Research Reports on Mathematical and Computing Sciences
Series B : Operations Research

Department of Mathematical and Computing Sciences
Tokyo Institute of Technology
2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152-8552 Japan

Sums of Squares Relaxations of Polynomial Semidefinite Programs

Masakazu Kojima†

Research Report B-397, November 2003

Abstract. A polynomial SDP (semidefinite program) minimizes a polynomial objective function over a feasible region described by a positive semidefinite constraint of a symmetric matrix whose components are multivariate polynomials. Sums of squares relaxations developed for polynomial optimization problems are extended to propose sums of squares relaxations for polynomial SDPs with an additional constraint for the variables to be in the unit ball. It is proved that optimal values of a sequence of sums of squares relaxations of the polynomial SDP, which correspond to duals of Lasserre's SDP relaxations applied to the polynomial SDP, converge to the optimal value of the polynomial SDP. The proof of the convergence is obtained by fully utilizing a penalty function and a generalized Lagrangian duals that were recently proposed by Kim *et al* for sparse polynomial optimization problems.

Key words.

Polynomial Optimization Problem, Sums of Squares Optimization, Semidefinite Program Relaxation, Lagrangian Relaxation, Lagrangian Dual, Penalty Function, Bilinear Matrix Inequality, Global Optimization.

† Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, 2-12-1 Oh-Okayama, Meguro-ku, Tokyo 152-8552 Japan. *kojima@is.titech.ac.jp*

1 Introduction

A polynomial SDP (semidefinite program) is a nonlinear and nonconvex optimization problem of minimizing a real valued polynomial objective function $a(\boldsymbol{x})$ subject to a matrix inequality $\boldsymbol{A}(\boldsymbol{x}) \succeq \boldsymbol{O}$ *i.e.*, a constraint for $\boldsymbol{A}(\boldsymbol{x})$ to be positive semidefinite. Here $\boldsymbol{x}$ denotes a vector variable in the n-dimensional Euclidean space and $\boldsymbol{A}(\boldsymbol{x})$ an $m \times m$ symmetric matrix whose (i, j)th component $A_{ij}(\boldsymbol{x})$ is a real valued polynomial in $\boldsymbol{x}$. The polynomial SDP is a generalization of the standard SDP (see, for example, [5, 9]) with a linear objective function and an LMI (linear matrix inequality) constraint to a polynomial objective function and a PMI (polynomial matrix inequality) constraint. It includes a wide class of problems, *e.g.*, POPs (polynomial optimization problems) where $\boldsymbol{A}(\boldsymbol{x})$ is a diagonal matrix and a BMI (bilinear matrix inequality) where $A_{ij}(\boldsymbol{x})$ is a quadratic function.

The purpose of this paper is to propose SOS (sum of squares) relaxation methods for a polynomial SDP with an additional ball constraint $\boldsymbol{x} \in B$ (the unit ball in the n-dimensional Euclidean space) by extending SOS relaxations introduced for POPs [6, 7]. We present a method of generating a sequence of SOS relaxation problems whose optimal values converge to the optimal value of the polynomial SDP. By applying a technique established in SOS relaxation methods, we can convert it into a sequence of standard SDPs.

Two related approaches provide a sequence of SDP relaxations whose optimal values converge to the optimal value of a given POP. The one is a dual approach and the other is a primal approach. The dual approach is based on SOS relaxations [6, 7]. In the recent paper [2], Kim *et al* presented a method to obtain a sequence of SDP relaxations by the dual approach. They also showed that the quality of the sequence of SDP relaxations was strengthened by applying a penalty function technique and a generalized Lagrangian dual. The use of the penalty function technique and the generalized Lagrangian dual provided a convenient way to exploit sparsity of polynomials in the POP and thus it was possible to introduce effective SOS relaxations for a sparse POP. The method proposed in this paper for the polynomial SDP is stemmed from those results in [2]. In particular, we introduce a penalty function and a generalized Lagrangian function for the polynomial SDP with the constraint $\boldsymbol{x} \in B$. The main emphasis is placed, however, on convergence analysis of the method but not on exploiting sparsity of the polynomial SDP.

A primal approach also produces a sequence of SDP relaxations for the polynomial SDP by extending Lasserre's SDP relaxation method [1, 4] for POPs to the polynomial SDP. These SDPs are duals of the ones derived in the dual approach mentioned above. A key idea behind the extension of Lasserre's SDP relaxation lies in the following fact. Let $\boldsymbol{h}(\boldsymbol{x})$ be a $(1 + \ell)$-dimensional column vector of a scalar constant 1 and real valued polynomials $h_1(\boldsymbol{x}), h_2(\boldsymbol{x}), \ldots, h_\ell(\boldsymbol{x})$ in $\boldsymbol{x}$. Then a PMI $\boldsymbol{A}(\boldsymbol{x}) \succeq \boldsymbol{O}$ is equivalent to a PMI to $\left(\boldsymbol{h}(\boldsymbol{x})\boldsymbol{h}(\boldsymbol{x})^T\right) \otimes \boldsymbol{A}(\boldsymbol{x}) \succeq \boldsymbol{O}$, where $\boldsymbol{M} \otimes \boldsymbol{N}$ denotes the Kronecker product of two matrices $\boldsymbol{M}$ and $\boldsymbol{N}$. This idea was presented as a technique to derive a valid constraint in the paper [3], but polynomial SDPs were not investigated. The extension presented here employs some techniques in the original SDP relaxation by Lasserre for a POP such as linearizing the polynomial objective function and the resulting PMI constraint to a standard SDP with an LMI.

We mention that the primal approach is more direct and easier to understand than the dual approach. However, the method in this paper is presented in terms of the dual approach instead of the primal approach because our theoretical analysis is based on the dual approach.

The remaining of the paper is organized as follows: In Section 2, we describe the definitions of polynomial matrices and sums of their squares after introducing some notation and symbols, and then show a characterization of sums of squares of polynomial matrices in terms of positive semidefinite matrices. In Section 3, we convert the polynomial SDP with the additional unit ball constraint $\boldsymbol{x} \in B$ into a sequence of POPs over the single constraint $\boldsymbol{x} \in B$ whose optimal values converge to the optimal value of the original polynomial SDP using a penalty function approach. Section 4 includes the extension of the sequence of penalized POPs over the single constraint $\boldsymbol{x} \in B$ given in Section 3 to a sequence of generalized Lagrangian duals, which provide better relaxations than the sequence of penalized POPs, and derivation of an equivalent sequence of SOS relaxations of the polynomial SDP. In Section 5, the primal approach to the polynomial SDP is presented to derive a sequence of its SDP relaxations, and a close relationship between the Lagrangian duals in Section 4 and the SDP relaxations in Section 5 is shown; the former correspond to the duals of the latter.

2 Polynomial matrices and sums of their squares

2.1 Symbols and notation

Let $\mathbb{R}^n$, $\mathbb{Z}_+$ and $\mathbb{Z}^n_+ \subset \mathbb{R}^n$ denote the n-dimensional Euclidean space, the set of nonnegative integers and the set of n-dimensional nonnegative integer vectors, respectively. We use the notation $\boldsymbol{x}^{\boldsymbol{\alpha}} = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$ for every $\boldsymbol{\alpha} \in \mathbb{Z}^n_+$ and every $\boldsymbol{x} = (x_1.x_2, \ldots, x_n)^T \in \mathbb{R}^n$. Here T denotes the transpose of a vector or a matrix. Let $\mathcal{M}^m$, $\mathcal{S}^m \subset \mathcal{M}^m$ and $\mathcal{S}^m_+ \subset \mathcal{S}^m$ denote the space of $m \times m$ real matrices, the space of $m \times m$ symmetric matrices and the cone of $m \times m$ positive semidefinite symmetric matrices, respectively. When $\boldsymbol{M} \in \mathcal{S}^m$, we often use the notation $\boldsymbol{M} \succeq \boldsymbol{O}$ to mean $\boldsymbol{M} \in \mathcal{S}^m_+$.

Let $\mathcal{F}$ be a nonempty finite subset of $\mathbb{Z}^n_+$, and $\boldsymbol{F}_{\boldsymbol{\alpha}} \in \mathcal{M}^m$ ($\boldsymbol{\alpha} \in \mathcal{F}$). We consider a polynomial $\boldsymbol{F}(\boldsymbol{x})$ in $\boldsymbol{x} \in \mathbb{R}^n$ with coefficients $\boldsymbol{F}_{\boldsymbol{\alpha}} \in \mathcal{M}^m$ ($\boldsymbol{\alpha} \in \mathcal{F}$) such that

$$\boldsymbol{F}(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha} \in \mathcal{F}} \boldsymbol{F}_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}}. \tag{1}$$

Let $\boldsymbol{\Xi}^{m \times m}$ denote the set of such polynomials in $\boldsymbol{x} \in \mathbb{R}^n$ with $m \times m$ coefficient matrices. We mention that each $\boldsymbol{F}(\boldsymbol{x}) \in \boldsymbol{\Xi}^{m \times m}$ is also characterized as an $m \times m$ matrix whose $(i, j)$the component $F_{ij}(\boldsymbol{x})$ is a real valued polynomial. We will call each $\boldsymbol{F}(\boldsymbol{x}) \in \boldsymbol{\Xi}^{m \times m}$ *an $m \times m$ polynomial matrix*, and $\mathcal{F}$ *a support* of $\boldsymbol{F}(\boldsymbol{x})$ if $\boldsymbol{F}(\boldsymbol{x})$ is represented as in (1).

We also consider special cases where all coefficient matrices $\boldsymbol{F}_{\boldsymbol{\alpha}} \in \mathcal{M}^m$ ($\boldsymbol{\alpha} \in \mathcal{F}$) are symmetric, *i.e.*, $\boldsymbol{F}_{\boldsymbol{\alpha}} \in \mathcal{S}^m$ ($\boldsymbol{\alpha} \in \mathcal{F}$) in (1). In this case, we will call $\boldsymbol{F}(\boldsymbol{x}) \in \boldsymbol{\Xi}^{m \times m}$ *an $m \times m$ symmetric polynomial matrix.* Let $\boldsymbol{\Xi}^{m \times m}_s$ denote the set of all $m \times m$ symmetric polynomial matrices. By definition, $\boldsymbol{\Xi}^{m \times m}_s \subset \boldsymbol{\Xi}^{m \times m}$. When $m = 1$, $\boldsymbol{\Xi}^{1 \times 1}_s = \boldsymbol{\Xi}^{1 \times 1}$. In this case, we write Ξ instead of $\boldsymbol{\Xi}^{1 \times 1}_s = \boldsymbol{\Xi}^{1 \times 1}$.

Let $a(\boldsymbol{x}) \in \Xi$, $\boldsymbol{A}(\boldsymbol{x}) \in \boldsymbol{\Xi}^{m \times m}_s$, and B the unit ball $\{\boldsymbol{x} \in \mathbb{R}^n : 1 - \boldsymbol{x}^T \boldsymbol{x} \geq 0\}$. Then the polynomial SDP (polynomial semidefinite program) that we deal with throughout the paper is described as

$$P_0 : \text{ minimize } \; a(\boldsymbol{x}) \text{ subject to } \; \boldsymbol{A}(\boldsymbol{x}) \succeq \boldsymbol{O} \; \text{ and } \boldsymbol{x} \in B.$$

We use the symbol $\mathcal{A}$ for a support of the $m \times m$ polynomial matrix $\boldsymbol{A}(\boldsymbol{x})$.

2.2 Sums of squares of polynomial matrices and their characterization

We define the set $\boldsymbol{\Sigma}^{m\times m}$ of sums of squares of $m \times m$ polynomial matrices as follows:

$$\boldsymbol{\Sigma}^{m\times m} \equiv \left\{ \sum_{p=1}^{q} (\boldsymbol{G}^p(\boldsymbol{x}))^T \boldsymbol{G}^p(\boldsymbol{x}) : \begin{array}{l} \boldsymbol{G}^p(\boldsymbol{x}) \in \boldsymbol{\Xi}^{m\times m} \ (p = 1, 2, \ldots, q), \\ q \text{ is a positive integer} \end{array} \right\}.$$

By definition, we know that $\boldsymbol{\Sigma}^{m\times m} \subset \boldsymbol{\Xi}_s^{m\times m}$ and that $\boldsymbol{F}(\boldsymbol{x}) \succeq \boldsymbol{O}$ for every $\boldsymbol{x} \in \mathbb{R}^n$ if $\boldsymbol{F}(\boldsymbol{x}) \in \boldsymbol{\Sigma}^{m\times m}$. We call each symmetric polynomial matrix in $\boldsymbol{\Sigma}^{m\times m}$ *a sum of squares of polynomial matrices*.

When $m = 1$, $\Sigma = \boldsymbol{\Sigma}^{1\times 1}$ is the set of sums of squares of real valued polynomials. It is well-known and easily shown that each $w(\boldsymbol{x}) \in \Sigma$ is represented as a positive semidefinite quadratic form of monomials and vice versa. This section generalizes this fact to the set $\boldsymbol{\Sigma}^{m\times m}$ of sums of squares of $m \times m$ polynomial matrices. We will associate each sum of squares of polynomial matrices with a positive semidefinite matrix.

Suppose that

$$\boldsymbol{F}(\boldsymbol{x}) \ = \ \sum_{p=1}^{q} (\boldsymbol{G}^p(\boldsymbol{x}))^T \boldsymbol{G}^p(\boldsymbol{x}) \in \boldsymbol{\Sigma}^{m\times m}.$$

We may assume that the polynomial matrices $\boldsymbol{G}^p(\boldsymbol{x}) \in \boldsymbol{\Xi}^{m\times m}$ $(p = 1, 2, \ldots, q)$ share a common support $\mathcal{G} \subset \mathbb{Z}^n_+$. Hence they can be represented as

$$\boldsymbol{G}^p(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha}\in\mathcal{G}} \boldsymbol{G}^p_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} \ (p = 1, 2, \ldots, q)$$

for some $\boldsymbol{G}^p_{\boldsymbol{\alpha}} \in \mathcal{M}^m$ $(\boldsymbol{\alpha} \in \mathcal{G},\ p = 1, 2, \ldots, q)$; we allow cases where some coefficient matrices $\boldsymbol{G}^p_{\boldsymbol{\alpha}}$ vanish. Let $p \in \{1, 2, \ldots, q\}$. Then

$$\begin{aligned} (\boldsymbol{G}^p(\boldsymbol{x}))^T \boldsymbol{G}^p(\boldsymbol{x}) &= \left(\sum_{\boldsymbol{\alpha}\in\mathcal{G}} \boldsymbol{G}^p_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} \right)^T \left(\sum_{\boldsymbol{\alpha}\in\mathcal{G}} \boldsymbol{G}^p_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} \right) \\ &= \sum_{\boldsymbol{\alpha}\in\mathcal{G}} \sum_{\boldsymbol{\beta}\in\mathcal{G}} (\boldsymbol{G}^p_{\boldsymbol{\alpha}})^T \boldsymbol{G}^p_{\boldsymbol{\beta}} \boldsymbol{x}^{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\beta}}. \end{aligned}$$

Let s denote the cardinality of $\mathcal{G}$. Let $\boldsymbol{V}^p$ denote the $sm \times sm$ symmetric matrix whose $(\boldsymbol{\alpha}, \boldsymbol{\beta})$th block $\boldsymbol{V}^p_{\boldsymbol{\alpha\beta}}$ is given by $(\boldsymbol{G}^p_{\boldsymbol{\alpha}})^T \boldsymbol{G}^p_{\boldsymbol{\beta}}$ $(\boldsymbol{\alpha} \in \mathcal{G},\ \boldsymbol{\beta} \in \mathcal{G})$. For example, if $\mathcal{G}$ consists of three elements $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, then

$$\begin{aligned} \boldsymbol{V}^p &= \begin{pmatrix} (\boldsymbol{G}^p_{\boldsymbol{\alpha}})^T \boldsymbol{G}^p_{\boldsymbol{\alpha}} & (\boldsymbol{G}^p_{\boldsymbol{\alpha}})^T \boldsymbol{G}^p_{\boldsymbol{\beta}} & (\boldsymbol{G}^p_{\boldsymbol{\alpha}})^T \boldsymbol{G}^p_{\boldsymbol{\gamma}} \\ \left(\boldsymbol{G}^p_{\boldsymbol{\beta}}\right)^T \boldsymbol{G}^p_{\boldsymbol{\alpha}} & \left(\boldsymbol{G}^p_{\boldsymbol{\beta}}\right)^T \boldsymbol{G}^p_{\boldsymbol{\beta}} & \left(\boldsymbol{G}^p_{\boldsymbol{\beta}}\right)^T \boldsymbol{G}^p_{\boldsymbol{\gamma}} \\ \left(\boldsymbol{G}^p_{\boldsymbol{\gamma}}\right)^T \boldsymbol{G}^p_{\boldsymbol{\alpha}} & \left(\boldsymbol{G}^p_{\boldsymbol{\gamma}}\right)^T \boldsymbol{G}^p_{\boldsymbol{\beta}} & \left(\boldsymbol{G}^p_{\boldsymbol{\gamma}}\right)^T \boldsymbol{G}^p_{\boldsymbol{\gamma}} \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{G}^p_{\boldsymbol{\alpha}} & \boldsymbol{G}^p_{\boldsymbol{\beta}} & \boldsymbol{G}^p_{\boldsymbol{\gamma}} \end{pmatrix}^T \begin{pmatrix} \boldsymbol{G}^p_{\boldsymbol{\alpha}} & \boldsymbol{G}^p_{\boldsymbol{\beta}} & \boldsymbol{G}^p_{\boldsymbol{\gamma}} \end{pmatrix}; \end{aligned}$$

hence $\boldsymbol{V}^p$ is a $3m \times 3m$ symmetric positive semidefinite matrix with rank at most m. In general, we see that $\boldsymbol{V}^p \in \mathcal{S}^{sm}_+$. We now let

$$\boldsymbol{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}} = \sum_{p=1}^{q} \boldsymbol{V}^p_{\boldsymbol{\alpha}\boldsymbol{\beta}} = \sum_{p=1}^{q} (\boldsymbol{G}^p_{\boldsymbol{\alpha}})^T \boldsymbol{G}^p_{\boldsymbol{\beta}} \quad (\boldsymbol{\alpha} \in \mathcal{G},\ \boldsymbol{\beta} \in \mathcal{G}) \text{ and } \boldsymbol{V} = \sum_{p=1}^{q} \boldsymbol{V}^p.$$

Note that each $\boldsymbol{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}}$ $(\boldsymbol{\alpha} \in \mathcal{G},\ \boldsymbol{\beta} \in \mathcal{G})$ corresponds to the $(\boldsymbol{\alpha}, \boldsymbol{\beta})$th block of $\boldsymbol{V} \in \mathcal{S}^{sm}_+$. We also see that

$$\boldsymbol{F}(\boldsymbol{x}) = \sum_{p=1}^{q} (\boldsymbol{G}^p(\boldsymbol{x}))^T \boldsymbol{G}^p(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha}\in\mathcal{G}} \sum_{\boldsymbol{\beta}\in\mathcal{G}} \boldsymbol{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}}\, \boldsymbol{x}^{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\beta}}.$$

Therefore we have shown that each $\boldsymbol{F}(\boldsymbol{x}) \in \boldsymbol{\Sigma}^{m\times m}$ is represented as

$$\boldsymbol{F}(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha}\in\mathcal{G}} \sum_{\boldsymbol{\beta}\in\mathcal{G}} \boldsymbol{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}}\, \boldsymbol{x}^{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\beta}} \tag{2}$$

for some $\boldsymbol{V} \in \mathcal{S}^{sm}_+$, where $\boldsymbol{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}}$ denotes the $(\boldsymbol{\alpha}, \boldsymbol{\beta})$th block of $\boldsymbol{V} \in \mathcal{S}^{sm}_+$.

Conversely, we assume that $\boldsymbol{F}(\boldsymbol{x}) \in \boldsymbol{\Xi}^{m\times m}_s$ is represented as in (2) for some $\boldsymbol{V} \in \mathcal{S}^{sm}_+$ to show that $\boldsymbol{F}(\boldsymbol{x})$ is a sum of squares of polynomial matrices. Since $\boldsymbol{V}$ is positive semidefinite, we can take a matrix $\boldsymbol{G} \in \mathcal{M}^{sm}$ such that $\boldsymbol{V} = \boldsymbol{G}^T\boldsymbol{G}$. Let $\boldsymbol{G}_{\boldsymbol{\alpha}\boldsymbol{\beta}}$ denote the $(\boldsymbol{\alpha}, \boldsymbol{\beta})$th block of $\boldsymbol{G}$. Then, for every pair $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ $(\boldsymbol{\alpha} \in \mathcal{G}, \boldsymbol{\beta} \in \mathcal{G})$, we have

$$\boldsymbol{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}} = \sum_{\boldsymbol{\gamma}\in\mathcal{G}} (\boldsymbol{G}_{\boldsymbol{\gamma}\boldsymbol{\alpha}})^T \boldsymbol{G}_{\boldsymbol{\gamma}\boldsymbol{\beta}}.$$

In view of (2), we then obtain that

$$\begin{aligned}
\boldsymbol{F}(\boldsymbol{x}) &= \sum_{\boldsymbol{\alpha}\in\mathcal{G}} \sum_{\boldsymbol{\beta}\in\mathcal{G}} \left(\sum_{\boldsymbol{\gamma}\in\mathcal{G}} (\boldsymbol{G}_{\boldsymbol{\gamma}\boldsymbol{\alpha}})^T \boldsymbol{G}_{\boldsymbol{\gamma}\boldsymbol{\beta}} \right) \boldsymbol{x}^{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\beta}} \\
&= \sum_{\boldsymbol{\gamma}\in\mathcal{G}} \sum_{\boldsymbol{\alpha}\in\mathcal{G}} \sum_{\boldsymbol{\beta}\in\mathcal{G}} (\boldsymbol{G}_{\boldsymbol{\gamma}\boldsymbol{\alpha}})^T \boldsymbol{G}_{\boldsymbol{\gamma}\boldsymbol{\beta}} \boldsymbol{x}^{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\beta}} \\
&= \sum_{\boldsymbol{\gamma}\in\mathcal{G}} \left(\sum_{\boldsymbol{\alpha}\in\mathcal{G}} \boldsymbol{G}_{\boldsymbol{\gamma}\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} \right)^T \left(\sum_{\boldsymbol{\beta}\in\mathcal{G}} \boldsymbol{G}_{\boldsymbol{\gamma}\boldsymbol{\beta}} \boldsymbol{x}^{\boldsymbol{\beta}} \right) \\
&= \sum_{\boldsymbol{\gamma}\in\mathcal{G}} (\boldsymbol{G}^{\boldsymbol{\gamma}}(\boldsymbol{x}))^T \boldsymbol{G}^{\boldsymbol{\gamma}}(\boldsymbol{x}),
\end{aligned}$$

where $\boldsymbol{G}^{\boldsymbol{\gamma}}(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha}\in\mathcal{G}} \boldsymbol{G}_{\boldsymbol{\gamma}\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} \in \boldsymbol{\Xi}^{m\times m}$ $(\gamma \in \mathcal{G})$. Thus we have shown that each $\boldsymbol{F}(\boldsymbol{x}) \in \boldsymbol{\Xi}^{m\times m}_s$ represented as in (2) for some $\boldsymbol{V} \in \mathcal{S}^{sm}_+$ is a sum of squares of polynomial matrices, *i.e.*, $\boldsymbol{F}(\boldsymbol{x}) \in \boldsymbol{\Sigma}^{m\times m}$.

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

$$\mathbf{\Sigma}^{m\times m}(\mathcal{G}) = \left\{ \sum_{\boldsymbol{\alpha}\in\mathcal{G}} \sum_{\boldsymbol{\beta}\in\mathcal{G}} \boldsymbol{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}}\, \boldsymbol{x}^{\boldsymbol{\alpha}}\boldsymbol{x}^{\boldsymbol{\beta}} : \begin{array}{l} \boldsymbol{V} \in \mathcal{S}^{sm}_+, \quad \text{where } s \text{ denotes} \\ \text{the cardinality of } \mathcal{G} \end{array} \right\}.$$

Then we can rewrite the entire set $\mathbf{\Sigma}^{m\times m}$ of sums of squares of polynomial matrices as the union of $\mathbf{\Sigma}^{m\times m}(\mathcal{G})$ over all nonempty finite subset $\mathcal{G}$ of $\mathbb{Z}^n_+$;

$$\mathbf{\Sigma}^{m\times m} = \bigcup_{\emptyset \neq \mathcal{G} \subset \mathbb{Z}^n_+} \mathbf{\Sigma}^{m\times m}(\mathcal{G}). \tag{3}$$

For a special case $m = 1$, we have that

$$\left.\begin{array}{l} \Sigma = \displaystyle\bigcup_{\emptyset \neq \mathcal{G} \subset \mathbb{Z}^n_+} \Sigma(\mathcal{G}), \\ \Sigma(\mathcal{G}) = \left\{ \displaystyle\sum_{\boldsymbol{\alpha}\in\mathcal{G}} \sum_{\boldsymbol{\beta}\in\mathcal{G}} V_{\boldsymbol{\alpha}\boldsymbol{\beta}}\, \boldsymbol{x}^{\boldsymbol{\alpha}}\boldsymbol{x}^{\boldsymbol{\beta}} : \begin{array}{l} \boldsymbol{V} \in \mathcal{S}^{t}_+, \\ \text{where } t \text{ denotes the cardinality of } \mathcal{G} \end{array} \right\}. \end{array}\right\} \tag{4}$$

3 A penalty function approach

The term "a penalty function" used below has a slightly different meaning from a conventional penalty function. The difference lies in that it imposes a penalty even for feasible points of the polynomial SDP P_0 although the penalty values for the feasible points tend to zero as the penalty parameter increases. We may regard it as a special case of the generalized Lagrangian function for the polynomial SDP P_0 given in the next section. We show convergence of the optimal values of a sequence of penalized polynomial optimization problems (Theorem 3.2), which will be used to establish convergence of the optimal values of a sequence of generalized Lagrangian duals of the polynomial SDP P_0 in the next section.

Let us first introduce an ideal penalty function ϕ_∞ of the polynomial SDP P_0 over B, defined by

$$\phi_\infty(\boldsymbol{x}) = \begin{cases} 0 & \text{if } \boldsymbol{x} \in C, \\ \infty & \text{if } \boldsymbol{x} \in B\backslash C. \end{cases}$$

Here C denotes the feasible region $\{\boldsymbol{x} \in B : \boldsymbol{A}(\boldsymbol{x}) \succeq \boldsymbol{O}\}$. Then the polynomial SDP P_0 is equivalent to the problem

$$\Psi_\infty : \text{ minimize } a(\boldsymbol{x}) + \phi_\infty(\boldsymbol{x}) \text{ subject to } \boldsymbol{x} \in B.$$

We will construct a sequence $\{\phi_p(\boldsymbol{x})\ (p \in \mathbb{Z}_+)\}$ of polynomial penalty functions that "converges" to the ideal one $\phi_\infty(\boldsymbol{x})$ on B.

Take an $\omega > 0$ such that $\|\boldsymbol{A}(\boldsymbol{x})\| \leq \omega$ for every $\boldsymbol{x} \in B$, where $\|\boldsymbol{N}\| \equiv \max_{\|\boldsymbol{z}\|=1} \|\boldsymbol{N}\boldsymbol{z}\|$ = the maximum absolute value of all eigenvalues of $\boldsymbol{N} \in \mathcal{S}^m$. (Note that such an $\omega > 0$ always exists since $\boldsymbol{A}(\boldsymbol{x})$ is continuous with respect to $\boldsymbol{x}$ in a compact set B). For every $\boldsymbol{x} \in \mathbb{R}^n$ and every $p \in \mathbb{Z}_+$, define

$$\phi_p(\boldsymbol{x}) = -\left(\boldsymbol{I} - \boldsymbol{A}(\boldsymbol{x})/\omega\right)^{2p} \bullet \boldsymbol{A}(\boldsymbol{x}).$$

Here $\boldsymbol{I}$ denotes the $m \times m$ identity matrix and we assume that $(\boldsymbol{I} - \boldsymbol{A}(\boldsymbol{x})/\omega)^0 = \boldsymbol{I}$. We consider the sequence of penalized POPs over the unit ball B

$$\Psi_p: \text{ minimize } a(\boldsymbol{x}) + \phi_p(\boldsymbol{x}) \text{ subject to } \boldsymbol{x} \in B$$

($p \in \mathbb{Z}_+$). Let ψ_p^* denote the optimal value of this problem. By definition,

$$\phi_p(\boldsymbol{x}) \leq 0 \text{ for every } \boldsymbol{x} \in C \ (p \in \mathbb{Z}_+).$$

Hence we see that $\psi_p^* \leq \zeta_0^*$ ($p \in \mathbb{Z}_+$). The lemma below shows that the polynomial function $\phi_p(\boldsymbol{x})$ "converges" to $\phi_\infty(\boldsymbol{x})$ on B as $p \to \infty$.

Lemma 3.1.

(a) For any $\epsilon > 0$, there exists a positive integer $\hat{p}$ such that $-\epsilon \leq \phi_p(\boldsymbol{x})$ for every $\boldsymbol{x} \in B$ and every $p \geq \hat{p}$.

(b) If $\tilde{\boldsymbol{x}} \in B \backslash C$ and $\kappa > 0$ then there exist a positive number $\tilde{\delta}$ and a positive integer $\tilde{p}$ such that $\kappa \leq \phi_p(\boldsymbol{x})$ for every $\boldsymbol{x} \in U_{\tilde{\delta}}(\tilde{\boldsymbol{x}}) \bigcap B$ and every $p \geq \tilde{p}$.

Proof: First we derive an inequality that will be used to show (a) and (b). Let $\boldsymbol{x} \in B$. Take an $m \times m$ orthogonal matrix $\boldsymbol{P}$ and an $m \times m$ diagonal matrix $\boldsymbol{M}$ with the eigenvalues μ_i of $\boldsymbol{A}(\boldsymbol{x})$ ($i = 1, 2, \ldots, m$) such that $\boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{P}\boldsymbol{M}\boldsymbol{P}^T$. Then

$$\begin{aligned}
\phi_p(\boldsymbol{x}) &= -(\boldsymbol{I} - \boldsymbol{A}(\boldsymbol{x})/\omega)^{2p} \bullet \boldsymbol{A}(\boldsymbol{x}) \\
&= -\text{Trace}(\boldsymbol{I} - \boldsymbol{P}\boldsymbol{M}\boldsymbol{P}^T/\omega)^{2p}\boldsymbol{P}\boldsymbol{M}\boldsymbol{P}^T \\
&= -\text{Trace}(\boldsymbol{I} - \boldsymbol{M}/\omega)^{2p}\boldsymbol{M} \\
&= -\sum_{i=1}^{m}(1 - \mu_i/\omega)^{2p}\mu_i \\
&= -\omega\left(\sum_{\mu_i \geq 0}(1 - \mu_i/\omega)^{2p}\mu_i/\omega + \sum_{\mu_i < 0}(1 - \mu_i/\omega)^{2p}\mu_i/\omega\right).
\end{aligned}$$

Since $\mu_i/\omega \in [0, 1]$ if $\mu_i \geq 0$, we obtain that

$$\phi_p(\boldsymbol{x}) \geq -m\omega\mu(p) - \omega\left(\sum_{\mu_i < 0}(1 - \mu_i/\omega)^{2p}\mu_i/\omega\right), \tag{5}$$

where $\mu(p) = \max\{(1 - \xi)^{2p}\xi : \xi \in [0, 1]\}$. To show (a), we let $\epsilon > 0$. Then we can find a positive integer $\hat{p}$ such that $\mu(p) \leq \epsilon/(m\omega)$ of every $p \geq \hat{p}$. Let $\boldsymbol{x} \in B$ and $p \geq \hat{p}$. It follows from (5) that

$$\phi_p(\boldsymbol{x}) \geq -m\omega\mu(p) \geq -\epsilon.$$

Now we prove (b). Suppose that $\tilde{\boldsymbol{x}} \in B \backslash C$ and $\kappa > 0$. Since $\tilde{\boldsymbol{x}} \in B \backslash C$, there exist $\delta > 0$ and $\bar{\mu} < 0$ such that if $\boldsymbol{x} \in U_{\tilde{\delta}}(\tilde{\boldsymbol{x}}) \bigcap B$ then the minimum eigenvalue of $\boldsymbol{A}(\boldsymbol{x})$ is not greater than $\bar{\mu}$. Since $(1 - \bar{\mu}/\omega) > 1$ and $\bar{\mu} < 0$, we can take a positive integer $\tilde{p}$ such that

$$-m\omega - (1 - \bar{\mu}/\omega)^{2p}\bar{\mu} \geq \kappa \text{ for every } p \geq \tilde{p}.$$

Suppose that $\boldsymbol{x} \in U_{\tilde{\delta}}(\tilde{\boldsymbol{x}}) \bigcap B$ in the inequality (5). Let $\mu_{\min}$ denote the minimum eigenvalue of $\boldsymbol{A}(\boldsymbol{x})$. Then $\mu_{\min} \leq \bar{\mu} < 0$. Hence

$$\phi_p(\boldsymbol{x}) \geq -m\omega\mu(p) - (1 - \mu_{\min}/\omega)^{2p}\mu_{\min} \geq -m\omega - (1 - \bar{\mu}/\omega)^{2p}\bar{\mu} \geq \kappa.$$

∎

Now we are ready to show the convergence of the optimal values ψ_p^* of Ψ_p $(p \in \mathbb{Z}_+)$ to the original polynomial SDP P_0 as $p \to \infty$.

Theorem 3.2. $\zeta_0^* \geq \psi_p^* \to \zeta_0^*$ *as* $p \to \infty$.

Proof: Let $\boldsymbol{x}^p$ be an optimal solution of Ψ_p $(\boldsymbol{p} \in \mathbb{Z}_+)$. Assume on the contrary that the corresponding optimal value $a(\boldsymbol{x}^p) + \phi_p(\boldsymbol{x}^p)$ does not converge to ζ_0^*. Then there is an $\epsilon > 0$ and a subsequence $\{\boldsymbol{x}^p \ (p \in J)\}$ for some $J \subset \mathbb{Z}_+$ such that

$$a(\boldsymbol{x}^p) + \phi_p(\boldsymbol{x}^p) \leq \zeta_0^* - \epsilon \ (p \in J). \tag{6}$$

Since the subsequence $\{\boldsymbol{x}^p \ (p \in J)\}$ is contained in the compact set B, we may assume without loss of generality that it converges to $\bar{\boldsymbol{x}} \in B$. By (6) and (b) of Lemma 3.1, $\bar{\boldsymbol{x}} \in C$. By (a) of Lemma 3.1, there exists a $\hat{p} \in \mathbb{Z}_+$ such that

$$a(\boldsymbol{x}^p) - \epsilon/2 \leq a(\boldsymbol{x}^p) + \phi_p(\boldsymbol{x}^p) \text{ if } p \in J \text{ and } p \geq \hat{p}.$$

If we take the limit of the subsequence $\{\boldsymbol{x}^p \ (p \in J)\}$, the the left hand side of the inequality above converges to $a(\bar{\boldsymbol{x}}) - \epsilon/2 \geq \zeta_0^* - \epsilon/2$. Hence we obtain that

$$\zeta_0^* - \epsilon < a(\boldsymbol{x}^p) + \phi_p(\boldsymbol{x}^p) \text{ for every sufficiently large } p \in J.$$

This contradicts to the inequality (6). ∎

4 A dual approach

4.1 A generalized Lagrangian dual

We introduce a (generalized) Lagrangian function

$$\lambda(\boldsymbol{x}, \boldsymbol{W}) = a(\boldsymbol{x}) - \boldsymbol{W}(\boldsymbol{x}) \bullet \boldsymbol{A}(\boldsymbol{x}) \text{ for every } \boldsymbol{W}(\boldsymbol{x}) \in \boldsymbol{\Sigma}^{m \times m} \text{ and } \boldsymbol{x} \in \mathbb{R}^n.$$

For every $\boldsymbol{W}(\boldsymbol{x}) \in \boldsymbol{\Sigma}^{m \times m}$, we consider a (generalized) Lagrangian relaxation

$$\Lambda(\boldsymbol{W}): \text{ minimize } \lambda(\boldsymbol{x}, \boldsymbol{W}) \text{ subject to } \boldsymbol{x} \in B.$$

Since B is a nonempty compact subset of $\mathbb{R}^n$, the Lagrangian relaxation $\Lambda(\boldsymbol{W})$ has an optimal solution. Let $\lambda^*(\boldsymbol{W})$ denote the optimal value of this problem;

$$\lambda^*(\boldsymbol{W}) = \min \{\lambda(\boldsymbol{x}, \boldsymbol{W}) \ : \ \boldsymbol{x} \in B\}.$$

If $\boldsymbol{W}(\boldsymbol{x}) \in \boldsymbol{\Sigma}^{m \times m}$ and $\boldsymbol{x} \in C$ then $\boldsymbol{A}(\boldsymbol{x}) \in \mathcal{S}_+^m$ and $\lambda(\boldsymbol{x}, \boldsymbol{W}) \leq a(\boldsymbol{x})$. Hence

$$\lambda^*(\boldsymbol{W}) \leq \zeta_0^* \text{ for every } \boldsymbol{W}(\boldsymbol{x}) \in \boldsymbol{\Sigma}^{m \times m}.$$

The penalized POP optimization problem Ψ_p in the previous section is a special case of the Lagrangian relaxation if we take $\boldsymbol{W}(\boldsymbol{x}) = (\boldsymbol{I} - \boldsymbol{A}(\boldsymbol{x})/\omega)^{2p} \in \boldsymbol{\Sigma}^{m\times m}$; Ψ_p is identical to $\Lambda((\boldsymbol{I} - \boldsymbol{A}(\boldsymbol{x})/\omega)^{2p})$.

For every nonempty subset $\mathcal{G}$ of $\mathbb{Z}^n_+$, we define a (generalized) Lagrangian dual

$$\Lambda(\mathcal{G}): \quad \text{maximize } \lambda^*(\boldsymbol{W}) \text{ subject to } \ \boldsymbol{W}(\boldsymbol{x}) \in \boldsymbol{\Sigma}^{m\times m}(\mathcal{G}).$$

Let $\lambda^*(\mathcal{G})$ denote the optimal value of this problem;

$$\lambda^*(\mathcal{G}) = \sup\left\{\lambda^*(\boldsymbol{W}) \ : \ \boldsymbol{W}(\boldsymbol{x}) \in \boldsymbol{\Sigma}^{m\times m}(\mathcal{G})\right\}.$$

We then see that

$$\begin{array}{l} \lambda^*(\boldsymbol{W}) \leq \lambda^*(\mathcal{G}) \leq \lambda(\mathcal{G}') \leq \lambda^*(\mathbb{Z}^n_+) \leq \zeta^*_0 \\ \qquad \text{if } \boldsymbol{W}(\boldsymbol{x}) \in \boldsymbol{\Sigma}^{m\times m}(\mathcal{G}) \ \text{ and } \mathcal{G} \subset \mathcal{G}' \subset \mathbb{Z}^n_+. \end{array} \tag{7}$$

Recall that $\mathcal{A}$ denotes the support of the $m \times m$ polynomial matrix $\boldsymbol{A}(\boldsymbol{x})$ involved in the polynomial SDP P_0. We define

$$\begin{array}{rcl} \mathcal{A}^{(0)} & = & \{\boldsymbol{0}\}, \ \mathcal{A}^{(1)} = \mathcal{A}\bigcup\{\boldsymbol{0}\} \subset \mathbb{Z}^n_+, \\ \mathcal{A}^{(p+1)} & = & \{\boldsymbol{\alpha}+\boldsymbol{\beta} : \boldsymbol{\alpha} \in \mathcal{A}^{(p)}, \ \boldsymbol{\beta} \in \mathcal{A}^{(1)}\} \subset \mathbb{Z}^n_+ \ (p = 1, 2, \dots, \ ). \end{array}$$

Theorem 4.1. $\zeta^*_0 \geq \lambda^*(\mathcal{A}^{(p)}) \to \zeta^*_0$ *as* $p \to \infty$.

Proof: It follows from (7) that $\lambda^*(\mathcal{A}^{(p)}) \leq \zeta^*_0$ $(p \in \mathbb{Z}_+)$. By construction, we know that $\mathcal{A}^{(p)}$ forms a support of the $m \times m$ polynomial matrix $(\boldsymbol{I} - \boldsymbol{A}(\boldsymbol{x})/\omega)^p$; hence $(\boldsymbol{I} - \boldsymbol{A}(\boldsymbol{x})/\omega)^{2p} \in \boldsymbol{\Sigma}^{m\times m}(\mathcal{A}^{(p)})$. Hence

$$\begin{array}{rcl} \psi^*_p & = & \min\{a(\boldsymbol{x}) + \phi_p(\boldsymbol{x}) : \boldsymbol{x} \in B\} \\ & = & \min\left\{a(\boldsymbol{x}) - (\boldsymbol{I} - \boldsymbol{A}(\boldsymbol{x})/\omega)^{2p} \bullet \boldsymbol{A}(\boldsymbol{x}) : \boldsymbol{x} \in B\right\} \\ & \leq & \displaystyle\sup_{\boldsymbol{W}(\boldsymbol{x}) \in \boldsymbol{\Sigma}^{m\times m}(\mathcal{A}^{(p)})} \min\{a(\boldsymbol{x}) - \boldsymbol{W}(\boldsymbol{x}) \bullet \boldsymbol{A}(\boldsymbol{x}) : \boldsymbol{x} \in B\} \\ & = & \lambda^*(\mathcal{A}^{(p)}) \ (p \in \mathbb{Z}_+). \end{array}$$

Therefore we have shown that $\psi^*_p \leq \lambda^*(\mathcal{A}^{(p)}) \leq \zeta^*_0$ $(p \in \mathbb{Z}_+)$. By Theorem 3.2, we obtain the desired results. ∎

Corollary 4.2. *Suppose that a sequence* $\left\{\mathcal{G}_p \subset \mathbb{Z}^n_+ \ (p \in \mathbb{Z}_+)\right\}$ *satisfies*

$$\emptyset \neq \mathcal{G}_p \subset \mathcal{G}_{p+1} \ (p \in \mathbb{Z}_+) \ \text{ and } \mathcal{A}^{(q)} \subset \bigcup_{p \in \mathbb{Z}_+} \mathcal{G}_p \ (q \in \mathbb{Z}_+). \tag{8}$$

Then $\zeta^*_0 \geq \lambda^*(\mathcal{G}_p) \to \zeta^*_0$ *as* $p \to \infty$.

Proof: By the first inclusion relation of (8) and (7), we have

$$\lambda^*(\mathcal{G}_p) \leq \lambda^*(\mathcal{G}_{p+1}) \leq \zeta^*_0 \ (p \in \mathbb{Z}_+).$$

For every $q \in \mathbb{Z}_+$, (8) ensures the existence of $p \in \mathbb{Z}_+$ such that $\mathcal{A}^{(q)} \subset \mathcal{G}_p$; hence $\lambda^*(\mathcal{A}^{(q)}) \leq \lambda^*(\mathcal{G}_p)$. Thus the result follows from Theorem 4.1. ∎

4.2 SOS relaxations of the Lagrangian duals

In this subsection, we present a numerical method for approximating ζ_0^* based on SOS relaxations of the Lagrangian duals $\Lambda(\mathcal{G}_p)$ $(p \in \mathbb{Z}_+)$, where $\left\{\mathcal{G}_p \in \mathbb{Z}_+^n \ (p \in \mathbb{Z}_+)\right\}$ is a sequence satisfying the assumption (8) of Corollary 4.2. For this purpose, we introduce an SOS relaxation of the Lagrangian dual $\Lambda(\mathcal{G})$ $(\mathcal{G} \in \mathbb{Z}_+^n)$. For every triplet of nonempty finite subsets $\mathcal{G}$, $\widetilde{\mathcal{G}}$ and $\widehat{\mathcal{G}}$ of $\mathbb{Z}_+^n$, we consider an SOS optimization problem

$$\begin{array}{lll} \widehat{\boldsymbol{\Lambda}}(\mathcal{G},\widetilde{\mathcal{G}},\widehat{\mathcal{G}}): & \text{maximize} & \eta \\ & \text{subject to} & \lambda(\boldsymbol{x},\boldsymbol{W}) - \widetilde{w}(\boldsymbol{x})(1-\boldsymbol{x}^T\boldsymbol{x}) - \eta \in \Sigma(\widehat{\mathcal{G}}), \\ & & \boldsymbol{W}(\boldsymbol{x}) \in \boldsymbol{\Sigma}^{m\times m}(\mathcal{G}),\ \widetilde{w}(\boldsymbol{x}) \in \Sigma(\widetilde{\mathcal{G}}). \end{array}$$

In addition to a sequence $\{\mathcal{G}_p \subset \mathbb{Z}_+^n \ (p \in \mathbb{Z}_+)\}$ satisfying (8), we prepare sequences $\left\{\widetilde{\mathcal{G}}_p \subset \mathbb{Z}_+^n \ (p \in \mathbb{Z}_+)\right\}$ and $\left\{\widehat{\mathcal{G}}_p \subset \mathbb{Z}_+^n \ (p \in \mathbb{Z}_+)\right\}$ such that

$$\left.\begin{array}{l} \emptyset \neq \widetilde{\mathcal{G}}_p \subset \widetilde{\mathcal{G}}_{p+1} \ (p \in \mathbb{Z}_+) \ \text{ and } \mathbb{Z}_+^n = \displaystyle\bigcup_{p \in \mathbb{Z}_+} \widetilde{\mathcal{G}}_p, \\ \emptyset \neq \widehat{\mathcal{G}}_p \subset \widehat{\mathcal{G}}_{p+1} \ (p \in \mathbb{Z}_+) \ \text{ and } \mathbb{Z}_+^n = \displaystyle\bigcup_{p \in \mathbb{Z}_+} \widehat{\mathcal{G}}_p. \end{array}\right\} \tag{9}$$

Let η_p^* denote the optimal value of the problem $\widehat{\boldsymbol{\Lambda}}(\mathcal{G}_p,\widetilde{\mathcal{G}}_p,\widehat{\mathcal{G}}_p)$;

$$\eta_p^* \;=\; \sup\left\{\eta : \begin{array}{l} \lambda(\boldsymbol{x},\boldsymbol{W}) - \widetilde{w}(\boldsymbol{x})(1-\boldsymbol{x}^T\boldsymbol{x}) - \eta \in \Sigma(\widehat{\mathcal{G}}_p), \\ \boldsymbol{W}(\boldsymbol{x}) \in \boldsymbol{\Sigma}^{m\times m}(\mathcal{G}_p),\ \widetilde{w}(\boldsymbol{x}) \in \Sigma(\widetilde{\mathcal{G}}_p) \end{array}\right\}.$$

Theorem 4.3. *$\zeta_0^* \geq \eta_p^* \to \zeta_0^*$ as $p \to \infty$.*

Proof: First we observe that $\eta_p^* \leq \zeta_0^*$ $(p \in \mathbb{Z}_+)$. Let ϵ be an arbitrary positive number. By Corollary 4.2, there exists a $\bar{p} \in \mathbb{Z}_+$ and a $\overline{\boldsymbol{W}}(\boldsymbol{x}) \in \boldsymbol{\Sigma}(\mathcal{G}_{\bar{p}})^{m\times m}$ such that

$$a(\boldsymbol{x}) - \overline{\boldsymbol{W}}(\boldsymbol{x}) \bullet \boldsymbol{A}(\boldsymbol{x}) - (\zeta_0^* - \epsilon) > 0 \ \text{ for every } \boldsymbol{x} \in B.$$

This implies that the polynomial $a(\boldsymbol{x}) - \overline{\boldsymbol{W}}(\boldsymbol{x}) \bullet \boldsymbol{A}(\boldsymbol{x}) - (\zeta_0^* - \epsilon)$ is positive on the ball $B = \{\boldsymbol{x} \in \mathbb{R}^n : 1 - \boldsymbol{x}^T\boldsymbol{x} \geq 0\}$. By Lemma 4.1 of [8], there exists a $\widetilde{w}(\boldsymbol{x}) \in \Sigma$ and $\widehat{w}(\boldsymbol{x}) \in \Sigma$ such that

$$a(\boldsymbol{x}) - \overline{\boldsymbol{W}}(\boldsymbol{x}) \bullet \boldsymbol{A}(\boldsymbol{x}) - \widetilde{w}(\boldsymbol{x})(1-\boldsymbol{x}^T\boldsymbol{x}) - (\zeta_0^* - \epsilon) = \widehat{w}(x) \ \text{ for every } \boldsymbol{x} \in \mathbb{R}^n.$$

By (9), we can take a nonnegative integer $\hat{p} \geq \bar{p}$ such that $\widetilde{w}(\boldsymbol{x}) \in \Sigma(\widetilde{\mathcal{G}}_p)$ and $\widehat{w}(\boldsymbol{x}) \in \Sigma(\widehat{\mathcal{G}}_p)$ for every $p \geq \hat{p}$. Since $\overline{\boldsymbol{W}}(\boldsymbol{x}) \in \boldsymbol{\Sigma}(\mathcal{G}_{\bar{p}})^{m\times m} \subset \boldsymbol{\Sigma}(\mathcal{G}_p)^{m\times m}$ for every $p \geq \hat{p}$, we obtain that $(\zeta_0^* - \epsilon) \leq \eta_p^* \leq \zeta_0^*$ for every $p \geq \hat{p}$.

∎

It is possible to convert the problem $\widehat{\boldsymbol{\Lambda}}(\mathcal{G},\widetilde{\mathcal{G}},\widehat{\mathcal{G}})$ into an SDP by applying the conventional technique commonly used in SOS optimization, but we do not describe the conversion here. Instead, we derive the SDP in the next section as the dual of the SDP that is obtained by a primal approach to the polynomial SDP P_0.

5 A primal approach

The purpose of this section is twofold. The one is to derive an SDP relaxation directly from the polynomial SDP P_0 without applying either a Lagrangian dual or its SOS relaxation. This part is an extension of Lasserre's SDP relaxation for POPs to polynomial SDPs. The other purpose is to show that the dual of the SDP relaxation derived is equivalent to the the SOS optimization problem $\widehat{\mathbf{\Lambda}}(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}})$.

5.1 Adding valid symmetric polynomial matrix inequality constraints

For every nonempty finite subset $\mathcal{G}$ of $\mathbb{Z}^n_+$ and $\boldsymbol{x} \in \mathbb{R}^n$, let $(\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \mathcal{G})$ denote a column vector consisting of elements $\boldsymbol{x}^{\boldsymbol{\alpha}}$ $(\boldsymbol{\alpha} \in \mathcal{G})$. Although the order of the elements $\boldsymbol{x}^{\boldsymbol{\alpha}}$ $(\boldsymbol{\alpha} \in \mathcal{G})$ in the vector is not important in the succeeding discussion, we assume $\boldsymbol{x}^{\boldsymbol{\alpha}}$ precedes $\boldsymbol{x}^{\boldsymbol{\beta}}$ if and only if $\boldsymbol{\alpha} \in \mathcal{G}$ is lexicographically smaller than $\boldsymbol{\beta} \in \mathcal{G}$. In particular, the vector $(\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \mathcal{G})$ begins with $\boldsymbol{x}^{\mathbf{0}} = 1$ when $\mathbf{0} \in \mathcal{G}$.

Let $\mathcal{G}$, $\widetilde{\mathcal{G}}$ and $\widehat{\mathcal{G}}$ be nonempty finite subsets of $\mathbb{Z}^n_+$. Let s, t and u denote the dimensions of the vectors $(\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \mathcal{G})$, $(\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \widetilde{\mathcal{G}})$ and $(\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \widehat{\mathcal{G}})$, respectively. We consider the polynomial SDP

$$
\begin{array}{lll}
P_1(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}}): & \text{minimize} & a(\boldsymbol{x}) \\
& \text{subject to} & \left((\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \mathcal{G})(\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \mathcal{G})^T\right) \otimes \boldsymbol{A}(\boldsymbol{x}) \in \mathcal{S}^{sm}_+, \\
& & \left((\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \widetilde{\mathcal{G}})(\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \widetilde{\mathcal{G}})^T\right)(1 - \boldsymbol{x}^T\boldsymbol{x}) \in \mathcal{S}^t_+, \\
& & (\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \widehat{\mathcal{G}})(\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \widehat{\mathcal{G}})^T \in \mathcal{S}^u_+.
\end{array}
$$

Here $\otimes$ denotes the Kronecker product of two matrices. If both $\mathcal{G}$ and $\widetilde{\mathcal{G}}$ contain $\mathbf{0} \in \mathbb{Z}^n_+$, the polynomial SDP $P_1(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}})$ turns out to be equivalent to the original polynomial SDP P_0.

5.2 Linearization leading to an SDP relaxation

Since the left hand sides of the inclusion relations in the constraint are symmetric polynomial matrices, we can rewrite the polynomial SDP $P_1(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}})$ as

$$
\begin{array}{lll}
P_2(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}}): & \text{minimize} & \displaystyle\sum_{\boldsymbol{\alpha} \in \mathcal{D}(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}})} d_{\boldsymbol{\alpha}} \, \boldsymbol{x}^{\boldsymbol{\alpha}} \\
& \text{subject to} & \displaystyle\sum_{\boldsymbol{\alpha} \in \mathcal{D}(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}})} \boldsymbol{D}_{\boldsymbol{\alpha}} \, \boldsymbol{x}^{\boldsymbol{\alpha}} - \boldsymbol{D}_{\mathbf{0}} \in \mathcal{S}^{sm}_+ \times \mathcal{S}^t_+ \times \mathcal{S}^u_+.
\end{array}
$$

Here

$$
\begin{array}{l}
\mathbf{0} \notin \mathcal{D}(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}}) \subset \mathbb{Z}^n_+, \ d_{\boldsymbol{\alpha}} \in \mathbb{R} \ (\boldsymbol{\alpha} \in \mathcal{D}(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}})), \\
\boldsymbol{D}_{\boldsymbol{\alpha}} \in \mathcal{S}^{sm}_+ \times \mathcal{S}^t_+ \times \mathcal{S}^u_+ \ (\boldsymbol{\alpha} \in \{\mathbf{0}\} \bigcup \mathcal{D}(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}})).
\end{array}
$$

By replacing each monomial $\boldsymbol{x}^{\boldsymbol{\alpha}}$ by a single variable $y_{\boldsymbol{\alpha}} \in \mathbb{R}$ ($\boldsymbol{\alpha} \in \mathcal{D}(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}})$) in the polynomial SDP $P_2(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}})$, we obtain an SDP relaxation of the polynomial SDP P_0

$$\begin{array}{lll} P_3(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}}): & \text{minimize} & \displaystyle\sum_{\boldsymbol{\alpha} \in \mathcal{D}(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}})} d_{\boldsymbol{\alpha}}\, y_{\boldsymbol{\alpha}} \\ & \text{subject to} & \displaystyle\sum_{\boldsymbol{\alpha} \in \mathcal{D}(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}})} \boldsymbol{D}_{\boldsymbol{\alpha}}\, y_{\boldsymbol{\alpha}} - \boldsymbol{D}_{\boldsymbol{0}} \in \mathcal{S}^{sm}_{+} \times \mathcal{S}^{t}_{+} \times \mathcal{S}^{u}_{+}. \end{array}$$

5.3 Dual of $P_3(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}})$

The dual of the SDP $P_3(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}})$ is given by

$$\begin{array}{lll} P_4(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}}): & \text{maximize} & \boldsymbol{D}_{\boldsymbol{0}} \bullet \boldsymbol{X} \\ & \text{subject to} & \boldsymbol{D}_{\boldsymbol{\alpha}} \bullet \boldsymbol{X} = d_{\boldsymbol{\alpha}} \ (\boldsymbol{\alpha} \in \mathcal{D}(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}})), \ \boldsymbol{X} \in \mathcal{S}^{sm}_{+} \times \mathcal{S}^{t}_{+} \times \mathcal{S}^{u}_{+}. \end{array}$$

We write each feasible solution $\boldsymbol{X} \in \mathcal{S}^{sm}_{+} \times \mathcal{S}^{t}_{+} \times \mathcal{S}^{u}_{+}$ of $P_4(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}})$ as

$$\begin{array}{rcl} \boldsymbol{X} & = & \operatorname{diag}(\boldsymbol{V}, \widetilde{\boldsymbol{V}}, \widehat{\boldsymbol{V}}) = \begin{pmatrix} \boldsymbol{V} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \widetilde{\boldsymbol{V}} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \widehat{\boldsymbol{V}} \end{pmatrix} \in \mathcal{S}^{sm}_{+} \times \mathcal{S}^{t}_{+} \times \mathcal{S}^{u}_{+}, \\ \boldsymbol{V} & = & \left(\boldsymbol{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}} : (\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{G} \times \mathcal{G}\right) \in \mathcal{S}^{sm}_{+} \\ & & \text{(an } sm \times sm \text{ matrix whose } (\boldsymbol{\alpha}, \boldsymbol{\beta})\text{th block is an } m \times m \text{ matrix } \boldsymbol{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}}), \\ \widetilde{\boldsymbol{V}} & = & \left(\widetilde{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}} : (\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \widetilde{\mathcal{G}} \times \widetilde{\mathcal{G}}\right) \in \mathcal{S}^{t}_{+} \\ & & \text{(an } t \times t \text{ matrix whose } (\boldsymbol{\alpha}, \boldsymbol{\beta})\text{th element is } \widetilde{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}} \in \mathbb{R}), \\ \widehat{\boldsymbol{V}} & = & \left(\widehat{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}} : (\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \widehat{\mathcal{G}} \times \widehat{\mathcal{G}}\right) \in \mathcal{S}^{u}_{+} \\ & & \text{(an } u \times u \text{ matrix whose } (\boldsymbol{\alpha}, \boldsymbol{\beta})\text{th element is } \widehat{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}} \in \mathbb{R}). \end{array}$$

Then we know that $\boldsymbol{X} = \operatorname{diag}(\boldsymbol{V}, \widetilde{\boldsymbol{V}}, \widehat{\boldsymbol{V}})$ is a feasible solution of the dual SDP $P_4(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}})$ with the objective function value $\zeta = \boldsymbol{D}_{\boldsymbol{0}} \bullet \boldsymbol{X}$ if and only if the identity

$$\begin{array}{l} a(\boldsymbol{x}) - \boldsymbol{V} \bullet \left((\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \mathcal{G})(\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \mathcal{G})^T \otimes \boldsymbol{A}(\boldsymbol{x})\right) \\ \quad -\widetilde{\boldsymbol{V}} \bullet \left((\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \widetilde{\mathcal{G}})(\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \widetilde{\mathcal{G}})^T (1 - \boldsymbol{x}^T\boldsymbol{x})\right) \\ \quad -\widehat{\boldsymbol{V}} \bullet \left((\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \widehat{\mathcal{G}})(\boldsymbol{x}^{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \widehat{\mathcal{G}})^T\right) = \zeta \ \text{ for every } \boldsymbol{x} \in \mathbb{R}^n \end{array}$$

holds. See Section 6 of [3]. We can rewrite the left hand side of the identity above as

$$\begin{array}{l} a(\boldsymbol{x}) - \left(\displaystyle\sum_{\boldsymbol{\alpha} \in \mathcal{G}} \sum_{\boldsymbol{\beta} \in \mathcal{G}} \boldsymbol{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}}\, \boldsymbol{x}^{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\beta}}\right) \bullet \boldsymbol{A}(\boldsymbol{x}) \\ \quad - \left(\displaystyle\sum_{\boldsymbol{\alpha} \in \widetilde{\mathcal{G}}} \sum_{\boldsymbol{\beta} \in \widetilde{\mathcal{G}}} \widetilde{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}}\, \boldsymbol{x}^{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\beta}}\right) (1 - \boldsymbol{x}^T\boldsymbol{x}) \ - \displaystyle\sum_{\boldsymbol{\alpha} \in \widehat{\mathcal{G}}} \sum_{\boldsymbol{\beta} \in \widehat{\mathcal{G}}} \widehat{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}}\, \boldsymbol{x}^{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\beta}}. \end{array}$$

Now, recall the relations (3) and (4) on the set $\boldsymbol{\Sigma}^{m\times m}$ of sums of squares of polynomial matrices and the set Σ of sums of real valued polynomials. Then we see that

$$\begin{aligned}
&\sum_{\boldsymbol{\alpha}\in\mathcal{G}}\sum_{\boldsymbol{\beta}\in\mathcal{G}} \boldsymbol{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}}\, \boldsymbol{x}^{\boldsymbol{\alpha}}\boldsymbol{x}^{\boldsymbol{\beta}} \in \boldsymbol{\Sigma}^{m\times m}(\mathcal{G}) \subset \boldsymbol{\Sigma}^{m\times m},\\
&\sum_{\boldsymbol{\alpha}\in\widetilde{\mathcal{G}}}\sum_{\boldsymbol{\beta}\in\widetilde{\mathcal{G}}} \widetilde{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}}\, \boldsymbol{x}^{\boldsymbol{\alpha}}\boldsymbol{x}^{\boldsymbol{\beta}} \in \Sigma(\widetilde{\mathcal{G}}) \subset \Sigma,\\
&\sum_{\boldsymbol{\alpha}\in\widehat{\mathcal{G}}}\sum_{\boldsymbol{\beta}\in\widehat{\mathcal{G}}} \widehat{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}}\, \boldsymbol{x}^{\boldsymbol{\alpha}}\boldsymbol{x}^{\boldsymbol{\beta}} \in \Sigma(\widehat{\mathcal{G}}) \subset \Sigma.
\end{aligned}$$

Thus each feasible solution $\boldsymbol{X} = \mathrm{diag}(\boldsymbol{V}, \widetilde{\boldsymbol{V}}, \widehat{\boldsymbol{V}})$ of the dual SDP $P_4(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}})$ with the objective function value $\zeta = \boldsymbol{D}_{\boldsymbol{0}} \bullet \boldsymbol{X}$ has induced

$$(\boldsymbol{W}(\boldsymbol{x}), \widetilde{w}(\boldsymbol{x}), \widehat{w}(\boldsymbol{x})) \in \boldsymbol{\Sigma}^{m\times m}(\mathcal{G}) \times \Sigma(\widetilde{\mathcal{G}}) \times \Sigma(\widehat{\mathcal{G}})$$

satisfying

$$a(\boldsymbol{x}) - \boldsymbol{W}(\boldsymbol{x}) \bullet \boldsymbol{A}(\boldsymbol{x}) - \widetilde{w}(\boldsymbol{x})(1 - \boldsymbol{x}^T\boldsymbol{x}) - \widehat{w}(\boldsymbol{x}) = \zeta \;\; \text{for every } \boldsymbol{x} \in \mathbb{R}^n \tag{10}$$

by the relations

$$\begin{aligned}
\boldsymbol{W}(\boldsymbol{x}) &= \sum_{\boldsymbol{\alpha}\in\mathcal{G}}\sum_{\boldsymbol{\beta}\in\mathcal{G}} \boldsymbol{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}}\, \boldsymbol{x}^{\boldsymbol{\alpha}}\boldsymbol{x}^{\boldsymbol{\beta}} \in \boldsymbol{\Sigma}^{m\times m}(\mathcal{G}) \subset \boldsymbol{\Sigma}^{m\times m},\\
\widetilde{w}(\boldsymbol{x}) &= \sum_{\boldsymbol{\alpha}\in\widetilde{\mathcal{G}}}\sum_{\boldsymbol{\beta}\in\widetilde{\mathcal{G}}} \widetilde{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}}\, \boldsymbol{x}^{\boldsymbol{\alpha}}\boldsymbol{x}^{\boldsymbol{\beta}} \in \Sigma(\widetilde{\mathcal{G}}) \subset \Sigma,\\
\widehat{w}(\boldsymbol{x}) &= \sum_{\boldsymbol{\alpha}\in\widehat{\mathcal{G}}}\sum_{\boldsymbol{\beta}\in\widehat{\mathcal{G}}} \widehat{V}_{\boldsymbol{\alpha}\boldsymbol{\beta}}\, \boldsymbol{x}^{\boldsymbol{\alpha}}\boldsymbol{x}^{\boldsymbol{\beta}} \in \Sigma(\widehat{\mathcal{G}}) \subset \Sigma.
\end{aligned}$$

Conversely, every $(\boldsymbol{W}(\boldsymbol{x}), \widetilde{w}(\boldsymbol{x}), \widehat{w}(\boldsymbol{x})) \in \boldsymbol{\Sigma}^{m\times m}(\mathcal{G}) \times \Sigma(\widetilde{\mathcal{G}}) \times \Sigma(\widehat{\mathcal{G}})$ satisfying the identity (10) induces a feasible solution $\boldsymbol{X} = \mathrm{diag}(\boldsymbol{V}, \widetilde{\boldsymbol{V}}, \widehat{\boldsymbol{V}})$ of the dual SDP $P_4(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}})$ with the objective function value $\zeta = \boldsymbol{D}_{\boldsymbol{0}} \bullet \boldsymbol{X}$. Therefore we can rewrite the dual SDP $P_4(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}})$ as

$$\begin{array}{lll}
P_5(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}}): & \text{maximize} & \zeta \\
& \text{subject to} & a(\boldsymbol{x}) - \boldsymbol{W}(\boldsymbol{x}) \bullet \boldsymbol{A}(\boldsymbol{x}) - \widetilde{w}(\boldsymbol{x})(1 - \boldsymbol{x}^T\boldsymbol{x}) - \zeta \in \Sigma(\widehat{\mathcal{G}}), \\
& & \boldsymbol{W}(\boldsymbol{x}) \in \Sigma^{m\times m}(\mathcal{G}) \;\text{ and } \widetilde{w}(\boldsymbol{x}) \in \Sigma(\widetilde{\mathcal{G}}),
\end{array}$$

which is identical to the SOS optimization problem $\widehat{\boldsymbol{\Lambda}}(\mathcal{G}, \widetilde{\mathcal{G}}, \widehat{\mathcal{G}})$.

6 Concluding discussions

Throughout this paper, we have focused on theoretical convergence of the optimal values of the sequence of SOS relaxations of the polynomial SDP P_0 based on analysis of the penalized problem Ψ_p and the Lagrangian dual $\Lambda(\mathcal{G})$ of P_0 (Theorem 4.3). We have also shown a relationship between the SOS relaxation and Lasserre's SDP relaxation applied to the polynomial SDP P_0.

Practical aspects of the SOS relaxation are remaining issues to be investigated. To solve an SOS relaxation of a polynomial optimization problem or a polynomial SDP, we need to convert it into a conventional SDP. The size of the resulting SDP increases very rapidly as the original problem becomes larger and/or the maximum degree of the polynomials involved there grows. This prevents the SDP relaxations from being used widely in practice. One way to reduce this difficulty is to utilize powerful computing resources for solving large scale SDPs. See [5] for example. Another way is to exploit sparsity of the data of the original problem to reduce the size of its SDP relaxation without losing the effectiveness. We can apply similar techniques proposed in the recent paper [2] for sparse polynomial optimization problems to sparse polynomial SDPs. But those techniques may not be sufficient to solve polynomial SDPs except very small size and/or low degree problems.

The importance of polynomial SDPs that can be observed with a key application such as bilinear matrix inequalities in system and control theory has lead us to investigate the SOS relaxations of polynomial SDPs in this paper. A further extension of SOS relaxations to a class of polynomial conic optimization problems may be a possibility. This will be a subject of future study.

Acknowledgments. The author would like to thank Sunyoung Kim and Hayato Waki for helpful discussions on SOS and SDP relaxations of polynomial optimization problems, which lead him to investigate the subject of this paper.

References

[1] 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 Scienti que, 7 Avenue du Colonel Roche, 31 077 Toulouse, cedex 4, France, February, 2002.

[2] 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.

[3] 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** (2003) 125-144.

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

[5] R. D. C. Monteiro, "First- and second-order methods for semidefinite programming," *Mathematical Programming*, **97** (2003) 209–244.

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

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

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

[9] H. Wolkowicz, R. Saigal and L. Vandenberghe, eds., *Handbook of Semidefinite Programming, Theory, Algorithms, and Applications* (Kluwer Academic Publishers, Massachusetts, USA, 2000).