Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 5, No. 2, pp. 295–310, May 2009

INVARIANCE UNDER AFFINE TRANSFORMATION IN SEMIDEFINITE PROGRAMMING RELAXATION FOR POLYNOMIAL OPTIMIZATION PROBLEMS

Hayato Waki, Masakazu Muramatsu and Masakazu Kojima

Abstract: Given a polynomial optimization problem (POP), any nonsingular affine transformation on its variable vector induces an equivalent POP. Applying Lasserre's SDP relaxation [SIAM J.Opt. 11:796–817, 2001] to the original and the transformed POPs, we have two SDPs. This paper shows that these two SDPs are isomorphic to each other under a nonsingular linear transformation, which maps the feasible region of one SDP onto that of the other isomorphically and preserves their objective values. This fact means that the SDP relaxation is invariant under any nonsingular affine transformation.

Key words: *polynomial optimization problem, semidefinite programming relaxation, sum of squares relaxation, invariance, affine transformation, polynomial SDP*

Mathematics Subject Classification: *65K05, 90C22, 99C30*

---

1 Introduction

A polynomial optimization problem (POP) is the problem of minimizing a polynomial objective function over a feasible region defined by polynomial equalities and inequalities. In recent years, intensive and extensive studies have been done on theoretical and practical aspects of semidefinite programming (SDP) relaxations for POPs since Lasserre's and Parrilo's pioneering works on this subject [11, 17]. See also [15, 19, 20] for earlier and more fundamental works. In theory, Lasserre's method constructs a sequence of SDP relaxation problems of a given POP, whose optimal values converge the optimal value of the POP under moderate assumptions [11, 14, 16]. In practice, some software packages [2, 18, 23] are available, and the sparse SDP relaxations [12, 22] can now be applied to large-scale POPs. Waki et al. [22] reported numerical results of the sparse SDP relaxations for large-scale POPs with sparse structure, including the minimization of the Broyden tridiagonal function with 1000 variables and a quadratic optimization problem with 1998 variables from the optimal control. The SDP relaxations also have been extended to polynomial SDPs [4, 5, 8] and POPs over symmetric cones [10, 22].

In this paper, we consider a POP (2.1) with an n-dimensional variable vector $\boldsymbol{x} \in \mathbb{R}^n$ and a POP (3.1) with a variable vector $\boldsymbol{w} \in \mathbb{R}^n$ transformed from (2.1) by a nonsingular affine transformation $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{w} + \boldsymbol{b}$, where $\boldsymbol{A}$ denotes an $n \times n$ nonsingular matrix and $\boldsymbol{b} \in \mathbb{R}^n$. Applying Lasserre's SDP relaxation, we obtain a pair of SDPs, one from the original POP (2.1) and the other from the transformed POP (3.1). We show that these two SDPs are isomorphic. More specially, there exists a nonsingular linear transformation between their feasible regions that preserves their objective values.

This paper is organized as follows. Section 2 describes the SDP relaxation proposed by Lasserre in a different way from [11]. Section 3 presents the main results, the invariant relations under a nonsingular affine transformation in the SDP relaxation illustrated in Figure 1. Section 4 is devoted to their proofs. Section 5 is devoted to some concluding remarks.

We introduce some symbols used in this paper. Let $\mathbb{R}$ denote the set of real numbers, $\mathbb{Z}_+$ the set of nonnegative integers, and $\mathbb{R}[\boldsymbol{x}]$ the set of polynomials in a variable vector $\boldsymbol{x} = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$. For every $\alpha \in \mathbb{Z}^n_+$, $\boldsymbol{x}^\alpha$ denotes the monomial $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ and $|\alpha| = \sum_{i=1}^n \alpha_i$. The degree $\deg(f)$ is the maximum value of $|\alpha|$ over all monomials $\boldsymbol{x}^\alpha$ whose coefficients f_α are nonzero.

2 Lasserre's SDP Relaxation

In this section, we present the SDP relaxation proposed by Lasserre [11]. Our description of the relaxation is, however, based on a general framework given in [9] for SDP relaxations of POPs over cones, and different from the original description using the moment theory by Lasserre [11] . The original of our description can be found in [7].

We consider the polynomial optimization problem

$$\text{minimize } \ f_0(\boldsymbol{x}) \text{ subject to } f_j(\boldsymbol{x}) \geq 0 \quad (j = 1, \ldots, m), \tag{2.1}$$

where $f_0, \ldots, f_m \in \mathbb{R}[\boldsymbol{x}]$. The SDP relaxation is composed of two steps. The first step is to replace the polynomial inequalities $f_j(\boldsymbol{x}) \geq 0 \quad (j = 1, \ldots, m)$ by a set of *valid polynomial matrix inequalities*. The resulting problem forms a polynomial SDP having the same polynomial objective function as POP (2.1) and polynomial matrix inequalities which are equivalent to the polynomial inequalities of POP (2.1). The second step is to *linearize* the polynomial SDP by replacing each monomial $\boldsymbol{x}^\alpha$ in the polynomial SDP with a variable y_α.

For every $r \in \mathbb{Z}_+$, let $G_r = \{\alpha \in \mathbb{Z}^n_+ \mid |\alpha| \leq r\}$ and let $\boldsymbol{u}_r(\boldsymbol{x})$ be the column vector of all monomials $\boldsymbol{x}^\alpha$ $(\alpha \in G_r)$: $\boldsymbol{u}_r(\boldsymbol{x}) = \left(\boldsymbol{x}^0, x_1, \ldots, x_n, x_1^2, x_1x_2, \ldots, x_n^2, \ldots, x_1^r, \ldots, x_n^r\right)^T$, where $\boldsymbol{x}^0$ is 1 for any $\boldsymbol{x} \in \mathbb{R}^n$. Let $s(r) = \binom{n+r}{r}$ denote the cardinality of G_r, which coincides with the size of the column vector $\boldsymbol{u}_r(\boldsymbol{x})$. We introduce the $s(r) \times s(r)$ symmetric matrix $\boldsymbol{u}_r(\boldsymbol{x})\boldsymbol{u}_r(\boldsymbol{x})^T$; the (β, γ)th element of the matrix is given by $\boldsymbol{x}^{\beta+\gamma}$ for each pair of row and column indices $\beta,\ \gamma \in G_r$. To represent $\boldsymbol{u}_r(\boldsymbol{x})\boldsymbol{u}_r(\boldsymbol{x})^T$ in terms of a polynomial in $\boldsymbol{x}$ with symmetric matrix coefficients, define an $s(r) \times s(r)$ matrix $\boldsymbol{E}_\alpha$ whose elements are given by

$$(E_\alpha)_{\beta,\gamma} = \begin{cases} 1 & \text{if } \alpha = \beta + \gamma, \quad \text{and } \beta, \gamma \in G_r, \\ 0 & \text{otherwise,} \end{cases} \tag{2.2}$$

for every $\alpha \in G_{2r}$. Then we can write $\boldsymbol{u}_r(\boldsymbol{x})\boldsymbol{u}_r(\boldsymbol{x})^T = \sum_{\alpha \in G_{2r}} \boldsymbol{x}^\alpha \boldsymbol{E}_\alpha$. We also deal with the $s(r) \times s(r)$ matrix $f(\boldsymbol{x})\boldsymbol{u}_r(\boldsymbol{x})\boldsymbol{u}_r(\boldsymbol{x})^T$ for each $f \in \mathbb{R}[\boldsymbol{x}]$. The (β, γ)th element of the matrix is $\boldsymbol{x}^{\beta+\gamma} f(\boldsymbol{x})$ for $\beta, \gamma \in G_r$. The matrix can be represented as $f(\boldsymbol{x})\boldsymbol{u}_r(\boldsymbol{x})\boldsymbol{u}_r(\boldsymbol{x})^T = \sum_{\alpha \in G_{2r+\deg(f)}} \boldsymbol{x}^\alpha \boldsymbol{B}_\alpha$, for some $s(r) \times s(r)$ matrices $\boldsymbol{B}_\alpha$ $(\alpha \in G_{2r+\deg(f)})$.

We observe that $\boldsymbol{u}_r(\boldsymbol{x})\boldsymbol{u}_r(\boldsymbol{x})^T$ is positive semidefinite for all $\boldsymbol{x} \in \mathbb{R}^n$, and that $f(\boldsymbol{x})\boldsymbol{u}_r(\boldsymbol{x})\boldsymbol{u}_r(\boldsymbol{x})^T$ is positive semidefinite for any $\boldsymbol{x}$ such that $f(\boldsymbol{x}) \geq 0$. As the first step of the SDP relaxation of POP (2.1), we will derive an equivalent polynomial SDP. Let $\bar{r}$ be the maximum value of $\lceil \deg(f_j)/2 \rceil$ over all $j = 0, 1, \ldots, m$. Choose a nonnegative integer $r \geq \bar{r}$, and let $r_j = r - \lceil \deg(f_j)/2 \rceil$ for all $j = 1, \ldots, m$. By definition, we see that $r, r_j \in \mathbb{Z}_+$

$(j = 1, 2, \ldots, m)$. Replacing each constraint $f_j(\boldsymbol{x}) \geq 0$ by $f_j(\boldsymbol{x})\boldsymbol{u}_{r_j}(\boldsymbol{x})\boldsymbol{u}_{r_j}(\boldsymbol{x})^T \succeq \boldsymbol{O}$ in POP (2.1) and adding $\boldsymbol{u}_r(\boldsymbol{x})\boldsymbol{u}_r(\boldsymbol{x})^T \succeq \boldsymbol{O}$ to POP (2.1), we now obtain a polynomial SDP

$$\left.\begin{array}{ll} \text{minimize} & f_0(\boldsymbol{x}) \\ \text{subject to} & f_j(\boldsymbol{x})\boldsymbol{u}_{r_j}(\boldsymbol{x})\boldsymbol{u}_{r_j}(\boldsymbol{x})^T \succeq \boldsymbol{O} \ (j = 1, \ldots, m), \\ & \boldsymbol{u}_r(\boldsymbol{x})\boldsymbol{u}_r(\boldsymbol{x})^T \succeq \boldsymbol{O}. \end{array}\right\} \tag{2.3}$$

Note that the $(0,0)$th element of the symmetric matrix $\boldsymbol{u}_{r_j}(\boldsymbol{x})\boldsymbol{u}_{r_j}(\boldsymbol{x})^T$ involved in the constraints is 1 for every $j = 1, \ldots, m$. This ensures that $f_j(\boldsymbol{x})\boldsymbol{u}_{r_j}(\boldsymbol{x})\boldsymbol{u}_{r_j}(\boldsymbol{x})^T \succeq \boldsymbol{O}$ if and only if $f_j(\boldsymbol{x}) \geq 0$. Therefore, POP (2.1) and polynomial SDP (2.3) are equivalent to each other. We further rewrite polynomial SDP (2.3) as

$$\left.\begin{array}{ll} \text{minimize} & \boldsymbol{c}_{2r}^T \boldsymbol{u}_{2r}(\boldsymbol{x}) \\ \text{subject to} & \sum_{\alpha \in G_{2r}} \boldsymbol{x}^\alpha \boldsymbol{B}_{j,\alpha} \succeq \boldsymbol{O} \ (j = 1, \ldots, m), \\ & \sum_{\alpha \in G_{2r}} \boldsymbol{x}^\alpha \boldsymbol{E}_\alpha \succeq \boldsymbol{O} \end{array}\right\} \tag{2.4}$$

for some $s(2r)$-dimensional column vector $\boldsymbol{c}_{2r}$ such that $f_0(\boldsymbol{x}) = \boldsymbol{c}_{2r}^T \boldsymbol{u}_{2r}(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathbb{R}^n$, some $s(r) \times s(r)$ symmetric matrices $\boldsymbol{E}_\alpha$ and some $s(r_j) \times s(r_j)$ symmetric matrices $\boldsymbol{B}_{j,\alpha}$ $(\alpha \in G_{2r}, j = 1, \ldots, m)$. By construction, we know that $\deg(f_0) \leq 2\bar{r} \leq 2r$. Hence, for any $\alpha \in G_{2r} \setminus G_{2\bar{r}}$, the αth element $(c_{2r})_\alpha$ of the column vector $\boldsymbol{c}_{2r}$ vanishes. This fact will be used later to see the monotonicity of the optimal value v_r^* of SDP (2.6) with respect to r.

Note that we use G_{2r} instead of $G_{2r_j+\deg(f_j)}$ to describe the matrices $\sum_{\alpha \in G_{2r}} \boldsymbol{x}^\alpha \boldsymbol{B}_{j,\alpha}$ in polynomial SDP (2.4) for the sake of simplicity. Indeed, we know that $G_{2r_j+\deg(f_j)} \subset G_{2r}$, and if $G_{2r} \setminus G_{2r_j+\deg(f_j)}$ is not empty, we set $\boldsymbol{B}_{j,\alpha} = \boldsymbol{O}$ for all $\alpha \in G_{2r} \setminus G_{2r_j+\deg(f_j)}$. Then $\sum_{\alpha \in G_{2r_j+\deg(f_j)}} \boldsymbol{x}^\alpha \boldsymbol{B}_{j,\alpha} = \sum_{\alpha \in G_{2r}} \boldsymbol{x}^\alpha \boldsymbol{B}_{j,\alpha}$ holds.

Before we proceed to the second step of the SDP relaxation, we show some examples to illustrate the symbols and notation used above.

Example 2.1. In the case of $n = 2$ and $r = 2$, we have

$$\begin{array}{rcl} G_r &=& \{(0,0),(1,0),(0,1),(2,0),(1,1),(0,2)\}, \\ \boldsymbol{u}_r(\boldsymbol{x}) &=& (\boldsymbol{x}^0, x_1, x_2, x_1^2, x_1x_2, x_2^2)^T, \\ \boldsymbol{u}_r(\boldsymbol{x})\boldsymbol{u}_r(\boldsymbol{x})^T &=& \left(\begin{array}{c|cc|ccc} \boldsymbol{x}^0 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 \\ \hline x_1 & x_1^2 & x_1x_2 & x_1^3 & x_1^2x_2 & x_1x_2^2 \\ x_2 & x_1x_2 & x_2^2 & x_1^2x_2 & x_1x_2^2 & x_2^3 \\ \hline x_1^2 & x_1^3 & x_1^2x_2 & x_1^4 & x_1^3x_2 & x_1^2x_2^2 \\ x_1x_2 & x_1^2x_2 & x_1x_2^2 & x_1^3x_2 & x_1^2x_2^2 & x_1x_2^3 \\ x_2^2 & x_1x_2^2 & x_2^3 & x_1^2x_2^2 & x_1x_2^3 & x_2^4 \end{array}\right), \\ G_{2r} &=& \left\{\begin{array}{c} (0,0),(1,0),(0,1),(2,0),(1,1),(0,2),(3,0),(2,1), \\ (1,2),(0,3),(4,0),(3,1),(2,2),(1,3),(0,4) \end{array}\right\}. \end{array}$$

Recall that $\boldsymbol{x}^0 = 1$ for any $\boldsymbol{x} \in \mathbb{R}^n$, so that the $(0,0)$th element of $\boldsymbol{u}_r(\boldsymbol{x})\boldsymbol{u}_r(\boldsymbol{x})^T$ is 1. If we take $\alpha = (2,0) \in G_{2r}$ and $\alpha = (3,1) \in G_{2r}$, we see

$$\boldsymbol{E}_{(2,0)} = \left(\begin{array}{c|cc|ccc} & & & 1 & & \\ \hline & 1 & & & & \\ & & & & & \\ \hline 1 & & & & & \\ & & & & & \\ & & & & & \end{array}\right) \text{ and } \boldsymbol{E}_{(3,1)} = \left(\begin{array}{c|cc|ccc} & & & & & \\ \hline & & & & & \\ & & & & & \\ \hline & & & & 1 & \\ & & & 1 & & \\ & & & & & \end{array}\right),$$

where each blank above means 0.

Example 2.2. Let $n = 2$, $f(\boldsymbol{x}) = 2 - x_1 + x_2$, and $r = 1$. Then

$$\begin{aligned}
G_r &= \{(0,0),(1,0),(0,1)\}, \\
\boldsymbol{u}_r(\boldsymbol{x}) &= (\boldsymbol{x}^0, x_1, x_2)^T, \\
& f(\boldsymbol{x})\boldsymbol{u}_1(\boldsymbol{x})\boldsymbol{u}_1(\boldsymbol{x})^T \\
&= \begin{pmatrix} 2\boldsymbol{x}^0 - x_1 + x_2 & 2x_1 - x_1^2 + x_1x_2 & 2x_2 - x_1x_2 + x_2^2 \\ 2x_1 - x_1^2 + x_1x_2 & 2x_1^2 - x_1^3 + x_1^2x_2 & 2x_1x_2 - x_1^2x_2 + x_1x_2^2 \\ 2x_2 - x_1x_2 + x_2^2 & 2x_1x_2 - x_1^2x_2 + x_1x_2^2 & 2x_2^2 - x_1x_2^2 + x_2^3 \end{pmatrix}, \quad (2.5) \\
& G_{2r+\deg(f)} \\
&= \{\ (0,0),(1,0),(0,1),(2,0),(1,1),(0,2),(3,0),(2,1),(1,2),(0,3)\ \}.
\end{aligned}$$

If we take $\alpha = (1,0),(1,1),(2,1) \in G_{2r+\deg(f)}$, we see

$$\boldsymbol{B}_{(1,0)} = \begin{pmatrix} -1 & 2 & \\ 2 & & \\ & & \end{pmatrix}, \ \boldsymbol{B}_{(1,1)} = \begin{pmatrix} & 1 & -1 \\ 1 & & 2 \\ -1 & 2 & \end{pmatrix} \text{ and } \boldsymbol{B}_{(2,1)} = \begin{pmatrix} & & \\ & 1 & -1 \\ & -1 & \end{pmatrix}.$$

If we replace each monomial $\boldsymbol{x}^\alpha$ on the right side of the identity (2.5), we have a linear mapping from the space of $s(2r+\deg(f))$-dimensional column vector $\boldsymbol{y}$ consisting of y_α ($\alpha \in G_{2r+\deg(f)}$) into the space of $s(r)\times s(r)$ symmetric matrices, which we will denote by $\boldsymbol{M}_r(f\boldsymbol{y})$ in the subsequent discussion;

$$\begin{aligned}
& \boldsymbol{M}_r(f\boldsymbol{y}) \\
&= \begin{pmatrix} 2y_0 - y_{(1,0)} + y_{(0,1)} & 2y_{(1,0)} - y_{(2,0)} + y_{(1,1)} & 2y_{(0,1)} - y_{(1,1)} + y_{(0,2)} \\ 2y_{(1,0)} - y_{(2,0)} + y_{(1,1)} & 2y_{(2,0)} - y_{(3,0)} + y_{(2,1)} & 2y_{(1,1)} - y_{(2,1)} + y_{(1,2)} \\ 2y_{(0,1)} - y_{(1,1)} + y_{(0,2)} & 2y_{(1,1)} - y_{(2,1)} + y_{(1,2)} & 2y_{(0,2)} - y_{(1,2)} + y_{(0,3)} \end{pmatrix} \\
&= \sum_{\alpha\in G_{2r+\deg(f)}} y_\alpha \boldsymbol{B}_\alpha.
\end{aligned}$$

Since the (β,γ)th element of the $s(r)\times s(r)$ matrix $f(\boldsymbol{x})\boldsymbol{u}_r(\boldsymbol{x})\boldsymbol{u}_r(\boldsymbol{x})^T$ is $\boldsymbol{x}^{\beta+\gamma}f(\boldsymbol{x})$ ($\beta,\gamma \in G_{2r}$), the corresponding element of the $s(r)\times s(r)$ matrix $\boldsymbol{M}_r(f\boldsymbol{y})$ is given by $\sum_{\alpha\in G_{\deg(f)}} f_\alpha y_{\alpha+\beta+\gamma}$ where f_α is the coefficient of the monomial $\boldsymbol{x}^\alpha$ in f. In this example, we have $f_{(0,0)} = 2, f_{(1,0)} = -1, f_{(0,1)} = 1$ and $f_\alpha = 0$ for all other α. If we take $\beta = (1,0), \gamma = (0,0)$, then we have

$$\begin{aligned}
\sum_{\alpha\in G_{\deg(f)}} f_\alpha y_{\alpha+\beta+\gamma} &= \sum_{\alpha\in G_{\deg(f)}} f_\alpha y_{\alpha+(1,0)+(0,0)} \\
&= f_{(0,0)}y_{(0,0)+(1,0)} + f_{(1,0)}y_{(1,0)+(1,0)} + f_{(0,1)}y_{(0,1)+(1,0)} \\
&= 2y_{(1,0)} - y_{(2,0)} + y_{(1,1)},
\end{aligned}$$

and we can see that the left-hand side is equal to the (β,γ)th element of the matrix $\boldsymbol{M}_r(f\boldsymbol{y})$.

Now we perform the second step of the SDP relaxation of POP (2.1). Recall that we have derived an equivalent polynomial SDP (2.4) from POP (2.1) in the first step. We apply the linearization to the objective polynomial function and the polynomial matrix inequality constraints of polynomial SDP (2.4) by replacing each $\boldsymbol{x}^\alpha$ by a single real variable y_α ($\alpha \in G_{2r}$). Then we obtain an SDP

$$\left.\begin{array}{ll} \text{minimize} & \boldsymbol{c}_{2r}^T\boldsymbol{y} \\ \text{subject to} & \boldsymbol{M}_{r_j}(f_j\boldsymbol{y}) \succeq \boldsymbol{O}\ (j=1,\ldots,m), \boldsymbol{M}_r(\boldsymbol{y}) \succeq \boldsymbol{O},\ y_0 = 1. \end{array}\right\} \quad (2.6)$$

Here

$$\boldsymbol{M}_{r_j}(f_j\boldsymbol{y}) = \sum_{\alpha\in G_{2r}} y_\alpha \boldsymbol{B}_{j,\alpha} \ (j=1,\dots,m), \ \boldsymbol{M}_r(\boldsymbol{y}) = \sum_{\alpha\in G_{2r}} y_\alpha \boldsymbol{E}_\alpha, \tag{2.7}$$

respectively. The size of variable vector $\boldsymbol{y}$ is $s(2r)$. For each $\beta,\gamma \in G_{r_j}$, the (β,γ)th element of $\boldsymbol{M}_{r_j}(f_j\boldsymbol{y})$ is $\sum_{\alpha\in G_{\deg(f_j)}} f_{j,\alpha} y_{\alpha+\beta+\gamma}$, where $f_{j,\alpha}$ the coefficient of the monomial $\boldsymbol{x}^\alpha$ of f_j.

We note that SDP (2.6) is defined for every nonnegative integer $r \geq \bar{r}$. Hence we obtain an infinite sequence of SDP relaxation problems of POP (2.1). Let v^* denote the optimal value of POP (2.1) and v^*_r the optimal value of SDP (2.6) with $r \geq \bar{r}$. Then $v^*_r \leq v^*_{r+1} \leq v^*$ for all $r \geq \bar{r}$. In fact, if $\boldsymbol{x} \in \mathbb{R}^n$ is a feasible solution of POP (2.1) (hence it is a feasible solution of polynomial SDP (2.4)), then $\boldsymbol{y} = \boldsymbol{u}_{2r}(\boldsymbol{x}) \in \mathbb{R}^{s(2r)}$ is a feasible solution of SDP (2.6) with the objective value $\boldsymbol{c}^T_{2r}\boldsymbol{y} = \boldsymbol{c}^T_{2r}\boldsymbol{u}_{2r}(\boldsymbol{x})$. This implies that if POP (2.1) attains an objective value at a feasible solution then so does SDP (2.6). Hence $v^*_r \leq v^*$. The monotonicity of v^*_r is proved as follows. Let $\bar{\boldsymbol{y}} \in \mathbb{R}^{s(2(r+1))}$ be a feasible solution of SDP (2.6) with $r = r+1$. Then $\boldsymbol{M}_{r_j+1}(\bar{\boldsymbol{y}}) \succeq \boldsymbol{O}$ $(j=1,\dots,m)$ and $\boldsymbol{M}_{r+1}(\bar{\boldsymbol{y}})$ hold from the feasibility. Let $\tilde{\boldsymbol{y}}$ denote the subvector of $\bar{\boldsymbol{y}}$ consisting of the elements $\bar{y}_\alpha$ with indices α restricted to the members of G_{2r}. Then $\boldsymbol{M}_{r_j}(\tilde{\boldsymbol{y}}) \succeq \boldsymbol{O}$ $(j=1,\dots,m)$ and $\boldsymbol{M}_r(\tilde{\boldsymbol{y}}) \succeq \boldsymbol{O}$ because $\boldsymbol{M}_{r_j}(\tilde{\boldsymbol{y}})$ $(j=1,\dots,m)$ and $\boldsymbol{M}_r(\tilde{\boldsymbol{y}})$ are leading principal submatrices of $\boldsymbol{M}_{r_j+1}(\tilde{\boldsymbol{y}})$ $(j=1,\dots,m)$ and $\boldsymbol{M}_{r+1}(\tilde{\boldsymbol{y}})$, respectively. Hence $\tilde{\boldsymbol{y}}$ is a feasible solution of SDP (2.6) with r. We also see that $\boldsymbol{c}^T_{2(r+1)}\bar{\boldsymbol{y}} = \boldsymbol{c}^T_{2r}\tilde{\boldsymbol{y}} = \sum_{\alpha\in G_{2\bar{r}}}(c_{2r})_\alpha \tilde{y}_\alpha$ because $(c_{2r})_\alpha = 0$ and $(c_{2(r+1)})_\alpha = 0$ for any $\alpha \in G_{2r} \setminus G_{2\bar{r}}$. As a result, we have $v^*_r \leq v^*_{r+1}$.

In [11], Lasserre showed the convergence of v^*_r $(r \geq \bar{r})$ to the optimal value v^* of POP (2.1) as $r \to \infty$ under a certain moderate condition (see Theorem 4.2 of [11]). He also demonstrated that the optimal value v^*_r of SDP (2.6) attains the optimal value v^* of POP (2.1) for a finite r, which is not much larger than $\bar{r}$, in all test problems reported there, and suggested that the finite convergence of v^*_r $(r \geq \bar{r})$ to v^* is expected in many practical problems. The following sufficient condition for the finite convergence, which we call *the rank condition*, was proved in [3, 13].

Proposition 2.3. *Let $\boldsymbol{y}^*$ be an optimal solution of SDP (2.6) and $d = \max_{j=1,\dots,m}\lceil \deg(f_j)/2 \rceil$. If rank $\boldsymbol{M}_r(\boldsymbol{y}^*)$ = rank $\boldsymbol{M}_{r-d}(\boldsymbol{y}^*)$, then v^*_r is equal to the optimal value v^* of POP (2.1).*

To check whether the optimal value of SDP (2.6) attains the optimal value of POP (2.1) or not, this condition was used in the software package GloptiPoly [2].

The dual problem of SDP (2.6) turns out to be

$$\left.\begin{array}{ll} \text{maximize} & p \\ \text{subject to} & \langle \boldsymbol{X}, \boldsymbol{E}_0\rangle + \sum_{j=1}^m \langle \boldsymbol{Y}_j, \boldsymbol{B}_{j,0}\rangle = (c_{2r})_0 - p, \\ & \langle \boldsymbol{X}, \boldsymbol{E}_\alpha\rangle + \sum_{j=1}^m \langle \boldsymbol{Y}_j, \boldsymbol{B}_{j,\alpha}\rangle = (c_{2r})_\alpha \quad (\alpha \in G_{2r}\setminus\{0\}), \\ & \boldsymbol{X}, \boldsymbol{Y}_j \succeq \boldsymbol{O} \quad (j=1,\dots,m), \end{array}\right\} \tag{2.8}$$

where $\langle \boldsymbol{A}, \boldsymbol{B}\rangle$ denotes the matrix inner product $\sum_k \sum_\ell A_{k\ell}B_{k\ell}$ for symmetric matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, and the size of the matrix variables $\boldsymbol{X}$ and $\boldsymbol{Y}_j$ are $s(r) \times s(r)$ and $s(r_j) \times s(r_j)$ $(j=1,\dots,m)$, respectively. We are also concerned with a sum of squares (SOS) problem induced from POP (2.1) (Lasserre [11])

$$\left.\begin{array}{ll} \text{maximize} & p \\ \text{subject to} & f_0(\boldsymbol{x}) - p = \boldsymbol{u}_r(\boldsymbol{x})^T \boldsymbol{X} \boldsymbol{u}_r(\boldsymbol{x}) + \sum_{j=1}^m f_j(\boldsymbol{x}) \boldsymbol{u}_{r_j}(\boldsymbol{x})^T \boldsymbol{Y}_j \boldsymbol{u}_{r_j}(\boldsymbol{x}) \quad (\forall \boldsymbol{x} \in \mathbb{R}^n), \\ & \boldsymbol{X}, \boldsymbol{Y}_j \succeq \boldsymbol{O} \quad (j=1,\dots,m). \end{array}\right\} \tag{2.9}$$

The equality condition of the problem (2.9) is the identity on $\boldsymbol{x}$. We can verify that the dual SDP (2.8) is equivalent with the problem (2.9). In fact, comparing coefficients of each monomial on the both sides of the identity, we obtain the equality constraints in SDP (2.8). See Lasserre [11] for more details.

3 Main Results

In this section, we first introduce a POP transformed from (2.1) by an affine transformation $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{w} + \boldsymbol{b}$, where $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ is a nonsingular matrix and $\boldsymbol{b} \in \mathbb{R}^n$.

$$\text{minimize } \tilde{f}_0(\boldsymbol{w}) \text{ subject to } \tilde{f}_j(\boldsymbol{w}) \geq 0 \quad (j = 1,\dots,m), \tag{3.1}$$

where $\tilde{f}_0(\boldsymbol{w}) = f_0(\boldsymbol{A}\boldsymbol{w}+\boldsymbol{b})$ and $\tilde{f}_j(\boldsymbol{w}) = f_j(\boldsymbol{A}\boldsymbol{w}+\boldsymbol{b})$. We derive Lasserre's SDP relaxation problem (3.2) for the transformed POP (3.1) and its dual (3.4) as we have done for POP (2.1), and we then describe details of the isomorphic relations illustrated in Figure 1.

By the definition of $\tilde{f}_j$, $\deg(f_j) = \deg(\tilde{f}_j)$ holds for all $j = 0, 1, \dots, m$. Thus, We can construct a sequence of SDP relaxation problems from POP (3.1) for all $r \in \mathbb{Z}_+$ satisfying $r \geq \bar{r}$. Moreover, the obtained SDP (3.2) has $s(r)\times s(r)$ and $s(r_j)\times s(r_j)$ coefficient matrices for all $r \geq \bar{r}$ as in (2.6).

To generate the SDP relaxation problem from POP (3.1), we use the monomial vector

$$\boldsymbol{u}_r(\boldsymbol{w}) = (\boldsymbol{w}^0, w_1, \dots, w_n, w_1^2, w_1w_2, \dots, w_n^2, \dots, w_1^r, \dots, w_n^r)^T$$

where $\boldsymbol{w}^0 = 1$ for any $\boldsymbol{w} \in \mathbb{R}^n$, and represent the matrix $\boldsymbol{u}_r(\boldsymbol{w})\boldsymbol{u}_r(\boldsymbol{w})^T$ in $\boldsymbol{w}$ as

$$\boldsymbol{u}_r(\boldsymbol{w})\boldsymbol{u}_r(\boldsymbol{w})^T = \sum_{\alpha\in G_{2r}} \boldsymbol{w}^\alpha \boldsymbol{E}_\alpha,$$

where $\boldsymbol{E}_\alpha$ is given by (2.2). By applying the discussion of Section 2 into POP(3.1), we obtain the following SDP relaxation problem of POP (3.1):

$$\left.\begin{array}{ll} \text{minimize} & \tilde{\boldsymbol{c}}_{2r}^T\boldsymbol{z} \\ \text{subject to} & \boldsymbol{M}_{r_j}(\tilde{f}_j\boldsymbol{z}) \succeq \boldsymbol{O}\ (j=1,\dots,m), \boldsymbol{M}_r(\boldsymbol{z}) \succeq \boldsymbol{O}, z_0 = 1, \end{array}\right\} \tag{3.2}$$

where $\tilde{\boldsymbol{c}}_{2r} \in \mathbb{R}^{s(2r)}$ is the column vector such that $\tilde{f}_0(\boldsymbol{w}) = \tilde{\boldsymbol{c}}_{2r}^T\boldsymbol{u}_{2r}(\boldsymbol{w})$ for all $\boldsymbol{w} \in \mathbb{R}^n$ and

$$\boldsymbol{M}_{r_j}(\tilde{f}_j\boldsymbol{z}) = \sum_{\alpha\in G_{2r}} z_\alpha \tilde{\boldsymbol{B}}_{j,\alpha}\ (j=1,\dots,m),\ \boldsymbol{M}_r(\boldsymbol{z}) = \sum_{\alpha\in G_{2r}} z_\alpha \boldsymbol{E}_\alpha \tag{3.3}$$

for some $s(r_j) \times s(r_j)$ real symmetric matrices $\tilde{\boldsymbol{B}}_{j,\alpha}$ $(j = 1,\dots,m; \alpha \in G_{2r})$. The size of the variable $\boldsymbol{z}$ is $s(2r)$. Note that the (β,γ)th element of the matrix $\boldsymbol{M}_{r_j}(\tilde{f}_j\boldsymbol{z})$ is $\sum_{\alpha\in G_{\deg(f_j)}} \tilde{f}_{j,\alpha} z_{\alpha+\beta+\gamma}$ for $\beta,\gamma \in G_{r_j}$, where $\tilde{f}_{j,\alpha}$ is the coefficient of the monomial $\boldsymbol{w}^\alpha$ of $\tilde{f}_j$.

The dual problem of SDP (3.2) is

$$\left.\begin{array}{ll} \text{maximize} & q \\ \text{subject to} & \langle \boldsymbol{W}, \boldsymbol{E}_0\rangle + \sum_{j=1}^m \langle \boldsymbol{Z}_j, \tilde{\boldsymbol{B}}_{j,0}\rangle = (\tilde{c}_{2r})_0 - q, \\ & \langle \boldsymbol{W}, \boldsymbol{E}_\alpha\rangle + \sum_{j=1}^m \langle \boldsymbol{Z}_j, \tilde{\boldsymbol{B}}_{j,\alpha}\rangle = (\tilde{c}_{2r})_\alpha \quad (\alpha \in G_{2r}\setminus\{0\}), \\ & \boldsymbol{W}, \boldsymbol{Z}_j \succeq \boldsymbol{O} \quad (j=1,\dots,m), \end{array}\right\} \tag{3.4}$$

where $\boldsymbol{W} \in \mathbb{R}^{s(r)\times s(r)}$ and $\boldsymbol{Z}_j \in \mathbb{R}^{s(r_j)\times s(r_j)}$. Note that SDP (3.4) is also equivalent with the SOS problem induced from POP (3.1):

$$\left.\begin{array}{ll} \text{maximize} & q \\ \text{subject to} & \tilde{f}_0(\boldsymbol{w}) - q = \boldsymbol{u}_r(\boldsymbol{w})^T \boldsymbol{W} \boldsymbol{u}_r(\boldsymbol{w}) + \sum_{j=1}^m \tilde{f}_j(\boldsymbol{w}) \boldsymbol{u}_{r_j}(\boldsymbol{w})^T \boldsymbol{Z}_j \boldsymbol{u}_{r_j}(\boldsymbol{w}) \ (\forall \boldsymbol{w} \in \mathbb{R}^n), \\ & \boldsymbol{W}, \boldsymbol{Z}_j \succeq \boldsymbol{O} \ (j = 1, \dots, m). \end{array}\right\} \tag{3.5}$$

Recall that a similar equivalent relation between SDP (2.8) and (2.9) was observed at the end of Section 2.

The following theorems are the main results of this paper. Their proofs will be given in Section 4.

Theorem 3.1. *There exists an $s(2r) \times s(2r)$ nonsingular matrix $\boldsymbol{P}_{s(2r)}$ satisfying the following properties.*

1. $(p, \boldsymbol{X}, \{\boldsymbol{Y}_j\}_{j=1}^m)$ *is a feasible (optimal) solution for SDP (2.8) if and only if*

$$(q, \boldsymbol{W}, \{\boldsymbol{Z}_j\}_{j=1}^m) = (p, \boldsymbol{P}_{s(r)}^T \boldsymbol{X} \boldsymbol{P}_{s(r)}, \{\boldsymbol{P}_{s(r_j)}^T \boldsymbol{Y}_j \boldsymbol{P}_{s(r_j)}\}_{j=1}^m),$$

is a feasible (optimal) solution for SDP (3.4), where $\boldsymbol{P}_{s(r)}$ and $\boldsymbol{P}_{s(r_j)}$ are the $s(r)\times s(r)$ and $s(r_j) \times s(r_j)$ leading principal matrices of $\boldsymbol{P}_{s(2r)}$.

2. $\boldsymbol{y}$ *is a feasible (optimal) solution for SDP (2.6) with an objective value $\boldsymbol{c}_{2r}^T \boldsymbol{y}$ if and only if $\boldsymbol{z} = \boldsymbol{P}_{s(2r)}^{-1} \boldsymbol{y}$ is a feasible (optimal) solution for SDP (3.2) with the same objective value $\tilde{\boldsymbol{c}}_{2r}^T \boldsymbol{z}$.*

3. *We have*

$$\begin{aligned} \tilde{\boldsymbol{c}}_{2r} &= \boldsymbol{P}_{s(2r)}^T \boldsymbol{c}_{2r}, \\ \tilde{\boldsymbol{B}}_{j,\alpha} &= \boldsymbol{P}_{s(r_j)}^{-1} \left(\sum_{\beta \in G_{2r}} \left(P_{s(2r)}\right)_{\beta,\alpha} \boldsymbol{B}_{j,\beta} \right) \boldsymbol{P}_{s(r_j)}^{-T}, \\ \boldsymbol{E}_\alpha &= \boldsymbol{P}_{s(r)}^{-1} \left(\sum_{\beta \in G_{2r}} \left(P_{s(2r)}\right)_{\beta,\alpha} \boldsymbol{E}_\beta \right) \boldsymbol{P}_{s(r)}^{-T}. \end{aligned}$$

Theorem 3.2. *Let $\boldsymbol{y}^*$ be a feasible solution of SDP (2.6), and let $\boldsymbol{z}^* = \boldsymbol{P}_{s(2r)}^{-1} \boldsymbol{y}^*$. If $\boldsymbol{y}^*$ satisfies the rank condition rank $\boldsymbol{M}_r(\boldsymbol{y}^*) =$ rank $\boldsymbol{M}_{r-d}(\boldsymbol{y}^*)$, then $\boldsymbol{z}^*$ satisfies the rank condition rank$\boldsymbol{M}_r(\boldsymbol{z}^*) =$ rank$\boldsymbol{M}_{r-d}(\boldsymbol{z}^*)$ with the same r.*

The POPs and SDPs which we deal with and the invariant relations which we establish are summarized in Figure 1.

4 Proofs

4.1 Basic lemmas

In this subsection, we construct matrices $\boldsymbol{P}_{s(2r)}$, $\boldsymbol{P}_{s(r)}$ and $\boldsymbol{P}_{s(r_j)}$ $(j = 1, \dots, m)$ involved in Theorem 3.1 from the affine transformation $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{w} + \boldsymbol{b}$, and show some basic properties on these matrices.

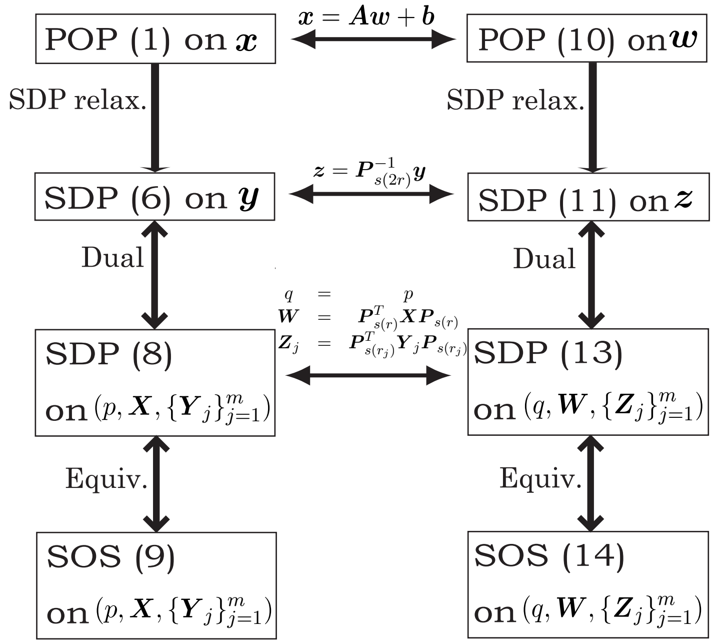

Figure 1: Invariance of Lasserre's SDP relaxation under an affine transformation

Lemma 4.1. *There exists a sequence of nonsingular matrices* $\boldsymbol{P}_{s(k)} \in \mathbb{R}^{s(k)\times s(k)}$ $(k \in \mathbb{Z}_+)$ *satisfying the following properties*:

1. $\boldsymbol{u}_k(\boldsymbol{x}) = \boldsymbol{P}_{s(k)}\boldsymbol{u}_k(\boldsymbol{w})$ *for every* $\boldsymbol{x}$ *and* $\boldsymbol{w}$ *such that* $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{w} + \boldsymbol{b}$.

2. *Let* $\ell < k$. *There exist matrices* $\boldsymbol{R} \in \mathbb{R}^{(s(k)-s(\ell))\times s(\ell)}$ *and* $\boldsymbol{S} \in \mathbb{R}^{(s(k)-s(\ell))\times(s(k)-s(\ell))}$ *such that*
$$\boldsymbol{P}_{s(k)} = \begin{pmatrix} \boldsymbol{P}_{s(\ell)} & \boldsymbol{O} \\ \boldsymbol{R} & \boldsymbol{S} \end{pmatrix}.$$

Proof. For every $k \in \mathbb{Z}_+$ and $\alpha \in G_k$, substituting $\boldsymbol{A}\boldsymbol{w} + \boldsymbol{b}$ for $\boldsymbol{x}$, we can represent the monomial $\boldsymbol{x}^\alpha$ as a polynomial in $\boldsymbol{w}$:

$$\boldsymbol{x}^\alpha = (\boldsymbol{A}\boldsymbol{w} + \boldsymbol{b})^\alpha = \prod_{i=1}^{n} (\boldsymbol{A}\boldsymbol{w} + \boldsymbol{b})_i^{\alpha_i} = \sum_{\beta \in G_k} P_{\alpha,\beta}\boldsymbol{w}^\beta$$

for some $P_{\alpha,\beta}$ $(\beta \in G_k)$. Defining $\boldsymbol{P}_{s(k)}$ to be an $s(k)\times s(k)$ matrix whose (α, β)th component is $P_{\alpha,\beta}$ for every $\alpha, \beta \in G_k$, we see that $\boldsymbol{u}_k(\boldsymbol{x}) = \boldsymbol{u}_k(\boldsymbol{A}\boldsymbol{w} + \boldsymbol{b}) = \boldsymbol{P}_{s(k)}\boldsymbol{u}_k(\boldsymbol{w})$. By a similar argument applied to the inverse affine transformations $\boldsymbol{w} = \boldsymbol{A}^{-1}\boldsymbol{x} - \boldsymbol{A}^{-1}\boldsymbol{b}$, there exists a nonsingular matrix $\boldsymbol{Q}_{s(k)} \in \mathbb{R}^{s(k)\times s(k)}$ such that $\boldsymbol{u}_k(\boldsymbol{w}) = \boldsymbol{u}_k(\boldsymbol{A}^{-1}\boldsymbol{x} - \boldsymbol{A}^{-1}\boldsymbol{b}) = \boldsymbol{Q}_{s(k)}\boldsymbol{u}_k(\boldsymbol{x})$. It follows from $\boldsymbol{u}_k(\boldsymbol{x}) = \boldsymbol{P}_{s(k)}\boldsymbol{Q}_{s(k)}\boldsymbol{u}_k(\boldsymbol{x})$ for every $\boldsymbol{x} \in \mathbb{R}^n$ that $\boldsymbol{P}_{s(k)}\boldsymbol{Q}_{s(k)} = \boldsymbol{I}$. We see that $\boldsymbol{P}_{s(k)}$ is nonsingular.

It remains to prove property 2 on $\boldsymbol{P}_{s(k)}$. We can partition $\boldsymbol{u}_k(\boldsymbol{x}) = (\boldsymbol{u}_\ell(\boldsymbol{x})^T, \boldsymbol{v}_\ell(\boldsymbol{x})^T)^T$ and $\boldsymbol{u}_k(\boldsymbol{w}) = (\boldsymbol{u}_\ell(\boldsymbol{w})^T, \boldsymbol{v}_\ell(\boldsymbol{w})^T)^T$, where $\boldsymbol{v}_\ell(\boldsymbol{x})$ and $\boldsymbol{v}_\ell(\boldsymbol{w})$ are column vectors of all monomials $\boldsymbol{x}^\alpha$ and $\boldsymbol{w}^\alpha$ for every $\alpha \in G_k \setminus G_\ell$. Let us write:

$$\boldsymbol{u}_k(\boldsymbol{x}) = \begin{pmatrix} \boldsymbol{u}_\ell(\boldsymbol{x}) \\ \boldsymbol{v}_\ell(\boldsymbol{x}) \end{pmatrix} = \begin{pmatrix} \boldsymbol{P}' & \boldsymbol{Q}' \\ \boldsymbol{R}' & \boldsymbol{S}' \end{pmatrix} \begin{pmatrix} \boldsymbol{u}_\ell(\boldsymbol{w}) \\ \boldsymbol{v}_\ell(\boldsymbol{w}) \end{pmatrix} = \boldsymbol{P}_{s(k)} \boldsymbol{u}_k(\boldsymbol{w}),$$

where $\boldsymbol{P}' \in \mathbb{R}^{s(\ell)\times s(\ell)}$, $\boldsymbol{Q}' \in \mathbb{R}^{s(\ell)\times(s(k)-s(\ell))}$, $\boldsymbol{R}' \in \mathbb{R}^{s(\ell)\times(s(k)-s(\ell))}$ and $\boldsymbol{S}' \in \mathbb{R}^{(s(k)-s(\ell))\times(s(k)-s(\ell))}$, respectively. It follows from this relation that $\boldsymbol{u}_\ell(\boldsymbol{x}) = \boldsymbol{P}'\boldsymbol{u}_\ell(\boldsymbol{w}) + \boldsymbol{Q}'\boldsymbol{v}_\ell(\boldsymbol{w})$. Because $\boldsymbol{u}_\ell(\boldsymbol{x}) = \boldsymbol{P}_{s(\ell)}\boldsymbol{u}_\ell(\boldsymbol{w})$ for all $\boldsymbol{x}, \boldsymbol{w}$ satisfying $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{w} + \boldsymbol{b}$, we obtain the identity on $\boldsymbol{w}$:

$$\boldsymbol{P}'\boldsymbol{u}_\ell(\boldsymbol{w}) + \boldsymbol{Q}'\boldsymbol{v}_\ell(\boldsymbol{w}) = \boldsymbol{P}_{s(\ell)}\boldsymbol{u}_\ell(\boldsymbol{w}) \text{ for all } \boldsymbol{w} \in \mathbb{R}^{\cdot}$$

Comparing the coefficients of each monomial $\boldsymbol{w}^\alpha$ on the both sides of this identity, we have $\boldsymbol{P}' = \boldsymbol{P}_{s(\ell)}$ and $\boldsymbol{Q}' = \boldsymbol{O}$.

Example 4.2. We consider the following affine transformation:

$$x_1 = \frac{w_1 + 1}{2} \text{ and } x_2 = \frac{w_2 + 1}{2}.$$

In this case,

$$\boldsymbol{A} = \begin{pmatrix} 1/2 & \\ & 1/2 \end{pmatrix} \text{ and } \boldsymbol{b} = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$$

Under this affine transformation, $\boldsymbol{u}_2(\boldsymbol{x}) = \boldsymbol{P}_{s(2)}\boldsymbol{u}_2(\boldsymbol{w})$ turns out to be

$$\boldsymbol{u}_2(\boldsymbol{x}) = \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \end{pmatrix} = \left(\begin{array}{c|cc|ccc} 1 & & & & & \\ \hline 1/2 & 1/2 & & & & \\ 1/2 & & 1/2 & & & \\ \hline 1/4 & 1/2 & & 1/4 & & \\ 1/4 & 1/4 & 1/4 & & 1/4 & \\ 1/4 & & 1/2 & & & 1/4 \end{array}\right) \begin{pmatrix} 1 \\ w_1 \\ w_2 \\ w_1^2 \\ w_1w_2 \\ w_2^2 \end{pmatrix} = \boldsymbol{P}_{s(2)}\boldsymbol{u}_2(\boldsymbol{w}).$$

Hence, $\boldsymbol{P}_{s(0)}$ and $\boldsymbol{P}_{s(1)}$ are

$$\boldsymbol{P}_{s(0)} = (1) \text{ and } \boldsymbol{P}_{s(1)} = \left(\begin{array}{c|cc} 1 & & \\ \hline 1/2 & 1/2 & \\ 1/2 & & 1/2 \end{array}\right).$$

We can see that $\boldsymbol{P}_{s(0)}$, $\boldsymbol{P}_{s(1)}$ and $\boldsymbol{P}_{s(2)}$ have all properties in Lemma 4.1.

Lemma 4.3. *Let $f \in \mathbb{R}[\boldsymbol{x}]$ and $k \geq \deg(f)$. Define a polynomial $\tilde{f} \in \mathbb{R}[\boldsymbol{w}]$ by $\tilde{f}(\boldsymbol{w}) = f(\boldsymbol{A}\boldsymbol{w} + \boldsymbol{b})$. Represent $f \in \mathbb{R}[\boldsymbol{x}]$ such that $f(\boldsymbol{x}) = \boldsymbol{f}^T\boldsymbol{u}_k(\boldsymbol{x})$ for some $\boldsymbol{f} \in \mathbb{R}^{s(k)}$ and $\tilde{f} \in \mathbb{R}[\boldsymbol{w}]$ such that $\tilde{f}(\boldsymbol{w}) = \tilde{\boldsymbol{f}}^T\boldsymbol{u}_k(\boldsymbol{w})$ for some $\tilde{\boldsymbol{f}} \in \mathbb{R}^{s(k)}$. Then $\tilde{\boldsymbol{f}} = \boldsymbol{P}_{s(k)}^T\boldsymbol{f}$.*

Proof. By definition, we see that $\tilde{\boldsymbol{f}}^T\boldsymbol{u}_k(\boldsymbol{w}) = \tilde{f}(\boldsymbol{w}) = f(\boldsymbol{x}) = \boldsymbol{f}^T\boldsymbol{u}_k(\boldsymbol{x})$ if $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{w} + \boldsymbol{b}$. By property 1 of Lemma 4.1, we know that $\boldsymbol{u}_k(\boldsymbol{x}) = \boldsymbol{P}_{s(k)}\boldsymbol{u}_k(\boldsymbol{w})$ if $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{w} + \boldsymbol{b}$. Hence $\tilde{\boldsymbol{f}}^T\boldsymbol{u}_k(\boldsymbol{w}) = (\boldsymbol{P}_{s(k)}^T\boldsymbol{f})^T\boldsymbol{u}_k(\boldsymbol{w})$ for all $\boldsymbol{w} \in \mathbb{R}^n$. Comparing the coefficients of all monomials $\boldsymbol{w}^\alpha$ on the both sides of this identity, we obtain the desired result.

4.2 Proof of Property 1 of Theorem 3.1

We only prove the "only if" part of property 1 of Theorem 3.1 because we can prove the "if" part similarly. Since SDP (2.8) is equivalent to SOS problem (2.9), any feasible solution $(p, \boldsymbol{X}, \{\boldsymbol{Y}_j\}_{j=1}^m)$ of SDP (2.8) satisfies the following identity on $\boldsymbol{x}$:

$$f_0(\boldsymbol{x}) - p = \boldsymbol{u}_r(\boldsymbol{x})^T \boldsymbol{X} \boldsymbol{u}_r(\boldsymbol{x}) + \sum_{j=1}^m f_j(\boldsymbol{x}) \boldsymbol{u}_{r_j}(\boldsymbol{x})^T \boldsymbol{Y}_j \boldsymbol{u}_{r_j}(\boldsymbol{x}) \quad \text{for every } \boldsymbol{x} \in \mathbb{R}^n.$$

By substituting $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{w} + \boldsymbol{b}$ into the both side of the identity above and by applying property 1 of Lemma 4.1, we obtain the following identity on $\boldsymbol{w}$:

$$\begin{aligned} &\tilde{f}_0(\boldsymbol{w}) - p \\ &= \boldsymbol{u}_r(\boldsymbol{w})^T \boldsymbol{P}_{s(r)}^T \boldsymbol{X} \boldsymbol{P}_{s(r)} \boldsymbol{u}_r(\boldsymbol{w}) + \sum_{j=1}^m \tilde{f}_j(\boldsymbol{w}) \boldsymbol{u}_{r_j}(\boldsymbol{w})^T \boldsymbol{P}_{s(r_j)}^T \boldsymbol{Y}_j \boldsymbol{P}_{s(r_j)} \boldsymbol{u}_{r_j}(\boldsymbol{w}) \\ &\quad \text{for every } \boldsymbol{w} \in \mathbb{R}^n. \end{aligned}$$

Note that $\boldsymbol{P}_{s(r)}^T \boldsymbol{X} \boldsymbol{P}_{s(r)}$ and $\boldsymbol{P}_{s(r_j)}^T \boldsymbol{Y}_j \boldsymbol{P}_{s(r_j)}$ $(j = 1, \ldots, m)$ are positive semidefinite matrices. These facts imply that $(q, \boldsymbol{W}, \{\boldsymbol{Z}_j\}_{j=1}^m) = (p, \boldsymbol{P}_{s(r)}^T \boldsymbol{X} \boldsymbol{P}_{s(r)}, \{\boldsymbol{P}_{s(r_j)}^T \boldsymbol{Y}_j \boldsymbol{P}_{s(r_j)}\}_{j=1}^m)$, is a feasible solution of SOS problem (3.5) induced from POP (3.1). Hence $(q, \boldsymbol{W}, \{\boldsymbol{Z}_j\}_{j=1}^m)$ is a feasible solution for SDP (3.4) because SDP (3.4) is equivalent to SOS problem (3.5).

4.3 Proof of Property 2 of Theorem 3.1

To prove property 2 of Theorem 3.1, we will use two lemmas below. Throughout this subsection, we assume that $r \geq \bar{r}$ is fixed, and we denote the (α, β)th element of $\boldsymbol{P}_{s(2r)}$ by $P_{\alpha,\beta}$ for simplicity of notation.

Lemma 4.4. *Let $h, k \in \mathbb{Z}_+$ and $\ell \in \mathbb{Z}_+$ satisfy $2r \geq h + k + \ell$. Then*

$$P_{\alpha+\beta+\gamma,\delta} = \sum_{\substack{\delta=\delta_1+\delta_2+\delta_3, \\ \delta_1 \in G_h, \delta_2 \in G_k, \delta_3 \in G_\ell}} P_{\alpha,\delta_1} P_{\beta,\delta_2} P_{\gamma,\delta_3} \text{ for every } \alpha \in G_h,\ \beta \in G_k,\ \gamma \in G_\ell \text{ and } \delta \in G_{2r}.$$

Proof. Let $\alpha \in G_h$, $\beta \in G_k$ and $\gamma \in G_\ell$ be fixed. Property 2 of Lemma 4.1 gives

$$\begin{aligned} &P_{\alpha,\delta} = 0 \quad (\forall \delta \in G_{2r} \setminus G_h),\ P_{\beta,\delta} = 0 \quad (\forall \delta \in G_{2r} \setminus G_k), \\ &P_{\gamma,\delta} = 0 \quad (\forall \delta \in G_{2r} \setminus G_\ell),\ P_{\alpha+\beta+\gamma,\delta} = 0 \quad (\forall \delta \in G_{2r} \setminus G_{h+k+\ell}). \end{aligned}$$

By property 1 of Lemma 4.1, we see that

$$\begin{aligned} \boldsymbol{x}^\alpha &= \sum_{\gamma \in G_{2r}} P_{\alpha,\delta} \boldsymbol{w}^\delta = \sum_{\delta \in G_h} P_{\alpha,\delta} \boldsymbol{w}^\delta, \quad \boldsymbol{x}^\beta = \sum_{\delta \in G_{2r}} P_{\beta,\delta} \boldsymbol{w}^\delta = \sum_{\delta \in G_k} P_{\beta,\delta} \boldsymbol{w}^\delta, \\ \boldsymbol{x}^\gamma &= \sum_{\delta \in G_{2r}} P_{\gamma,\delta} \boldsymbol{w}^\delta = \sum_{\delta \in G_\ell} P_{\gamma,\delta} \boldsymbol{w}^\delta, \quad \boldsymbol{x}^{\alpha+\beta+\gamma} = \sum_{\delta \in G_{2r}} P_{\alpha+\beta+\gamma,\delta} \boldsymbol{w}^\delta = \sum_{\delta \in G_{h+k+\ell}} P_{\alpha+\beta+\gamma,\delta} \boldsymbol{w}^\delta. \end{aligned}$$

It follows from these relations that

$$\begin{aligned}\sum_{\delta \in G_{h+k+\ell}} P_{\alpha+\beta+\gamma,\delta} \boldsymbol{w}^{\delta} &= \boldsymbol{x}^{\alpha+\beta+\gamma} = \boldsymbol{x}^{\alpha}\boldsymbol{x}^{\beta}\boldsymbol{x}^{\gamma} \\ &= \left(\sum_{\delta_1 \in G_h} P_{\alpha,\delta_1}\boldsymbol{w}^{\delta_1}\right)\left(\sum_{\delta_2 \in G_k} P_{\beta,\delta_2}\boldsymbol{w}^{\delta_2}\right)\left(\sum_{\delta_3 \in G_\ell} P_{\gamma,\delta_3}\boldsymbol{w}^{\delta_3}\right) \\ &= \sum_{\delta \in G_{h+k+\ell}} \left(\sum_{\substack{\delta=\delta_1+\delta_2+\delta_3, \\ \delta_1 \in G_h, \delta_2 \in G_k, \delta_3 \in G_\ell}} P_{\alpha,\delta_1}P_{\beta,\delta_2}P_{\gamma,\delta_3} \right) \boldsymbol{w}^{\delta}.\end{aligned}$$

Comparing the coefficients of each monomial $\boldsymbol{w}^{\delta}$, we obtain the desired result.

Lemma 4.5. *Assume that* $f \in \mathbb{R}[\boldsymbol{x}]$, $r \geq \lceil \deg(f)/2 \rceil$ *and* $\hat{\boldsymbol{z}} \in \mathbb{R}^{s(2r)}$. *Let* $\hat{\boldsymbol{y}} = \boldsymbol{P}_{s(2r)}\hat{\boldsymbol{z}}$, $r' = r - \lceil \deg(f)/2 \rceil$ *and* $\tilde{f}(\boldsymbol{w}) = f(\boldsymbol{A}\boldsymbol{w} + \boldsymbol{b})$. *Then* $\boldsymbol{P}_{s(r')}\boldsymbol{M}_{r'}(\tilde{f}\hat{\boldsymbol{z}})\boldsymbol{P}^T_{s(r')} = \boldsymbol{M}_{r'}(f\hat{\boldsymbol{y}})$ *holds.*

Proof. Because the size of $\boldsymbol{P}_{s(r')}\boldsymbol{M}_{r'}(\tilde{f}\hat{\boldsymbol{z}})\boldsymbol{P}^T_{s(r')}$ is the same as that of $\boldsymbol{M}_{r'}(f\hat{\boldsymbol{y}})$, it is enough to show that the (α,β)th element $\tilde{m}_{\alpha,\beta}$ of $\boldsymbol{P}_{s(r')}\boldsymbol{M}_{r'}(\tilde{f}\hat{\boldsymbol{z}})\boldsymbol{P}^T_{s(r')}$ is equal to the (α,β)th element $m_{\alpha,\beta}$ of $\boldsymbol{M}_{r'}(f\hat{\boldsymbol{y}})$ for all $\alpha, \beta \in G_{r'}$. Substituting $\hat{\boldsymbol{y}} = \boldsymbol{P}_{s(2r)}\hat{\boldsymbol{z}}$ into $\boldsymbol{M}_{r'}(f\hat{\boldsymbol{y}})$, and using Lemma 4.1, we see

$$\begin{aligned} m_{\alpha,\beta} &= \sum_{\gamma \in G_{\deg(f)}} f_\gamma \hat{y}_{\alpha+\beta+\gamma} = \sum_{\gamma \in G_{\deg(f)}} f_\gamma \left(\sum_{\delta \in G_{2r}} P_{\alpha+\beta+\gamma,\delta}\hat{z}_\delta \right) \\ &= \sum_{\gamma \in G_{\deg(f)}} f_\gamma \left(\sum_{\delta \in G_{\deg(f)+2r'}} P_{\alpha+\beta+\gamma,\delta}\hat{z}_\delta \right), \end{aligned}$$

where f_γ is the coefficient in f with respect to $\boldsymbol{x}^\gamma$. On the other hand, we obtain by the definition of $\tilde{m}_{\alpha,\beta}$ and $\deg(f) = \deg(\tilde{f})$ that

$$\tilde{m}_{\alpha,\beta} = \sum_{\delta_1 \in G_{r'}} \sum_{\delta_2 \in G_{r'}} P_{\alpha,\delta_1} \left(\sum_{\gamma' \in G_{\deg(f)}} \tilde{f}_{\gamma'} \hat{z}_{\gamma'+\delta_1+\delta_2} \right) P_{\beta,\delta_2},$$

where $\tilde{f}_{\gamma'}$ is the coefficient in $\tilde{f}$ with respect to $\boldsymbol{w}^{\gamma'}$. We also see from Lemma 4.3 and

$\deg(f) = \deg(\tilde{f})$ that $\tilde{f}_{\gamma'} = \sum_{\gamma \in G_{\deg(f)}} f_\gamma P_{\gamma,\gamma'}$ for all $\gamma' \in G_{\deg(f)}$. Now we obtain:

$$\begin{aligned}
\tilde{m}_{\alpha,\beta} &= \sum_{\delta_1 \in G_{r'}} \sum_{\delta_2 \in G_{r'}} P_{\alpha,\delta_1} \left(\sum_{\gamma' \in G_{\deg(f)}} \left(\sum_{\gamma \in G_{\deg(f)}} f_\gamma P_{\gamma,\gamma'} \right) \hat{z}_{\gamma'+\delta_1+\delta_2} \right) P_{\beta,\delta_2} \\
&= \sum_{\gamma \in G_{\deg(f)}} f_\gamma \left(\sum_{\delta_1 \in G_{r'}} \sum_{\delta_2 \in G_{r'}} \sum_{\gamma' \in G_{\deg(f)}} P_{\alpha,\delta_1} P_{\beta,\delta_2} P_{\gamma,\gamma'} \hat{z}_{\gamma'+\delta_1+\delta_2} \right) \\
&= \sum_{\gamma \in G_{\deg(f)}} f_\gamma \sum_{\delta \in G_{\deg(f)+2r'}} \left(\sum_{\substack{\delta=\delta_1+\delta_2+\gamma', \\ \delta_1,\delta_2 \in G_{r'}, \gamma' \in G_{\deg(f)}}} P_{\alpha,\delta_1} P_{\beta,\delta_2} P_{\gamma,\gamma'} \right) \hat{z}_\delta \\
&= \sum_{\gamma \in G_{\deg(f)}} f_\gamma \left(\sum_{\delta \in G_{\deg(f)+2r'}} P_{\alpha+\beta+\gamma,\delta} \hat{z}_\delta \right) \quad \text{(by Lemma 4.4)} \\
&= m_{\alpha,\beta}.
\end{aligned}$$

This completes the proof.

Now we are ready to prove property 2 of Theorem 3.1. We only prove the "only if" part since we can prove the "if" part similarly. Letting $f = 1$ in Lemma 4.5, we obtain $\boldsymbol{P}_{s(r)} \boldsymbol{M}_r(\boldsymbol{z}) \boldsymbol{P}_{s(r)}^T = \boldsymbol{M}_r(\boldsymbol{y})$. Since $\boldsymbol{y}$ is feasible for SDP (2.6) and $\boldsymbol{z} = \boldsymbol{P}_{s(2r)}^{-1} \boldsymbol{y}$, we obtain that

$$\begin{aligned}
\boldsymbol{M}_r(\boldsymbol{z}) &= \boldsymbol{P}_{s(r)}^{-1} \boldsymbol{M}_r(\boldsymbol{y}) \boldsymbol{P}_{s(r)}^{-T} \succeq \boldsymbol{O}, \\
\boldsymbol{M}_{r_j}(\tilde{f}_j \boldsymbol{z}) &= \boldsymbol{P}_{s(r_j)}^{-1} \boldsymbol{M}_{r_j}(f_j \boldsymbol{y}) \boldsymbol{P}_{s(r_j)}^{-T} \succeq \boldsymbol{O} \quad (j = 1, \ldots, m).
\end{aligned}$$

These imply that $\boldsymbol{z}$ is feasible for SDP (3.2). By Lemma 4.3 and the definitions of $\boldsymbol{c}_{2r}$ and $\tilde{\boldsymbol{c}}_{2r}$, we also see that $\tilde{\boldsymbol{c}}_{2r} = \boldsymbol{P}_{s(2r)}^T \boldsymbol{c}_{2r}$ and $\tilde{\boldsymbol{c}}_{2r}^T \boldsymbol{z} = \boldsymbol{c}_{2r}^T \boldsymbol{P}_{s(2r)} \boldsymbol{P}_{s(2r)}^{-1} \boldsymbol{y} = \boldsymbol{c}_{2r}^T \boldsymbol{y}$. This shows that the objective value of SDP (3.2) coincides with that of SDP (2.6).

4.4 Proof of Property 3 of Theorem 3.1

Recall that we have already proved $\tilde{\boldsymbol{c}}_{2r} = \boldsymbol{P}_{s(2r)}^T \boldsymbol{c}_{2r}$, which is the first identity of property 3 of Theorem 3.1, in Lemma 4.3. To prove the second identity, let $j \in \{1, \ldots, m\}$ be fixed arbitrarily. Then we observe that

$$\begin{aligned}
\sum_{\alpha \in G_{2r}} z_\alpha \tilde{\boldsymbol{B}}_{j,\alpha} &= \boldsymbol{M}_{r_j}(\tilde{f}_j \boldsymbol{z}) \quad \text{(by (3.3))} \\
&= \boldsymbol{P}_{s(r_j)}^{-1} \boldsymbol{M}_{r_j}(f_j \boldsymbol{y}) \boldsymbol{P}_{s(r_j)}^{-T} \quad \text{(by Lemma 4.5)} \\
&= \boldsymbol{P}_{s(r_j)}^{-1} \left(\sum_{\beta \in G_{2r}} \boldsymbol{B}_{j,\beta} y_\beta \right) \boldsymbol{P}_{s(r_j)}^{-T} \quad \text{(by (2.7))} \\
&= \boldsymbol{P}_{s(r_j)}^{-1} \left(\sum_{\beta \in G_{2r}} \boldsymbol{B}_{j,\beta} \sum_{\alpha \in G_{2r}} P_{\beta,\alpha} z_\alpha \right) \boldsymbol{P}_{s(r_j)}^{-T} \quad \text{(by property 2 of Theorem 3.1)} \\
&= \boldsymbol{P}_{s(r_j)}^{-1} \left(\sum_{\alpha \in G_{2r}} z_\alpha \sum_{\beta \in G_{2r}} P_{\beta,\alpha} \boldsymbol{B}_{j,\beta} \right) \boldsymbol{P}_{s(r_j)}^{-T} \\
&= \sum_{\alpha \in G_{2r}} z_\alpha \boldsymbol{P}_{s(r_j)}^{-1} \left(\sum_{\beta \in G_{2r}} P_{\beta,\alpha} \boldsymbol{B}_{j,\beta} \right) \boldsymbol{P}_{s(r_j)}^{-T}.
\end{aligned}$$

Comparing the both sides of the above equality, we have the second identity of property 3 of Theorem 3.1.

Taking $f_j(\boldsymbol{x}) = 1$ in the above argument, we can similarly show the third identity of property 3 of Theorem 3.1. The details are omitted here.

4.5 Proof of Theorem 3.2

It suffices to show that $\text{rank}\boldsymbol{M}_k(\boldsymbol{z}^*) = \text{rank}\boldsymbol{M}_k(\boldsymbol{y}^*)$ for every $k \le r$. To show this, let $k \in \{0, 1, \dots, r\}$ be fixed. By property 2 of Lemma 4.1, we can express

$$\boldsymbol{P}_{s(r)} = \begin{pmatrix} \boldsymbol{P}_{s(k)} & \boldsymbol{O} \\ \boldsymbol{R} & \boldsymbol{S} \end{pmatrix}$$

for some $\boldsymbol{R} \in \mathbb{R}^{(s(r)-s(k)) \times s(k)}$ and $\boldsymbol{S} \in \mathbb{R}^{(s(r)-s(k)) \times (s(r)-s(k))}$. Substituting this into the relation $\boldsymbol{P}_{s(r)} \boldsymbol{M}_r(\boldsymbol{z}^*) \boldsymbol{P}_{s(r)}^T = \boldsymbol{M}_r(\boldsymbol{y}^*)$, and taking the $s(k) \times s(k)$ principal submatrices of the both sides, we obtain $\boldsymbol{P}_{s(k)} \boldsymbol{M}_k(\boldsymbol{z}^*) \boldsymbol{P}_{s(k)}^T = \boldsymbol{M}_k(\boldsymbol{y}^*)$. Because $\boldsymbol{P}_{s(k)}$ is nonsingular due to Lemma 4.1, we see that $\text{rank}\boldsymbol{M}_k(\boldsymbol{z}^*) = \text{rank}\boldsymbol{M}_k(\boldsymbol{y}^*)$ for all $k \le r$.

5 Concluding Remarks

We have shown that Lasserre's SDP relaxation [11] is invariant under any nonsingular affine transformation on the variable space $\mathbb{R}^n$. We can also say that $\boldsymbol{P}_{s(2r)}$ has the invariance property between the polynomial SDP and its linear SDP relaxation, and that the affine transformation induces such linear transformation on $\mathbb{R}[\boldsymbol{x}]$. In fact, one of the key observations was the block lower triangular structure of $\boldsymbol{P}_{s(2r)}$ of the linear transformation on $\mathbb{R}[\boldsymbol{x}]$. See property 2 of Lemma 4.1. We can hardly imagine that any linear transformation on $\mathbb{R}[\boldsymbol{x}]$ that does not have this property will have the good invariance property. On the other hand, we can consider some other linear transformations on $\mathbb{R}[\boldsymbol{x}]$ having the same block lower triangular structure property. Such a linear transformation is natural in the sense that it maps any polynomial of degree r to a polynomial of the same degree. A linear transformation which maps the standard monomials $\{\boldsymbol{x}^\alpha\}_{\alpha\in G_r}$ to Chebyshev polynomials of degree up to r is one of such examples. This linear transformation produces the so-called Chebyshev basis. In [1], de Klerk et al. use Chebyshev basis, and report that the numerical stability was improved for POP constructed from univariate polynomials. However, the formulation into the SDP problem is different from ours, and its invariance property is not mentioned. Whether a linear transformation which maps any polynomial of degree r to a polynomial of the same degree has the same invariance property or not will be the subject to further research.

An important usage of an affine transformation of a POP is to increase numerical stability. We can scale a POP to be solved by applying an appropriate affine transformation in advance so that the transformed POP could be solved more stably. Suppose that a POP to be solved involves a higher degree monomials in variables $x_1, \ldots, x_n$, and that they are expected to take large numerical values at optimal solutions. Then it is likely that optimal solutions of its SDP relaxation contain huge numerical values, which causes a numerical instability. It was reported in [22] that scaling those variables within $[0,1]^n$ is very effective to avoid this type of numerical instability, and this technique was incorporated in SparsePOP [23]. One of the problems where we can see such a remedy in numerical stability is Bex3_1_1:

$$\left\{\begin{array}{ll} \min & x_1+x_2+x_3 \\ \text{subject to} & 0.0025x_4+0.0025x_6 \leq 1, -0.0025x_4+0.0025x_5+0.0025x_7 \leq 1 \\ & -0.01x_5+0.01x_8 \leq 1, 100x_1 - x_1x_6 + 833.33252x_4 \leq 83333.333, \\ & x_2x_4 - x_2x_7 - 1250x_4 + 1250x_5 \leq 0, x_3x_5 - x_3x_8 - 2500x_5 \leq -1250000, \\ & 100 \leq x_1 \leq 700, 1000 \leq x_2 \leq 2000, 1000 \leq x_3 \leq 6000, \\ & 10 \leq x_4 \leq 300, 10 \leq x_5 \leq 400, 10 \leq x_6 \leq 400, \\ & 10 \leq x_7 \leq 400, 10 \leq x_8 \leq 500. \end{array}\right.$$

For Bex3_1_1, when we apply the SDP relaxation method as is, the SDP solver SeDuMi [21] terminates by reporting a message on the numerical error. On the other hand, if we rescale each variable between $[0,1]$ by applying the following affine transformation $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{w} + \boldsymbol{b}$:

$$\boldsymbol{A} = \begin{pmatrix} 600 & & & & & & & \\ & 1000 & & & & & & \\ & & 5000 & & & & & \\ & & & 290 & & & & \\ & & & & 390 & & & \\ & & & & & 390 & & \\ & & & & & & 390 & \\ & & & & & & & 490 \end{pmatrix}, \boldsymbol{b} = \begin{pmatrix} 100 \\ 1000 \\ 1000 \\ 10 \\ 10 \\ 10 \\ 10 \\ 10 \end{pmatrix},$$

then SeDuMi terminates normally and returns the accurate optimal value and solution of the SDP relaxation problem.

In this paper, we have not paid any attention to the sparsity of the polynomials involved in a POP, and we have restricted ourselves to the "dense" SDP relaxation of a POP. Another important usage of an affine transformation on the space of a variable vector of a POP is to improve its sparsity so that we can apply the "sparse" SDP relaxation [12, 22] to the transformed POP. This issue is discussed in the recent paper [6]. We should mention, however, that when we apply the "sparse" SDP relaxation to POPs the invariance under affine transformation does not hold any more in general.

References

[1] E. de Klerk, G. Elabwabi and D. den Hertog, Optimization of univariate functions on bounded intervals by interpolation and semidefinite programming, Discussion paper, 2006-026. Tilburg University. Center for Economic Research, 2006.

[2] D. Henrion and J. B. Lasserre, GloptiPoly: Global optimization over polynomials with Matlab and SeDuMi, *ACM Transactions on Mathematical Software* 29 (2003) 165–194.

[3] D. Henrion and J. B. Lasserre, Detecting global optimality and extracting solutions in GloptiPoly, in *Positive Polynomials in Control* D. Henrion and A. Garulli(eds.), Lecture Notes on Control and Information Sciences, Springer Verlag 2005.

[4] D. Henrion and J. B. Lasserre, Convergent relaxations of polynomial matrix inequalities and static output feedback, *IEEE Transactions on Automatic Control* 51 (2006) 191–202.

[5] C.W.J. Hol and C.W. Scherer, Sum of squares relaxations of polynomial semidefinite programming, in *Proc. Symp. on Mathematical Theory and Networks and Systems* (MTNS), Leuven, Belgium, 2004.

[6] S. Kim, M. Kojima and Ph. L. Toint, Recognizing underlying sparsity in optimization, to appear in *Mathematical Programming* 119 (2009) 273–303.

[7] S. Kim, M. Kojima and H.Waki, Generalized Lagrangian duals and sums of squares relaxations of sparse polynomial optimization problems, *SIAM Journal on Optimization* 15 (2005) 697–719.

[8] M. Kojima, Sums of squares relaxations of polynomial semidefinite programs, Research Report B-397, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro-ku, Tokyo 152-8552, Nov. 2003.

[9] M. Kojima, S. Kim and H. Waki, A general framework for convex relaxation of polynomial optimization problems over cones, *Journal of the Operations Research Society of Japan* 46 (2003) 125–144.

[10] M. Kojima and M. Muramatsu, An extension of sums of squares relaxations to polynomial optimization problems over symmetric cones, *Mathematical programming* 110 (2007) 315–336.

[11] J.B. Lasserre, Global optimization with polynomials and the problems of moments, *SIAM Journal on Optimization* 11 (2001) 796–817.

[12] J.B. Lasserre, Convergent semidefinite relaxations in polynomial optimization with sparsity, *SIAM Journal on Optimization* 17 (2006) 822–843.

[13] M. Laurent, Revisiting two theorems of Curto and Fialkow on moment matrices *American Mathematical Society* 133 (2005) 2965–2976.

[14] M. Laurent, Semidefinite representations for finite varieties, *Mathematical programming* 109 (2007) 1–26.

[15] Y. Nesterov, Structure of non-negative polynomials and optimization problems, *CORE Discussion Paper NO. 9749*, 1997

[16] J. Nie, W. Demmel and B. Sturmfels, Minimizing polynomials via sum of squares over the gradient ideal, *Mathematical programming* 106 (2006) 587–606.

[17] P. A. Parrilo, Semidefinite programming relaxations for semialgebraic problems, *Mathematical programming* 96 (2003) 293–320.

[18] S. Prajna, A. Papachristodoulou and P.A. Parrilo, SOSTOOLS: Sum of Squares Optimization Toolbox for MATLAB – User's Guide, Control and Dynamical Systems, California Institute of Technology, Pasadena, CA 91125 USA, 2002.

[19] N.Z. Shor, Class of global minimum bounds of polynomial functions, *Cybernetics* 23 (1987) 731–734.

[20] N.Z. Shor and P.I. Stetsyuk, The use of a modification of the r-algorithm for finding the global minimum of polynomial functions, *Cybernetics and Systems Analysys* 33 (1997) 482–497.

[21] J.F. Sturm, SeDuMi 1.02, a MATLAB toolbox for optimizatin over symmetric cones, *Optim. Meth. Softw.* 11 & 12 (1999) 625–653.

[22] H. Waki, S. Kim, M. Kojima and M. Muramatsu, Sums of squares and semidefinite programming relaxations for polynomial optimization problems with structured sparsity, *SIAM Journal on Optimization* 17 (2006) 218–242.

[23] H. Waki, S. Kim, M. Kojima, M. Muramatsu and H. Sugimoto, Algorithm 883: sparsePOP : a sparse semidefinite programming relaxation of polynomial optimization problem, *ACM Transactions on Mathematical Software* 35 (2008) 15:1–15:13.

Manuscript received 23 October 2007
revised 4 April 2008, 24 July 2008, 8 August 2008
accepted for publication 19 August 2008

Hayato Waki
Department of Computer Science, The University of Electro-Communications
1-5-1 Chofugaoka, Chofu-Shi, Tokyo 182-8585 Japan
E-mail address: `hayato.waki@jsb.cs.uec.ac.jp`

Masakazu Muramatsu
Department of Computer Science, The University of Electro-Communications
1-5-1 Chofugaoka, Chofu-Shi, Tokyo 182-8585 Japan
E-mail address: `muramatu@cs.uec.ac.jp`

Masakazu Kojima
Department of Mathematical and Computing Sciences, Tokyo Institute of Technology
2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152-8552 Japan
E-mail address: `kojima@is.titech.ac.jp`