(\mathbf{a}) x^\mathbf{a} \) for some finite subset \( \mathcal{A} \) of \( \mathbb{Z}_+^n \) and some \( c(\mathbf{a}) \in \mathbb{R} \ (\mathbf{a} \in \mathcal{A}) \). We call \( \mathcal{A} \) the support of the polynomial \( f(x) \). Replacing each \( x^\mathbf{a} \) of the polynomial \( f(x) \) by a single variable \( y_\mathbf{a} \in \mathbb{R} \), we define a linearization \( F((y_\mathbf{a} : \mathbf{a} \in \mathcal{A})) \) of \( f(x) \) as \( F((y_\mathbf{a} : \mathbf{a} \in \mathcal{A})) = \sum_{\mathbf{a} \in \mathcal{A}} c(\mathbf{a}) y_\mathbf{a} \). Here \( (y_\mathbf{a} : \mathbf{a} \in \mathcal{A}) \) denotes a vector consisting of real variables \( y_\mathbf{a} \ (\mathbf{a} \in \mathcal{A}) \). We can naturally extend the definition of the linearization \( F((y_\mathbf{a} : \mathbf{a} \in \mathcal{A})) \) of a polynomial \( f(x) \) to a system of polynomials \( f(x) = (f_1(x), f_2(x), \ldots, f_m(x)) \).

Let \( F_0((y_\mathbf{a} : \mathbf{a} \in \mathcal{A})) \) and \( F((y_\mathbf{a} : \mathbf{a} \in \mathcal{A})) \) denote the linearizations of the polynomial \( f_0(x) \) and the polynomial system \( f(x) \), respectively. Here we assume for simplicity of notation that all polynomials \( f_0(x), f_1(x), \ldots, f_m(x) \) share a common support \( \mathcal{A} \); if the
term $x^a$ does not appear in a polynomial for some $a \in A$, let the corresponding coefficient $c(a)$ of the term $x^a$ be zero. A linearization of the POP (2) is

$$\text{maximize } F_0((y_a : a \in A)) \text{ subject to } F((y_a : a \in A)) \in K.$$ (9)

Since $F_0$ and $F$ are linear functions in variables $y_a (a \in A)$, the problem (9) forms an LOP over the cone $K$. By construction, we know that the LOP (9) serves as a convex relaxation of the POP (2) over cones. In fact, for any feasible solution $x$ of the POP (2) over $K$, $(y_a : a \in A) = (x^a : a \in A)$ gives a feasible solution of the LOP (9) at which the objective value of the function $F_0((y_a : a \in A))$ takes the same objective value $f_0(x)$ as the POP (2) over $K$.

Similarly we can derive a linearization of the POP (3)

$$\text{maximize } F_0((y_a : a \in A)) \text{ subject to } F_i((y_a : a \in A)) \in K_i \text{ (i = 1, 2, \ldots, k).}$$ (10)

Here $F_0((y_a : a \in A))$ and $F_i((y_a : a \in A))$ mean the linearizations of the polynomial $f_0(x)$ and the polynomial system $f_i(x) \text{ (i = 1, 2, \ldots, k)}$, respectively. When each cone $K_i$ is either the nonnegative orthant of the Euclidean space, a positive semidefinite matrix cone, or a second order cone, interior-point methods can be applied to solve the LOP (10) effectively.

The linearization of the the POP (8) leads to an SOCP

$$\begin{align*}
\text{maximize} & \quad -2y_{10} + y_{01} \\
\text{subject to} & \quad y_{10} \in \mathbb{R}_+, \ y_{01} \in \mathbb{R}_+, \ y_{20} + y_{02} - 2y_{01} \in \mathbb{R}_+, \\
& \quad y_{20} \in \mathbb{R}_+, \ y_{11} \in \mathbb{R}_+, \ y_{02} \in \mathbb{R}_+, \\
& \quad \begin{pmatrix} 2y_{10} \\ y_{20} + y_{10} \end{pmatrix} \in \mathcal{N}_2^3, \ \begin{pmatrix} 2y_{01} \\ y_{11} + y_{01} \end{pmatrix} \in \mathcal{N}_2^3,
\end{align*}$$ (11)

which serves as a convex relaxation of the POP (1).

## 3 Universally valid polynomial constraints over cones

Given a POP of the form (3) over cones, there are two types of valid polynomial constraints over cones which we can use to add to the original constraints for strengthening a resulting convex relaxation. The one is uvp-constraints (universally valid polynomial constraints) over cones that can be added to any POPs, and the other is polynomial constraints over cones that can be derived from the original polynomial constraints over cones and the uvp-constraints. We say that $f(x) \in K \subset \mathbb{R}^\ell$ is a uvp-constraint if it holds for every $x \in \mathbb{R}^n$. This section contains some representative and useful examples of valid polynomial constraints of the first type, and the next section the second type.

### 3.1 Positive semidefinite matrix cones

Let $u$ be a mapping from $\mathbb{R}^n$ into $\mathbb{R}^\ell$ whose $j$th component $u_j$ is a polynomial in $x_1, x_2, \ldots, x_n$. Then the $\ell \times \ell$ symmetric matrix $u(x)u(x)^T$ is positive semidefinite for any $x \in \mathbb{R}^n$, $u(x)u(x)^T \in S_+^\ell$. Thus we can add the polynomial constraint $u(x)u(x)^T \in S_+^\ell$ over cones to any POP over cones.
The popular SDP relaxation of nonconvex QOPs in the variable vector $x \in \mathbb{R}^n$ is derived by taking $u(x) = (1, x_1, x_2, \ldots, x_n)^T$. In this case, the constraint $u(x)u(x)^T \in S^n_+$ is written as

$$\begin{pmatrix} 1 & x_1 & x_2 & \cdots & x_n \\ x_1 & x_1x_1 & x_1x_2 & \cdots & x_1x_n \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ x_n & x_nx_1 & x_nx_2 & \cdots & x_nx_n \end{pmatrix} \in S^{1+n}_+, \quad (12)$$

and the corresponding linearization yields the standard positive semidefinite constraint for nonconvex QOPs. See, for example, [2, 7].

Specifically we have

$$\begin{pmatrix} 1 & x_1 \\ x_1 & x_1x_1 \\ x_2 & x_2x_1 \end{pmatrix} \in S^3_+, \quad \text{when } n = 2.$$

We can add this positive semidefinite constraint to the POP (8). The corresponding linearization becomes

$$\begin{pmatrix} 1 & y_{10} & y_{01} \\ y_{10} & y_{20} & y_{11} \\ y_{01} & y_{11} & y_{02} \end{pmatrix} \in S^3_+, \quad \text{(13)}$$

which we can add to the SOCP relaxation (11) of the POP (8).

In his paper [9, 10], Lasserre (implicitly) used a family of more stronger uvp-constraints over positive semidefinite matrix cones than the standard positive semidefinite constraint of the form (12). He took $u(x) = u^r(x)$ to be the column vector consisting of a basis for real-valued polynomials of degree $r$

$$1, x_1, x_2, \ldots, x_n, x_1^2, x_1x_2, \ldots, x_1x_n, x_2^2, x_2x_3, \ldots, x_n^2, \ldots, x_1^r, \ldots, x_n^r. \quad \text{(14)}$$

The linearization of the matrix $u^r(x)(u^r(x))^T$ is often called the moment matrix. See (15) below for the case that $n = 2$ and $r = 2$. Although the class of uvp-constraints $u^r(x)(u^r(x))^T \in S^n_{+}$, where $\ell$ denotes the dimension of the vector $u^r(x)$, may be theoretically powerful enough to attain better bounds for optimal values and/or the exact optimal values of a wide class of polynomial programs as shown in Theorem 4.2 of [9], other types of uvp-constraints are of importance in practice mainly because SDPs with large scale positive semidefinite matrix variables are difficult to solve.

When $n = 2$ and $r = 2$, the constraint $u^2(x)(u^2(x))^T \in S^6_+$ becomes

$$\begin{pmatrix} 1 \\ x_1 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \\ x_2x_2 \end{pmatrix} \begin{pmatrix} 1 & x_1 & x_1^2 & x_1x_2 & x_2^2 \\ x_1 & x_1x_1 & x_1x_2 & x_1^2x_2 & x_1x_2^2 \\ x_2 & x_1x_2 & x_2^2 & x_1x_2x_2 & x_2^2x_2 \\ x_1^2 & x_1x_1 & x_1^2x_2 & x_1^3x_2 & x_1^2x_2^2 \\ x_1x_2 & x_1x_2^2 & x_2^2x_2 & x_1^2x_2^2 & x_2^3x_2 \\ x_1^3 & x_1^2x_2 & x_1^2x_2^2 & x_2^3x_2 & x_1^3x_2^2 \end{pmatrix} \in S^{6}_{++},$$

$$\equiv \begin{pmatrix} 1 & x_1 & x_1^2 & x_1x_2 & x_2^2 \\ x_1 & x_1x_1 & x_1x_2 & x_1^2x_2 & x_1x_2^2 \\ x_2 & x_1x_2 & x_2^2 & x_1x_2x_2 & x_2^2x_2 \\ x_1^2 & x_1x_1 & x_1^2x_2 & x_1^3x_2 & x_1^2x_2^2 \\ x_1x_2 & x_1x_2^2 & x_2^2x_2 & x_1^2x_2^2 & x_2^3x_2 \\ x_1^3 & x_1^2x_2 & x_1^2x_2^2 & x_2^3x_2 & x_1^3x_2^2 \end{pmatrix} \in S^{6}_{++},$$
which we can add to the problem (8), and its linearization is

\[
\begin{pmatrix}
1 & y_{10} & y_{11} & y_{12} \\
y_{10} & y_{20} & y_{21} & y_{22} \\
y_{11} & y_{21} & y_{31} & y_{32} \\
y_{12} & y_{22} & y_{32} & y_{02}
\end{pmatrix} \in S^6_+, \quad (15)
\]

which can be used for the SOCP relaxation (11) of the problem (8).

### 3.2 Second order cones

This section presents another class of uvp-constraints over cones. Let \(f_1\) and \(f_2\) be mappings from \(\mathbb{R}^n\) into \(\mathbb{R}^\ell\) whose components are polynomials in \(x_1, x_2, \ldots, x_n\). By the Cauchy-Schwartz inequality, we see that \((f_1(x)^T f_2(x))^2 \leq (f_1(x)^T f_1(x))(f_2(x)^T f_2(x))\) holds. We can rewrite this inequality as a constraint over \(\mathcal{N}_2^3\):

\[
\begin{pmatrix}
f_1(x)^T f_1(x) + f_2(x)^T f_2(x) \\
f_1(x)^T f_1(x) - f_2(x)^T f_2(x) \\
2f_1(x)^T f_2(x)
\end{pmatrix} \in \mathcal{N}_2^3.
\]

In particular, if we take \(f_1(x) = x_i\) and \(f_2(x) = x_j\) for some \(1 \leq i < j \leq n\), we obtain

\[
\begin{pmatrix}
x_i^2 + x_j^2 \\
x_i^2 - x_j^2 \\
2x_i x_j
\end{pmatrix} \in \mathcal{N}_2^3.
\]

To derive another useful uvp-constraints, we consider a trivial relation

\[
f(x)f(x)^T - f(x)f(x)^T \in S^\ell_+
\]

for a mapping \(f\) from \(\mathbb{R}^n\) into \(\mathbb{R}^\ell\) whose components are polynomials in \(x_1, x_2, \ldots, x_n\). Let \(C\) be an arbitrary \(\ell \times \ell\) symmetric positive semidefinite matrix, \(L\) a \(p \times \ell\) matrix satisfying \(C = L^T L\), where \(1 \leq p \leq \ell\). Then the inequality \((Lf(x))^T (Lf(x)) \leq (f(x)^T C f(x))\) is true for any \(x \in \mathbb{R}^\ell\). We can convert this inequality to

\[
\begin{pmatrix}
1 + C \bullet f(x)f(x)^T \\
1 - C \bullet f(x)f(x)^T \\
2Lf(x)
\end{pmatrix} \in \mathcal{N}_2^{2+p}, \quad (17)
\]

which forms a uvp-constraint over \(\mathcal{N}_2^{2+p}\), and we obtain

\[
\begin{pmatrix}
1 + C \bullet H((ya : a \in A)) \\
1 - C \bullet H((ya : a \in A)) \\
2LF((ya : a \in A))
\end{pmatrix} \in \mathcal{N}_2^{2+p} \quad (18)
\]

as its linearization. Here \(H((ya : a \in A))\) and \(F((ya : a \in A))\) denote the linearizations of \(f(x)f(x)^T\) and \(f(x)\), respectively, and \(A \bullet B\) stands for the inner product of two matrices \(A, B \in S^p\); \(A \bullet B = \sum_{i=1}^p \sum_{j=1}^p A_{ij} B_{ij}\).
Notice that any polynomial constraint over a second order cone
\[
\begin{pmatrix}
  h_0(x) \\
  h(x)
\end{pmatrix} \in \mathcal{N}_{2}^{\ell+1},
\]
can be rewritten as
\[
\begin{pmatrix}
  h_0(x) & h(x)^T \\
  h(x) & h_0(x)I
\end{pmatrix} \in S^{1+\ell}_+,
\]
where \( h_i(x) \) denotes a polynomial in \( x_1, x_2, \ldots, x_n \) \((i = 0, 1, 2, \ldots, \ell)\), \( h(x) \) the polynomial system \((h_1(x), h_2(x), \ldots, h_\ell(x))^T \in \mathbb{R}^\ell \) and \( I \) the \( \ell \times \ell \) identity matrix. Specifically we can rewrite (16) as a valid polynomial inequality over \( S^{1+\ell}_+ \)
\[
\begin{pmatrix}
  1 & f(x)^T \\
  f(x) & f(x)f(x)^T
\end{pmatrix} \in S^{1+\ell}_+.
\]
As for the linearization of the constraint above, we obtain
\[
\begin{pmatrix}
  1 \\
  F((y_a : a \in A))^T \\
  H((y_a : a \in A))
\end{pmatrix} \in S^{1+\ell}_+.
\]
Then we may regard (18) as a further relaxation of (20). In view of the effectiveness of convex relaxations, adding the valid constraint (17) over \( \mathcal{N}_{2}^{2+p} \) does not result in a stronger convex relaxation than adding the valid constraint (19). Concerning their computational costs, however, (18) is much cheaper than (20). This fact was presented in the paper [5] for a special case in which \( f(x) = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n \) was taken.

4 Deriving valid polynomial constraints over cones

4.1 Kronecker products of positive semidefinite matrix cones

For every pair of an \( \ell \times \ell \) matrix \( A \) and an \( m \times m \) matrix \( B \), we use the notation \( A \otimes B \) for their Kronecker product. It is well-known that \( \nu \) is an eigenvalue of the \( \ell m \times \ell m \) matrix \( A \otimes B \) if and only if \( \nu \) is represented as the product of an eigenvalue \( \lambda \) of \( A \) and an eigenvalue \( \mu \) of \( B \) (see, for example, [4]). Hence, if \( A \) and \( B \) are positive semidefinite then so is their Kronecker product \( A \otimes B \).

Suppose that the POP (3) has two constraints over positive semidefinite matrix cones
\[
f_i(x) \in \mathcal{K}_i \equiv S^{m_i}_+ \quad \text{and} \quad f_j(x) \in \mathcal{K}_j \equiv S^{m_j}_+.
\]
Here we assume that every component of mappings \( f_i : \mathbb{R}^n \rightarrow S^{m_i} \) and \( f_j : \mathbb{R}^n \rightarrow S^{m_j} \) is a polynomial in \( x_1, x_2, \ldots, x_n \). In view of the discussion above, if \( x \in \mathbb{R}^n \) satisfies the constraint (21) then it satisfies
\[
f_i(x) \otimes f_j(x) \in S^{m_im_j}_+.
\]
As a result, we can add (22) to the POP (3) as a valid polynomial constraint over \( S^{m_im_j}_+ \).
We describe two special cases of interests as follows. The first one is the simplest case of the constraints (21) for \(m_i = m_j = 1\). In this case, both \(S_i^m\) and \(S_j^m\) become the set \(\mathbb{R}_+\) of nonnegative numbers. It follows that (21) and (22) can be rewritten as

\[
f_i(x) \geq 0 \quad \text{and} \quad f_j(x) \geq 0
\]

and

\[
f_i(x) f_j(x) \geq 0,
\]

respectively. A higher order polynomial inequality can always be derived in this way from two polynomial inequalities in the problem (3). This is an essential technique used in the RLT [17, 18, 19, 20].

Let us show the second case, which played an essential role in the papers [9, 10]. We let \(\mathbf{u}^r(x)\) denote the column vector consisting of a basis given in (14) for real-valued polynomials of degree \(r\), for some positive integer \(r\). Take \(m_i = 1\), \(f_j(x) = \mathbf{u}^r(x)(\mathbf{u}^r(x))^T\) and \(m_j\) to be the dimension of the vector \(\mathbf{u}^r(x)\). Then (21) and (22) becomes

\[
f_i(x) \geq 0 \quad \text{and} \quad \mathbf{u}^r(x)(\mathbf{u}^r(x))^T \in S_i^{m_j}
\]

and

\[
f_i(x) \mathbf{u}^r(x)(\mathbf{u}^r(x))^T \in S_j^{m_j},
\]

respectively.

### 4.2 Hadamard products of \(p\)-order cones \((p \geq 1)\)

We use the symbol \(\circ\) to denote the Hadamard product of two vectors \(\mathbf{v}\) and \(\mathbf{w}\) in \(\mathbb{R}^{1+\ell}\), \(\mathbf{v} \circ \mathbf{w} = (v_0 w_0, v_1 w_1, v_2 w_2, \ldots, v_{\ell} w_{\ell})^T\), and to indicate the Hadamard product of two closed convex cones \(\mathcal{V}\) and \(\mathcal{W}\) in \(\mathbb{R}^{1+\ell}\), \(\mathcal{V} \circ \mathcal{W} = \{\mathbf{v} \circ \mathbf{w} : \mathbf{v} \in \mathcal{V}, \mathbf{w} \in \mathcal{W}\}\). If \(1 \leq p \leq q \leq +\infty\), then \(\mathcal{N}_p^{1+\ell} \circ \mathcal{N}_q^{1+\ell} \subset \mathcal{N}_p^{1+\ell}\). In fact, if \(z = \mathbf{v} \circ \mathbf{w} \in \mathcal{N}_p^{1+\ell} \circ \mathcal{N}_q^{1+\ell}\) then

\[
\sum_{j=1}^{\ell} |z_j|^p = \sum_{j=1}^{\ell} |v_j w_j|^p \\
\leq \quad w_0^p \sum_{j=1}^{\ell} |v_j|^p \quad (\text{since } |w_j| \leq w_0 \quad (j = 1, 2, \ldots, \ell)) \\
\leq \quad (w_0 v_0)^p \quad (\text{since } \mathbf{v} \in \mathcal{N}_p^{1+\ell}) \\
= \quad z_0^p.
\]

Hence \(\mathcal{N}_p^{1+\ell} \circ \mathcal{N}_q^{1+\ell} \subset \mathcal{N}_p^{1+\ell}\). Recall that \(\mathcal{N}_p^{1+\ell} \subset \mathcal{N}_q^{1+\ell}\) when \(1 \leq p \leq q \leq +\infty\).

Suppose that \(\mathcal{K}_i = \mathcal{N}_p^{1+\ell}\) and \(\mathcal{K}_j = \mathcal{N}_q^{1+\ell}\) hold for some \(p\), \(q\) such that \(1 \leq p \leq q \leq +\infty\) in problem (3); either or both of \(p\) and \(q\) can be \(+\infty\). Then \(f_i(x) \circ f_j(x) \in \mathcal{N}_p^{1+\ell}\) provides a valid polynomial constraint for the problem (3).

It should be also noted that the cone \(\mathcal{N}_q^{1+\ell}\) is symmetric with respect to the coordinates \(1, 2, \ldots, \ell\). Thus, for any permutation \((s_1, s_2, \ldots, s_{\ell})\) of \((1, 2, \ldots, \ell)\),

\[
f_i(x) \circ ((f_j(x))_{s_0}, (f_j(x))_{s_1}, (f_j(x))_{s_2}, \ldots, (f_j(x))_{s_{\ell}})^T \in \mathcal{N}_p^{1+\ell}
\]
remains a valid polynomial constraint for the POP (3).

If \( h(x) \in \mathbb{R}_+ \) and \( f_j(x) \in \mathcal{N}_{p}^{1+\ell} \) then we have \( h(x)f_j(x) \in \mathcal{N}_{p}^{1+\ell} \). This relation is straightforward, and can be also derived from the above discussion if we take

\[
f_i(x) = (h(x), h(x), \ldots, h(x))^T \in \mathbb{R}^{1+\ell}.
\]

### 4.3 Linear transformation of cones

Any feasible solution \( x \) of the POP (3) satisfies

\[
T(f_1(x), f_2(x), \ldots, f_j(x)) \in \mathcal{K}_0
\]

if \( T \) is a linear mapping from \( \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \times \cdots \times \mathbb{R}^{m_j} \) into \( \mathbb{R}^{m_0} \) satisfying \( T(\mathcal{K}_1 \times \mathcal{K}_2 \times \cdots \times \mathcal{K}_j) \subset \mathcal{K}_0 \). Here \( j \geq 1 \) and \( \mathcal{K}_0 \) stands for a closed convex cone in \( \mathbb{R}^{m_0} \). Thus (23) serves as a valid polynomial constraint over \( \mathcal{K}_0 \) for the POP (3). As a result, we can add (23) to the POP (3) or replace some of the constraints \( f_i(x) \in \mathcal{K}_i \) \((i = 1, 2, \ldots, j))\) by (23) before producing other candidates of valid polynomial constraints over cones for the POP (3). If we compare the linearization of the constraint (23) with the set of the linearizations of the original constraints \( f_i(x) \in \mathcal{K}_i \) \((i = 1, 2, \ldots, j))\), however, the former is weaker than the latter; therefore we gain nothing for the effectiveness of convex relaxation. The purpose of applying a linear transformation to the constraints of the POP (3) is to create a smaller size or more tractable valid polynomial constraint over a cone combining some constraints of the POP (3) to save the computational cost. Recall that we have derived the uvp-constraint (17) and its linearization (18) by applying a linear transformation \( Y \in \mathcal{S}^\ell \rightarrow \mathbf{C} \cdot Y \in \mathbb{R} \) to the uvp-constraint (16) over \( \mathcal{S}^\ell_+ \).

### 4.4 Quadratic convexity

Let \( Q \in \mathcal{S}^n_+ \), \( q \in \mathbb{R}^n \) and \( \gamma \in \mathbb{R} \). We can transform a convex quadratic inequality \( x^T Q x + q^T x + \gamma \leq 0 \) to a linear constraint over the second order cone \( \mathcal{N}_{2}^{2+p} \)

\[
\begin{pmatrix}
1 - q^T x - \gamma \\
1 + q^T x + \gamma \\
2Lx
\end{pmatrix} \in \mathcal{N}_{2}^{2+p},
\]

where \( Q = L^T L \) indicates a factorization of \( Q \) for some \( p \times n \) matrix \( L \). Conversions like this are often used when we transform a quadratically constrained convex quadratic program into an SOCP to which interior-point methods can be applied effectively. In this section, we emphasize another advantage of this conversion. That is, after obtaining valid constraints over second order cones, we can apply the technique using the Hadamard product presented in Section 4.2 to create valid higher order polynomial constraints over second order cones.

We now consider the inequality constraint

\[
x^T Q x - h_1(x)h_2(x) \leq 0, \quad h_1(x) \geq 0 \text{ and } h_2(x) \geq 0.
\]

Here we assume that the \( n \times n \) symmetric matrix \( Q \) is positive semidefinite, and that \( h_1(x) \) and \( h_2(x) \) are polynomials in \( x_1, x_2, \ldots, x_n \). Since \( Q \) is positive semidefinite, we can find
an $p \times n$ matrix $L$ such that $Q = L^T L$. Then (24) turns to be an equivalent polynomial second order cone constraint

$$
\begin{pmatrix}
  h_1(x) + h_2(x) \\
  h_1(x) - h_2(x) \\
  2Lx
\end{pmatrix} \in \mathbb{N}^{2+p}_2.
$$

4.5 Adding valid polynomial inequality constraints from numerical computation

Let $e(x)$ be a polynomial in real variables $x_1, x_2, \ldots, x_n$ and $x = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$. We consider a valid polynomial inequality constraint of the form

$$
\lambda - e(x) \geq 0 \quad (25)
$$

for the POP (3). Here $\lambda \in \mathbb{R}$ serves as a parameter to be chosen such that the constraint (25) is valid inequality constraint for the POP (3). This leads to another POP

$$
\text{maximize } e(x) \text{ subject to } \mathcal{F} \equiv \{ x \in \mathbb{R}^n : f_i(x) \in \mathcal{K} (i = 1, 2, \ldots, k) \}. \quad (26)
$$

If $\lambda$ is not less than the optimal value $\lambda^*$ of this problem, then (25) provides a valid inequality constraint for the POP (3). Ideally we want to choose the optimal value $\lambda^*$ itself for $\lambda$. The computation $\lambda^*$ is usually as difficult as the original POP (3). Therefore, we apply convex relaxation to the POP (3), and formulate a LOP over cones

$$
\text{maximize } E((y_a : a \in \mathcal{A})) \text{ subject to } \mathcal{F}_i((y_a : a \in \mathcal{A})) \in \mathcal{K}(i = 1, 2, \ldots, k). \quad (27)
$$

Here $E((y_a : a \in \mathcal{A}))$ means the linearization of the polynomial $e(x)$. Let $\hat{\lambda}$ be the optimal value of the relaxed LOP (27). Obviously $\lambda^* \leq \hat{\lambda}$, and it follows that (25) with $\lambda = \hat{\lambda}$ is ensured to be a valid inequality constraint for the POP (3).

Adding a single valid inequality constraint (25) may not be enough to strengthen the convex relaxation of the POP (3). In general, it is necessary to generate multiple valid inequality constraints of the type $\hat{\lambda}_i - e_i(x) \geq 0$ with different polynomials $e_i(x)$ and their upper bounds $\hat{\lambda}_i$ over $\mathcal{F}$, and then apply the techniques given in the previous sections to them and some of the original polynomial constraints of the POP (3) to create new valid polynomial constraints over cones.

The discussion above brings the idea of a successive convex relaxation of the feasible region $\mathcal{F}$ of the POP (3) [1, 7, 8, 12]. We assume below that the objective function $f_0(x)$ of the POP (3) is a linear function. Then, the maximal objective value of the POP (3) coincides with the maximal value of $f_0(x)$ over the convex hull $\text{c.hull}(\mathcal{F})$ of $\mathcal{F}$. As a result, computing (an approximation of) the maximal objective value of the POP (3) is reduced to a tractable description (of an approximation) of the convex hull $\text{c.hull}(\mathcal{F})$ of $\mathcal{F}$. Suppose that we have a description $\mathcal{P}_q$ of the feasible region of the POP (3) at the $q$th iteration. Each member of $\mathcal{P}_q$ is a pair of a polynomial system $h(x)$ and a closed convex cone $\mathcal{H}$; specifically $\mathcal{P}_0 = \{(f_i, \mathcal{K}_i) : i = 1, 2, \ldots, k\}$, and we can rewrite the POP (3) as

$$
\text{maximize } f_0(x) \text{ subject to } h(x) \in \mathcal{H} (\langle h(x), \mathcal{H} \rangle \in \mathcal{P}_0).
$$
Figure 1: Problem (1): $n = 2, m = 4$ and the optimal value $0.0000$.

In each iteration, we compute valid inequality constraints for the feasible region $\mathcal{F}$ of the POP (3) with the description $\mathcal{P}_q$ by solving relaxed problems as presented above. Then we refine the description $\mathcal{P}_q$ of $\mathcal{F}$ by incorporating the newly computed valid inequality constraints into the description $\mathcal{P}_q$ to generate a better description $\mathcal{P}_{q+1}$ of $\mathcal{F}$. Here “better” means that the convex relaxation $C_q \equiv \{ x \in \mathbb{R}^n : H((y_a : a \in A)) \in \mathcal{H} ((h(x), \mathcal{H}) \in \mathcal{P}_q) \}$ shrinks; ideally to the convex hull $\text{c.hull}(\mathcal{F})$ of $\mathcal{F}$ such that $C_q \supset C_{q+1}$ and $\bigcap_{i=0}^{\infty} C_q = \text{c.hull}(\mathcal{F})$.

Some existing successive convex relaxation methods use linear functions for $e_i(x)$'s, and quadratic valid inequalities of the form $(\hat{\lambda}_i - e_i(x)) (\hat{\lambda}_j - e_j(x)) \geq 0$ are generated and added to $\mathcal{P}_q$. See [1, 7, 8, 12] for more details.

5 Numerical experiment on the POP (1)

We investigate and compare three different convex relaxations of the problem (1), an SDP relaxation, an SOCP relaxation and a combination of them. See Figure 1.

We transform the problem (1) to a QOP

Maximize $-2x_1 + x_2$

subject to $x_1 \geq 0, x_2 \geq 0, x_1 x_2 \geq 0,$

$x_1^2 + x_2^2 - 2x_2 \geq 0, x_1^2 + x_2^2 + 2x_1 - 3 \leq 0.$

Here we have added a valid quadratic constraint $x_1 x_2 \geq 0$. The standard application of the SDP relaxation to this QOP leads us to an SDP

Maximize $-2y_{10} + y_{01}$

subject to $y_{10} \geq 0, y_{01} \geq 0, y_{11} \geq 0,$

$y_{20} + y_{02} - 2y_{01} \geq 0, y_{20} + y_{02} + 2y_{10} - 3 \leq 0,$

$y \equiv \begin{pmatrix}
1 & y_{10} & y_{01} \\
y_{10} & y_{20} & y_{11} \\
y_{01} & y_{11} & y_{02}
\end{pmatrix} \in \mathcal{S}_+^3$.

(28)
which has an approximate optimal solution

\[
y_{\text{sdp}} = \begin{pmatrix}
1 & y_{10} & y_{01} \\
y_{10} & y_{20} & y_{11} \\
y_{01} & y_{11} & y_{02}
\end{pmatrix} = \begin{pmatrix}
1.0000 & 0.0000 & 1.5000 \\
0.0000 & 0.3830 & 0.0000 \\
1.5000 & 0.0000 & 2.6170
\end{pmatrix},
\]

and an approximate optimal value 

\[-2y_{10} + y_{01} = 1.5000.\]

Now we consider an SOCP relaxation of the problem (1). Recall that we have converted the problem (1) into a POP (8) over cones \(R_+^*\) and \(N_2^3\) by adding some valid constraints given in (5), (6) and (7) to (1), and that we have derived an SOCP relaxation (11) of the POP (8). Solving the SOCP (11), we obtain an approximate optimal solution

\[
y_{\text{socp}} = \begin{pmatrix}
1 & y_{10} & y_{01} \\
y_{10} & y_{20} & y_{11} \\
y_{01} & y_{11} & y_{02}
\end{pmatrix} = \begin{pmatrix}
1.0000 & 0.3863 & 1.4416 \\
0.3863 & 0.3863 & 0.0000 \\
1.4416 & 0.0000 & 2.4969
\end{pmatrix},
\]

and an approximate optimal value 

\[-2y_{10} + y_{01} = 0.6691.\]

Comparing the optimal value 1.5000 of the SDP relaxation (28) with the optimal value 0.6691 of the SOCP relaxation (11), we know that the the SOCP relaxation provides a better bound for the optimal value 0.0000 of the original problem (1).

It is easily verified that:

- The optimal solution \(y_{\text{sdp}}\) of the SDP (28) does not satisfy the constraint

\[
\begin{pmatrix}
2y_{10} \\
y_{20} + y_{10} \\
y_{11}
\end{pmatrix} \in N_2^3
\]

of the SOCP (11),

- The optimal solution \(y_{\text{socp}}\) of the SOCP (11) does not satisfy the positive semidefinite constraint

\[
y = \begin{pmatrix}
1 & y_{10} & y_{01} \\
y_{10} & y_{20} & y_{11} \\
y_{01} & y_{11} & y_{02}
\end{pmatrix} \in S_+^3
\]

of the SDP (28).

This suggests to combine the SDP and SOCP relaxations to get a tighter bound for the optimal objective value. Indeed if we add the positive semidefinite constraint above to the SOCP (11), the resulting SOCP-SDP problem attains a better bound 0.5490 at \(y_{10} = 0.3170\) and \(y_{01} = 1.1830\). See Figure 1. We conclude that both SDP and SOCP relaxations are important in this example.

6 Characterization of the linearization of POPs over cones

Throughout this section, we deal with the POP of the form (2) over the cone \(K \subseteq \mathbb{R}^m\), and we present characterization of the upper bound that is attained by its linearization (9) for its optimal objective value.
6.1 Basic theory

We denote each \( f_j(x) \) in (2) as \( f_j(x) = \gamma_j + d_j^T x + \sum_{a \in A^N} c_j(a)x^a \), and its linearization as \( F_j(x, (y_a : a \in A^N)) = \gamma_j + d_j^T x + \sum_{a \in A^N} c_j(a)y_a \). Here \( \gamma_j \in \mathbb{R}, d_j \in \mathbb{R}^n \), and \( A^N \) means the subset of the support \( A \) corresponding to the nonlinear terms; \( A^N = \{ a \in A : \sum_{i=1}^n a_i \geq 2 \} \). We assume that \( \gamma_0 = 0 \). Let

\[
F(x, (y_a : a \in A^N)) = (F_1(x, (y_a : a \in A^N)), \ldots, F_m(x, (y_a : a \in A^N)))^T.
\]

Then a linearization of the POP (2) is defined as an LOP over the cone \( K \):

\[
\text{maximize} \quad d_0^T x + \sum_{a \in A^N} c_0(a)y_a \quad \text{subject to} \quad F(x, (y_a : a \in A^N)) \in K, \tag{29}
\]

which is a convex relaxation of the POP (2). For the succeeding discussions, we also present the dual of the LOP (29)

\[
\text{minimize} \quad \sum_{j=1}^m \gamma_j v_j \quad \text{subject to} \quad v \in G, \tag{30}
\]

where \( G = \{ v \in K^* : d_0 + \sum_{j=1}^m d_j v_j = 0, c_0(a) + \sum_{j=1}^m c_j(a)v_j = 0 (a \in A^N) \} \), and the Lagrangian function \( L(x, v) = f_0(x) + v^T f(x) \) for every \( x \in \mathbb{R}^n \) and every \( v \in K^* \). Then we know that

\[
\begin{align*}
\text{the supremum objective value of the POP (2)} & \leq \text{the supremum objective value of the primal LOP (29)} \\
& \leq \sum_{j=1}^m \gamma_j v_j \text{ at any } v \in G. \tag{31}
\end{align*}
\]

The lemma below characterizes the bound \( \sum_{j=1}^m \gamma_j v_j \) in the last line of (31).

**Lemma 6.1.**

(i) A pair of \( v \in K^* \) and \( \zeta \in \mathbb{R} \) satisfies

\[
L(x, v) = \zeta \text{ for every } x \in \mathbb{R}^n \tag{32}
\]

if and only if \( v \in G \) and \( \zeta = \sum_{j=1}^m \gamma_j v_j \).

(ii) Suppose that there exist \( v \in K^* \) and \( \zeta \in \mathbb{R} \) satisfying the identity (32). Then the objective value of the primal LOP (29) is bounded by \( \zeta \) from above.

**Proof:** The assertion (i) follows from the observation that (32) holds if and only if

\[
\sum_{j=1}^m \gamma_j v_j = \zeta, \quad d_0 + \sum_{j=1}^m d_j v_j = 0, \quad c_0(a) + \sum_{j=1}^m c_j(a)v_j = 0 (a \in A^N)
\]
hold. The assertion (ii) follows from (31).

From the lemma, we see that

\[
\inf_{v \in K} \sup_{x \in \mathbb{R}^n} L(x, v) \leq \inf_{v \in G} \sup_{x \in \mathbb{R}^n} L(x, v) = \inf_{v \in G} \sum_{j=1}^m \gamma_j v_j.
\]

This shows that the upper bound given by the Lagrangian dual of the POP (2) for the optimal value of the POP (2) is at least as effective as the one by the dual LOP (30). In general, there is a gap between these two upper bounds. For example, consider a problem

\[
\text{maximize } x^4 - x^2 \text{ subject to } -x^4 + 1 \geq 0.
\]

The Lagrangian dual of this problem

\[
\inf_{v \in \mathbb{R}^+} \sup_{x \in \mathbb{R}} x^4 - x^2 + v(-x^4 + 1)
\]

attains the upper bound +1 for the optimal value 0 of the problem, while the linear relaxation

\[
\text{maximize } y_4 - y_2 \text{ subject to } -y_4 + 1 \geq 0
\]

forms an unbounded linear program; hence its dual is infeasible.

The theorem below provides a sufficient condition for the primal LOP (29) to attain the supremum objective value of the POP (2).

**Theorem 6.2.** Let \( \zeta^* < \infty \) be the supremum objective value of the POP (2). Suppose that the relation (32) holds for \( \zeta = \zeta^* \) and some \( v = v^* \in K^* \). Then \( v^* \) is an optimal solution of the dual LOP (30) and the identities

\[
\zeta^* = \text{the supremum objective values of the primal LOP (29)}
\]

\[
= \text{the optimal value of the dual LOP (30) at the optimal solution } v^*
\]

hold.

**Proof:** In general,

\[
\zeta^* = \sup \left\{ d_0^T x + \sum_{a \in A^N} c_0(a) x^a : f(x) \in K \right\}
\]

\[
\leq \sup \left\{ d_0^T x + \sum_{a \in A^N} c_0(a) y_a : F(x, (y_a : a \in A^N)) \in K \right\}
\]

\[
\leq \inf \left\{ \sum_{j=1}^m \gamma_j v_j : v \in G \right\}.
\]

On the other hand, we know in view of Lemma 6.1 that \( v^* \) is a feasible solution of the dual LOP (30) with the objective value \( \zeta^* \). Therefore the conclusion follows.

The theorem below characterizes the supremum objective value of the primal LOP (29) under the existence of an interior feasible solution of the POP (2).
Theorem 6.3. Assume that there is an \( \bar{x} \in \mathbb{R}^n \) such that \( f(\bar{x}) \) lies in the interior of \( K \) and that the objective values of the primal LOP (29) is bounded from above. Let \( \zeta \) denote the supremum objective values of (29). Then there exists \( v \in K^* \) satisfying (32).

Proof: Let \( (\bar{y}_a : a \in A^N) = (\bar{x}_a : a \in A^N) \). Then \( F(\bar{x}, (\bar{y}_a : a \in A^N)) = f(\bar{x}) \), which lies in the interior of the cone \( K \). By the duality theorem (see, for example, [16]), there exists an optimal solution \( v \) of the dual LOP (30) with the objective value \( \zeta = \sum_{j=1}^{m} \gamma_j v_j \). Thus the desired result follows from Lemma 6.1.

6.2 Convex relaxations of the feasible region of the POP (2)

We consider the projection \( \tilde{F} \) of the feasible region of the primal LOP (29) on the \( x \)-space

\[
\tilde{F} = \{ x \in \mathbb{R}^n : F(x, (y_a : a \in A^N)) \in K \text{ for some } (y_a : a \in A^N) \}.
\]

Then we may regard \( \tilde{F} \) as a convex relaxation of the feasible region \( F \) of the original POP (2). Characterization of \( \tilde{F} \) is given in this subsection. Such characterization has a meaning when the objective function \( f_0(x) \) of the POP (2) is a linear function \( c^T_0 x \) because its convex relaxation of (29) can be rewritten as

\[
\begin{align*}
\text{maximize} & \quad c^T_0 x \\
\text{subject to} & \quad x \in \tilde{F}.
\end{align*}
\]

Theoretically, we may assume without loss of generality that the objective function of the POP (2) is linear; otherwise replace the nonlinear objective function \( f_0(x) \) by \( x_{n+1} \) and \( f_0(x) - x_{n+1} \geq 0 \) to the constraint.

Let

\[
L = \{ v \in K^* : v^T f(x) \text{ is linear in } x \}^*
\]

or equivalently \( L^* = \{ v \in K^* : \sum_{j=1}^{m} v_j c_j(a) = 0 \ (a \in A^N) \} \). Then \( G \subset L^* \); hence if \( v \in K^* \) satisfies (32) then \( v \in L^* \).

We now introduce another convex relaxation of the feasible region of the POP (2):

\[
\tilde{F} \equiv \{ x \in \mathbb{R}^n : f(x) \in L \}.
\]

Note that each \( x \in \tilde{F} \) is characterized as

\[
v^T f(x) = \sum_{j=1}^{m} \gamma_j v_j + \left( \sum_{j=1}^{m} d_j v_j \right)^T x \geq 0 \text{ for every } v \in L^* \subset K^*.
\]

Hence \( \tilde{F} \) is represented in terms of an infinite number of linear inequalities.

Theorem 6.4.

(i) \( \tilde{F} \subset \tilde{F} \).
(ii) Assume that there is an \( \bar{x} \in \mathbb{R}^n \) such that \( f(\bar{x}) \) lies in the interior of \( K \). Then the closure of \( \tilde{F} \) coincides with \( F \).

Proof: (i) Suppose that \( x \in \tilde{F} \). Then there exists a \( (y_a : a \in A^N) \) such that \( F(x, (y_a : a \in A^N)) \in K \). Then, for every \( v \in L^* \subset K^* \), we see that \( v^T f(x) = v^T F(x, (y_a : a \in A^N)) \geq 0 \). Hence \( x \in F \).

(ii) We have shown above that \( \tilde{F} \subset \tilde{F} \). Since \( \tilde{F} \) is closed, we know that the the closure of \( \tilde{F} \) is contained in \( \tilde{F} \). Hence it suffices to show that the closure of \( \tilde{F} \) contains \( \tilde{F} \). Assume on the contrary that there is an \( \tilde{x} \in \tilde{F} \) which does not lie in the closure of \( \tilde{F} \). Then, by the separation theorem (see, for example [13]), there exist \( \tilde{c} \in \mathbb{R}^n \) and \( \tilde{\zeta} \in \mathbb{R} \) such that

\[
\tilde{c}^T \tilde{x} > \tilde{\zeta} \geq \tilde{c}^T x \quad \text{for every } x \in \tilde{F}.
\]  

Now, replace the objective function \( f_0(x) \) of the POP (2) and the primal LOP (29) by the linear function \( \tilde{c}^T x \) and apply Theorem 6.3. Then there exists a \( v \in L^* \) such that \( \tilde{c}^T x + v^T f(x) = \zeta \) for every \( x \in \mathbb{R}^n \). Here \( \zeta = \sup_{x \in \tilde{F}} \tilde{c}^T x \). Therefore \( \tilde{c}^T \tilde{x} \leq \tilde{c}^T x + v^T f(\tilde{x}) = \zeta \leq \tilde{\zeta} \). This contradicts to the first strict inequality in (34).

The result above is an extension of Theorem 2.1 of Fujie-Kojima [2] for QOPs to POPs over cones. Also the relation \( \tilde{F} \subset \tilde{F} \) played an essential role in the convergence analysis of the Kojima-Tuncel [7, 8] successive convex relaxations of nonconvex sets.

6.3 Application to Lasserre’s SDP Relaxation

We derive Theorem 4.2 (b) of [9], one of the main results shown by Lasserre as a special case of Theorem 6.2. Consider a polynomial program of the form

\[
\text{maximize } f_0(x) \text{ subject to } \tilde{f}_j(x) \geq 0 \ (j = 1, 2, \ldots, m).
\]  

Here we assume that \( f_0(x) \) and \( \tilde{f}_j(x) \) are polynomials in \( x_1, x_2, \ldots, x_n \) of degree at most \( \omega_0 \) and \( \omega_j \ (j = 1, 2, \ldots, m) \). Let \( \tilde{\omega}_j = \lfloor \omega_j/2 \rfloor \) and choose a nonnegative integer \( N \) not less than \( \omega_0/2 \) and \( \tilde{\omega}_j \ (j = 1, 2, \ldots, m) \). Define

\[
\begin{align*}
\mathbf{u}^T(x) &= \text{the column vector consisting of a basis for real-valued polynomials of degree } r \text{ given in (14)}, \\
\ell_j &= \text{the dimension of the column vector } \mathbf{u}^{N-\tilde{\omega}_j}(x) \ (j = 1, 2, \ldots, m), \\
\ell_{m+1} &= \text{the dimension of the column vector } \mathbf{u}^N(x), \\
K_j &= S^\ell_j_+ \ (j = 1, 2, \ldots, m + 1), \\
\mathbf{f}_j(x) &= \tilde{f}_j(x) \mathbf{u}^{N-\tilde{\omega}_j}(x) (\mathbf{u}^{N-\tilde{\omega}_j}(x))^T \in S^\ell_j_+ \ (j = 1, 2, \ldots, m), \\
\mathbf{f}_{m+1}(x) &= \mathbf{u}^N(x) (\mathbf{u}^N(x))^T \in S^m_{++}.
\end{align*}
\]

Then we can derive

\[
\text{maximize } f_0(x) \text{ subject to } \mathbf{f}_j(x) \in K_j \ (j = 1, 2, \ldots, m + 1)
\]  

17
as a POP over $\mathcal{K} = \mathcal{K}_1 \times \mathcal{K}_2 \times \cdots \times \mathcal{K}_{m+1}$, which is equivalent to the polynomial program 
(35), and its linearization

\[
\begin{aligned}
\text{maximize} & \quad F_0(x, (y_a : a \in \mathcal{A}^N)) \\
\text{subject to} & \quad F_j(x, (y_a : a \in \mathcal{A}^N)) \in \mathcal{K}_j \ (j = 1, 2, \ldots, m+1),
\end{aligned}
\tag{37}
\]

where $F_0(x, (y_a : a \in \mathcal{A}^N))$ and $F_j(x, (y_a : a \in \mathcal{A}^N))$ denote the linearization of $f_0(x)$ and $f_j(x)$ $(j = 1, 2, \ldots, m+1)$, respectively. We note that the LOP (37) is corresponding to “the LMI problem” (4.5) of Lasserre [9], but any problem corresponding to (36) was not described in [9] explicitly.

Since the dual $\mathcal{K}_j^*$ of each $\mathcal{K}_j = \mathcal{S}_{+}^{\ell_j}$ is itself, the Lagrangian function is

\[
L(x, V_1, V_2, \ldots, V_{m+1}) = f_0(x) + \sum_{j=1}^{m+1} V_j \cdot f_j(x)
\]

for every $x \in \mathbb{R}^n$ and $V_j \in \mathcal{S}_{+}^{\ell_j} \ (j = 1, 2, \ldots, m+1)$.

Let $\zeta \in \mathbb{R}$. Suppose that the identity

\[
L(x, V_1, V_2, \ldots, V_{m+1}) = \zeta \text{ for all } x \in \mathbb{R}^n.
\tag{38}
\]

holds for some $V_j \in \mathcal{S}_{+}^{\ell_j} \ (j = 1, 2, \ldots, m)$. Then we know by Lemma 6.1 that $(V_1, V_2, \ldots, V_{m+1})$ is a feasible solution of the dual LOP of (37). We represent each $V_j \in \mathcal{S}_{+}^{\ell_j}$ as

\[
V_j = \sum_{k=1}^{\ell_j} \lambda_{jk} w_{jk} w_{jk}^T,
\]

where $w_{jk}$ denotes an eigenvector of $V_j$ and $\lambda_{jk}$ the eigenvalue of $V_j$ associated with $w_{jk}$.

Then

\[
\begin{aligned}
L(x, V_1, \ldots, V_{m+1}) &= f_0(x) + \sum_{j=1}^{m} \left( \lambda_{jk} w_{jk} w_{jk}^T \right) \cdot \left( \tilde{f}_j(x) u^{N-\tilde{w}_j}(x) (u^{N-\tilde{w}_j}(x))^T \right) \\
&\quad + \left( \sum_{k=1}^{\ell_{m+1}} \lambda_{m+1,k} w_{m+1,k} w_{m+1,k}^T \right) \cdot u^{N}(x) (u^{N}(x))^T \\
&= f_0(x) + \sum_{j=1}^{m} \tilde{f}_j(x) \left( \sum_{k=1}^{\ell_j} \lambda_{jk} (w_{jk}^T u^{N-\tilde{w}_j}(x))^2 \right) + \sum_{k=1}^{\ell_{m+1}} \lambda_{m+1,k} (w_{m+1,k}^T u^{N}(x))^2.
\end{aligned}
\]

Letting

\[
t_j(x) = \sum_{k=1}^{\ell_j} (\sqrt{\lambda_{jk}} w_{jk}^T u^{N-\tilde{w}_j}(x))^2 \ (j = 1, 2, \ldots, m),
\]

\[
q(x) = \sum_{k=1}^{\ell_{m+1}} (\sqrt{\lambda_{m+1,k}} w_{m+1,k}^T u^{N}(x))^2,
\]

18
we obtain \( L(x, V_1, V_2, \ldots, V_{m+1}) = f_0(x) + \sum_{j=1}^{m} \tilde{f}_j(x) t_j(x) + q(x) \), for some polynomial \( q(x) \) of degree at most \( 2N \) and some polynomials \( t_j(x) \) of degree at most \( 2N - \omega_j \) \((j = 1, 2, \ldots, m)\), all sums of squares.

Conversely, suppose that the identity

\[
q(x) = \sum_{k=1}^{\ell_{m+1}} (w_{m+1,k}^T u^N(x))^2 \quad \text{and} \quad t_j(x) = \sum_{k=1}^{\ell_j} (w_{jk}^T u^{N-\omega_j}(x))^2.
\]

holds for some polynomial \( q(x) \) of degree at most \( 2N \) and some polynomials \( t_j(x) \) of degree at most \( 2N - \omega_j \) \((j = 1, 2, \ldots, m)\), all sums of squares such that

\[
q(x) = \sum_{k=1}^{\ell_{m+1}} (w_{m+1,k}^T u^N(x))^2 \quad \text{and} \quad t_j(x) = \sum_{k=1}^{\ell_j} (w_{jk}^T u^{N-\omega_j}(x))^2.
\]

Let \( V_j = \sum_{k=1}^{\ell_j} w_{jk} w_j^T \) \((j = 1, 2, \ldots, m+1)\). Then \( L(x, V_1, V_2, \ldots, V_{m+1}) = \zeta \).

Let \( \zeta \in \mathbb{R} \). We have shown that there exist \( V_j \in \mathcal{S}_{m+1}^j \) \((j = 1, 2, \ldots, m+1)\) satisfying (38) if and only if there exist a polynomial \( q(x) \) of degree at most \( 2N \) and polynomials \( t_j(x) \) of degree at most \( 2N - \omega_j \) \((j = 1, 2, \ldots, m)\), all sums of squares, satisfying (39). Consequently we obtain the corollary below by Lemma 6.1 and Theorem 6.2.

**Corollary 6.5.** Let \( \zeta^* < \infty \) be the supremum objective values of the POP (36) over the cone \( K = \mathcal{S}_+^{l_1} \times \mathcal{S}_+^{l_2} \times \ldots \times \mathcal{S}_+^{l_{m+1}} \). Suppose that the identity (39) holds for \( \zeta = \zeta^* \) and some polynomial \( q(x) \) of degree at most \( 2N \) and some polynomials \( t_j(x) \) of degree at most \( 2N - \omega_j \) \((j = 1, 2, \ldots, m)\), all sums of squares. Then

\[
\zeta^* = \text{the supremum of the objective values of the primal LOP (37)} = \text{the optimal value of the dual LOP of (37)}
\]

and if \( x^* \) is an optimal solution of the POP (36), then \((x^*, (y^*_a : a \in \mathcal{A}^N)) = (x^*, ((x^*)^a : a \in \mathcal{A}^N)) \) is an optimal solution of (37). In addition, if \((V_1, V_2, \ldots, V_{m+1}) \) is an optimal solution of the dual LOP of (37), then

\[
f_0(x) + \sum_{j=1}^{m+1} V_j \cdot f_j(x)
\]

\[
= f_0(x) + \sum_{j=1}^{m} \tilde{f}_j(x) \left( \sum_{k=1}^{\ell_j} \lambda_{jk} (w_{jk}^T u^{N-\omega_j}(x))^2 \right) + \sum_{k=1}^{\ell_{m+1}} \lambda_{m+1,k} (w_{m+1,k}^T u^N(x))^2
\]

\[
= \zeta^* \quad \text{for every } x \in \mathbb{R}^n
\]

holds. Here \( w_{jk} \) denotes an eigenvector of \( V_j \) and \( \lambda_{jk} \) the eigenvalue of \( V_j \) associated with \( w_{jk} \).

The corollary above corresponds to Theorem 4.2 (b) of [9].
7 Concluding remarks

We have presented a new framework for convex relaxation of POPs over cones in terms of LOPs over cones. Although this framework is quite general and flexible, there remain various important theoretical and practical issues.

From the practical viewpoint, the new framework certainly provides various ways of convex relaxation using LOPs over cones; in particular, the SOCP relaxation is likely to play a more important role. Further theoretical and numerical investigation is necessary to resolve the issue of how we create inexpensive valid polynomial constraints over cones that yield effective convex relaxation. Combining the branch-and-bound method and the new framework is also an interesting approach to utilize the framework.

Exact optimal values of a few special cases [6, 25, 26] of QOPs can be attained by constructing a single SDP relaxation. For some other cases [3, 23, 15] of QOPs, approximate optimal values generated through their SDP relaxations attain some guaranteed percentages of their exact optimal values. However, a single application of SDP relaxation to a general POP (or even a general QOP) neither attain the optimal value nor any guaranteed percentage of the optimal value. To get better bounds for the optimal value, it is necessary to use convex relaxation techniques repeatedly or to incorporate convex relaxation into some other optimization methods such as the branch-and-bound method. The former methodology includes the successive convex relaxation technique [1, 7, 8, 12] presented in Section 4.5, and the RLT (Reformulation-Linearization Technique) [17, 20] which constructs a sequence or a hierarchy of convex relaxation problems and then solves them sequentially (see also [9, 10]). The new framework proposed in this paper covers those techniques.

References


