数理解析研究所講究録 1409 巻 2005 年 87-100

疎な多項式計画問題に対する半正定値計画緩和[1]

SDP relaxations for sparse Polynomial Optimization Problems

東京工業大学大学院 情報理工学研究科　脇 隼人 (Hayato Waki)
Department of Mathematical and Computing Sciences, Tokyo Institute of Technology
東京工業大学大学院 情報理工学研究科　小島 政和 (Masakazu Kojima)
Department of Mathematical and Computing Sciences, Tokyo Institute of Technology
Department of Mathematics, Ewha Women's University　Sunyoung Kim

1 Introduction

POPs (Polynomial optimization problems or optimization problems with polynomial objective and constraints) represent a broad range of applications in science and engineering.

Recently, an important theoretical development has been made by Lasserre [6] toward achieving optimal values of POPs. According to the paper [4], his method to obtain a sequence of SDP relaxations can be considered as a primal approach. He proved that when the feasible region of the POP is compact, its optimal value can be approximated within any accuracy by the sequence of SDP relaxations. However, the size of an SDP relaxation to be solved in the sequence increases very rapidly as a higher accuracy for an approximation to the optimal value of the POP is required.

The purpose of this paper is to present generalized Lagrangian duals and their SOS (sums of squares) relaxations [7] for sparse POPs. This approach may be regarded as dual of Lasserre's SDP relaxations [6] mentioned above. Instead of selecting nonnegative numbers for Lagrangian multipliers, we choose Lagrangian multipliers to be SOS polynomials satisfying similar sparsity to associated constraint polynomials. Then, we define a generalized Lagrangian dual for a POP over such SOS polynomial multipliers. After a sequence of sets of SOS polynomials is constructed, *e.g.* SOS polynomials of increasing degree, for Lagrangian multipliers, a sequence of Lagrangian duals is obtained. We derive a sufficient condition for the sequence of Lagrangian duals to attain the optimal value of the POP, based on the idea of the penalty function method. For practical purposes, each Lagrangian dual in the sequence is relaxed to an SOS optimization problem, which is further converted into an equivalent SDP. Thus we have sequences of SOS relaxations and SDP relaxations of the POP. The resulting sequence of SDP relaxations corresponds to dual of the sequence of SDP relaxations obtained from the primal approach[6].

An advantage of the dual approach in this paper is that sparsity of objective and constraint polynomials in a POP can be exploited to reduce the size of the SDP relaxations. The size of the SDP relaxations depends on the supports of the polynomials in the dual approach, whereas the size of the SDP relaxations from the primal approach depends on the degree of the polynomials.

Throughout the paper, we use the following notation: Let $\mathbb{R}^n$ and $\mathbb{Z}^n_+$ denote the n-dimensional Euclidean space and the set of n-dimensional nonnegative integer vectors, respectively. Let $f_j : \mathbb{R}^n \to \mathbb{R}$ be a real valued polynomial in $\boldsymbol{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ $(j = 0, 1, 2, \dots, m)$. We denote each polynomial $f_j(\boldsymbol{x})$ as $f_j(\boldsymbol{x}) = \sum_{\boldsymbol{a} \in \mathcal{F}_j} c_j(\boldsymbol{a})\boldsymbol{x}^{\boldsymbol{a}}$, where a nonempty finite subset $\mathcal{F}_j$ of $\mathbb{Z}^n_+$ denotes a support of the polynomial $f_j(\boldsymbol{x})$, $c_j(\boldsymbol{a}) \in \mathbb{R}$

[1]This article is a short version of the paper [3].

and $x^{a} = x_1^{a_1}x_2^{a_2}\cdots x_n^{a_n}$ for every $a = (a_1, a_2, \ldots, a_n) \in \mathcal{F}_j$ $(j = 0, 1, 2, \ldots, m)$ and $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. Let r_j denote the degree of each polynomial $f_j(x)$ $(j = 0, 1, 2, \ldots, m)$; $r_j = \max\{\sum_{i=1}^{n} a_j : a \in \mathcal{F}_j\}$.

2 Polynomial optimization problems and sparsity

We consider the POP (polynomial optimization problem):

$$\text{minimize } f_0(x) \ \text{ subject to } f_j(x) \geq 0 \ (j = 1, 2, \ldots, m). \tag{1}$$

Let us focus on the support $\mathcal{F}_j$ of the polynomial $f_j(x)$ $(j = 0, 1, 2, \ldots, m)$ to describe sparsity of the POP (1). A polynomial $f(x)$ of degree r or its support $\mathcal{F}$ is called sparse if the number of elements in the support $\mathcal{F}$ is much smaller than the number of elements in the support $\mathcal{G}(\xi) \equiv \{a \in \mathbb{Z}_+^n : \sum_{i=1}^{m} a_i \leq \xi\}$ of a general fully dense polynomial of degree ξ. In particular, if the number of indices in $I_+(\mathcal{F}) \equiv \{i : a_i > 0 \text{ for some } a \in \mathcal{F}\}$ is much smaller than n, then $f(x)$ is sparse. We present an example below.

Example 2.1 A box constraint POP. Let $m = n$ and $f_j(x) = 1 - x_j^2$ $(j = 1, 2, \ldots, n)$. In this case, we have $\mathcal{F}_j = \{0, \ 2e^j\}$ $(j = 1, 2, \ldots, n)$. Each $\mathcal{F}_j$ has two elements and $I_+(\mathcal{F}_j) = \{j\}$. Here e^j denotes the jth unit coordinate vector of $\mathbb{R}^n$ with 1 in the jth component and 0 elsewhere.

Another example is given in Section 6 with some preliminary numerical results.
Let F denote the feasible region of the POP (1);

$$F = \{x \in \mathbb{R}^n : f_j(x) \geq 0 \ (j = 1, 2, \ldots, m)\}.$$

Throughout the paper, we assume that F is nonempty and bounded. Then, the POP (1) has a finite optimal value ζ^* at an optimal solution $x^* \in F$. In what it follows, we further need a bound $\rho > 0$ to be known explicitly for the feasible region. We are concerned with the following cases:

- $F \subset C_\rho \equiv \{x \in \mathbb{R}^n : \rho^2 - x_i^2 \geq 0 \ (i = 1, 2, \ldots, n)\}$ or
- $F \subset B_\rho \equiv \{x \in \mathbb{R}^n : \rho^2 - x^T x \geq 0\}$.

If $F \subset C_\rho$ holds then the POP (1) is equivalent to

$$\text{minimize } f_0(x) \ \text{ subject to } f_j(x) \geq 0 \ (j = 1, 2, \ldots, m) \ \text{ and } x \in C_\rho. \tag{2}$$

Example 2.1 is a special case of the POP (2) where we take $m = 0$ and $\rho = 1$. If $F \subset B_\rho$ is satisfied, then we have $F \subset C_\rho$; hence the two POPs (1) and (2) are equivalent to

$$\text{minimize } f_0(x) \ \text{ subject to } f_j(x) \geq 0 \ (j = 1, 2, \ldots, m) \ \text{ and } x \in B_\rho. \tag{3}$$

In the next section, we present a (generalized) Lagrangian function, a Lagrangian relaxation and a Lagrangian dual for each of the POPs (2) and (3). For both POPs, the Lagrangian dual converges to the optimal value ζ^* of the original POP (1) under a moderate assumption. But, only the SOS relaxation derived from the POP (3) is guaranteed to converge to ζ^* [3], while the SOS relaxation from the POP (2) inherits more sparsity of the original POP (1) than the one from the POP (3).

3 Generalized Lagrangian duals

3.1 Lagrangian functions

Let $\overline{\Sigma}$ denote the set of sums of squares of polynomials in $\boldsymbol{x} \in \mathbb{R}^n$;

$$\overline{\Sigma} = \left\{ \sum_{i=1}^{k} \chi_i(\boldsymbol{x})^2 : \begin{array}{l} \chi_i \text{ is a polynomial in } \boldsymbol{x} \in \mathbb{R}^n \ (i = 1, 2, \ldots, k) \\ \text{and } k \text{ is any finite positive integer} \end{array} \right\}.$$

We define two types of (generalized) Lagrangian functions $L_B : B_\rho \times \overline{\Sigma}^m \to \mathbb{R}$ for POP (3) and $L_C : \mathbb{R}^n \times \overline{\Sigma}^{m+n} \to \mathbb{R}$ for POP (2):

$$\begin{aligned} L_B(\boldsymbol{x}, \varphi) &= f_0(\boldsymbol{x}) - \sum_{j=1}^{m} \varphi_j(\boldsymbol{x}) f_j(\boldsymbol{x}) \\ L_C(\boldsymbol{x}, \varphi, \psi) &= f_0(\boldsymbol{x}) - \sum_{j=1}^{m} \varphi_j(\boldsymbol{x}) f_j(\boldsymbol{x}) - \sum_{i=1}^{n} \psi_i(\boldsymbol{x})(\rho^2 - x_i^2). \end{aligned}$$

Here $\overline{\Sigma}^\ell$ denotes the Cartesian product of ℓ-tuples of $\overline{\Sigma}$;

$$\overline{\Sigma}^\ell = \left\{ (\varphi_1, \varphi_2, \ldots, \varphi_\ell) : \varphi_j \in \overline{\Sigma} \ (j = 1, 2, \ldots, \ell) \right\} \ (\ell = m \ \text{ or } m + n).$$

Let $(\varphi, \psi) \in \overline{\Sigma}^{m+n}$. Then

$$\left.\begin{array}{lll} L_C(\boldsymbol{x}, \varphi, \psi) \leq f_0(\boldsymbol{x}) & \text{if } \boldsymbol{x} \in F \bigcap C_\rho, \\ L_B(\boldsymbol{x}, \varphi) \leq f_0(\boldsymbol{x}) & \text{if } \boldsymbol{x} \in F \bigcap B_\rho, \\ L_C(\boldsymbol{x}, \varphi, \psi) \leq L_B(\boldsymbol{x}, \varphi) & \text{if } \boldsymbol{x} \in B_\rho. \end{array}\right\} \tag{4}$$

3.2 Lagrangian relaxations and duals

We introduce a (generalized) Lagrangian relaxation of the POP (2):

$$L_C^*(\varphi, \psi) = \inf \{ L_C(\boldsymbol{x}, \varphi, \psi) : \boldsymbol{x} \in \mathbb{R}^n \}$$

for each fixed $(\varphi, \psi) \in \overline{\Sigma}^{m+n}$, and a (generalized) Lagrangian relaxation of the POP (3):

$$L_B^*(\varphi) = \inf \{ L_B(\boldsymbol{x}, \varphi) : \boldsymbol{x} \in B_\rho \}$$

for each fixed $\varphi \in \overline{\Sigma}^m$. Let $(\varphi, \psi) \in \overline{\Sigma}^{m+n}$. By (4), we see that

$$\left.\begin{array}{l} L_C^*(\varphi, \psi) \leq \zeta^* = \min\{ f_0(\boldsymbol{x}) : \boldsymbol{x} \in F \} \text{ if } F \subset C_\rho, \\ L_C^*(\varphi, \psi) \leq L_B^*(\varphi) \leq \zeta^* \text{ if } F \subset B_\rho. \end{array}\right\} \tag{5}$$

For every $(\boldsymbol{\Sigma}, \boldsymbol{\Xi}) \subset \overline{\Sigma}^{m+n}$, we define a (generalized) Lagrangian dual of the POP (2):

$$\text{maximize} \quad L_C^*(\varphi, \psi) \quad \text{subject to} \quad (\varphi, \psi) \in \boldsymbol{\Sigma} \times \boldsymbol{\Xi}. \tag{6}$$

For every $\boldsymbol{\Sigma} \subset \overline{\Sigma}^m$, we define a (generalized) Lagrangian dual of the POP (3):

$$\text{maximize} \quad L_B^*(\varphi) \quad \text{subject to} \quad \varphi \in \boldsymbol{\Sigma}. \tag{7}$$

Let $L^*_C(\Sigma \times \Xi)$ and $L^*_B(\Sigma)$ denote the optimal values of (6) and (7), respectively;

$$L^*_C(\Sigma \times \Xi) = \sup_{(\varphi,\psi)\in\Sigma\times\Xi} L^*_C(\varphi,\psi) \text{ and } L^*_B(\Sigma) = \sup_{\varphi\in\Sigma} L^*_B(\varphi).$$

It follows from (5) that

$$\left.\begin{array}{l} L^*_C(\Sigma \times \Xi) \leq \zeta^* \text{ if } F \subset C_\rho, \\ L^*_C(\Sigma \times \Xi) \leq L^*_B(\Sigma) \leq \zeta^* \text{ if } F \subset B_\rho \end{array}\right\} \tag{8}$$

holds for every $(\Sigma, \Xi) \subset \overline{\Sigma}^{m+n}$.

Assume that $F \subset C_\rho$. Then, the two POPs (1) and (2) are equivalent. If we restrict (φ, ψ) to the nonnegative orthant $\mathbb{R}^{m+n}_+$ of $\mathbb{R}^{m+n}$, $L_C(x, \varphi, \psi)$ becomes the standard Lagrangian function for the POP (2).

As we take a larger set $\Sigma \times \Xi \subset \overline{\Sigma}^{m+n}$, the duality gap between $L^*_C(\Sigma \times \Xi)$ and ζ^* is expected to decrease. We regard $(\varphi, \psi) \in \overline{\Sigma}^{m+n}$ as "a penalty parameter", and

$$\Phi_C(x, \varphi, \psi) = -\sum_{j=1}^{m} \varphi_j(x) f_j(x) - \sum_{i=1}^{n} \psi_i(x)(\rho^2 - x_i^2)$$

(the terms added to the objective function $f_0(x)$
in the construction of the Lagrangian function $L_C(x, \varphi, \psi)$)

as "a penalty function" with a choice of penalty parameters $(\varphi, \psi) = (\varphi^p, \psi^p) \in \overline{\Sigma}^{m+n}$ $(p \in \mathbb{Z}_+)$ such that

$$\begin{array}{l} \text{if } x \in F \text{ then } \Phi(x, \varphi^p, \psi^p) \to 0 \text{ as } p \to \infty, \\ \text{if } x \notin F \text{ then } \Phi(x, \varphi^p, \psi^p) \to \infty \text{ as } p \to \infty. \end{array}$$

It is shown that the Lagrangian function $L_C(x, \varphi, \psi) = f_0(x) + \Phi_C(x, \varphi, \psi)$ with the penalty parameter $(\varphi, \psi) = (\varphi^p, \psi^p) \in \overline{\Sigma}^{m+n}$ has a global minimizer x^p over $\mathbb{R}^n$ with the objective value $f_0(x^p) \to \zeta^*$ as $p \to \infty$ and that $\{x^p\}$ has an accumulation point in the optimal solution set of the POP (2).

In order to describe this observation precisely, we use some notation. Take a real number $\gamma \geq 1$ such that

$$\begin{array}{rcl} |f_j(x)/\gamma| & \leq & 1 \text{ if } \|x\|_\infty \leq \sqrt{2}\rho, \\ |f_j(x)/\gamma| & \leq & \|x/\rho\|^r_\infty \text{ if } \|x\|_\infty \geq \sqrt{2}\rho \end{array}$$

$(j = 0, 1, 2, \ldots, m)$, where $r = \max\{r_0, r_1, \ldots, r_m\}$.

Letting (φ^p, ψ^p) be

$$\begin{array}{rcl} \varphi^p_j(x) & = & (1 - f_j(x)/\gamma)^{2p} \ (j = 1, 2, \ldots, m, \ p \in \mathbb{Z}_+), \\ \varphi^p(x) & = & (\varphi^p_1(x), \varphi^p_2(x), \ldots, \varphi^p_m(x)) \ (p \in \mathbb{Z}_+), \\ \psi^p_i(x) & = & ((m+2)\gamma/\rho^2) \, (x_i/\rho)^{2r(p+1)} \ (i = 1, 2, \ldots, n, \ p \in \mathbb{Z}_+), \\ \psi^p(x) & = & (\psi^p_1(x), \psi^p_2(x), \ldots, \psi^p_n(x)) \ (p \in \mathbb{Z}_+), \end{array}$$

the following theorem holds:

Theorem 3.1 *[3] Assume that* $\Sigma \times \Xi \subset \overline{\Sigma}^{m+n}$ *contains an infinite subsequence of* $\{(\varphi^p, \psi^p)\ (p \in \mathbb{Z}_+)\}$. *Then* $L^*_C(\Sigma \times \Xi) = \zeta^*$ *if* $F \subset C_\rho$, *and* $L^*_C(\Sigma \times \Xi) = L^*_B(\Sigma) = \zeta^*$ *if* $F \subset B_\rho$.

3.3 Construction of $\Sigma \times \Xi$ satisfying the assumption of Theorem 3.1

For every nonempty subset $\mathcal{A}$ of $\mathbb{Z}^n_+$, we define

$$\Sigma(\mathcal{A}) = \left\{ \sum_{i=1}^{k} \chi_i(\boldsymbol{x})^2 : \begin{array}{l} \chi_i \text{ is a polynomial in } \boldsymbol{x} \in \mathbb{R}^n \text{ with a support in } \mathcal{A} \\ (i = 1, 2, \ldots, k) \text{ and } k \text{ is any finite positive integer} \end{array} \right\}.$$

Suppose that $\mathcal{A}^q_j$ $(j = 1, 2, \ldots, m,\ q \in \mathbb{Z}_+)$ and $\mathcal{B}^q_i$ $(i = 1, 2, \ldots, n,\ q \in \mathbb{Z}_+)$ are nonempty finite subsets of $\mathbb{Z}^n_+$ such that

$$\varphi_j^{\lambda(q)} \in \Sigma(\mathcal{A}^q_j) \text{ and } \psi_i^{\lambda(q)}(\boldsymbol{x}) \in \Sigma(\mathcal{B}^q_i) \text{ if } q \geq q^* \tag{9}$$

for some $q^* \in \mathbb{Z}_+$ and some mapping λ from $\mathbb{Z}_+$ into itself satisfying

$$\lambda(q) \leq \lambda(q+1) \ (q \in \mathbb{Z}_+) \text{ and } \lim_{q\to\infty} \lambda(q) = \infty.$$

Let

$$\boldsymbol{\Sigma}^q = \prod_{j=1}^{m} \Sigma(\mathcal{A}^q_j),\ \boldsymbol{\Xi}^q = \prod_{i=1}^{n} \Sigma(\mathcal{B}^q_i),\ \boldsymbol{\Sigma}^\infty = \bigcup_{q\in\mathbb{Z}_+} \boldsymbol{\Sigma}^q \text{ and } \boldsymbol{\Xi}^\infty = \bigcup_{q\in\mathbb{Z}_+} \boldsymbol{\Xi}^q.$$

By construction, we see that

$$(\boldsymbol{\varphi}^{\lambda(q)}, \boldsymbol{\psi}^{\lambda(q)}) \in \boldsymbol{\Sigma}^q \times \boldsymbol{\Xi}^q \ (q \geq q^*).$$

Hence $\boldsymbol{\Sigma}^\infty \times \boldsymbol{\Xi}^\infty$ contains an infinite subsequence $\{(\boldsymbol{\varphi}^{\lambda(q)}, \boldsymbol{\psi}^{\lambda(q)})\ (q \geq q^*)\}$. Therefore $\boldsymbol{\Sigma} \times \boldsymbol{\Xi} = \boldsymbol{\Sigma}^\infty \times \boldsymbol{\Xi}^\infty$ satisfies the assumption of Theorem 3.1.

We give some examples of $\mathcal{A}^q_j$ $(j = 1, 2, \ldots, m,\ q \in \mathbb{Z}_+)$ and $\mathcal{B}^q_i$ $(i = 1, 2, \ldots, n,\ q \in \mathbb{Z}_+)$ satisfying the assumption (9).

Example 3.2 For every $j = 1, 2, \ldots, m,\ i = 1, 2, \ldots, n,\ q \in \mathbb{Z}_+$, let

$$\begin{aligned} &\mathcal{A}^0_j = \{\boldsymbol{0}\},\ \mathcal{A}^1_j = \{\boldsymbol{0}\} \bigcup \mathcal{F}_j,\ \mathcal{A}^{q+1}_j = \left\{\boldsymbol{a} + \boldsymbol{b} : \boldsymbol{a} \in \mathcal{A}^q_j,\ \boldsymbol{b} \in \mathcal{A}^1_j\right\} \ (q \geq 1), \\ &\mathcal{B}^q_i = \{k\boldsymbol{e}^i : k = 0, 1, 2, \ldots, (q+1)r\}. \end{aligned}$$

Example 3.3 For every $j = 1, 2, \ldots, m,\ i = 1, 2, \ldots, n,\ q \in \mathbb{Z}_+$, let

$$\begin{aligned} &\mathcal{A}^0_j = \{\boldsymbol{0}\},\ \mathcal{A}^1_j = \{\boldsymbol{0}\} \bigcup \left\{\boldsymbol{e}^k : k \in I_+(\mathcal{F}_j)\right\}, \\ &\mathcal{A}^{q+1}_j = \left\{\boldsymbol{a} + \boldsymbol{b} : \boldsymbol{a} \in \mathcal{A}^q_j,\ \boldsymbol{b} \in \mathcal{A}^1_j\right\} \ (q \geq 1), \\ &\mathcal{B}^q_i = \{k\boldsymbol{e}^i : k = 0, 1, 2, \ldots, q+1\} \ (q \in \mathbb{Z}_+). \end{aligned}$$

Here $I_+(\mathcal{F}_j) = \{i : a_i > 0 \text{ for some } \boldsymbol{a} \in \mathcal{F}_j\}$.

In all the examples, both $\mathcal{A}^q_j$ and $\mathcal{B}^q_i$ expand monotonically as q increases, and for any $\bar{p} \in \mathbb{Z}_+$ there exists $\bar{q} \in \mathbb{Z}_+$ such that $\varphi^p_j(\boldsymbol{x}) \in \Sigma(\mathcal{A}^q_j)$ and $\psi^p_i(\boldsymbol{x}) \in \Sigma(\mathcal{B}^q_i)$ for all $p \leq \bar{p}$ and $q \geq \bar{q}$. It should be noted that if $\mathcal{F}_j$ is sparse then $\mathcal{A}^q_j$ remains sparse in Example 3.2. The choice of $\mathcal{A}^q_j$ in Example 3.3 may be also reasonable when the number of the indices $I_+(\mathcal{F}_j)$ is smaller than n.

4 Numerical methods for generalized Lagrangian duals

In the following three subsections, we discuss numerical methods for generalized Lagrangian duals (6) and those for (7) are discussed in [3]. These subsections include how $L_C^*(\Sigma^\infty \times \Xi^\infty)$ can be approximated numerically. Throughout this section, we assume that $F \subset C_\rho$, so that $L_C^*(\Sigma^\infty \times \Xi^\infty) = \zeta^*$.

4.1 Approximation of generalized Lagrangian duals

We introduce a sequence of subproblems of the Lagrangian dual (6) for $q \in \mathbb{Z}_+$.

$$\text{maximize} \quad L_C^*(\varphi, \psi) \quad \text{subject to} \quad (\varphi, \psi) \in \Sigma^q \times \Xi^q. \tag{10}$$

We donote an optimal value of (10) as $L_C^*(\Sigma^q, \Xi^q)$.

Lemma 4.1 *[3]*

(a) $L_C^*(\Sigma^q \times \Xi^q) \leq L_C^*(\Sigma^\infty \times \Xi^\infty)$ $(q \in \mathbb{Z}_+)$.

(b) For any $\epsilon > 0$, there exists a nonnegative integer $\bar{q}$ such that

$$L_C^*(\Sigma^\infty \times \Xi^\infty) - \epsilon \leq L_C^*(\Sigma^q \times \Xi^q) \ (q \geq \bar{q}).$$

4.2 Sums of square relaxations

Let $q \in \mathbb{Z}_+$ be fixed throughout this subsection. We can rewrite the problems in (10) as

$$\left.\begin{array}{ll} \text{maximize} & \zeta \\ \text{subject to} & L_C(\boldsymbol{x}, \varphi, \psi) - \zeta \geq 0 \ (\forall \boldsymbol{x} \in \mathbb{R}^n), \\ & (\varphi, \psi) \in \Sigma^q \times \Xi^q. \end{array}\right\} \tag{11}$$

Note that $\boldsymbol{x} \in \mathbb{R}^n$ is not a vector variable but it serves as an index vector for infinite number of inequality constrains $L_C(\boldsymbol{x}, \varphi, \psi) - \zeta \geq 0$ $(\forall \boldsymbol{x} \in \mathbb{R}^n)$. Replacing the inequality constraints $L_C(\boldsymbol{x}, \varphi, \psi) - \zeta \geq 0$ $(\forall \boldsymbol{x} \in \mathbb{R}^n)$ by a sum of squares condition $L_C(\boldsymbol{x}, \varphi, \psi) - \zeta \in \overline{\Sigma}$ in (11), we obtain an SOSOP (sums of squares optimization problem).

$$\left.\begin{array}{ll} \text{maximize} & \zeta \\ \text{subject to} & L_C(\boldsymbol{x}, \varphi, \psi) - \zeta = \varphi_0(\boldsymbol{x}) \ (\forall \boldsymbol{x} \in \mathbb{R}^n), \\ & (\varphi, \psi) \in \Sigma^q \times \Xi^q, \ \varphi_0(\boldsymbol{x}) \in \overline{\Sigma}. \end{array}\right\} \tag{12}$$

Let ζ_C^q denote the optimal value of the SOSOP (12);

$$\zeta_C^q = \sup\left\{\zeta : \begin{array}{l} L_C(\boldsymbol{x}, \varphi, \psi) - \zeta = \varphi_0(\boldsymbol{x}) \ (\forall \boldsymbol{x} \in \mathbb{R}^n), \\ (\varphi, \psi) \in \Sigma^q \times \Xi^q, \ \varphi_0(\boldsymbol{x}) \in \overline{\Sigma} \end{array}\right\}.$$

If $(\zeta, \varphi, \psi, \varphi_0)$ is a feasible solution of the SOSOP (12), then (ζ, φ, ψ) is a feasible solution of the problem (11). It follows that $\zeta_C^q \leq L_C^*(\Sigma^q \times \Xi^q)$. Although neither $\zeta_C^q = L_C^*(\Sigma^q \times \Xi^q)$ nor the convergence of ζ_C^q to $L_C^*(\Sigma^\infty \times \Xi^\infty)$ as $q \to \infty$ is guaranteed, we can solve the SOSOP (12) as we show in the next subsection while the problem (11) is difficult to solve in general.

4.3 Reduction to SDPs

Let us fix $q \in \mathbb{Z}_+$ throughout this subsection. We show how to solve the SOSOP (12) as an SDP (semidefinite program). If we rewrite the constraint $(\varphi, \psi) \in \Sigma^q \times \Xi^q$ for each component, we have

$$\varphi_j(\boldsymbol{x}) \in \Sigma(\mathcal{A}_j^q) \ (j = 1, 2, \ldots, m) \ \text{ and } \psi_i(\boldsymbol{x}) \in \Sigma(\mathcal{B}_i^q) \ (i = 1, 2, \ldots, n).$$

Notice that finite supports $\mathcal{A}_j^q$ and $\mathcal{B}_i^q$ are given for generating variable polynomials $\varphi_j(\boldsymbol{x})$ and $\psi_i(\boldsymbol{x})$ $(j = 1, 2, \ldots, m,\ i = 1, 2, \ldots, n)$. But, no finite support is specified for the variable polynomial $\varphi_0(\boldsymbol{x})$. The first step for constructing an SDP is to find an appropriate finite set $\mathcal{G} \subset \mathbb{Z}_+^n$ so that $\varphi_0(\boldsymbol{x})$ can be chosen from $\Sigma(\mathcal{G})$.

To choose such a $\mathcal{G} \subset \mathbb{Z}_+^n$, we focus on the support of the left hand side polynomial $L_C(\boldsymbol{x}, \varphi, \psi) - \zeta$ of the equality constraint in the SOSOP (12). From the support $\mathcal{F}_0$ of the objective polynomial function $f_0(\boldsymbol{x})$, the support of each term $\varphi_j(\boldsymbol{x}) f_j(\boldsymbol{x})$

$$\widehat{\mathcal{A}}_j^q \equiv \left\{ \boldsymbol{a} + \boldsymbol{b} + \boldsymbol{c} : \boldsymbol{a} \in \mathcal{A}_j^q,\ \boldsymbol{b} \in \mathcal{A}_j^q,\ \boldsymbol{c} \in \mathcal{F}_j \right\}$$

$(j = 1, 2, \ldots, m)$ and the support of term $\psi_i(\boldsymbol{x})(\rho^2 - x_i^2)$

$$\widehat{\mathcal{B}}_i^q \equiv \left\{ \boldsymbol{a} + \boldsymbol{b} + \boldsymbol{c} : \boldsymbol{a} \in \mathcal{B}_i^q,\ \boldsymbol{b} \in \mathcal{B}_i^q,\ \boldsymbol{c} \in \{\boldsymbol{0}, 2\boldsymbol{e}^i\} \right\}$$

$(i = 1, 2, \ldots, n)$, we know that the support of $L_C(\boldsymbol{x}, \varphi, \psi) - \zeta$ becomes

$$\mathcal{F}_L = \mathcal{F}_0 \bigcup \{\boldsymbol{0}\} \bigcup \left(\bigcup_{j=1}^{m} \widehat{\mathcal{A}}_j^q \right) \bigcup \left(\bigcup_{i=1}^{n} \widehat{\mathcal{B}}_i^q \right).$$

Here $\{\boldsymbol{0}\}$ stands for the support of the term ζ.

By Theorem 1 of [9], we can use

$$\mathcal{G}^0 = \left(\text{the convex hull of } \left\{ \boldsymbol{a}/2 : \begin{array}{l} \boldsymbol{a} \in \mathcal{F}_L, \\ \text{every } a_i \text{ is even } (i = 1, 2, \ldots, n) \end{array} \right\} \right) \bigcap \mathbb{Z}_+^n.$$

for such a support $\mathcal{G}$ that $\varphi_0(\boldsymbol{x})$ can be chosen from $\Sigma(\mathcal{G})$. We can further apply a method proposed recently by the authors [5] for reducing the size of $\mathcal{G}^0$ to obtain the smallest support $\mathcal{G}^*$ in a class of supports including $\mathcal{G}^0$. See the paper [5] for more details.

To transform the SOSOP (12) into an SDP, we need some notations and symbols. Let $\mathcal{F} \in \mathbb{Z}_+^n$ be a nonempty finite set. Let $|\mathcal{F}|$ denote the cardinality of $\mathcal{F}$ and $\mathbb{R}(\mathcal{F})$ the $|\mathcal{F}|$-dimensional Euclidean space whose coordinates are indexed by $\boldsymbol{a} \in \mathcal{F}$. Although the order of the coordinates is not relevant in the succeeding discussions, we may assume that the coordinates are arranged according to the lexicographical order. Each element of $\mathbb{R}(\mathcal{F})$ is denoted as $\boldsymbol{v} = (v_{\boldsymbol{a}} : \boldsymbol{a} \in \mathcal{F})$. We use the symbol $\mathcal{S}(\mathcal{F})_+$ for the set of $|\mathcal{F}| \times |\mathcal{F}|$ symmetric positive semidefinite matrices with coordinates $\boldsymbol{a} \in \mathcal{F}$; each $\boldsymbol{V} \in \mathcal{S}(\mathcal{F})_+$ has elements $V_{\boldsymbol{ab}}$ $(\boldsymbol{a} \in \mathcal{F},\ \boldsymbol{b} \in \mathcal{F})$ such that $V_{\boldsymbol{ab}} = V_{\boldsymbol{ba}}$ and that $\boldsymbol{w}^T \boldsymbol{V} \boldsymbol{w} = \sum_{\boldsymbol{a} \in \mathcal{F}} \sum_{\boldsymbol{b} \in \mathcal{F}} V_{\boldsymbol{ab}} w_{\boldsymbol{a}} w_{\boldsymbol{b}} \geq 0$ for every $\boldsymbol{w} = (w_{\boldsymbol{a}} : \boldsymbol{a} \in \mathcal{F}) \in \mathbb{R}(\mathcal{F})$. For every $\boldsymbol{x} \in \mathbb{R}^n$, let $\boldsymbol{u}(\boldsymbol{x}, \mathcal{F}) = (\boldsymbol{x}^{\boldsymbol{a}} : \boldsymbol{a} \in \mathcal{F})$ be a column vector consisting of elements $\boldsymbol{x}^{\boldsymbol{a}}$ $(\boldsymbol{a} \in \mathcal{F})$.

Lemma 4.2 *[1, 5, 7, 8] Let $\mathcal{F}$ be a nonempty finite subset of $\mathbb{Z}_+^n$. A polynomial $\varphi(\boldsymbol{x})$ is contained in $\Sigma(\mathcal{F})$ if and only if there exists a $\boldsymbol{V} \in \mathcal{S}(\mathcal{F})_+$ such that*

$$\varphi(\boldsymbol{x}) = \boldsymbol{u}(\boldsymbol{x}, \mathcal{F})^T \boldsymbol{V} \boldsymbol{u}(\boldsymbol{x}, \mathcal{F}) = \sum_{\boldsymbol{a} \in \mathcal{F}} \sum_{\boldsymbol{b} \in \mathcal{F}} V_{\boldsymbol{ab}} \boldsymbol{x}^{\boldsymbol{a}+\boldsymbol{b}}. \tag{13}$$

Applying Lemma 4.2 to the polynomials $\varphi_j(\boldsymbol{x}) \in \Sigma(\mathcal{A}_j^{(\lambda_j)})$ $(j = 1, 2, \ldots, m)$, $\psi_i(\boldsymbol{x}) \in \Sigma(\mathcal{B}_i^{(\mu_i)})$ $(i = 1, 2, \ldots, n)$ and $\varphi_0(\boldsymbol{x}) \in \Sigma(\mathcal{G}^*)$, we represent as follows:

$$\begin{aligned}
\varphi_j(\boldsymbol{x}) &= \boldsymbol{u}(\boldsymbol{x}, \mathcal{A}_j^q)^T \boldsymbol{V}^j \boldsymbol{u}(\boldsymbol{x}, \mathcal{A}_j^q), \ \boldsymbol{V}^j \in \mathcal{S}(\mathcal{A}_j^q)_+, \\
\psi_i(\boldsymbol{x}) &= \boldsymbol{u}(\boldsymbol{x}, \mathcal{B}_i^q)^T \boldsymbol{V}^{m+i} \boldsymbol{u}(\boldsymbol{x}, \mathcal{B}_i^q), \ \boldsymbol{V}^{m+i} \in \mathcal{S}(\mathcal{B}_i^q)_+, \\
\varphi_0(\boldsymbol{x}) &= \boldsymbol{u}(\boldsymbol{x}, \mathcal{G}^*)^T \boldsymbol{V}^0 \boldsymbol{u}(\boldsymbol{x}, \mathcal{G}^*), \ \boldsymbol{V}^0 \in \mathcal{S}(\mathcal{G}^*)_+.
\end{aligned}$$

Substituting these functions in the SOSOP (12) leads to

$$\left.\begin{array}{ll}
\text{maximize} & \zeta \\
\text{subject to} & f_0(\boldsymbol{x}) - \sum_{j=1}^m f_j(\boldsymbol{x}) \boldsymbol{u}(\boldsymbol{x}, \mathcal{A}_j^q)^T \boldsymbol{V}^j \boldsymbol{u}(\boldsymbol{x}, \mathcal{A}_j^q) \\
& - \sum_{i=1}^n (\rho - x_i^2) \boldsymbol{u}(\boldsymbol{x}, \mathcal{B}_i^q)^T \boldsymbol{V}^{m+i} \boldsymbol{u}(\boldsymbol{x}, \mathcal{B}_i^q) \\
& - \boldsymbol{u}(\boldsymbol{x}, \mathcal{G}^*)^T \boldsymbol{V}^0 \boldsymbol{u}(\boldsymbol{x}, \mathcal{G}^*) - \zeta = 0 \ (\forall \boldsymbol{x} \in \mathbb{R}^n), \\
& \boldsymbol{V}^j \in \mathcal{S}(\mathcal{A}_j^q)_+ \ (j = 1, 2, \ldots, m), \\
& \boldsymbol{V}^{m+i} \in \mathcal{S}(\mathcal{B}_i^q)_+ \ (i = 1, 2, \ldots, n), \boldsymbol{V}^0 \in \mathcal{S}(\mathcal{G}^*)_+.
\end{array}\right\}$$

Since the left hand side of the equality constraint in the problem above is a polynomial with the support

$$\mathcal{F}_C = \mathcal{F}_0 \bigcup \{\boldsymbol{0}\} \bigcup \left(\bigcup_{j=1}^m \widehat{\mathcal{A}}_j^q \right) \bigcup \left(\bigcup_{i=1}^n \widehat{\mathcal{B}}_i^q \right) \bigcup \{\boldsymbol{a} + \boldsymbol{b} : \boldsymbol{a} \in \mathcal{G}^*, \ \boldsymbol{b} \in \mathcal{G}^*\},$$

and the coefficients are linear functions of matrix variable $\boldsymbol{V}^j$ $(j = 1, 2, \ldots, m)$, $\boldsymbol{V}^{m+i}$ $(i = 1, 2, \ldots, n)$, $\boldsymbol{V}^0$ and ζ, we can rewrite equality constraint of the problem above as

$$\sum_{\boldsymbol{a} \in \mathcal{F}_C} d(\boldsymbol{a}, \boldsymbol{V}, \zeta) \boldsymbol{x}^{\boldsymbol{a}} = 0,$$

where $d(\boldsymbol{a}, \boldsymbol{V}, \zeta)$ is a linear function in the matrix variables $\boldsymbol{V}^j$ $(j = 0, 1, 2, \ldots, m + n)$ and a real variable ζ for each $\boldsymbol{a} \in \mathcal{F}_C$. This identity needs to be satisfied for all $\boldsymbol{x} \in \mathbb{R}^n$ in the problem, and the equality constraint is equivalent to a system of linear equations

$$d(\boldsymbol{a}, \boldsymbol{V}, \zeta) = 0 \ (\boldsymbol{a} \in \mathcal{F}_C).$$

Consequently, we obtain the following SDP which is equivalent to the SOSOP (12).

$$\left.\begin{array}{ll}
\text{maximize} & \zeta \\
\text{subject to} & d(\boldsymbol{a}, \boldsymbol{V}, \zeta) = 0 \ (\boldsymbol{a} \in \mathcal{F}_C), \\
& \boldsymbol{V}^j \in \mathcal{S}(\mathcal{A}_j^q)_+ \ (j = 1, 2, \ldots, m), \\
& \boldsymbol{V}^{m+i} \in \mathcal{S}(\mathcal{B}_i^q)_+ \ (i = 1, 2, \ldots, n), \boldsymbol{V}^0 \in \mathcal{S}(\mathcal{G}^*)_+.
\end{array}\right\} \quad (14)$$

The numerical efficiency of solving an SDP depends mainly on its size. In the SDP (14) above, the number of equality constraint and the sizes of matrix variables are determined by the supports $\mathcal{F}_j$ of the polynomial functions $f_j(\boldsymbol{x})$ $(j = 0, 1, \ldots, m)$ in the original POP (1) to be solved and $q \in \mathbb{Z}_+$. When the supports are sparse, the size of the resulting SDP becomes small. As we take a larger $q \in \mathbb{Z}_+$, we can expect to have a more accurate lower bound for the unknown optimal value ζ^* of the POP (1), but the number of equality constraint and the size of the matrix variables increases.

5 Preliminary numerical results

We provide an illustrative example of structured and sparse POPs and show how the choice of SOS polynomials in SOS relaxations can enhance the efficiency of the proposed relaxations greatly while preserving the effectiveness.

As mentioned in Remark 4.2 of [3], the support set $\mathcal{G}^*$ in the proposed SOS relaxation of (2) becomes dense even for sparse $\mathcal{F}_j$ ($j = 0, 1, 2, \ldots, m$) of the POP (2) because a polynomial $\varphi_0(x)$ is determined from the support of $L_C(x, \varphi) - \xi$. The convergence result shown in Section 4 is based on this choice of φ_0. In view of practical implementation of the proposed SOS relaxations, however, it may be more important to obtain a good lower bound with relatively small size SDP relaxations. We show the formulation of SOS relaxation presented in this paper can be easily adapted for the consideration in practice with the following example. The aim of the illustrative example is not to propose a practical method for general structured and sparse POPs, but to show how the SOS relaxation with convergent property can be modified for a specific problem in practice.

We consider an example

$$\left.\begin{array}{ll} \text{minimize} & f_0(\boldsymbol{x}) \equiv \sum_{i=1}^{n-1} f_{0i}(x_i, x_{i+1}) \\ \text{subject to} & f_{ij}(x_i, x_{i+1}) \geq 0 \ (i = 1, 2, \ldots, n-1, \ j = 1, 2, \ldots, m). \end{array}\right\} \tag{15}$$

Here $m \in \{1,\ n,\ n^2\}$, each $f_{0i}(x_i, x_{i+1})$ denotes a (fully dense) polynomial with degree 6 in two variables x_i, and x_{i+1} whose coefficients are chosen randomly from the interval $(-1, 1)$ ($i = 1, 2, \ldots, n-1$), and each $f_{ij}(x_i, x_{i+1})$ denotes a polynomial in two variables x_i and x_{i+1} of the form

$$1 - \left(x_i^\ell, x_{i+1}\right)\left(\frac{1}{\lambda_1^2}\begin{pmatrix} -a_2 \\ a_1 \end{pmatrix}(-a_2, a_1) + \frac{1}{\lambda_2^2}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}(a_1, a_2)\right)\begin{pmatrix} x_i^\ell \\ x_{i+1} \end{pmatrix}$$

for some $\boldsymbol{a} = (a_1, a_2)^T$ chosen from the unit circle, λ_1, λ_2 chosen randomly from the interval $(0.5, 2)$ and $\ell \in \{1,\ 3\}$ ($i = 1, 2, \ldots, n-1, j = 1, 2, \ldots, m$). When $\ell = 1$, each constraint $f_{ij}(x_i, x_{i+1}) \geq 0$ forms an ellipsoid in the (x_i, x_{i+1}) space with the center at the origin; if $\lambda_1 > \lambda_2$, $(a_1, a_2)^T$ corresponds to the major axis and $(-a_2, a_1)^T$ the minor axis.

Let us derive three relaxations of (15): the dual of Lasserre's SDP relaxation, the SOS relaxation presented in Section 4.2, and a practical version of the SOS relaxation. If we want to have the POP (15) in the form of (2) and to follow the theory described so far literally, the redundant inequalities

$$1 - x_i^2 \geq 0 \ (i = 1, 2, \ldots, n)$$

need to be added to the POP (15). However, for simplicity of discussion, we consider the problem without these inequalities. Notice that if these inequalities are added, stronger relaxations for the three relaxations result in. As far as the size of the relaxations is concerned, adding the inequalities increases the size of all three relaxations. The biggest increase in the size occurs in case of the dual of Lasserre's SDP relaxation given in (16). We also note that all the SOS relaxations presented below remain effective without the inequalities.

Define the Lagrangian function

$$L(\boldsymbol{x},\varphi) = f_0(\boldsymbol{x}) - \sum_{i=1}^{n-1}\sum_{j=1}^{m}\varphi_{ij}(\boldsymbol{x})f_{ij}(x_i,x_{i+1}) \text{ for every } \boldsymbol{x}\in\mathbb{R}^n$$
$$\text{and every } \varphi \equiv (\varphi_{ij})_{i=1,\dots,n-1,j=1,\dots,m} \in \overline{\Sigma}^{m(n-1)}$$

For every $i = 1,2,\dots,n-1$ and $q = 0,1,\dots$, let

$$\mathcal{A}_i^q = \{\mu e^i + \nu e^{i+1} : \mu\in\mathbb{Z}_+,\ \nu\in\mathbb{Z}_+,\ \mu+\nu\le q\},$$
$$\mathcal{A}_0^q = \left\{a\in\mathbb{Z}_+^n : \sum_{i=1}^n a_i \le q\right\}.$$

Then we have two types of SOS relaxations. The one is

$$\left.\begin{array}{ll} \text{maximize} & \zeta \\ \text{subject to} & L(\boldsymbol{x},\varphi)-\zeta = \varphi_0(\boldsymbol{x})\ (\forall \boldsymbol{x}\in\mathbb{R}^n), \\ & \varphi_{ij}\in\Sigma(\mathcal{A}_0^q)\ (i=1,2,\dots,n-1, j=1,2,\dots,m), \\ & \varphi_0\in\Sigma(\mathcal{A}_0^{q+\ell}), \end{array}\right\} \tag{16}$$

which corresponds to the dual of Lasserre's SDP relaxation applied to the POP (15). The other is

$$\left.\begin{array}{ll} \text{maximize} & \zeta \\ \text{subject to} & L(\boldsymbol{x},\varphi)-\zeta = \varphi_0(\boldsymbol{x})\ (\forall \boldsymbol{x}\in\mathbb{R}^n), \\ & \varphi_{ij}\in\Sigma(\mathcal{A}_i^q)\ (i=1,2,\dots,n-1, j=1,2,\dots,m), \\ & \varphi_0\in\Sigma(\mathcal{A}_0^{q+\ell}), \end{array}\right\} \tag{17}$$

which exploits the sparsity of the constraint inequalities of (15). In both relaxations, we take nonnegative integers q and ℓ such that $q+\ell\ge 3$; hence $q = 2,3,\dots$ if $\ell = 1$, and $q = 0,1,2,\dots$ if $\ell = 3$.

The SDP relaxation (16) uses

$$m(n-1) \text{ copies of support sets } \mathcal{A}_0^q \text{ of size } \#\mathcal{A}_0^q = \binom{n+q}{n},$$
$$\text{a support set } \mathcal{A}_0^{q+\ell} \text{ of size } \#\mathcal{A}_0^{q+\ell} = \binom{n+q+\ell}{n},$$

while the SDP relaxation (17) uses

$$m \text{ copies of support sets } \mathcal{A}_i^q \text{ of size } \#\mathcal{A}_i^q = \binom{2+q}{2}\ (i=1,2,\dots,n-1),$$
$$\text{a support set } \mathcal{A}_0^{q+\ell} \text{ of size } \#\mathcal{A}_0^{q+\ell} = \binom{n+q+\ell}{n}.$$

The two SOS relaxations (16) and (17) share the support set $\mathcal{A}_0^{q+\ell}$ with size $\binom{n+q+\ell}{n}$. The difference between them lies in the support sets $\mathcal{A}_0^q$ and $\mathcal{A}_i^q$. We can see that the size of the SOS relaxation (17) is smaller than the size of the SOS relaxation (16). When q is fixed, the advantage of the SOS relaxation (17) in the size of the problem over the SOS relaxation (16) becomes larger as m increases. This will be shown in Tables 1, 2 and 3. In the case of m fixed, the common support set $\mathcal{A}_0^{q+\ell}$ dominates all other support sets in both SOS relaxations in terms of size. As a result, the advantage from the size when increasing q is not as much as the case of q fixed in Table 4.

Exploiting the structure of polynomials may improve the weakness of the SOS relaxation (17) of the POP (15). We focus on " tridiagonal structure" of the support of the left hand side polynomial $L(\boldsymbol{x}, \boldsymbol{\varphi}) - \zeta$ of the equality constraint of the SOS relaxation (17), where φ_{ij} is assumed to be chosen from $\Sigma(\mathcal{A}_i^q)$. Specifically, the support of the polynomial $L(\boldsymbol{x}, \boldsymbol{\varphi}) - \zeta$ is covered by $\bigcup_{i=1}^{n-1}\left(\mathcal{A}_i^{q+\ell} + \mathcal{A}_i^{q+\ell}\right)$. Here we assume that $q+\ell \geq 3$. From this observation, we can expect that the polynomial $L(\boldsymbol{x}, \boldsymbol{\varphi}) - \zeta$ is represented as sums of squares of polynomials, each of which has a support in one of $\mathcal{A}_i^{q+\ell}$ $(i = 1, 2, \ldots, n-1)$. We replace $\varphi_0(\boldsymbol{x})$ and $\varphi_0 \in \Sigma(\mathcal{A}_0^{q+\ell})$ by $\sum_{i=1}^{n-1} \psi_i(\boldsymbol{x})$ and $\psi_i \in \Sigma(\mathcal{A}_i^{q+\ell})$ $(i = 1, 2, \ldots, n-1)$ in the SOS relaxation (17), respectively, to obtain a new SOS relaxation of the POP (15):

$$\left.\begin{array}{ll} \text{maximize} & \zeta \\ \text{subject to} & L(\boldsymbol{x}, \boldsymbol{\varphi}) - \zeta = \displaystyle\sum_{i=1}^{n-1} \psi_i(\boldsymbol{x}) \ (\forall \boldsymbol{x} \in \mathbb{R}^n), \\ & \varphi_{ij} \in \Sigma(\mathcal{A}_i^q) \ (i = 1, 2, \ldots, n-1, j = 1, 2, \ldots, m), \\ & \psi_i \in \Sigma(\mathcal{A}_i^{q+\ell}) \ (i = 1, 2, \ldots, n-1). \end{array}\right\} \tag{18}$$

It should be noted that the size of every support set in the SOS relaxation (18) is independent of the dimension n of the POP (15). When m and q are fixed, the total size of support sets in the SOS relaxation (18) grows linearly with the dimension n while the growth rate of the total size of support sets in the SOS relaxation (16) as well as that in the SOS relaxation (17) are of $O(n^{q+\ell})$. This shows that the SOS relaxation (18) has a considerable computational advantage in solving the POP (15) with large dimension n.

The numerical experiment was done using SDPA 6.0 [10] on Pentium IV (XEON) 2.4 GHz with 6GB memory, and the optimal values of the POP (15) with $m \in \{1,\ n,\ n^2\}$, $\ell \in \{1,\ 3\}$, and $n \in \{4,\ 5,\ 6,\ 7\}$ were computed by GloptiPoly [2]. Tables 1, 2 and 3 show some numerical results from the three SOS relaxations (16), (17) and (18) of the POP (15) with $m \in \{1,\ n,\ n^2\}$, $\ell = 1$ and dimension $n \in \{4,\ 5,\ 6,\ 7\}$. We observe that:

- All the SOS relaxations (16), (17) and (18) attain optimal values of the POP (15) with the lowest order $q = 2$.
- The SOS relaxation (17) requires less cpu time than the SOS relaxation (16), and the difference in cpu time becomes larger as m increases.

Table 4 shows some numerical results from the three SOS relaxations (16), (17) and (18) of the POP (15) with $m = n$, $\ell = 3$ and dimension $n \in \{3,\ 4,\ 5,\ 6\}$. In this case:

- The SOS relaxations (17) and (18) attain optimal values of the POP (15) or their lower bounds of (almost) the same quality as the SOS relaxation (16).
- The SOS relaxation (17) requires less cpu time than the SOS relaxation (16), but the difference is small.
- When $n = 6$, the SOS relaxation (16) with the order $q = 2$ attains the optimal value, but the other two SOS relaxations with the same order $q = 2$ provide only lower

bounds for the optimal value. It should also be noted that the SOS relaxation (18) with the higher order $q = 3$ attains the optimal value.

In all cases reported in Tables 1, 2, 3 and 4:

- The SOS relaxation (18) has a clear advantage over the other SOS relaxations.

Table 1: Numerical results on the POP (15) with $m = 1$, $\ell = 0$ and $q = 2$

POP (15)	cpu time in second		
n	relaxation (16)	relaxation (17)	relaxation (18)
4	0.6	0.4	0.1
5	4.9	2.2	0.1
6	22.3	21.5	0.1
7	153.8	98.6	0.2

Table 2: Numerical results on the POP (15) with $m = n$, $\ell = 0$ and $q = 2$

POP (15)	cpu time in second		
n	relaxation (16)	relaxation (17)	relaxation (18)
4	1.6	0.5	0.1
5	11.2	2.6	0.2
6	83.3	13.2	0.4
7	607.1	64.4	0.5

Table 3: Numerical results on the POP (15) with $m = n^2$, $\ell = 0$ and $q = 2$

POP (15)	cpu time in second		
n	relaxation (16)	relaxation (17)	relaxation (18)
4	6.6	1.0	0.5
5	83.7	5.1	1.1
6	717.1	23.2	2.7
7	7402.5	135.6	4.1

6 Concluding discussions

Considering two types of POPs (2) and (3) obtained from different characterizations of the feasible region of the POP (1), we have proposed a sequence of SOS relaxations from generalized Lagrangian duals of POP (2).

Theoretically, It is known in [3] that the SOS relaxation of the Lagrangian dual of (3) attains the optimal value ζ^* of the POP (3). But there remains a gap between the

Table 4: Numerical results on the POP (15) with $m = n$ and $\ell = 3$

POP (15)		cpu time in second (optimal value)		
n (optimal value)	q	relaxation (16)	relaxation (17)	relaxation (18)
3 (-1.782266)	0	0.1 (-148.0654)	0.1 (-148.0654)	0.1 (-148.0654)
	1	0.4 (-1.872454)	0.4 (-1.888884)	0.1 (-1.888890)
	2	1.9 (-1.782266)	1.6 (-1.782266)	0.2 (-1.782266)
4 (-2.244005)	0	0.4 (-129.5713)	0.4 (-129.5713)	0.1 (-129.5713)
	1	11.7 (-2.277639)	5.6 (-2.277639)	0.2 (-2.277844)
	2	46.2 (-2.244005)	36.1 (-2.244005)	0.4 (-2.244005)
5 (-3.848386)	0	2.6 (-120.1503)	2.5 (-120.2150)	0.1 (-120.1503)
	1	65.3 (-3.888779)	61.7 (-3.888779)	0.3 (-3.892605)
	2	787.1 (-3.848386)	644.6 (-3.848386)	0.8 (-3.848386)
6 (-3.531009)	0	13.3 (-120.2150)	13.4 (-120.2150)	0.1 (-120.2168)
	1	500.4 (-3.603920)	469.2 (-3.696910)	0.5 (-3.698462)
	2	11,912.4 (-3.531009)	11,718.1 (-3.535911)	1.4 (-3.537123)
	3	not solved	not solved	2.8 (-3.531009)

Lagrangian dual of (2) and its SOS relaxation in Section 4.2; the former attains ζ^* but the latter is not guaranteed to attain ζ^*. Thus it is interesting to prove or disprove that the SOS relaxation of the Lagrangian dual of (2) attains ζ^*. This will be a subject of future study.

The size of the SOS relaxation or the SDP relaxation obtained from the Lagrangian dual approach by exploiting sparsity is smaller than the size of Lasserre's SDP relaxation. This is of course a nice feature, but this may not necessarily mean that the former SDP relaxation is as effective as the latter in practice. To attain an approximation to the optimal value ζ^* of the POP (1) with as high accuracy as the one from Lasserre's SDP relaxation, we may need higher degree SOS polynomials in our dual approach, which makes the size of the resulting SDP relaxation larger.

One of the advantages of the proposed method is that we have much flexibility in implementation of the SOS relaxation and the SDP relaxation of the POP (1); sets of Lagrangian multiplier SOS polynomials satisfying the assumption of Theorem 3.1 can be freely chosen to strengthen the resulting relaxations. We have presented an illustrative example of how the framework of the proposed SOS relaxation can be used to have a practical SOS relaxation exploiting a structured sparsity. Numerical results of the example have indicated that it is possible to drastically improve computational efficiency of SOS relaxations by making proper heuristic choices of supports, depending on problems.

Additional numerical experiments on the SDP relaxation with heuristically chosen supports were performed for various types of polynomial optimization problems with certain types of sparsity. We have not included the numerical results from the additional numerical experiments in this paper because we believe that the discussion of heuristics is beyond the scope of this paper. The main purpose of this paper has been proposing general methods for sparsity in SDP relaxations for polynomial optimization and introducing Lagrangian dual and penalty function approaches into SDP relaxations for polynomial optimization. Although the numerical results supported the claim that the SOS relaxations could improve the efficiency, it would be necessary to address issues such as (i) a reasonable definition of structured sparsity in polynomial optimization problems, (ii) technical details of heuristic choices of supports, and (iii) extensive numerical experiments on various problems with structured sparsity. These will consist of a paper on practical performance of the heuristics, which we hope to present in near future.

References

[1] M. D. Choi, T. Y. Lam, and B. Reznick, "Sums of squares of real polynomials", *Proceedings of Symposia in Pure Mathematics*, **58** 2 (1995) 103-126.

[2] D. Henrion and J. B. Lasserre, "GloptiPoly: Global optimization over polynomials with Matlab and SeDuMi", Laboratoire d ' Analyse et d ' Architecture des Syst'emes, Centre National de la Recherche Scientifique, 7 Avenue du Colonel Roche, 31 077 Toulouse, cedex 4, France, February 2002.

[3] S. Kim, M. Kojima and H. Waki, "Generalized Lagrangian duals and sums of squares relaxations of sparse polynomial optimization problems", Research Report B-395, Dept. of Mathematical and computing Sciences, Tokyo Institute of Technology, Meguro, Tokyo 152-8552, September 2003. Revised July 2004. To appear in *SIAM Journal on Optimization.*

[4] M. Kojima, S. Kim and H. Waki, "A general framework for convex relaxation of polynomial optimization problems over cones", *Journal of Operations Research Society of Japan*, **46** 2 (2003) 125-144.

[5] M. Kojima, S. Kim and H. Waki, "Sparsity in sums of squares of polynomials", Research Report B-391, Dept. of Mathematical and computing Sciences, Tokyo Institute of Technology, Meguro, Tokyo 152-8552, June 2003. Revised August 2003. To appear in *Mathematical Programming.*

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

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

[8] V. Powers and T. Wörmann, "An algorithm for sums of squares of real polynomials", *Journal of Pure and Applied Algebra*, **127** (1998) 99-104.

[9] B. Reznick, "Extremal psd forms with few terms", *Duke Mathematical Journal*, **45** (1978) 363-374.

[10] M. Yamashita, K. Fujisawa and M. Kojima, "Implementation and evaluation of SDPA 6.0 (SemiDefinite Programming Algorithm 6.0)", *Optimization Methods and Software*, **18** (2003) 491-505.