Journal of the Operations Research
Society of Japan
2008, Vol. 51, No. 1, 111-125

EQUALITY BASED CONTRACTION OF SEMIDEFINITE PROGRAMMING RELAXATIONS IN POLYNOMIAL OPTIMIZATION

Cong Vo
DaiTri Joint Stock Company

Masakazu Muramatsu
The University of Electro-Communications

Masakazu Kojima
Tokyo Institute of Technology

(Received December 25, 2006)

Abstract The SDP (semidefinite programming) relaxation for general POPs (polynomial optimization problems), which was proposed as a method for computing global optimal solutions of POPs by Lasserre, has become an active research subject recently. We propose a new heuristic method exploiting the equality constraints in a given POP, and strengthen the SDP relaxation so as to achieve faster convergence to the global optimum of the POP. We can apply this method to both of the dense SDP relaxation which was originally proposed by Lasserre, and the sparse SDP relaxation which was later proposed by Kim, Kojima, Muramatsu and Waki. Especially, our heuristic method incorporated into the sparse SDP relaxation method has shown a promising performance in numerical experiments on large scale sparse POPs. Roughly speaking, we induce valid equality constraints from the original equality constraints of the POP, and then use them to convert the dense or sparse SDP relaxation into a new stronger SDP relaxation. Our method is enlightened by some strong theoretical results on the convergence of SDP relaxations for POPs with equality constraints provided by Lasserre , Parrilo and Laurent, but we place the main emphasis on the practical aspect to compute more accurate lower bounds of larger sparse POPs.

Keywords: Nonlinear programming, optimization, polynomial optimization, global optimization, semidefinite programming, equality constraints, exploiting sparsity

1. Introduction

We consider the following POP (polynomial optimization problem):

$$\langle\text{POP}\rangle \begin{cases} \text{minimize} & f(\boldsymbol{x}), \\ \text{subject to} & g_i(\boldsymbol{x}) \geq 0 \ (i = 1, \dots, p), \\ & h_i(\boldsymbol{x}) = 0 \ (i = 1, \dots, k) \end{cases} \tag{1.1}$$

where $f(\boldsymbol{x})$, $g_1(\boldsymbol{x}), \dots, g_p(\boldsymbol{x})$ and $h_1(\boldsymbol{x}), \dots, h_k(\boldsymbol{x})$ are polynomials in $\boldsymbol{x} \in \mathbb{R}^n$ with real coefficients.

Example 1.1 *Throughout this paper we use the example below:*

$$\begin{cases} \text{minimize} & x_1^2 - x_2^2, \\ \text{subject to} & x_1 \geq 0, \ x_2 \geq 0, \\ & x_1 + x_2 = 1. \end{cases} \tag{1.2}$$

The SDP relaxation (semidefinite programming relaxation, or sums of squares relaxation) for POPs, which was proposed as a method for computing global optimal solutions of general POPs by Lasserre [5] (see also Parrilo [9, 11]), has become an active research subject recently. He introduced a hierarchy of SDP (semidefinite programming) relaxations for a POP. Here we assume that the POP is a minimization problem of the form ⟨POP⟩ in (1.1). Then a sequence of SDP relaxation problems is constructed so that the optimal value of each problem serves as a lower bound for the global minimum of ⟨POP⟩ and is nondecreasing along the sequence. Under a mild assumption which requires the compactness of the feasible region, the sequence of lower bounds converges to the global minimum of ⟨POP⟩ ([5, Theorem 4.2]).

Although the theoretical global convergence property on the SDP relaxation mentioned above is very attractive, the sizes of the SDPs grow rapidly with the size of a POP to be solved. This discourages engineering applications of the SDP relaxation. Therefore, it is necessary from a practical point of view to revise the SDP relaxation so as to construct a hierarchy of smaller size SDP relaxation problems without sacrificing the quality of the lower bounds they provide. The heuristic proposed in this paper aims at such a revision effectively utilizing equality constraints involved in a POP.

The research on the SDP relaxation method spreads widely from fundamental theory to practical implementation of the method for large scale POPs. Important theoretical results include the theories on the convergence of the optimal values of the SDP relaxations to the optimal value of the POP (see Lasserre [5], Parrilo [11]). Henrion and Lasserre [2] showed by numerical experiments that the method is powerful for small size POPs. Applications of the method to POPs from engineering, which are of larger scales but sparse, is the main subject of practical approaches by Kim, Kojima, Muramatsu and Waki [3, 15]. We call the SDP relaxation method originally proposed by Lasserre the *dense* SDP relaxation method, and the one proposed later by Kim et al. the *sparse* SDP relaxation method. Recently, Lasserre [7] proved convergence of a hierarchy of sparse SDP relaxations in the spirit of the paper [15] by Waki et al. .

We propose a new heuristic method exploiting the equality constraints in ⟨POP⟩, which we can apply to both of the dense and sparse SDP relaxations, and strengthen the SDP relaxations so as to achieve faster convergence to the global minimum of ⟨POP⟩. Roughly speaking, we first add valid equality constraints induced from the original equality constraints of ⟨POP⟩. Secondly we use them to convert the dense or sparse SDP relaxation into an SDP satisfying the following properties:

- The psd (positive semidefinite) constraints of the new SDP becomes smaller than that of the original SDP.
- The new SDP may have more equality constraints.
- The lower bound given as the optimal value of the new SDP for ⟨POP⟩ is expected better (at least not worse) than that given by the original SDP.

Our method incorporated into the sparse SDP relaxation method has shown a promising performance in numerical experiments on large scale sparse POPs. The experiments have shown that the new SDPs are easier to solve, and the optimal values of the new SDPs converge faster to the optimal value of a given POP.

Our method is enlightened by the theoretical results referred below, but it places the main emphasis on the practical aspect to compute more accurate lower bounds of larger sparse POPs. Lasserre [6, 4] showed that optimization of a polynomial on a grid of finite points in $\mathbb{R}^n$: $\{\boldsymbol{x} \in \mathbb{R}^n : h_i(\boldsymbol{x}) = 0, \ (i = 1, \ldots, k)\}$, where $h_i(\boldsymbol{x}) = \prod_{j=1}^{r_i}(x_j - a_{i,j})$, reduces to an SDP problem of fixed size. In particular, every 0-1 POP in n variables is equivalent to an SDP problem in $2^n - 1$ variables (generated by the n-th SDP relaxation). Later, Parrilo [10] extends Lasserre's results to the case where the feasible region of ⟨POP⟩ is an arbitrary finite set. He required that the set $\{\boldsymbol{x} \in \mathbb{C}^n : h_i(\boldsymbol{x}) = 0 \ (i = 1, \ldots, k)\}$ to be finite and that the ideal $\mathcal{I}$ generated by $h_1(\boldsymbol{x})$,..., $h_k(\boldsymbol{x})$ to be radical, and proved that every polynomial nonnegative on the feasible region belongs to the quadratic module generated by the equality constraints and inequality constraints. As a consequence, the sum of squares relaxation is exact for some proper finite supports. However, a prerequisite for obtaining the SDP representation is to completely enumerate all feasible solutions of the given POP. Laurent [8] presented a new SDP relaxation which involves combinatorial moment matrices. When the polynomial equality constraints have a finite set of complex solutions, without assuming that the ideal $\mathcal{I}$ generated by the equality constraints is radical, she extended Parrilo's results furthermore and proved that the POP can be reformulated as a finite SDP problem. She then proved the convergence of the sum of squares relaxations in the case where $\mathcal{I}$ is radical (i.e. the result of [10]) and in the case where the equality constraints construct a Groebner basis of $\mathcal{I}$. The combinatorial moment matrices involved in her method are indexed by a basis of the quotient space $\mathbb{R}[\boldsymbol{x}]/\mathcal{I}$. Hence they have smaller size comparing to the size of the classical moment matrices, though, the formulation of combinatorial moment matrices requires a considerable computational cost.

After some preparation of the dense and sparse SDP relaxation method for $\langle\text{POP}\rangle$ in Section 2, we show how to reduce the dimensions of psd matrices in these relaxations in Section 3, how to add valid equality constraints to $\langle\text{POP}\rangle$ in Section 4, and report numerical results in Section 5. Concluding discussions are given in Section 6.

2. Preliminaries

Let $\mathbb{R}^n$ denote the n dimensional Euclidean space, and $\mathbb{Z}^n_+$ the set of nonnegative integer vectors in $\mathbb{R}^n$. A real-valued polynomial $t : \mathbb{R}^n \to \mathbb{R}$ is defined as follows. Let $S \subset \mathbb{Z}^n_+$ be a nonempty finite set. Assuming that for each $\boldsymbol{\alpha} \in S$, a real value $t_{\boldsymbol{\alpha}}$ is given, we have $t(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha}\in S} t_{\boldsymbol{\alpha}}\boldsymbol{x}^{\boldsymbol{\alpha}}$, where $\boldsymbol{x}^{\boldsymbol{\alpha}} = x_1^{\alpha_1}\cdots x_n^{\alpha_n}$. The support of t is defined by $\operatorname{supp} t = \{\boldsymbol{\alpha} \in S, t_{\boldsymbol{\alpha}} \neq 0\}$, and the degree of t by $\deg t = \max_{\boldsymbol{\alpha}\in\operatorname{supp} t} |\boldsymbol{\alpha}|$, where $|\boldsymbol{\alpha}| = \sum_{i=1}^n \alpha_i$. Let $\deg\langle\text{POP}\rangle = \max(\deg f, \deg g_1, \ldots, \deg g_p, \deg h_1, \ldots, \deg h_k)$. Let $\mathbb{R}[\boldsymbol{x}]$ denote the set of real-valued polynomials of $\boldsymbol{x} \in \mathbb{R}^n$.

We say that $\boldsymbol{\alpha} <_{\text{lex}} \boldsymbol{\beta}$ and $\boldsymbol{x}^{\boldsymbol{\alpha}} <_{\text{lex}} \boldsymbol{x}^{\boldsymbol{\beta}}$ for $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{Z}^n_+, \boldsymbol{x} \in \mathbb{R}^n$ if, in the vector $\boldsymbol{\alpha} - \boldsymbol{\beta}$ the left-most nonzero entry is negative. Let $S = \{\boldsymbol{\alpha}^1, \boldsymbol{\alpha}^2, \ldots, \boldsymbol{\alpha}^k\} \subset \mathbb{Z}^n_+$ with $\boldsymbol{\alpha}^1 <_{\text{lex}} \boldsymbol{\alpha}^2 <_{\text{lex}} \cdots <_{\text{lex}} \boldsymbol{\alpha}^k$ and $t_{\boldsymbol{\alpha}^i} \in \mathbb{R}$ $(i = 1, \ldots, k)$. We define $(\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in S) := \left(\boldsymbol{x}^{\boldsymbol{\alpha}^1}, \boldsymbol{x}^{\boldsymbol{\alpha}^2}, \ldots, \boldsymbol{x}^{\boldsymbol{\alpha}^k}\right)^T$ and $(t_{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in S) := (t_{\boldsymbol{\alpha}^1}, t_{\boldsymbol{\alpha}^2}, \ldots, t_{\boldsymbol{\alpha}^k})^T$. A polynomial $t(\boldsymbol{x})$ is then identified with its coefficient vector $(t_{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \operatorname{supp} t)$, since $t(\boldsymbol{x}) \equiv (t_{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \operatorname{supp} t)^T (\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \operatorname{supp} t)$. In the following, the elements of a finite set $S \subset \mathbb{Z}^n_+$ are always arranged in the "lexicographical" order mentioned above.

For a symmetric square matrix A, $A \succcurlyeq \mathbf{O}$ means that A is positive semidefinite. Let

$$M_S(\boldsymbol{x}) := \ (\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in S) \ (\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in S)^T$$

for a finite set $S \subset \mathbb{Z}^n_+$. From $\langle\text{POP}\rangle$ we introduce the following PSDP (polynomial SDP) problem which is to lead to a dense SDP relaxation or a sparse SDP relaxation:

$$\langle\text{PSDP}\rangle \begin{cases} \text{minimize} & f(\boldsymbol{x}), \\ \text{subject to} & M_{C_i}(\boldsymbol{x}) \succcurlyeq \mathbf{O} \ (i = 1, \ldots, r), \\ & g_i(\boldsymbol{x}) \ M_{G_i}(\boldsymbol{x}) \succcurlyeq \mathbf{O} \ (i = 1, \ldots, p), \\ & h_i(\boldsymbol{x}) \ M_{H_i}(\boldsymbol{x}) = \mathbf{O} \ (i = 1, \ldots, k) \end{cases} \tag{2.1}$$

where C_i, G_i, H_i are finite subsets of $\mathbb{Z}^n_+$ such that, for some integer $N \geq \frac{1}{2}\deg\langle\text{POP}\rangle$:

$$\begin{cases} \mathbf{0} \in C_i, \deg M_{C_i}(\boldsymbol{x}) = 2N \ (i = 1, \ldots, r), \\ \mathbf{0} \in G_i, \deg g_i(\boldsymbol{x}) \ M_{G_i}(\boldsymbol{x}) = 2N \text{ or } 2N - 1 \ (i = 1, \ldots, p), \\ \mathbf{0} \in H_i, \deg h_i(\boldsymbol{x}) \ M_{H_i}(\boldsymbol{x}) = 2N \text{ or } 2N - 1 \ (i = 1, \ldots, k). \end{cases} \tag{2.2}$$

Clearly, if $\boldsymbol{x} \in \mathbb{R}^n$ is a feasible solution of $\langle$POP$\rangle$, then it is feasible to $\langle$PSDP$\rangle$, since $M_S(\boldsymbol{x}) \succcurlyeq \mathbf{O}$ for any finite set $S \subset \mathbb{Z}^n_+$. Furthermore, the elements in the upper left corners of $M_{C_i}(\boldsymbol{x}), M_{G_j}(\boldsymbol{x}), M_{H_l}(\boldsymbol{x})$ always take value 1, since $\mathbf{0} \in C_i, \mathbf{0} \in G_j, \mathbf{0} \in H_l$ $(i = 1, \ldots, r, j = 1, \ldots, p, l = 1, \ldots, k)$, therefore any feasible solution of $\langle$PSDP$\rangle$ will be feasible to $\langle$POP$\rangle$. Moreover, the objective functions of $\langle$POP$\rangle$ and $\langle$PSDP$\rangle$ are the same. Consequently, $\langle$PSDP$\rangle$ is equivalent to $\langle$POP$\rangle$.

We can rewrite $\langle$PSDP$\rangle$ as:

$$\begin{cases} \text{minimize} & f(\boldsymbol{x}), \\ \text{subject to} & \sum_{\boldsymbol{\alpha} \in S_{C_i}} \widehat{M}_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} \succcurlyeq \mathbf{O} \ (i = 1, \ldots, r), \\ & \sum_{\boldsymbol{\alpha} \in S_{G_i}} \widetilde{M}_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} \succcurlyeq \mathbf{O} \ (i = 1, \ldots, p), \\ & \sum_{\boldsymbol{\alpha} \in S_{H_i}} \overline{M}_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} = \mathbf{O} \ (i = 1, \ldots, k) \end{cases} \tag{2.3}$$

where $S_{C_1}, \ldots, S_{C_r}, S_{G_1}, \ldots, S_{G_p}, S_{H_1}, \ldots, S_{H_k}$ are finite subsets of $\mathbb{Z}^n_+$ and $\widehat{M}_{\boldsymbol{\alpha}}$, $\widetilde{M}_{\boldsymbol{\alpha}}$, $\overline{M}_{\boldsymbol{\alpha}}$ are symmetric square matrices, deriving from the coefficients of polynomials $g_i(\boldsymbol{x}), h_i(\boldsymbol{x})$ and matrices of monomials $M_{C_i}(\boldsymbol{x})$, $M_{G_i}(\boldsymbol{x}), M_{H_i}(\boldsymbol{x})$ in $\langle$PSDP$\rangle$. Let $\langle$SDP$\rangle$ denote the SDP resulting from $\langle$PSDP$\rangle$ by linearization, i.e. by replacing each monomial $\boldsymbol{x}^{\boldsymbol{\alpha}}$ in (2.3) by a new variable $y_{\boldsymbol{\alpha}} \in \mathbb{R}$:

$$\langle \text{SDP} \rangle \begin{cases} \text{minimize} & \left(f_{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \operatorname{supp} f\right)^T \ \left(y_{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \operatorname{supp} f\right), \\ \text{subject to} & \sum_{\boldsymbol{\alpha} \in S_{C_i}} \widehat{M}_{\boldsymbol{\alpha}} y_{\boldsymbol{\alpha}} \succcurlyeq \mathbf{O} \ (i = 1, \ldots, r), \\ & \sum_{\boldsymbol{\alpha} \in S_{G_i}} \widetilde{M}_{\boldsymbol{\alpha}} y_{\boldsymbol{\alpha}} \succcurlyeq \mathbf{O} \ (i = 1, \ldots, p), \\ & \sum_{\boldsymbol{\alpha} \in S_{H_i}} \overline{M}_{\boldsymbol{\alpha}} y_{\boldsymbol{\alpha}} = \mathbf{O} \ (i = 1, \ldots, k). \end{cases} \tag{2.4}$$

We say that $\langle$SDP$\rangle$ is a SDP relaxation of $\langle$POP$\rangle$ in *relaxation order* N. This $\langle$SDP$\rangle$ represents a dense SDP relaxation ([5]) or a sparse SDP relaxation ([3, 15]) depending on how we choose C_i, G_i, H_i in $\langle$PSDP$\rangle$.

Example 2.1 *For the instance* (1.2)*, if we take the relaxation order* $= 1$*, then we have a polynomial SDP*

$$\begin{cases} \text{minimize} & x_1^2 - x_2^2, \\ \text{subject to} & x_1 + x_2 = 1, \ x_1 \geq 0, \ x_2 \geq 0, \\ & x_1^2 + x_1 x_2 - x_1 = 0, \ x_1 x_2 + x_2^2 - x_2 = 0, \\ & \begin{pmatrix} 1 \\ x_1 \\ x_2 \end{pmatrix} \begin{pmatrix} 1 & x_1 & x_2 \end{pmatrix} \equiv \begin{pmatrix} 1 & x_1 & x_2 \\ x_1 & x_1^2 & x_1 x_2 \\ x_2 & x_1 x_2 & x_2^2 \end{pmatrix} \succcurlyeq \mathbf{O}. \end{cases} \tag{2.5}$$

Hence we obtain an SDP relaxation problem

$$\begin{cases} \text{minimize} & y_{20} - y_{02}, \\ \text{subject to} & y_{10} + y_{01} = 1,\ y_{10} \geq 0, y_{01} \geq 0, \\ & y_{20} + y_{11} - y_{10} = 0,\ y_{11} + y_{02} - y_{01} = 0, \\ & \begin{pmatrix} 1 & y_{10} & y_{01} \\ y_{10} & y_{20} & y_{11} \\ y_{01} & y_{11} & y_{02} \end{pmatrix} \succcurlyeq \mathbf{O}. \end{cases}$$

Proposition 2.1 *If ⟨POP⟩ has an optimal solution $\bar{\boldsymbol{x}}$, then ⟨SDP⟩ has a nonempty feasible solution set. Moreover, the optimal value of ⟨SDP⟩ is a lower bound for that of ⟨POP⟩.*

Proof Let $\bar{y}_{\boldsymbol{\alpha}} = \bar{\boldsymbol{x}}^{\boldsymbol{\alpha}}$, then $(\bar{y}_{\boldsymbol{\alpha}})$ is a feasible solution of ⟨SDP⟩. Moreover, the objective value of ⟨SDP⟩ at $(\bar{y}_{\boldsymbol{\alpha}})$ is equivalent to the optimal value of ⟨POP⟩. The conclusion follows since ⟨SDP⟩ is a minimization problem. ∎

The dimensions of the symmetric square matrices appearing in ⟨SDP⟩ are determined by the cardinality of C_i, G_i and H_i as follows:

- $M_{\boldsymbol{\alpha}}$ $(\boldsymbol{\alpha} \in S_{C_i})$ have the same dimensions with $M_{C_i}(\boldsymbol{x})$, i. e. $|C_i| \times |C_i|$, for $i = 1, \dots, r$.
- $M_{\boldsymbol{\alpha}}$ $(\boldsymbol{\alpha} \in S_{G_i})$ have the same dimensions with $M_{G_i}(\boldsymbol{x})$, i. e. $|G_i| \times |G_i|$, for $i = 1, \dots, p$.
- $M_{\boldsymbol{\alpha}}$ $(\boldsymbol{\alpha} \in S_{H_i})$ have the same dimensions with $M_{H_i}(\boldsymbol{x})$, i. e. $|H_i| \times |H_i|$, for $i = 1, \dots, k$.

Lasserre introduces in [5] a hierarchy of SDP relaxations, which we call *dense* relaxations in this paper, since C_i, G_i, H_i are set as the largest possible ones, i. e. given N, referring to (2.1):

$$\begin{cases} C_1 = \{\boldsymbol{\alpha} \in \mathbb{Z}^n_+ \ :\ |\boldsymbol{\alpha}| \leq N\}, & r = 1, \\ G_i = \{\boldsymbol{\alpha} \in \mathbb{Z}^n_+ \ :\ |\boldsymbol{\alpha}| \leq N - \frac{1}{2}\deg g_i(\boldsymbol{x})\} & (i = 1, \dots, p), \\ H_i = \{\boldsymbol{\alpha} \in \mathbb{Z}^n_+ \ :\ |\boldsymbol{\alpha}| \leq N - \frac{1}{2}\deg h_i(\boldsymbol{x})\} & (i = 1, \dots, k). \end{cases} \tag{2.6}$$

The hierarchy of dense relaxations provides a sequence of ⟨SDP⟩ whose associated optimal values asymptotically converge to the global minimum of ⟨POP⟩, under a mild assumption ([5, Theorem 4.2]). In practice, the convergence is usually fast, and often finite (up to machine precision); see e.g. [2]. However, despite these nice features, the size of the dense relaxations grows rapidly with the size of the original POP. Typically, the dense relaxation of relaxation order N has to handle at least one psd matrix of size $\binom{n+N}{N} \times \binom{n+N}{N}$, and $\binom{n+2N}{2N}$ variables. The large size of the dense relaxations limits the applicability of the method to problems with small to medium size only. Following Kim, Kojima, Muramatsu and Waki [3, 15], we introduce the correlative sparsity. We first observe that the number of possible monomials in a polynomial $t \in \mathbb{R}[\boldsymbol{x}]$ of a degree d is $\binom{n+d}{d}$, however, a polynomial in practical cases often consists of a much smaller number of monomials; that is when the polynomial is *sparse*. Let $V = \{1, \dots, n\}$ and for a polynomial $t(\boldsymbol{x})$, $\operatorname{ind} t = \{j \in V \ :\ \max_{\boldsymbol{\alpha} \in \operatorname{supp} t} \alpha_j > 0\}$ is to denote indices of the variables appearing in $t(\boldsymbol{x})$.

- a csp (correlative sparsity pattern) graph $G = \langle V, E \rangle$ is built as: $\{i, j\} \in E$ if and only if
  - either x_i and x_j appear simultaneously in a monomial of $f(\boldsymbol{x})$
  - or they appear in an inequality (or equality) constraint;

- let $\overline{C_1}, \dots, \overline{C_r}$ be the maximum cliques of a chordal extension of G (a chordal graph is a simple graph possessing no chordless cycles, see [16]);
- let $\overline{G_i} = \bigcup_{\overline{C_j} \supseteq \mathrm{ind}\, g_i} \overline{C_j}$ for $i = 1, \dots, p$;
- let $\overline{H_i} = \bigcup_{\overline{C_j} \supseteq \mathrm{ind}\, h_i} \overline{C_j}$ for $i = 1, \dots, k$.

Finally, referring to (2.1):

$$\begin{cases} C_i = \{\boldsymbol{\alpha} \in \mathbb{Z}_+^n \ : \ |\boldsymbol{\alpha}| \le N, \alpha_j = 0 \text{ if } j \notin \overline{C_i}\} & (i = 1, \dots, r), \\ G_i = \{\boldsymbol{\alpha} \in \mathbb{Z}_+^n \ : \ |\boldsymbol{\alpha}| \le N - \frac{1}{2} \deg g_i(\boldsymbol{x}), \alpha_j = 0 \text{ if } j \notin \overline{G_i}\} & (i = 1, \dots, p), \\ H_i = \{\boldsymbol{\alpha} \in \mathbb{Z}_+^n \ : \ |\boldsymbol{\alpha}| \le N - \frac{1}{2} \deg h_i(\boldsymbol{x}), \alpha_j = 0 \text{ if } j \notin \overline{H_i}\} & (i = 1, \dots, k). \end{cases} \tag{2.7}$$

Let $\kappa = \max\{|C_1|, \dots, |C_r|, |G_1|, \dots, |G_p|, |H_1|, \dots, |H_k|\}$. The sparse relaxation of order N consists of symmetric square matrices of size $\mathcal{O}\left(\begin{pmatrix} \kappa + N \\ N \end{pmatrix} \times \begin{pmatrix} \kappa + N \\ N \end{pmatrix}\right)$, and about $\mathcal{O}\left(\begin{pmatrix} \kappa + 2N \\ 2N \end{pmatrix}\right)$ variables. When κ is small, a frequent case in practical applications of interest, the sparse SDP relaxation method above succeeds in exploiting the sparsity of the POP to obtain impressive gains in the size of the resulting SDP relaxations, as well as in the computational time needed for obtaining an optimal solution.

Lasserre [7] has made some (computationally) slight modification to the SDP relaxations defined in [15] and has proved theoretical convergence of the sparse SDP relaxations under a certain condition on the sparsity pattern.

3. Contraction of Semidefinite Matrices

A typical psd matrix $M_{C_i}(\boldsymbol{x})$ or $g_i(\boldsymbol{x})\ M_{G_i}(\boldsymbol{x})$ in $\langle$PSDP$\rangle$ (2.1) has the form

$$t(\boldsymbol{x})\ M_S(\boldsymbol{x}) \succcurlyeq \mathbf{O} \tag{3.1}$$

where $t(\boldsymbol{x})$ may be 1. We will show in this section how to translate this psd constraint into a smaller one and additional equality constraints when we are given valid equality constraints on the monomial set S, i. e. given a constant matrix K such that:

$$K\ \ (\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in S) = \mathbf{0} \tag{3.2}$$

is satisfied by any feasible solution $\boldsymbol{x}$ of $\langle$PSDP$\rangle$. In Section 4, we will explain how to find such a matrix K, given a finite set $S \subset \mathbb{Z}_+^n$. Assume K is of full row rank. There exists a matrix J such that

$$L = \begin{pmatrix} J \\ K \end{pmatrix} \text{ is a non-singular square matrix.} \tag{3.3}$$

Then (3.1) is equivalent to

$$Lt(\boldsymbol{x})\ M_S(\boldsymbol{x}) L^T \succcurlyeq \mathbf{O} \tag{3.4}$$

or

$$\begin{pmatrix} t(\boldsymbol{x}) J M_S(\boldsymbol{x}) J^T & t(\boldsymbol{x}) J M_S(\boldsymbol{x}) K^T \\ t(\boldsymbol{x}) K M_S(\boldsymbol{x}) J^T & t(\boldsymbol{x}) K M_S(\boldsymbol{x}) K^T \end{pmatrix} \succcurlyeq \mathbf{O}. \tag{3.5}$$

However, by (3.2), among the four submatrices of the psd matrix above, the Northeast, the Southeast and the Southwest must be zeros, and the Northwest psd, or equivalently:

$$\begin{cases} t(\boldsymbol{x}) J M_S(\boldsymbol{x}) J^T & \succcurlyeq \mathbf{O}, \\ t(\boldsymbol{x}) K M_S(\boldsymbol{x}) & = \mathbf{O}. \end{cases} \tag{3.6}$$

Example 3.1 *For the concrete example* (2.5)*, we rewrite the equality constraint* $x_1+x_2=1$ *as*

$$K\begin{pmatrix}1\\x_1\\x_2\end{pmatrix}=0$$

for

$$K=\begin{pmatrix}-1 & 1 & 1\end{pmatrix}.$$

Choose

$$J=\begin{pmatrix}1 & 0 & 0\\0 & 1 & 0\end{pmatrix}$$

so that

$$\begin{pmatrix}J\\K\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\0 & 1 & 0\\-1 & 1 & 1\end{pmatrix}$$

forms a 3×3 *nonsingular matrix. Then, for any* $(x_1.x_2)$ *satisfying the equality constraint* $x_1+x_2=1$*, it satisfies the positive semidefinite constraint in the polynomial SDP* (2.5)

$$\begin{pmatrix}1 & x_1 & x_2\\x_1 & x_1^2 & x_1x_2\\x_2 & x_1x_2 & x_2^2\end{pmatrix}\succcurlyeq \mathbf{O}$$

if and only if it satisfies

$$\begin{pmatrix}J\\K\end{pmatrix}\begin{pmatrix}1\\x_1\\x_2\end{pmatrix}\begin{pmatrix}1 & x_1 & x_2\end{pmatrix}\begin{pmatrix}J\\K\end{pmatrix}^T\succcurlyeq \mathbf{O},$$

or equivalently

$$\begin{pmatrix}1\\x_1\\0\end{pmatrix}\begin{pmatrix}1 & x_1 & 0\end{pmatrix}\equiv\begin{pmatrix}1 & x_1 & 0\\x_1 & x_1^2 & 0\\0 & 0 & 0\end{pmatrix}\succcurlyeq \mathbf{O}. \tag{3.7}$$

Now let us write (3.1) as

$$\sum_{\boldsymbol{\alpha}\in T}(\hat{M}_S)_{\boldsymbol{\alpha}}\boldsymbol{x}^{\boldsymbol{\alpha}}\succcurlyeq \mathbf{O}. \tag{3.8}$$

Then its linearization becomes

$$\sum_{\boldsymbol{\alpha}\in T}(\hat{M}_S)_{\boldsymbol{\alpha}}y_{\boldsymbol{\alpha}}\succcurlyeq \mathbf{O}. \tag{3.9}$$

Proposition 3.1 *We consider the following system:*

$$\begin{cases}\sum_{\boldsymbol{\alpha}\in T}J(\hat{M}_S)_{\boldsymbol{\alpha}}J^Ty_{\boldsymbol{\alpha}}\succcurlyeq \mathbf{O},\\ \sum_{\boldsymbol{\alpha}\in T}K(\hat{M}_S)_{\boldsymbol{\alpha}}y_{\boldsymbol{\alpha}}=\mathbf{O}.\end{cases} \tag{3.10}$$

1. (3.10) *is the linearization of* (3.6).
2. *Any solution* $\boldsymbol{y}$ *satisfying* (3.10) *also satisfy* (3.9).

Proof Premultiplying and multiplying L and L^T respectively to (3.8), we see that (3.4) and (3.5) can be rewrote as

$$\sum_{\boldsymbol{\alpha}\in T} L(\hat{M}_S)_{\boldsymbol{\alpha}} L^T \boldsymbol{x}^{\boldsymbol{\alpha}} = \begin{pmatrix} \sum_{\boldsymbol{\alpha}\in T} J(\hat{M}_S)_{\boldsymbol{\alpha}} J^T \boldsymbol{x}^{\boldsymbol{\alpha}} & \sum_{\boldsymbol{\alpha}\in T} J(\hat{M}_S)_{\boldsymbol{\alpha}} K^T \boldsymbol{x}^{\boldsymbol{\alpha}} \\ \sum_{\boldsymbol{\alpha}\in T} K(\hat{M}_S)_{\boldsymbol{\alpha}} J^T \boldsymbol{x}^{\boldsymbol{\alpha}} & \sum_{\boldsymbol{\alpha}\in T} K(\hat{M}_S)_{\boldsymbol{\alpha}} K^T \boldsymbol{x}^{\boldsymbol{\alpha}} \end{pmatrix} \succcurlyeq \mathbf{O}, \tag{3.11}$$

and its linearization

$$\begin{pmatrix} \sum_{\boldsymbol{\alpha}\in T} J(\hat{M}_S)_{\boldsymbol{\alpha}} J^T y_{\boldsymbol{\alpha}} & \sum_{\boldsymbol{\alpha}\in T} J(\hat{M}_S)_{\boldsymbol{\alpha}} K^T y_{\boldsymbol{\alpha}} \\ \sum_{\boldsymbol{\alpha}\in T} K(\hat{M}_S)_{\boldsymbol{\alpha}} J^T y_{\boldsymbol{\alpha}} & \sum_{\boldsymbol{\alpha}\in T} K(\hat{M}_S)_{\boldsymbol{\alpha}} K^T y_{\boldsymbol{\alpha}} \end{pmatrix} \succcurlyeq \mathbf{O}. \tag{3.12}$$

It is clear from (3.11), (3.12) and (3.5) that (3.10) is exactly the linearization of (3.6). Because L is nonsingular due to (3.3), clearly (3.9) and (3.12) are equivalent each other.

If $\boldsymbol{y}$ satisfy (3.10), then the Northeast, the Southeast, and the Southwest submatrices of (3.12) are all zeros, and the Northwest submatrix of (3.12) is positive semidefinite. Therefore the positive semidefiniteness of the whole matrix follows. ∎

Moreover, when K in (3.2) is given, we can choose J to satisfy (3.3) such that $J\ (\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in S)$ becomes a sub-vector of $(\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in S)$, as follows. Let $k = |S|$, $\boldsymbol{e}_i$ denote the i-th row of the $k \times k$ identity matrix. Since the vector set $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_k\}$ is linearly independent, we can select a subset $V = \left\{\boldsymbol{e}_{i_1}, \boldsymbol{e}_{i_2}, \ldots, \boldsymbol{e}_{i_{k-\operatorname{rank} K}}\right\} \subset \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_k\}$ such that

$$\operatorname{rank} \begin{pmatrix} K \\ \boldsymbol{e}_{i_1} \\ \boldsymbol{e}_{i_2} \\ \vdots \\ \boldsymbol{e}_{i_\ell} \end{pmatrix} = \ell + \operatorname{rank} K$$

for $\ell = 1, 2, \ldots, k - \operatorname{rank} K$. Finally, we use vectors in the vector set V as row vectors to build the matrix J.

Now, since J is a sub-matrix of the identity matrix, we have $\left(\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \bar{S}\right) = J\ (\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in S)$ for some $\bar{S} \subset S$. Therefore (3.6) becomes

$$\begin{cases} t(\boldsymbol{x})\ M_{\bar{S}}(\boldsymbol{x}) & \succcurlyeq \mathbf{O}, \\ t(\boldsymbol{x}) K\ M_S(\boldsymbol{x}) & = \mathbf{O}. \end{cases} \tag{3.13}$$

Note that if for the monomial set S we do not have K available to satisfy (3.2), then in (3.13) we suppose $\bar{S} = S$ and omit the equality constraint (or let $K = \mathbf{O}$).

Using the technique described above, corresponding to ⟨PSDP⟩ (2.1), if we have the following valid equality constraints:

$$\begin{aligned} K_{c_i}\ (\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in C_i) &= \mathbf{0} \quad (i = 1, \ldots, r), \\ K_{g_i}\ (\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in G_i) &= \mathbf{0} \quad (i = 1, \ldots, p), \end{aligned}$$

then we can translate ⟨PSDP⟩ into the following optimization problems:

$$\overline{\langle \text{PSDP} \rangle} \begin{cases} \text{minimize} & f(\boldsymbol{x}), \\ \text{subject to} & M_{\bar{C}_i}(\boldsymbol{x}) \succcurlyeq \mathbf{O}\ (i = 1, \ldots, r), \\ & g_i(\boldsymbol{x})\ M_{\bar{G}_i}(\boldsymbol{x}) \succcurlyeq \mathbf{O}\ (i = 1, \ldots, p), \\ & h_i(\boldsymbol{x})\ M_{H_i}(\boldsymbol{x}) = \mathbf{O}\ (i = 1, \ldots, k), \\ & K_{c_i} M_{C_i}(\boldsymbol{x}) = \mathbf{O}\ (i = 1, \ldots, r), \\ & K_{g_i} g_i(\boldsymbol{x})\ M_{G_i}(\boldsymbol{x}) = \mathbf{O}\ (i = 1, \ldots, p). \end{cases} \tag{3.14}$$

Note that $\bar{G}_i$ and $\bar{C}_i$ are subsets of G_i and C_i, respectively. Though $\overline{\langle\text{PSDP}\rangle}$ is equivalent to $\langle\text{PSDP}\rangle$, the sizes of the psd matrices in $\overline{\langle\text{PSDP}\rangle}$ are smaller than the sizes of respective psd matrices in $\langle\text{PSDP}\rangle$. Let the linearizations of $\overline{\langle\text{PSDP}\rangle}$ and $\langle\text{PSDP}\rangle$ be $\overline{\langle\text{SDP}\rangle}$ and $\langle\text{SDP}\rangle$, respectively, we have the following propositions.

Proposition 3.2 *If $\langle POP\rangle$ has an optimal solution $\bar{\boldsymbol{x}}$, then $\overline{\langle SDP\rangle}$ has a nonempty feasible solution set. Moreover, the optimal value of $\overline{\langle SDP\rangle}$ is a lower bound for that of $\langle POP\rangle$.*

Proof Similar to the proof of Proposition 2.1. ▮

Proposition 3.3 *If $\langle POP\rangle$ has an optimal solution $\bar{\boldsymbol{x}}$, then $\overline{\langle SDP\rangle}$ has an optimal value larger than, or at least equal to that of $\langle SDP\rangle$.*

Proof First, the objective functions of $\overline{\langle\text{SDP}\rangle}$ and $\langle\text{SDP}\rangle$ are the same. Second, the linearization of the constraints $h_i(\boldsymbol{x})\ M_{H_i}(\boldsymbol{x}) = \mathbf{O}$ $(i = 1, \dots, k)$ are the same in the two systems. Finally, each psd constraint of the form (3.1) in $\langle\text{PSDP}\rangle$ (where it may happen that $t(\boldsymbol{x}) = 1$) corresponds to a sub-system of the form (3.13) in $\overline{\langle\text{PSDP}\rangle}$, and the linearization of the latter sub-system is not looser than the linearization of the former psd constraint (Proposition 3.1). The conclusion follows. ▮

Example 3.2 *Since* (3.7), *we obtain the following reduced version of the polynomial SDP* (2.5)

$$\begin{cases} \text{minimize} & x_1^2 - x_2^2, \\ \text{subject to} & x_1 + x_2 = 1,\ x_1 \geq 0,\ x_2 \geq 0, \\ & x_1^2 + x_1x_2 - x_1 = 0,\ x_1x_2 + x_2^2 - x_2 = 0, \\ & \begin{pmatrix} 1 & x_1 \\ x_1 & x_1^2 \end{pmatrix} \succcurlyeq \mathbf{O}. \end{cases}$$

and an SDP relaxation problem

$$\begin{cases} \text{minimize} & y_{20} - y_{02}, \\ \text{subject to} & y_{10} + y_{01} = 1,\ y_{10} \geq 0, y_{01} \geq 0, \\ & y_{20} + y_{11} - y_{10} = 0,\ y_{11} + y_{02} - y_{01} = 0, \\ & \begin{pmatrix} 1 & y_{10} \\ y_{10} & y_{20} \end{pmatrix} \succcurlyeq \mathbf{O}. \end{cases}$$

as its linearization.

4. Valid Equality Constraints

The main idea in producing new valid equality constraints boils down to recognizing that the more valid equality constraints we have, the more we can reduce the size of psd matrices $M_{C_i}(\boldsymbol{x})$ $(i = 1, \dots, r)$ and $M_{G_j}(\boldsymbol{x})$ $(j = 1, \dots, p)$ in $\langle\text{PSDP}\rangle$ as we have done in Section 3. We will show how to conduct valid equality constraints for a given finite monomial set $(\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in S)$, $S \subset \mathbb{Z}_+^n$. In other words, we find a coefficient matrix K to satisfy (3.2), i. e.

$$K\ \ (\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in S) = \mathbf{0},$$

but first we need to conduct more valid equality constraints for $\langle\text{POP}\rangle$. An equality constraint $h(\boldsymbol{x}) = 0$ forms a valid equality constraint of $\langle\text{POP}\rangle$ if $h(\boldsymbol{x}) = \sum_{i=1}^{k} q_i(\boldsymbol{x})h_i(\boldsymbol{x})$, $q_i \in \mathbb{R}[\boldsymbol{x}]$ for $i = 1, \dots, k$.

Let the degree of $\langle\text{PSDP}\rangle$ be $2N$ as it is set in (2.1) and (2.2). We have the following system of valid equality constraints for $\langle\text{PSDP}\rangle$:

$$h_i(\boldsymbol{x})\boldsymbol{x}^{\boldsymbol{\alpha}} = 0 \quad (|\boldsymbol{\alpha}| \leq N - \deg h_i, i = 1, \ldots, k), \tag{4.1}$$

which we rewrite as

$$C \ (\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in A) = \mathbf{0} \tag{4.2}$$

for a finite set $A \subset \mathbb{Z}_+^n$ and a coefficient matrix C. Note that the degree of (4.2) is exactly N; if $\deg h_j > N$ for some j $(1 \leq j \leq k)$, then we do not use that h_j in the system of valid equality constraints (4.1) above.

To find a coefficient matrix K satisfying (3.2), separate A into the set B which is contained in S, and the set $\bar{B}$, the other:

$$B = S \cap A,$$
$$\bar{B} = A \backslash B.$$

Instead of finding valid equality constraints on $(\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in S)$, we can find valid equality constraints on $(\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in B)$, since $B \subseteq S$. For appropriate submatrices C_B and $C_{\bar{B}}$ of the coefficient matrix C we rewrite (4.2) as follows.

$$C_B \ (\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in B) + C_{\bar{B}} \ \left(\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \bar{B}\right) = \mathbf{0}. \tag{4.3}$$

We need a matrix Q such that $QC_{\bar{B}} = \mathbf{O}$, since premultiplying such Q to (4.3) we obtain the necessary valid equality constraints:

$$QC_B \ (\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in B) = \mathbf{0}.$$

We adopt a heuristic to compute the necessary valid equality constraints faster. Let $\bar{C}_B$ denote the submatrix of C_B, which consists of rows of C_B where corresponding rows of $C_{\bar{B}}$ are zeros. Immediately we have:

$$\bar{C}_B \ (\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in B) = \mathbf{0}.$$

We use this heuristic in the numerical experiments in Section 5. It works significantly well when the equality constraints of $\langle\text{POP}\rangle$ are sparse, whence C in (4.2) is sparse and the sparsity of C_B and $C_{\bar{B}}$ in (4.3) follows.

5. Numerical Experiments

We have incorporated our method into the sparse SDP relaxation method [3, 15] to test its performance on large scale sparse POPs. The method is implemented on a 2.4 GHz Linux workstation with 7.5GB memory, using SparsePOP [14] version 1.20 and SeDuMi [13, 12] version 1.10. SparsePOP is used to solve $\langle\text{PSDP}\rangle$ or $\overline{\langle\text{PSDP}\rangle}$; internally, it calls SeDuMi to solve $\langle\text{SDP}\rangle$ or $\overline{\langle\text{SDP}\rangle}$. We have originally created the Matlab code to convert $\langle\text{PSDP}\rangle$ into $\overline{\langle\text{PSDP}\rangle}$.

Some test problems are selected from the literature and the others are generated randomly. Table 1 explains the notation used subsequently. Note that `SDPobj` is an effective lower bound of the optimal objective value of the POP, due to Proposition 2.1. Proper box constraints $\boldsymbol{u} \geq \boldsymbol{x} \geq \boldsymbol{l}$ $(\boldsymbol{u} \in \mathbb{R}^n, \boldsymbol{l} \in \mathbb{R}^n)$ are added to all test problems for computational efficiency. Typical parameters of SparsePOP and SeDuMi are explained in Table 2 and Table 3. The numerical solutions computed by SeDuMi are accepted only if SeDuMi's indicators satisfy the constraints in Table 4. For more details on parameters and indicators of SparsePOP and SeDuMi, see the manuals [14, 12, 13].

Table 1: Notation

Notation	Explanation
N	relaxation order, see (2.2)
`EBC/STD`	EBC is to indicate that our method – equalities based contraction – is applied. STD is to indicate that the standard sparse semidefinite relaxation method is applied.
`SDPobj`	optimal objective value of $\langle \mathrm{SDP} \rangle$ or $\overline{\langle \mathrm{SDP} \rangle}$
`POPobj`	objective value at an approximate solution $\boldsymbol{\alpha}$ of $\langle \mathrm{POP} \rangle$
`r.acc`	`‖SDPobj-POPobj‖/max(1,‖SDPobj‖)`
`r.err`	$\left(-\sum_{i=1}^{p} \min(0, g_i(\boldsymbol{\alpha})) + \sum_{i=1}^{k} \lvert h_i(\boldsymbol{\alpha}) \rvert \right)/a$, where a is the maximum absolute values of the coefficients of the polynomials g_i, h_i
`sTime`	CPU time consumed by SeDuMi to solve $\langle \mathrm{SDP} \rangle$ or $\overline{\langle \mathrm{SDP} \rangle}$
`eTime`	CPU time used to convert $\langle \mathrm{PSDP} \rangle$ into $\overline{\langle \mathrm{PSDP} \rangle}$
`nY`	number of variables in $\langle \mathrm{SDP} \rangle$ or $\overline{\langle \mathrm{SDP} \rangle}$
`nEQ`	number of equality constraints in $\langle \mathrm{SDP} \rangle$ or $\overline{\langle \mathrm{SDP} \rangle}$
`nPSD`	number of elements of psd matrices in $\langle \mathrm{SDP} \rangle$ or $\overline{\langle \mathrm{SDP} \rangle}$

Table 2: Parameters for SparsePOP

Parameter	Value	Explanation
`perturbation`	1.0e-5	perturb the objective polynomial
`scalingSW`	1	scale the polynomials to improve numerical stability

Table 3: Parameters for SeDuMi

Parameter	Value	Explanation
`stepdif`	`0`	primal-dual step differentiation is disabled
`par.eps`	`1.e-5`	desired accuracy
`free`	`1`	free variables are placed inside a Lorentz cone

Table 4: Indicators of SeDuMi

Indicator	Value	Explanation
`pinf, dinf`	`0`	solution is primal and dual feasible
`feasratio`	between 1±`1.e-3`	value of the feasibility indicator
`numerr`	`0`	desired accuracy is achieved

Table 5: Test problems from GlobalLib

Problem	#InEqns	#Eqns	n
ex9_1_1	0	13	14
ex9_1_2	0	10	11
ex9_1_8	1	12	15
ex9_2_8	0	6	7

Table 6: Test problems from GlobalLib

Problem	N	method	SDPobj	r.acc	r.err	sTime	eTime
ex9_1_1	2	STD	-1.30E+01	1.37E-16	-6.13E-05	1.40	0
ex9_1_1	2	EBC	-1.30E+01	0.00E+00	-2.36E-04	0.76	0.17
ex9_1_2	2	STD	-1.60E+01	2.22E-16	-8.42E-05	0.52	0
ex9_1_2	2	EBC	-1.60E+01	0.00E+00	-3.99E-05	0.34	0.08
ex9_1_8	2	STD	-3.25E+00	2.73E-16	-1.65E-05	0.55	0
ex9_1_8	2	EBC	-3.25E+00	2.73E-16	-2.75E-05	0.37	0.09
ex9_2_8	2	STD	1.50E+00	1.83E-07	-5.50E-07	0.11	0
ex9_2_8	1	EBC	1.50E+00	8.88E-07	-1.09E-07	0.04	0.01

5.1. Test problems from GlobalLib

We have selected from GlobalLib [17] quadratic optimization problems with a sufficient number of equality constraints, see Table 5. In this table, `#InEqns` denotes the number of inequality constraints (except additional box constraints), `#Eqns` the number of equality constraints, and n the number of variables. These problems also appear in [1]. Numerical results are given in Table 6. We see that the speed-up ratio is about $1.5 \sim 2.8$. In particular, for problem `ex9_2_8`, our method helps solving the problem with relaxation order 1, while the original sparse (or even dense) relaxations only solve the problem with relaxation order 2.

5.2. Linearly constrained quadratic problems

We have generated linearly constrained quadratic optimization problems with structured sparsity, following [15].

$$\begin{aligned} &\text{minimize } \sum_{i=1}^{m} t_i(\boldsymbol{x}), \\ &\text{subject to } \boldsymbol{g}_{i,j}^T \boldsymbol{v}_i(\boldsymbol{x}) = b_{i,j} \; (i = 1, \dots, m, j = 1, \dots, m_i), \\ &\qquad\qquad \boldsymbol{l} \le \boldsymbol{x} \le \boldsymbol{u} \end{aligned}$$

where, given n large enough then m is chosen properly such that:

- $n_1 = 1 < n_2 < n_3 < \cdots < n_m < n_{m+1} = n$,
- $n_{i+1} - n_i = 40$ for $i = 1, \dots, m-1$,
- $\boldsymbol{v}_i(\boldsymbol{x}) = (x_{n_i}, \dots, x_{n_{i+1}})$ for $i = 1, \dots, m$,
- $m_i = n_{i+1} - n_i - 3$ for $i = 1, \dots, m$,
- each $\boldsymbol{g}_{i,j}^T$ has only 3 nonzeros, for $i = 1, \dots, m, j = 1, \dots, m_i$,
- $t_i(\boldsymbol{x})$ is a quadratic polynomial in $x_{n_i}, \dots, x_{n_{i+1}}$ for $i = 1, \dots, m$.

Table 7: Linearly constrained quadratic problems, N = 1

n	method	r.acc	r.err	sTime	eTime	nY	nEQ	nPSD
110	STD	3.15E-07	-2.31E-10	18.87	0	2276	100	4451
	EBC	1.93E-06	-1.89E-11	2.88	0.18	2276	3976	162
210	STD	1.88E-07	-1.82E-10	66.76	0	4425	191	8661
	EBC	1.16E-06	-4.89E-11	9.20	0.42	4425	7747	447
470	STD	3.64E-07	-2.52E-10	411.88	0	10301	432	20181
	EBC	1.36E-05	-1.30E-11	52.28	2.08	10301	18144	669
530	STD	1.11E-08	-7.03E-12	320.92	0	11453	487	22429
	EBC	1.22E-06	-6.71E-12	40.35	2.61	11453	20131	997

Besides, the constants $b_{i,j}$ and bounds $\boldsymbol{l}$ and $\boldsymbol{u}$ are set properly for the existence of a feasible solution. Numerical results are given in Table 7. We see the speed-up ratio is about 7, which is remarkable.

5.3. Concave problems with transportation constraints

We consider the following transportation problem (test problem 8 in [1, Chapter 2])

$$\begin{aligned} \text{minimize} \ & \sum_{i=1}^{m}\sum_{j=1}^{k}(a_{ij}x_{ij} + b_{ij}x_{ij}^2), \\ \text{subject to} \ & \sum_{i=1}^{m} x_{ij} = c_j \ (j = 1, \dots, k), \\ & \sum_{j=1}^{k} x_{ij} = d_i \ (i = 1, \dots, m), \\ & x_{ij} \geq 0 \ (i = 1, \dots, m, j = 1, \dots, k). \end{aligned}$$

The coefficients a_{ij}, b_{ij}, c_j, d_i are integers satisfying

$$\left\{ \begin{array}{c} m \leq k, a_{ij} \in \{200, 201, \dots, 800\}, \\ b_{ij} \in \{-6, -5, -4, -3, -2\}, \\ d_i \in \{3, \dots, 9\}, c_j \in \{2, 3, \dots\}, \sum_{i=1}^{m} d_i = \sum_{j=1}^{k} c_j. \end{array} \right.$$

Our method can solve some problems of this type with relaxation order 1 and satisfying accuracy, while the ordinary relaxation method can not. The numerical results are given in Table 8. We see that for the first test problem with $m = 4, k = 5$, the speed-up ratio is about 2825, since our method solves the problem with relaxation order 1 in 0.13 seconds, though the ordinary method requires relaxation order 2 and 367.28 seconds . For the other test problems, the ordinary relaxation method runs out of memory when trying to solve the problems with relaxation order 2. In contrast, our method can solve the problems with relaxation order 1. These results are conclusive.

6. Concluding Discussions

The equality constraints of a POP can be used to strengthen SDP relaxations for the POP. We first (i) induce valid equality constraints from the original equality constraints of the

Table 8: Concave quadratic optimization problem with transportation constraints

m	k	N	method	r.acc	r.err	sTime	eTime	nY	nEQ	nPSD
4	5	2	STD	1.96E-08	-1.47E-09	367.28	0	4530	389	42042
4	5	1	EBC	1.33E-07	-9.92E-08	0.12	0.01	176	78	661
5	5	1	STD	5.09E+00	-4.49E-10	0.18	0	260	9	1239
5	5	1	EBC	4.10E-06	-2.42E-07	0.19	0.02	260	94	1059
8	8	1	STD	3.98E+00	-3.32E-11	12.55	0	1500	16	8326
8	8	1	EBC	5.07E-06	-2.72E-08	18.31	0.08	1500	272	7846
10	10	1	STD	2.79E+00	-3.97E-09	214.59	0	3545	20	20848
10	10	1	EBC	9.33E-06	-3.68E-07	298.27	2.31	3545	389	20088

POP, and then (ii) use them to contract the psd matrices in the SDP relaxations. Given that the POP is sparse, these tasks are computationally cheap. This leads to notable speedup for some test cases, especially those with a large number of sparse linear equality constraints.

Theoretically, our heuristic method should work more effectively as the POP (1.1) involves more equality constraints. We should mention, however, that if the POP has too many equality constraints then the number of equality constraints in the SDP relaxation problem generated by our heuristic method get very large, causing instability for SDP solvers currently available. On the other hand, if the given POP has only a very few equality constraints, then the number of valid equality constraints added is not enough, and our heuristic does not work effectively. Therefore, SDP solvers more stable to handle many equality constraints would be necessary for our heuristic method to demonstrate its real power.

Acknowledgment

This research was partially supported by Freshwind Information Technology Corporation, HaNoi, Vietnam and Hermes Systems Inc., Tokyo, Japan.

References

[1] C.A. Floudas, P.M. Pardalos, C.S. Adjiman, W.R. Esposito, Z.H. Gümüş, S.T. Harding, J.L. Klepeis, C.A. Meyer, and C.A. Schweiger: *Handbook of test problems in local and global optimization*, vol. 33 of Nonconvex Optimization and its Applications, (Kluwer Academic Publishers, Dordrecht, 1999).

[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] 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.

[4] J.B. Lasserre: An explicit exact SDP relaxation for nonlinear 0-1 programs. In *Integer programming and combinatorial optimization* (Utrecht, 2001), vol. 2081 of Lecture Notes in Computer Science, (Springer, Berlin, 2001), 293–303.

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

[6] J.B. Lasserre: Polynomials nonnegative on a grid and discrete optimization. *Transactions of the American Mathematical Society*, **354** (2002), 631–649.

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

[8] M. Laurent: Semidefinite representations for finite varieties. to appear in *Mathematical Programming*, (2004).

[9] P.A. Parrilo: Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization, PhD thesis, California Institute of Technology, May 2000. Available at `http://www.cds.caltech.edu/~pablo/`.

[10] P.A. Parrilo: An explicit construction of distinguished representations of polynomials nonnegative over finite sets. IfA Technical Report AUT02-02, (2002).

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

[12] I. Polik: *Addendum to the SeDuMi user guide version 1.1.* (2005). Available from `http://sedumi.mcmaster.ca/`.

[13] J.F. Sturm: Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, **11/12** (1999), 625–653.

[14] H. Waki, S. Kim, M. Kojima, and M. Muramatu: SparsePOP: a Sparse Semidefinite Programming Relaxation of Polynomial Optimization Problems, Research Report B-414, (2005). Available from `http://www.is.titech.ac.jp/~kojima/SparsePOP`.

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

[16] D.B. West: *Introduction to graph theory*, (Prentice Hall Upper Saddle River, NJ, 1996).

[17] *GLOBAL Library*. Available from `http://www.gamsworld.org/global/globallib.htm`.

Cong Vo
DaiTri Joint Stock Company,
02/64 AnDuongVuong,
QuiNhon, BinhDinh, Vietnam.
URL: `http://daitri.biz`
Email: `congvo@ams.kuramae.ne.jp`