ISSN 1342-2804

Research Reports on Mathematical and Computing Sciences

A Note on Sparse SOS and SDP Realxations for Polynomial Optimization Problems over Symmetric Cones

Masakazu Kojima and Masakazu Muramatsu

January 2006, B–421

Department of Mathematical and Computing Sciences
Tokyo Institute of Technology

SERIES B: Operations Research

B-421 A Note on Sparse SOS and SDP Relaxations for Polynomial Optimization Problems over Symmetric Cones

Masakazu Kojima⋆ and Masakazu Muramatsu†

January 2006

Abstract.

This short note extends the sparse SOS (sum of squares) and SDP (semidefinite programming) relaxation proposed by Waki, Kim, Kojima and Muramatsu for normal POPs (polynomial optimization problems) to POPs over symmetric cones, and establishes its theoretical convergence based on the recent convergence result by Lasserre on the sparse SOS and SDP relaxation for normal POPs. A numerical example is also given to exhibit its high potential.

Key words.

Polynomial Optimization Problem, Conic Program, Symmetric Cone, Euclidean Jordan Algebra, Sum of Squares, Global Optimization, Semidefinite Program

⋆ Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, 2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152-8552 Japan. Research supported by Grant-in-Aid for Scientific Research on Priority Areas 16016234. *kojima@is.titech.ac.jp*

† Department of Computer Science, The University of Electro-Communications, Chofugaoka, Chofu-Shi, Tokyo 182-8585 Japan. Research supported in part by Grant-in-Aid for Young Scientists (B) 15740054. *muramatu@cs.uec.ac.jp*

1 Main result

Let $\mathbb{R}$ denote the set of real numbers, $(\mathcal{E}, \circ)$ *a Euclidean Jordan algebra* and $\mathcal{E}_+ = \{\boldsymbol{y} \circ \boldsymbol{y} : \boldsymbol{y} \in \mathcal{E}\}$ *the associated symmetric cone*. We use the symbols $\mathbb{R}[\boldsymbol{x}]$ and $\mathcal{E}[\boldsymbol{x}]$ for the set of real valued polynomials and the set of *$\mathcal{E}$-valued polynomials* in a vector variable $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ (or the set of polynomials in $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ with coefficients in $\mathcal{E}$), respectively. Precise definitions on Euclidean Jordan algebras and $\mathcal{E}$-valued polynomials are given in Section 2.1. Given $f \in \mathbb{R}[\boldsymbol{x}]$ and $g \in \mathcal{E}[\boldsymbol{x}]$, an optimization problem

$$\text{minimize} \quad f(\boldsymbol{x}) \text{ subject to } g(\boldsymbol{x}) \in \mathcal{E}_+, \tag{1}$$

is called *a POP (polynomial optimization problem) over a symmetric cone $\mathcal{E}_+$*. This problem was introduced by the authors in the paper [8] as a unified framework to extend the hierarchies of SOS (sum of squares) and SDP (semidefinite programming) relaxation which was proposed by Lasserre [9]. See also Parrilo [11]. The POP (1) over $\mathcal{E}_+$ covers not only a normal POP over the m-dimensional nonnegative orthant

$$\mathbb{R}^m_+ = \left\{\boldsymbol{y} \circ \boldsymbol{y} = (y_1^2, y_2^2, \ldots, y_m^2) : \boldsymbol{y} = (y_1, y_2, \ldots, y_m) \in \mathbb{R}^m\right\},$$

but also a polynomial SDP (semidefinite programming) problem (or a POP over $\mathcal{S}^\ell_+$) where $\mathcal{E} = \mathcal{S}^\ell$ (the set of $\ell \times \ell$ real symmetric matrices) and

$$\begin{aligned} \mathcal{E}_+ \;\; = \;\; & \mathcal{S}^\ell_+ = \left\{\boldsymbol{Y} \circ \boldsymbol{Y} = \boldsymbol{Y}^2 : \boldsymbol{Y} \in \mathcal{S}^\ell\right\} \\ & \text{(the set of } \ell \times \ell \text{ positive semidefinite real symmetric matrices).} \end{aligned}$$

The SOS and SDP relaxation of Lasserre [9] was extended to a polynomial SDP problem by [3, 4, 7], and further to a POP (1) over $\mathcal{E}_+$ by Kojima and Muramatsu [8]. The aim of this short note is to extend a sparse variant of the SOS and SDP relaxation proposed by Waki, Kim, Kojima and Muramatsu [13] (see also [5, 6]) for a normal POP over $\mathbb{R}^m$ to a sparse POP over $\mathcal{E}_+$, and to prove its theoretical convergence based on the recent convergence result by Lasserre [10] on the sparse SOS and SDP relaxation for a normal POP over $\mathbb{R}^m_+$.

Suppose that the Euclidean Jordan algebra $(\mathcal{E}, \circ)$ involved in the POP (1) is represented as the product of m Euclidean Jordan algebras $(\mathcal{E}_1, \circ), (\mathcal{E}_2 \circ), \cdots, (\mathcal{E}_m, \circ)$. Let $\mathcal{E}_{j+} = \{\boldsymbol{y} \circ \boldsymbol{y} : \boldsymbol{y} \in \mathcal{E}_j\}$ be the symmetric cone associated with the Jordan algebra $(\mathcal{E}_j, \circ)$ $(j = 1, 2, \ldots, m)$. For any subset I of $\{1, 2, \ldots, n\}$, we use the notation $\boldsymbol{x}_I = (x_i : i \in I)$ to denote the subvector of $\boldsymbol{x}$ consisting of elements x_i $(i \in I)$. Let I_j be a nonempty subset of $\{1, 2, \ldots, n\}$ $(j = 1, 2, \ldots, m)$. Given $f_j \in \mathbb{R}[\boldsymbol{x}_{I_j}]$ and $g_j \in \mathcal{E}_j[\boldsymbol{x}_{I_j}]$ $(j = 1, 2, \ldots, m)$, we consider the following POP throughout this note:

$$\langle POP \rangle \text{ minimize } \quad \sum_{j=1}^m f_j(\boldsymbol{x}_{I_j}) \quad \text{subject to} \quad g_j(\boldsymbol{x}_{I_j}) \in \mathcal{E}_{j+} \ (j = 1, 2, \ldots, m).$$

If we let $\mathcal{E}_+ = \mathcal{E}_{1+} \times \mathcal{E}_{2+} \times \cdots \times \mathcal{E}_{m+}$ and $g(\boldsymbol{x}) = (g_1(\boldsymbol{x}_{I_1}), g_2(\boldsymbol{x}_{I_2}), \ldots, g_m(\boldsymbol{x}_{I_m}))$, we could reduce $\langle POP \rangle$ to the POP (1) over $\mathcal{E}_+$. In practical computation, however, we can take full advantage in exploiting structured sparsity of $\langle POP \rangle$.

Let

$$
\begin{aligned}
K_j &= \{\boldsymbol{z} \in \mathbb{R}^{|I_j|} : g_j(\boldsymbol{z}) \in \mathcal{E}_{j+}\} \ (j = 1, 2, \ldots, m), \\
K &= \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x}_{I_j} \in K_j \ (j = 1, 2, \ldots, m)\} \\
&= \{\boldsymbol{x} \in \mathbb{R}^n : g_j(\boldsymbol{x}_{I_j}) \in \mathcal{E}_{j+} \ (j = 1, 2, \ldots, m)\} \\
&\quad \text{(the feasible region of } \langle POP \rangle).
\end{aligned}
$$

We impose the following three assumption on $\langle POP \rangle$:

$$
K_j \text{ is nonempty and compact } (j = 1, 2, \ldots, m), \tag{2}
$$

$$
K \text{ is nonempty,} \tag{3}
$$

$$
\left.
\begin{array}{l}
\cup_{j=1}^m I_j = \{1, 2, \ldots, m\}, \\
\forall k \in \{1, 2, \ldots, m-1\} \ \exists s \geq k+1; I_k \cap \left(\cup_{j=k+1}^m I_j\right) \subset I_s.
\end{array}
\right\} \tag{4}
$$

We note that the conditions (2) and the first relation of (4) imply that K is compact.

We now show a simple example of sparse POPs over symmetric cones.

$$
\left.
\begin{array}{ll}
\text{minimize} & \displaystyle\sum_{i=1}^n a_i x_i \\
\text{subject to} & \boldsymbol{A}(x_j, x_{j+1}) \in \mathcal{S}^2_+, \\
& \left(0.3(x_j^3 + x_n) + 1\right) - \|(x_j + \beta_i, x_n)\| \geq 0, \\
& 1 - x_j^2 - x_{j+1}^2 - x_n^2 \geq 0 \ (j = 1, \ldots, n-2).
\end{array}
\right\} \tag{5}
$$

Here

$$
\begin{aligned}
&\boldsymbol{A}(x_j, x_{j+1}) \\
&\quad = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} b_j & c_j \\ c_j & d_j \end{pmatrix} x_j + \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} x_j x_{j+1} + \begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix} x_{j+1},
\end{aligned}
$$

and $a_i, b_j, d_j \in (-1, 0)$, $c_j, \beta_j \in (0, 1)$ are random numbers. We can reformulate this problem as $\langle POP \rangle$ by letting

$$
\begin{aligned}
m &= n - 2, \ I_j = \{j, j+1, n\} \ (j = 1, 2, \ldots, m), \\
\mathcal{E}_j &= \mathcal{S}^2 \times \mathbb{R}^{1+2} \times \mathbb{R} \ (j = 1, 2, \ldots, m), \ \mathcal{E} = \prod_{j=1}^m \mathcal{E}_j, \\
\mathcal{E}_{j+} &= \mathcal{S}^2_+ \times \mathcal{Q}^{1+2}_+ \times \mathbb{R}_+ \ (j = 1, 2, \ldots, m), \ \mathcal{E}_+ = \prod_{j=1}^m \mathcal{E}_{j+}, \\
&g_j(x_j, x_{j+1}, x_n) \\
&\quad = \left(\boldsymbol{A}(x_j, x_{j+1}), \left(0.3(x_j^3 + x_n) + 1, x_j + \beta_i, x_n\right), (1 - x_j^2 - x_{j+1}^2 - x_n^2)\right).
\end{aligned}
$$

Here $\mathcal{Q}^{1+2}_+$ denotes the 3-dimensional second-order cone. We can easily verify that the conditions (2), (3) and (4) are satisfied for the resulting POP over $\mathcal{E}_+$. Specifically, we see

$$
I_k \cap \left(\cup_{j=k+1}^m I_j\right) \subset I_{k+1} \quad \text{for } \forall k \in \{1, 2, \ldots, m-1\}.
$$

In Section 3, we show some numerical results on this example.

The condition (4) is essentially equivalent to the correlative sparsity condition presented in the paper [13] by Waki, Kim, Kojima and Muramatsu for developing a sparse variant of the SOS and SDP relaxation of Lasserre [9] for a normal POP over $\mathbb{R}^m$. This condition was explicitly used in the paper [10] to prove the convergence of the sparse SOS and SDP relaxation. We will add some remarks on the condition (4) in Section 4.

Define

$$\mathcal{C}_j \;=\; \mathbb{R}[\boldsymbol{x}_{I_j}]^2 + g_j \bullet \mathcal{E}_j[\boldsymbol{x}_{I_j}]^2 \subset \mathbb{R}[\boldsymbol{x}_{I_j}] \ (j = 1, 2, \ldots, m) \ \text{ and } \mathcal{C} = \sum_{j=1}^{m} \mathcal{C}_j \subset \mathbb{R}[\boldsymbol{x}],$$

which form cones in $\mathbb{R}[\boldsymbol{x}]$. Here $\mathbb{R}[\boldsymbol{x}_{I_j}]^2$ ($\mathcal{E}_j[\boldsymbol{x}_{I_j}]^2$) is the set of sums of squares of $\mathbb{R}$-valued ($\mathcal{E}_j$-valued) polynomials in $\boldsymbol{x}_{I_j}$ ($j = 1, 2, \ldots, m$) and $\varphi \bullet \psi$ denotes the inner product of $\varphi, \psi \in \mathcal{E}_j[\boldsymbol{x}_{I_j}]$ which is naturally induced from the inner product in $\mathcal{E}_j$. More details will be given in Section 2.1.

Now we present the major result of this note.

Theorem 1 *Assume that the conditions (2), (3), (4) and*

$$\forall j \in \{1, 2, \ldots, m\} \ \exists p_j \in \mathcal{C}_j; \quad \{\boldsymbol{x}_{I_j} : p_j(\boldsymbol{x}_{I_j}) \geq 0\} \ \textit{ is compact} \tag{6}$$

hold. Then any positive polynomial $a \in \sum_{j=1}^{m} \mathbb{R}[\boldsymbol{x}_{I_j}]$ *on the set* K *belongs to the cone* $\mathcal{C}$.

If we take the Euclidean space $\mathbb{R}^{m_j}$ and its nonnegative orthant $\mathbb{R}^{m_j}_{+}$ with some positive integer m_j for $\mathcal{E}_j$ and $\mathcal{E}_{j+}$, respectively, $\langle POP \rangle$ becomes a normal POP. In this case, Theorem 1 is comparable to Corollary 3.8 of Lasserre [10]. We note that the conditions assumed in Theorem 1 are slightly weaker than those in Corollary 3.8. Specifically, it is assumed in Corollary 3.8 that K has nonempty interior.

The theorem above may be regarded as a partial generalization of Putinar's lemma, Lemma 4.1 of [12]; in the original Putinar's lemma, a necessary and sufficient condition for any positive polynomial $a \in \mathbb{R}[\boldsymbol{x}]$ on a compact set $\{\boldsymbol{x} \in \mathbb{R}^n : g'_j(\boldsymbol{x}) \geq \boldsymbol{0} \ (j = 1, 2, \ldots, m)\}$ to have an SOS representation is given, where $g'_j \in \mathbb{R}[\boldsymbol{x}]$ ($j = 1, 2, \ldots, m$), while Theorem 1 provides only a sufficient condition for any positive polynomial $a \in \sum_{j=1}^{m} \mathbb{R}[\boldsymbol{x}_{I_j}]$ on the set K to have an SOS representation.

After giving definitions of some basic materials which have been used above and also necessary for the succeeding discussions, we will prove the major result in Section 2. In Section 3, we will briefly present a sparse SOS relaxation of $\langle POP \rangle$ based on Theorem 1 and numerical results on the sparse SOS relaxation applied to example (5) to show its high potential.

2 Proof

2.1 Euclidean Jordan algebras, $\mathcal{E}$-valued polynomials and their sums of squares

A finite dimensional vector space $\mathcal{E}$ over the field $\mathbb{R}$ of real numbers is called a *Jordan algebra* if a bilinear mapping (multiplication) $\mathcal{E} \times \mathcal{E} \to \mathcal{E}$ denoted by $\circ$ is defined satisfying

(J1) $\boldsymbol{f} \circ \boldsymbol{g} = \boldsymbol{g} \circ \boldsymbol{f}$,

(J2) $[L(\boldsymbol{f} \circ \boldsymbol{f}), L(\boldsymbol{f})] = O$,

where $L(\boldsymbol{f})$ is a linear transformation of $\mathcal{E}$ defined by $L(\boldsymbol{f})\boldsymbol{g} = \boldsymbol{f} \circ \boldsymbol{g}$, and $[A, B] = AB - BA$ for a pair of linear transformations A and B on $\mathcal{E}$. Note that associativity does not hold for $\circ$, *i.e.,* $\boldsymbol{f} \circ (\boldsymbol{g} \circ \boldsymbol{h}) \neq (\boldsymbol{f} \circ \boldsymbol{g}) \circ \boldsymbol{h}$ in general. A Jordan algebra $\mathcal{E}$ is *Euclidean* if an *associative* inner product $\bullet$ is defined, *i.e.,* $(\boldsymbol{f} \circ \boldsymbol{g}) \bullet \boldsymbol{h} = \boldsymbol{f} \bullet (\boldsymbol{g} \circ \boldsymbol{h})$ holds for $\forall\ \boldsymbol{f}, \boldsymbol{g}, \boldsymbol{h} \in \mathcal{E}$. We assume that $\mathcal{E}$ is a Euclidean Jordan algebra having an identity element $\boldsymbol{e}$; $\boldsymbol{e} \circ \boldsymbol{f} = \boldsymbol{f} \circ \boldsymbol{e} = \boldsymbol{f}$ for all $\boldsymbol{f} \in \mathcal{E}$. Such an identity element is unique. We define $\boldsymbol{f}^2 = \boldsymbol{f} \circ \boldsymbol{f}$ and $\boldsymbol{f}^p = \boldsymbol{f}^{p-1} \circ \boldsymbol{f}$ recursively for $p \geq 3$. For more details, see textbooks of Euclidean Jordan algebras, for example, [2].

We denote by $\mathbb{Z}_+$ the set of nonnegative integers. Let $\mathcal{G} \subset \mathbb{Z}_+^n$ be a nonempty finite set. For each $\boldsymbol{\alpha} \in \mathcal{G}$, we assume that a vector $\boldsymbol{f}_{\boldsymbol{\alpha}} \in \mathcal{E}$ is given. Then an $\mathcal{E}$-valued polynomial $f : \mathbb{R}^n \to \mathcal{E}$ is defined by $f(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha} \in \mathcal{G}} \boldsymbol{f}_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}}$, where $\boldsymbol{x}^{\boldsymbol{\alpha}} = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. The set of $\mathcal{E}$-valued polynomials is denoted by $\mathcal{E}[\boldsymbol{x}]$. For example, when $\mathcal{E} = \mathbb{R}$, $\mathbb{R}[\boldsymbol{x}]$ is the set of real-valued polynomials. The *support* of f is defined by $\operatorname{supp} f = \{\, \boldsymbol{\alpha} \in \mathcal{G} : \boldsymbol{f}_{\boldsymbol{\alpha}} \neq \boldsymbol{0} \,\}$. Then f can be expressed uniquely as $f(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha} \in \operatorname{supp} f} \boldsymbol{f}_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}}$. For $r \in \mathbb{Z}_+$, we denote by $\mathcal{E}[\boldsymbol{x}]_r$ a finite dimensional linear subspace of the $\mathcal{E}$-valued polynomials whose degree is less than or equal to r: $\mathcal{E}[\boldsymbol{x}]_r = \{\, f \in \mathcal{E}[\boldsymbol{x}] \ : \ \deg(f) \leq r \,\}$. Specifically, we assume that $\mathcal{E}[\boldsymbol{x}]_0 = \mathcal{E}$.

For $f, g \in \mathcal{E}[\boldsymbol{x}]$, we define a bilinear mapping $\circ$ by

$$(f \circ g)(\boldsymbol{x}) = \left(\sum_{\boldsymbol{\alpha} \in \operatorname{supp} f} \boldsymbol{f}_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} \right) \circ \left(\sum_{\boldsymbol{\beta} \in \operatorname{supp} g} \boldsymbol{g}_{\boldsymbol{\beta}} \boldsymbol{x}^{\boldsymbol{\beta}} \right) = \sum_{\boldsymbol{\alpha} \in \operatorname{supp} f} \sum_{\boldsymbol{\beta} \in \operatorname{supp} g} \left(\boldsymbol{f}_{\boldsymbol{\alpha}} \circ \boldsymbol{g}_{\boldsymbol{\beta}} \right) \boldsymbol{x}^{\boldsymbol{\alpha} + \boldsymbol{\beta}},$$

where $\circ$ on the right-hand side is the multiplication of Jordan algebra $\mathcal{E}$. We denote by e the function of identity: $e(\boldsymbol{x}) = \boldsymbol{e}$ for $\forall\ \boldsymbol{x} \in \mathbb{R}^n$. Then for any $f \in \mathcal{E}[\boldsymbol{x}]$, $e \circ f = f \circ e = f$.

Let $\mathcal{D}$ be a linear subspace of $\mathcal{E}[\boldsymbol{x}]$. Using $\circ$, we define the sums of squares of $\mathcal{E}$-valued polynomials in $\mathcal{D}$ by

$$\mathcal{D}^2 = \left\{ \sum_{i=1}^{q} f_i \circ f_i \ : \ \exists \text{integer } q \geq 1, \ f_i \in \mathcal{D} \right\}.$$

It is easy to verify that $\mathcal{D}^2$ is a convex cone.

Notice that when $\mathcal{D} = \mathcal{E}[\boldsymbol{x}]$, we have the sums of squares of $\mathcal{E}$-valued polynomials

$$\mathcal{E}[\boldsymbol{x}]^2 = \left\{ \sum_{i=1}^{q} f_i \circ f_i \ : \ \exists \text{integer } q \geq 1, \ f_i \in \mathcal{E}[\boldsymbol{x}] \right\},$$

and that when $\mathcal{D} = \mathbb{R}[\boldsymbol{x}]$, we have the sums of squares of real-valued polynomials

$$\mathbb{R}[\boldsymbol{x}]^2 = \left\{ \sum_{i=1}^{q} f_i \circ f_i \ : \ \exists \text{integer } q \geq 1, \ f_i \in \mathbb{R}[\boldsymbol{x}] \right\}.$$

2.2 Proof of Theorem 1

.

Throughout this section, we assume that the conditions (2), (3), (4) and (6) hold. We first observe that each K_j is contained in the compact set $\{\boldsymbol{x}_{I_j} : p_j(\boldsymbol{x}_{I_j}) \geq 0\}$. Hence we can take a positive number M_j such that

$$K_j \subset \{\boldsymbol{x}_{I_j} : p_j(\boldsymbol{x}_{I_j}) \geq 0\} \subset \{\boldsymbol{x}_{I_j} : h_j(\boldsymbol{x}_{I_j}) > 0\} \subset B_j \equiv \{\boldsymbol{x}_{I_j} : h_j(\boldsymbol{x}_{I_j}) \geq 0\}, \tag{7}$$

where $h_j(\boldsymbol{x}_{I_j}) = M_j - \sum_{i \in I_j} x_i^2$. Since h_j is positive on the compact set $\{\boldsymbol{x}_{I_j} : p_j(\boldsymbol{x}_{I_j}) \geq 0\}$, we see, by Putinar's lemma, that

$$h_j \;\in\; \mathbb{R}[\boldsymbol{x}_{I_j}]^2 + \mathbb{R}[\boldsymbol{x}_{I_j}]^2 p_j \subseteq \mathbb{R}[\boldsymbol{x}_{I_j}]^2 + \mathcal{E}_j[\boldsymbol{x}_{I_j}]^2 \bullet g_j. \tag{8}$$

Here the second inclusion above follows from the condition (6). The relations (7) and (8) will be used in the succeeding discussion. Let

$$\begin{aligned} B &= \{\boldsymbol{x} \in \mathbb{R}^n : h_j(\boldsymbol{x}_{I_j}) \geq 0 \ (j = 1, 2, \ldots, m)\} \\ &= \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x}_{I_j} \in B_j \ (j = 1, 2, \ldots, m)\}. \end{aligned}$$

Then $K \subset B$.

Let

$$\lambda_{\max} = \sup \left\{ \begin{array}{l} \text{the maximum absolute eigenvalue of } g_j(\boldsymbol{x}_{I_j}) : \\ \hfill j = 1, \ldots, m, \ \boldsymbol{x} \in B \end{array} \right\}$$

($\lambda_{\max}$ is finite because B is compact). See [2, 8] for the definition of and some properties of eigenvalues of $\boldsymbol{y} \in \mathcal{E}$. We define $\psi_r \in -\sum_{j=1}^{m} g_j \bullet \mathcal{E}_j[\boldsymbol{x}_{I_j}]^2 \subset \sum_{j=1}^{m} \mathbb{R}[\boldsymbol{x}_{I_j}]$ by

$$\psi_r = -\sum_{j=1}^{m} g_j \bullet (\boldsymbol{e}_j - g_j/\lambda_{\max})^{2r},$$

for any nonnegative integer r. Here $\boldsymbol{e}_j$ denote the unit element of the Euclidean Jordan algebra $(\mathcal{E}_j, \circ)$; $\boldsymbol{e}_j \circ \boldsymbol{y} = \boldsymbol{y} \circ \boldsymbol{e}_j = \boldsymbol{y}$ for $\forall \boldsymbol{y} \in \mathcal{E}_j$.

The proof of Theorem 1 relies on the following two lemmas.

Lemma 2 *Suppose that $a \in \mathbb{R}[\boldsymbol{x}]$ is positive on K. Then there exists a positive integer $\bar{r}$ such that $a + \psi_r$ is positive on B for $\forall\ r \geq \bar{r}$.*

Proof : A proof is given in Section 2.3. □

Lemma 3 *Suppose that $a \in \sum_{j=1}^{m} \mathbb{R}[\boldsymbol{x}_{I_j}]$ is positive on B. Then*

$$a \in \sum_{j=1}^{m} \left(\mathbb{R}[\boldsymbol{x}_{I_j}]^2 + \mathbb{R}[\boldsymbol{x}_{I_j}]^2 h_j \right).$$

Proof : The lemma follows directly from Corollary 3.8 of [10]. □

Suppose that $a \in \sum_{j=1}^m \mathbb{R}[\boldsymbol{x}_{I_j}]$ is positive on the set K. By Lemma 2, there exists a positive integer r such that $a + \psi_r \in \sum_{j=1}^m \mathbb{R}[\boldsymbol{x}_{I_j}]$ is positive on B. Now, applying Lemma 3, we see that

$$a + \psi_r \in \sum_{j=1}^m \left(\mathbb{R}[\boldsymbol{x}_{I_j}]^2 + \mathbb{R}[\boldsymbol{x}_{I_j}]^2 h_j\right).$$

By $\psi_r \in -\sum_{j=1}^m g_j \bullet \mathcal{E}_j[\boldsymbol{x}_{I_j}]^2$ and (8), we obtain that

$$a \;\in\; \sum_{j=1}^m \left(\mathbb{R}[\boldsymbol{x}_{I_j}]^2 + \mathbb{R}[\boldsymbol{x}_{I_j}]^2 h_j + g_j \bullet \mathcal{E}_j[\boldsymbol{x}_{I_j}]^2\right) \subseteq \sum_{j=1}^m \left(\mathbb{R}[\boldsymbol{x}_{I_j}]^2 + g_j \bullet \mathcal{E}_j[\boldsymbol{x}_{I_j}]^2\right).$$

This completes the proof of Theorem 1.

2.3 Proof of Lemma 2

For $\forall\, j = 1, 2, \ldots, m$, let $\psi_{rj} = -g_j \bullet (\boldsymbol{e}_j - g_j/\lambda_{\max})^{2r}$. Then $\psi_r = \displaystyle\sum_{j=1}^m \psi_{rj}$.

Lemma 4 *Let $j \in \{1, 2, \ldots, m\}$.*

1. *For any $\epsilon > 0$, there exists a nonnegative integer $\hat{r}$ such that $\psi_{rj}(\boldsymbol{x}_{I_j}) \geq -\epsilon$ for $\forall$ $\boldsymbol{x}_{I_j} \in B_j$ and $r \geq \hat{r}$.*

2. *Suppose that $\tilde{\boldsymbol{x}}_{I_j} \in B_j - K_j$. Then for any $\kappa > 0$, there exist a positive number $\tilde{\delta}$ and a nonnegative integer $\tilde{r}$ such that $\psi_{rj}(\boldsymbol{x}_{I_j}) \geq \kappa$ for $\forall$ $\boldsymbol{x}_{I_j} \in B_j(\tilde{\boldsymbol{x}}_{I_j}, \tilde{\delta}) \cap B_j$ and $\forall$ $r \geq \tilde{r}$, where $B_j(\tilde{\boldsymbol{x}}_{I_j}, \tilde{\delta}) = \{\, \boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x}_{I_j} - \tilde{\boldsymbol{x}}_{I_j}\| < \tilde{\delta} \,\}$.*

Proof : The lemma follows directly from Lemma 4 of [8]. □

Let $\tilde{\boldsymbol{x}}^r$ be a minimizer of $a + \psi_r$ on the compact set B. We show Lemma 2 by proving that there exists a positive integer $\bar{r}$ such that $a(\tilde{\boldsymbol{x}}^r) + \psi_r(\tilde{\boldsymbol{x}}^r) > 0$ for $\forall\, r \geq \bar{r}$. Assume on the contrary that the set $L = \{\, r : a(\tilde{\boldsymbol{x}}^r) + \psi_r(\tilde{\boldsymbol{x}}^r) \leq 0 \,\}$ is infinite. Since $\{\, \tilde{\boldsymbol{x}}^r : r \in L \,\} \subseteq B$, we can take an accumulation point $\tilde{\boldsymbol{x}}^* \in B$ of $\{\, \tilde{\boldsymbol{x}}^r : r \in L \,\}$, and a subsequence $\{\, \tilde{\boldsymbol{x}}^r : r \in L' \,\}$ ($L' \subseteq L$) which converges to $\tilde{\boldsymbol{x}}^* \in B$. In the following, we will prove that there exists $\tilde{r} > 0$ and $\tilde{\delta} > 0$ such that $a(\boldsymbol{x}) + \psi_r(\boldsymbol{x}) > 0$ for $\forall$ $\boldsymbol{x} \in B(\tilde{\boldsymbol{x}}^*, \tilde{\delta}) \cap B$ and $r \geq \tilde{r}$. Because $\tilde{\boldsymbol{x}}^r \in B(\tilde{\boldsymbol{x}}^*, \tilde{\delta}) \cap B$ for sufficiently large $r \in L'$, this contradicts that $\tilde{\boldsymbol{x}}^*$ is an accumulation point of $\{\, \tilde{\boldsymbol{x}}^r : r \in L \,\}$, which establishes the lemma.

We first consider the case where $\tilde{\boldsymbol{x}}^* \in K$. Since K is compact, we can take a positive number ϵ such that $a(\boldsymbol{x}) \geq \epsilon$ for $\forall$ $\boldsymbol{x} \in K$. Then there exists a positive number $\tilde{\delta}$ such that $a(\boldsymbol{x}) \geq \epsilon/2$ for $\forall$ $\boldsymbol{x} \in B(\tilde{\boldsymbol{x}}^*, \tilde{\delta})$. On the other hand, 1 of Lemma 4 implies that there exists a positive number $\tilde{r}$ such that $\psi_{rj}(\boldsymbol{x}_{I_j}) \geq -\epsilon/(4m)$ for $\forall\, r \geq \tilde{r}$ and $\boldsymbol{x} \in B$ ($j = 1, 2, \ldots, m$). Therefore, if $r \geq \tilde{r}$ and $\boldsymbol{x} \in B(\tilde{\boldsymbol{x}}^*, \tilde{\delta}) \cap B$, then $a(\boldsymbol{x}) + \psi_r(\boldsymbol{x}) = a(\boldsymbol{x}) + \sum_{j=1}^m \psi_r(\boldsymbol{x}_{I_j}) \geq \epsilon/4 > 0$.

Next we consider the case where $\tilde{\boldsymbol{x}}^* \in B - K$. Let $\kappa^* = \inf\{\, a(\boldsymbol{x}) : \boldsymbol{x} \in B \,\}$, which is finite because B is compact. By 1 of Lemma 4, a positive number $\tilde{r}$ such that $\psi_{rj}(\boldsymbol{x}_{I_j}) \geq -1/(2m)$ for $\forall\, r \geq \tilde{r}$ and $\boldsymbol{x} \in B$ ($j = 1, 2, \ldots, m$). Since $\tilde{\boldsymbol{x}}^* \in B - K$, there exits $\ell \in \{1, 2, \ldots, n\}$

such that $\tilde{\boldsymbol{x}}^*_{I_\ell} \in B_\ell - K_\ell$. By 2 of Lemma 4, there exists a positive number $\tilde{\delta}$ and a positive integer $\hat{r} \geq \tilde{r}$ such that $\psi_{\ell r}(\boldsymbol{x}_{I_\ell}) \geq -\kappa^* + 1$ for $\forall\ \boldsymbol{x} \in B(\tilde{\boldsymbol{x}}^*, \tilde{\delta}) \cap B$ and $r \geq \hat{r}$. For such $\boldsymbol{x}$ and r, we have

$$a(\boldsymbol{x}) + \psi_r(\boldsymbol{x}) = a(\boldsymbol{x}) + \sum_{j=1}^{m} \psi_{jr}(\boldsymbol{x}) \geq \kappa^* - 1/2 - \kappa^* + 1 = 1/2 > 0$$

This completes the proof of Lemma 2.

3 A sparse SOS and SDP relaxation for $\langle POP \rangle$

We briefly present a sparse variant of Lasserre's SOS and SDP relaxation [9] for $\langle POP \rangle$, and show its theoretical convergence using Theorem 1. Let $r_f = \lceil \deg(f)/2 \rceil$ and $r_j = \lceil \deg g_j/2 \rceil$ $(j = 1, 2, \ldots, m)$. For $r \geq \max\{r_f, r_1, \ldots, r_m\}$, we consider an SOS optimization problem:

$$\langle SOS \rangle_r \left\{ \begin{array}{ll} \text{maximize} & \zeta \\ \text{subject to} & f - \zeta \in \displaystyle\sum_{j=1}^{m} \left(\mathbb{R}[\boldsymbol{x}_{I_j}]^2_r + g_j \bullet \mathcal{E}_j[\boldsymbol{x}_{I_j}]^2_{r-r_j} \right). \end{array} \right.$$

Here the nonnegative integers r_f, r_j $(j = 1, 2, \ldots, m)$ and r, which appear as the subscripts of $\mathcal{E}_j[\boldsymbol{x}_{I_j}]_{r-r_j}$ and $\mathbb{R}[\boldsymbol{x}_{I_j}]_r$, respectively, have been chosen so that the degrees of the polynomials involved in the constraint are balanced. We denote the optimal value of $\langle POP \rangle$ by ζ^*, and the optimal value of $\langle SOS \rangle_r$ by ζ_r.

Theorem 5 *Under the same assumption as Theorem 1,*

$$\zeta_r \leq \zeta_{r+1} \leq \zeta^* \ (r \geq \max\{r_f, r_1, \ldots, r_m\}) \ \textit{ and } \zeta_r \to \zeta^* \ \textit{ as } r \to \infty.$$

Proof : We first note the monotonicity relation

$$\mathcal{E}_j[\boldsymbol{x}_{I_j}]^2_{r-r_j} \subset \mathcal{E}_j[\boldsymbol{x}_{I_j}]^2_{r+1-r_j} \ \text{ and } \mathbb{R}[\boldsymbol{x}_{I_j}]^2_r \subset \mathbb{R}[\boldsymbol{x}_{I_j}]^2_{r+1} \ (j = 1, 2, \ldots, m), \tag{9}$$

which implies that $\zeta_r \leq \zeta_{r+1}$. Let $r \geq \max\{r_f, r_1, \ldots, r_m\}$, and let ζ be a feasible solution of $\langle SOS \rangle_r$. Then there exist $w_j \in \mathcal{E}_j[\boldsymbol{x}_{I_j}]^2_{r-r_j}$ $(j = 1, 2, \ldots, m)$ and $\tilde{w}_j \in \mathbb{R}[\boldsymbol{x}_{I_j}]^2_r$ $(j = 1, 2, \ldots, m)$ such that

$$f(\boldsymbol{x}) - \zeta = \sum_{j=1}^{m} g_j(\boldsymbol{x}_{I_j}) \bullet w_j(\boldsymbol{x}_{I_j}) + \sum_{j=1}^{m} \tilde{w}_j(\boldsymbol{x}_{I_j}) \ \text{ for } \forall \boldsymbol{x} \in \mathbb{R}^n.$$

We also know that

$$\sum_{j=1}^{m} g_j(\boldsymbol{x}_{I_j}) \bullet w_j(\boldsymbol{x}_{I_j}) + \sum_{j=1}^{m} \tilde{w}_j(\boldsymbol{x}_{I_j}) \geq 0 \ \text{ for } \forall \boldsymbol{x} \in K,$$

which implies that $f(\boldsymbol{x}) - \zeta \geq 0$ for $\forall \boldsymbol{x} \in K$. This inequality holds at $\zeta = \zeta_r$. Thus we have shown that $\zeta_r \leq \zeta^*$. Finally we prove $\zeta_r \to \zeta^*$ as $r \to \infty$. Let $\epsilon > 0$. Then

Table 1: Numerical results on the SOS relaxation applied to example (5)

n	cpu time	r	ϵ_{obj}	ϵ_{feas}	the size of $\boldsymbol{A}$ in SeDuMi	# of nonzeros in $\boldsymbol{A}$
200	8.3	2	9.6e-12	0.0	3,974 × 37422	78,012
400	16.0	2	1.5e-11	0.0	7,974 × 75,222	156,812
600	25.7	2	4.0e-12	0.0	11,974 × 113,022	235,612
800	34.8	2	3.2e-12	0.0	15,974 × 150,822	314,412
1000	44.5	2	1.6e-12	0.0	19,974 × 188,622	393,212

$f - \zeta^* + \epsilon \in \sum_{j=1}^m \mathbb{R}[\boldsymbol{x}_{I_j}]$ is positive on K. By Theorem 1 and the monotonicity relation (9), there exists a positive integer p such that

$$f - \zeta^* + \epsilon \in \sum_{j=1}^{m} \left(\mathbb{R}[\boldsymbol{x}_{I_j}]_p^2 + g_j \bullet \mathcal{E}_j[\boldsymbol{x}_{I_j}]_p^2\right).$$

Take $r \geq \max\{r_f, p + r_1, \ldots, p + r_m\})$. Then

$$f - \zeta^* + \epsilon \in \sum_{j=1}^{m} \left(\mathbb{R}[\boldsymbol{x}_{I_j}]_r^2 + g_j \bullet \mathcal{E}_j[\boldsymbol{x}_{I_j}]_{r-r_j}^2\right).$$

Hence $\zeta = \zeta^* - \epsilon$ is a feasible solution of $\langle SOS \rangle_r$. Hence $\zeta^* - \epsilon \leq \zeta_r$. □

We can reformulate $\langle SOS \rangle_r$ as an SDP problem, and we can also apply a sparse SDP relaxation to $\langle POP \rangle$ to derive its dual. We refer to the paper [8] for derivation of those SDP problems, and we only show numerical results on the sparse SOS relaxation applied to example (5) in Table 1. We solved the resulting SDP problems by SeDuMi on Macintosh with 2.5GHz PowerPC G5. The symbols in Table 1 are:

$$\begin{aligned}
\text{cpu time} &= \text{the computational time in seconds for SeDuMi to solve the SDP,}\\
\epsilon_{\text{feas}} &= -\min\{\text{the left side (min.eigen)values over all constraints, } 0\},\\
\epsilon_{\text{obj}} &= \frac{|\text{the lower bound for opt. value} - \text{the approx. opt. value}|}{\max\{1, |\text{the lower bound for opt. value}|\}},\\
r &= \text{the parameter } r \text{ used in } \langle SOS \rangle_r,\\
\boldsymbol{A} &= \text{the constraint matrix of the SDP in SeDuMi format.}
\end{aligned}$$

In Table 1, we observe:

- The sparse SOS relaxation can solve large size problems with dimension up to 1000 less than 1 minute, which we can not solve without exploiting sparsity.
- The data matrix $\boldsymbol{A}$ is large, but very sparse; about 20 nonzero elements in each row of $\boldsymbol{A}$ in average.
- The cpu time and the number of nonzero elements in $\boldsymbol{A}$ increase linearly with n.

4 Concluding remarks

In the condition (4), some I_k can be a subset of I_j $(k \neq j)$ or that some I_k can be non-maximal among the family I_j $(j = 1, 2, \ldots, m)$; hence even $I_k = I_j$ is allowed for some $k \neq j$. If I_k is a subset of I_j $(k \neq j)$, we can redefine

$$\mathcal{E}_j \ \leftarrow \ \mathcal{E}_j \times \mathcal{E}_k, \ \mathcal{E}_{j+} \leftarrow \mathcal{E}_{j+} \times \mathcal{E}_{k+}, \ g_j(\boldsymbol{x}_{I_j}) \leftarrow (g_j(\boldsymbol{x}_{I_j}), g_j(\boldsymbol{x}_{I_k}))$$

so that the resulting POP over $\mathcal{E}_+$ is not only equivalent to $\langle POP \rangle$ but also remains to satisfy all the conditions (2), (3) and (4). Thus we can choose the smallest family with deleting all non-maximal I_j and reconstruct a POP over $\mathcal{E}_+$ which is equivalent to $\langle POP \rangle$. In this case, the resulting family I_j $(j \in J)$ for some $J \subset \{1, 2, \ldots, m\})$ satisfies

$$\text{each } I_k \text{ is maximal among the family, i.e., } I_j \not\subseteq I_k \text{ if } j \neq k. \tag{10}$$

We may impose the condition (10) in addition to (2), (3) and (4) to describe a sparse SOS relaxation in theory, but then we may loose some effectiveness in the sparse SOS relaxation $\langle SOS \rangle_r$ in practice. For example, consider a case where $I_k \subset I_j$ and the degree of g_k is much smaller than the degree of g_j. If we combine I_k into I_j then the SOS relaxation $\langle SOS \rangle_r$ of the resulting POP over $\mathcal{E}_+$ is weaker than the one derived from the original POP over $\mathcal{E}_+$ because the degree of the combined polynomial (g_j, g_k) is much larger than the degree of g_j.

Let $G = (V, E)$ be a graph having a node set $V = \{1, 2, \ldots, n\}$ and an edge set $E = \{\{i, j\} : \{i, j\} \subset I_j \text{ for } \exists j\}$. Then the condition (4) together with (10) implies that the graph G is a chordal graph and that each I_j $(j \in J)$ is corresponding to a maximal clique of G. If we define an $n \times n$ symmetric symbolic matrix $\boldsymbol{M} = (M_{ij})$ with $\star$ designating an unspecified nonzero number and 0 such that $M_{ij} = \star$ iff either $\{i, j\} \in I_k$ for some $k = 1, 2, \ldots, m$ or $i = j$ and $M_{ij} = 0$ otherwise, the condition (4) holds if and only if $\boldsymbol{M}$ allows a sparse Cholesky factorization with no fill-in. The condition (4) is called as the running intersection property in graph theory (see *e.g* [1] for more details).

References

[1] J. R. S. Blair and B. Peyton, *An introduction to chordal graphs and clique trees, in Graph Theory and Sparse Matrix Computation*, A. George, J. R. Gilbert and J. W. H. Liu, eds., Springer-Verlag, New York, 1993, pp. 1?29.

[2] J. Faraut and A. Korányi, *Analysis on Symmetric Cones,* Oxford University Press, New York, NY, 1994.

[3] D. Henrion and J. B. Lasserre, "Convergent relaxations of polynomial matrix inequalities and static output feedback", LAAS-CNRS, 7 Avenue du Colonel Roche, 31077 Toulouse, France (2004).

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

[5] S. Kim, M. Kojima and H.Waki, "Generalized Lagrangian duals and sums of squares relaxations of sparse polynomial optimization problems", *SIAM J. Optim.*, **15** (2005) 697-719 .

[6] M. Kojima, S. Kim and H. Waki, "Sparsity in Sums of Squares of Polynomials", *Mathematical Programming*, **103** (2005) 45-62.

[7] M. Kojima, "Sums of squares relaxations of polynomial semidefinite programs," B-397, Dept. of Mathematical and Computing Sciences Tokyo Institute of Technology, Tokyo 152-8552, Nov. 2003.

[8] M. Kojima and M. Muramatsu, "An Extension of Sums of Squares Relaxations to Polynomial Optimization Problems over Symmetric Cones", Research Report on Mathematical and Computing Sciences B-406, Tokyo Institute of Technology, (2004; Revised 2005).

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

[10] J. B. Lasserre, "Convergent Semidefinite Relaxation in Polynomial Optimization with Sparsity", LAAS-CNRS, 7 Avenue du Colonel Roche, 31077 Toulouse, France (2005).

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

[12] M. Putinar, "Positive polynomials on compact semi-algebraic sets," *Indiana University Mathematics Journal*, **42** (1993) 969-984.

[13] H. Waki, S. Kim, M. Kojima and M. Muramatsu, "Sums of Squares and Semidefinite Programming Relaxations for Polynomial Optimization Problems with Structured Sparsity". To appear in SIAM Journal on Optimization.