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

Sparsity in Sums of Squares of Polynomials

Masakazu Kojima[†], Sunyoung Kim[‡] and Hayato Waki[⋆]

Research Report B-391, June 2003. Revised July 2004

Abstract.
Representation of a given nonnegative multivariate polynomial in terms of a sum of squares of polynomials has become an essential subject in recent developments of sums of squares optimization and SDP (semidefinite programming) relaxation of polynomial optimization problems. We discuss effective methods to obtain a simpler representation of a "sparse" polynomial as a sum of squares of sparse polynomials by eliminating redundancy.

Key words.
Sums of Squares of Polynomial, Polynomial Optimization Problem, Semidefinite Program, Sparsity

† 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*

‡ Department of Mathematics, Ewha Women's University, 11-1 Dahyun-dong, Sudaemoon-gu, Seoul 120-750 Korea. A considerable part of this work was conducted while this author was visiting Tokyo Institute of Technology. Research supported by Kosef R004-000-2001-00200. *skim@ewha.ac.kr*.

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

1 Introduction

Determining global nonnegativity of a (multivariate) polynomial has been an important issue in many applications. If a polynomial can be represented as a sum of squares (of polynomials), global nonnegativity of the polynomial is guaranteed. Representing a polynomial as a sum of squares has gained a lot of attention in recent developments of sum of squares optimization [10, 11] and SDP (semidefinite programming) relaxation of polynomial optimization problems [5, 6, 7, 8].

When we aim to represent a nonnegative polynomial in terms of a sum of squares of polynomials, we need to address two issues of whether such representation is possible and how it can be computed. The first issue is studied by many researchers starting from Hilbert. See [15]. The second computational issue has been dealt with from various viewpoints as in [3, 13, 16]. The focus of this paper is on increasing computational efficiency to represent sparse polynomials as a sum of squares. It has been known [3, 13] that the problem of representing a nonnegative polynomial as a sum of squares can be converted to an LMI (linear matrix inequality) [2], whose size determines the computational efficiency. We present effective methods to obtain a simpler representation of a sparse polynomial as a sum of squares of sparse polynomials by eliminating redundancy and, therefore, reducing the size of the resulting LMI.

We consider a polynomial

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

in the variable vector $\boldsymbol{x} \in \mathbb{R}^n$ with a support $\mathcal{F} \subset \mathbb{Z}^n_+$ and real coefficients $c_\alpha \neq 0$ ($\boldsymbol{\alpha} \in \mathcal{F}$). Here, for every variable vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$ and every $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n)^T \in \mathbb{Z}^n_+$, we use the notation $\boldsymbol{x}^{\boldsymbol{\alpha}}$ for the term $x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$. We write any polynomial in a variable vector $\boldsymbol{x} \in \mathbb{R}^n$ as $\sum_{\boldsymbol{\alpha}\in\mathcal{A}} d_\alpha \boldsymbol{x}^{\boldsymbol{\alpha}}$ for some nonempty finite subset $\mathcal{A}$ of $\mathbb{Z}^n_+$ and some $d_\alpha \in \mathbb{R}$ ($\boldsymbol{\alpha} \in \mathcal{A}$). We call $\mathcal{A}$ the support of the polynomial. In general, we allow some d_α to be zero.

Assume that $f(\boldsymbol{x})$ is a nonnegative polynomial; $f(\boldsymbol{x}) \geq 0$ for every $\boldsymbol{x} \in \mathbb{R}^n$. Then, the degree of the polynomial $f(\boldsymbol{x})$, *i.e.*, the highest degree of monomials $\boldsymbol{x}^{\boldsymbol{\alpha}}$ ($\boldsymbol{\alpha} \in \mathcal{F}$) of the polynomial $f(\boldsymbol{x})$, must be a nonnegative even integer. Furthermore the following fact is known (Lemma in Section 3 of [14]). Let $\mathcal{F}^e$ denote the set of integer vectors $\boldsymbol{\alpha} \in \mathcal{F}$ with coordinates α_k ($k = 1, 2, \ldots, n$) of even nonnegative integers; $\mathcal{F}^e = \mathcal{F} \cap \left(2\mathbb{Z}^n_+\right)$, where μA denotes the set $\{\mu \boldsymbol{a} : \boldsymbol{a} \in A\}$ for every $A \subset \mathbb{R}^n$ and every $\mu \geq 0$. Then $\mathcal{F}$ is contained in $\mathrm{co}(\mathcal{F}^e)$, the convex hull of $\mathcal{F}^e$.

Now assume that $f(\boldsymbol{x})$ can be represented as a sum of squares of polynomials $g^1(\boldsymbol{x}), g^2(\boldsymbol{x}), \ldots, g^r(\boldsymbol{x})$ such that

$$f(\boldsymbol{x}) = \sum_{i=1}^{r} g^i(\boldsymbol{x})^2, \tag{2}$$

where both r and polynomials $g^1(\boldsymbol{x}), g^2(\boldsymbol{x}), \ldots, g^r(\boldsymbol{x})$ are unknowns to be found. It is necessary to estimate and decide supports of unknown polynomial $g^i(\boldsymbol{x})$ ($i = 1, 2, \ldots, r$) prior to computing them. Let $\mathcal{G}_i$ denote an unknown support of the polynomial $g^i(\boldsymbol{x})$

$(i = 1, 2, \ldots, r)$. Then each polynomial $g^i(\boldsymbol{x})$ is represented as

$$g^i(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha} \in \mathcal{G}_i} v^i_\alpha \boldsymbol{x}^{\boldsymbol{\alpha}}$$

with some nonzero or zero coefficients v^i_α $(\boldsymbol{\alpha} \in \mathcal{G}_i)$ $(i = 1, 2, \ldots, r)$.

Since $f(\boldsymbol{x})$ is a nonnegative polynomial, we can apply the above fact, $\mathcal{F} \subset \mathrm{co}(\mathcal{F}^e)$ to the polynomial $f(\boldsymbol{x})$. The following relation is also known (Theorem 1 of [14]):

$$\left\{\boldsymbol{\alpha} \in \mathbb{Z}^n_+ : \boldsymbol{\alpha} \in \mathcal{G}_i \text{ and } v^i_\alpha \neq 0 \text{ for some } i \in \{1, 2, \ldots, r\}\right\} \subset \frac{1}{2}\mathrm{co}(\mathcal{F}^e).$$

Hence we can confine effective supports of unknown polynomials $g^1(\boldsymbol{x}), g^2(\boldsymbol{x}), \ldots, g^r(\boldsymbol{x})$ to subsets of

$$\mathcal{G}^0 = \left(\frac{1}{2}\mathrm{co}(\mathcal{F}^e)\right) \cap \mathbb{Z}^n_+. \tag{3}$$

Without loss of generality, we may further assume that all polynomials $g^1(\boldsymbol{x}), g^2(\boldsymbol{x}), \ldots, g^r(\boldsymbol{x})$ share a common support $\mathcal{G} \subset \mathcal{G}^0$; take the union of supports $\mathcal{G}_i$ $(i = 1, 2, \ldots, r)$ of all polynomials for $\mathcal{G}$ and define the coefficient of the monomial $\boldsymbol{x}^{\boldsymbol{\alpha}}$ of $g^i(\boldsymbol{x})$ to be zero if $\boldsymbol{\alpha} \in \mathcal{G}$ is not contained in the original support $\mathcal{G}_i$ of $g^i(\boldsymbol{x})$. Then each polynomial $g^i(\boldsymbol{x})$ is represented as

$$g^i(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha} \in \mathcal{G}} v^i_\alpha \boldsymbol{x}^{\boldsymbol{\alpha}} \tag{4}$$

with some nonzero or zero coefficients v^i_α $(\boldsymbol{\alpha} \in \mathcal{G})$ $(i = 1, 2, \ldots, r)$.

Reducing the size of $\mathcal{G}$ helps us achieve computational efficiency. After a common support $\mathcal{G}$ is determined correctly, the problem of computing unknown polynomials $g^1(\boldsymbol{x}), g^2(\boldsymbol{x}), \ldots, g^r(\boldsymbol{x})$ can be converted to an LMI (linear matrix inequality) [10, 11]. The size of a symmetric matrix variable of the resulting LMI is $s \times s$, where s denotes the cardinality of $\mathcal{G}$ or the number of integer vectors of $\mathcal{G}$. When we apply interior-point methods to the resulting LMI, the numerical efficiency of solving an LMI depends heavily on its size. Although it is safe to take the maximal possible support $\mathcal{G}^0$ for unknown polynomials $g^1(\boldsymbol{x}), g^2(\boldsymbol{x}), \ldots, g^r(\boldsymbol{x})$, a smaller support $\mathcal{G} \subset \mathcal{G}^0$ provides practical computational efficiency. We present this conversion in detail in Section 2.2.

We call a polynomial $f(\boldsymbol{x})$ sparse if the number of elements in the support $\mathcal{F}$ is much less than the number of the maximal support $\left\{\boldsymbol{\alpha} \in \mathbb{Z}^n_+ : \boldsymbol{e}^T\boldsymbol{\alpha} \leq 2\rho\right\}$ over polynomials of degree 2ρ in the variable vector $\boldsymbol{x} \in \mathbb{R}^n$. Sparse polynomials can provide a small and sparse support $\mathcal{G}$ for unknown polynomials $g^1(\boldsymbol{x}), g^2(\boldsymbol{x}), \ldots, g^r(\boldsymbol{x})$ compared with the maximal support $\left\{\boldsymbol{\alpha} \in \mathbb{Z}^n_+ : \boldsymbol{e}^T\boldsymbol{\alpha} \leq \rho\right\}$ over polynomials of degree r in the variable vector $\boldsymbol{x} \in \mathbb{R}^n$. In a numerical method to construct $\mathcal{G}$, the set $\mathcal{G}^0$ given in (3) serves as an initial big guess of a legitimate support for unknown polynomials $g^1(\boldsymbol{x}), g^2(\boldsymbol{x}), \ldots, g^r(\boldsymbol{x})$ for which the identity (2) holds. The next step is to eliminate unnecessary integer vectors from $\mathcal{G}^0$ in a continued manner to finally arrive at a small $\mathcal{G}$. We present numerical methods to reduce the size of $\mathcal{G}$ as much as possible.

It is common to have a polynomial with the coefficients $c_\alpha(\boldsymbol{w})$ $(\boldsymbol{\alpha} \in \mathcal{F})$ of $f(\boldsymbol{x})$ as linear functions of a parameter vector $\boldsymbol{w} \in \mathbb{R}^m$: $f(\boldsymbol{x}, \boldsymbol{w}) = \sum_{\boldsymbol{\alpha} \in \mathcal{F}} c_\alpha(\boldsymbol{w})\boldsymbol{x}^{\boldsymbol{\alpha}}$, in many applications arising from sum of squares optimization problems [10, 11] and SDP (semidefinite

programming) relaxation of polynomial optimization problems [5, 6, 7, 8]. It is an extension of (1) where the coefficients $c_\alpha(\boldsymbol{w})$ ($\boldsymbol{\alpha} \in \mathcal{F}$) of the polynomial $f(\boldsymbol{x})$ are constant. The goal is to generate a small subset $\mathcal{G}$ of $\mathbb{Z}^n_+$ such that for each fixed $\boldsymbol{w} \in \mathbb{R}^m$ $f(\boldsymbol{x}, \boldsymbol{w})$ is sum of squares of a finite number of polynomials $g^1(\boldsymbol{x}, \boldsymbol{w}), g^2(\boldsymbol{x}, \boldsymbol{w}), \ldots, g^r(\boldsymbol{x}, \boldsymbol{w})$ having a support $\mathcal{G}$. Because the proposed methods to generate an effective small support $\mathcal{G}$ for unknown polynomials $g^1(\boldsymbol{x}, \boldsymbol{w}), g^2(\boldsymbol{x}, \boldsymbol{w}), \ldots, g^r(\boldsymbol{x}, \boldsymbol{w})$ do not depend on the coefficients of $f(\boldsymbol{x}, \boldsymbol{w})$ but only on the support $\mathcal{F}$ of $f(\boldsymbol{x}, \boldsymbol{w})$, they can be applied to handle polynomials with parameterized coefficients mentioned above.

Throughout the paper, we consider a simple unconstrained polynomial optimization problem as an illustrative example,

$$\text{minimize } f_0(\boldsymbol{x}) \equiv -4x_1^3x_2^4 + 2x_1^4x_2^3 + 5x_1^6x_2^8 - 2x_1^7x_2^7 + 2x_1^8x_2^6.$$

We can rewrite this problem as

$$\text{maximize } \zeta \text{ subject to } f_0(\boldsymbol{x}) - \zeta \geq 0.$$

Replacing the inequality constraint $f_0(\boldsymbol{x}) - \zeta \geq 0$ by the condition that $f_0(\boldsymbol{x}) - \zeta$ is a sum of squares polynomials, we obtain a sum of squares optimization problem [10, 11]

$$\text{maximize } \zeta \text{ subject to } f_0(\boldsymbol{x}) - \zeta \text{ is a sum of squares of polynomials.}$$

Let $f(\boldsymbol{x}, \zeta) = f_0(\boldsymbol{x}) - \zeta$. Then, for each $\zeta \in \mathbb{R}$ such that $f(\boldsymbol{x}, \zeta) = f_0(\boldsymbol{x}) - \zeta$ is a sum of squares of polynomials, we want to represent $f(\boldsymbol{x}, \zeta) = \sum_{i=1}^r g^i(\boldsymbol{x}, \zeta)^2$ for some polynomials $g^1(\boldsymbol{x}, \zeta), g^2(\boldsymbol{x}, \zeta), \ldots, g^r(\boldsymbol{x}, \zeta)$. In this case, we see

$$\left.\begin{array}{rcl} \mathcal{F} &=& \left\{\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}3\\4\end{pmatrix}, \begin{pmatrix}4\\3\end{pmatrix}, \begin{pmatrix}6\\8\end{pmatrix}, \begin{pmatrix}7\\7\end{pmatrix}, \begin{pmatrix}8\\6\end{pmatrix}\right\}, \\ 2\rho &=& 14 \text{ (the degree of the polynomial } f(\boldsymbol{x}, \zeta)) \\ \mathcal{F}^e &=& \left\{\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}6\\8\end{pmatrix}, \begin{pmatrix}8\\6\end{pmatrix}\right\}, \\ \mathcal{G}^0 &=& \left\{\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}2\\2\end{pmatrix}, \begin{pmatrix}3\\3\end{pmatrix}, \begin{pmatrix}3\\4\end{pmatrix}, \begin{pmatrix}4\\3\end{pmatrix}\right\} \end{array}\right\} \quad (5)$$

By applying our method, which we will present in Section 4, we can eliminate 3 unnecessary integer vectors from the initial guess $\mathcal{G}^0$ to generate

$$\mathcal{G} = \left\{\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}3\\4\end{pmatrix}, \begin{pmatrix}4\\3\end{pmatrix}\right\}.$$

Since we can represent $f(\boldsymbol{x}, \zeta)$ as

$$f(\boldsymbol{x}, \zeta) = \left(\sqrt{-1-\zeta}\right)^2 + \left(1 - 2x_1^3x_2^4 + x_1^4x_2^3\right)^2 + \left(x_1^3x_2^4 + x_1^4x_2^3\right)^2$$

for every $\zeta \leq -1$, we know $\mathcal{G}$ computed is the minimal legitimate support for unknown polynomials $g^1(\boldsymbol{x}, \zeta), g^2(\boldsymbol{x}, \zeta), \ldots, g^r(\boldsymbol{x}, \zeta)$ for which the identity $f(\boldsymbol{x}, \zeta) = \sum_{i=1}^r g^i(\boldsymbol{x}, \zeta)^2$ holds in this case.

This paper is organized as follows: In Section 2, we discuss how polynomials represented in a sum of squares can be converted to an LMI. In Section 3, we examine theoretical properties of supports to express a nonnegative polynomial as a sum of squares. These theoretical properties on supports lead to algorithms to compute the smallest support in a certain class of common effective supports of unknown polynomial $g^1(\boldsymbol{x}), g^2(\boldsymbol{x}), \ldots, g^r(\boldsymbol{x})$. We present numerical methods in Section 4 for the smallest support. In Section 5, we present some basic properties on the smallest support. It is followed by numerical experiments in Section 6. Section 7 is devoted to concluding remarks.

2 Preliminaries

2.1 Notation and symbols

Suppose that $\mathcal{A}$ is a nonempty finite subset of $\mathbb{Z}^n_+$. Let $|\mathcal{A}|$ denote the cardinality of $\mathcal{A}$ and $\mathbb{R}^{\mathcal{A}}$ the $|\mathcal{A}|$-dimensional Euclidean space whose coordinates are indexed by $\boldsymbol{\alpha} \in \mathcal{A}$. 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 $\mathbb{R}^{\mathcal{A}}$ is denoted as $\boldsymbol{v} = (v_\alpha : \boldsymbol{\alpha} \in \mathcal{A})$. We use the symbol $\mathcal{S}^{\mathcal{A}}_+$ for the set of $|\mathcal{A}| \times |\mathcal{A}|$ symmetric matrices with coordinates $\boldsymbol{\alpha} \in \mathcal{A}$; each $\boldsymbol{V} \in \mathcal{S}^{\mathcal{A}}_+$ has elements $V_{\alpha\beta}$ ($\boldsymbol{\alpha} \in \mathcal{A},\ \boldsymbol{\beta} \in \mathcal{A}$) such that $V_{\alpha\beta} = V_{\beta\alpha}$ and that $\boldsymbol{w}^T\boldsymbol{V}\boldsymbol{w} = \sum_{\boldsymbol{\alpha}\in\mathcal{A}}\sum_{\boldsymbol{\beta}\in\mathcal{A}} V_{\alpha\beta}w_\alpha w_\beta \geq 0$ for every $\boldsymbol{w} = (w_\alpha : \boldsymbol{\alpha} \in \mathcal{A})$. For every $\boldsymbol{x} \in \mathbb{R}^n$, let $\boldsymbol{u}(\boldsymbol{x}, \mathcal{A}) = (\boldsymbol{x}^\alpha : \boldsymbol{\alpha} \in \mathcal{A})$, a column vector consisting of elements $\boldsymbol{x}^\alpha$ ($\boldsymbol{\alpha} \in \mathcal{A}$).

Using the notation and symbols introduced above, we can rewrite (1) as $f(\boldsymbol{x}) = \boldsymbol{c}^T\boldsymbol{u}(\boldsymbol{x}, \mathcal{F})$, where $\boldsymbol{c} = (c_\alpha : \boldsymbol{\alpha} \in \mathcal{F}) \in \mathbb{R}^{\mathcal{F}}$, and (4) as

$$g^i(\boldsymbol{x}) = (\boldsymbol{v}^i)^T\boldsymbol{u}(\boldsymbol{x}, \mathcal{G}), \tag{6}$$

where $\boldsymbol{v}^i = (v^i_\alpha : \boldsymbol{\alpha} \in \mathcal{G}) \in \mathbb{R}^{\mathcal{G}}$ $(i = 1, 2, \ldots, r)$.

For every pair of nonempty subsets $\mathcal{A}$ and $\mathcal{B}$ of $\mathbb{Z}^n_+$, we use the notation $\mathcal{A} + \mathcal{B}$ for their Minkovski sum:

$$\mathcal{A} + \mathcal{B} = \{\boldsymbol{a} + \boldsymbol{b} : \boldsymbol{a} \in \mathcal{A},\ \boldsymbol{b} \in \mathcal{B}\}.$$

2.2 Reduction of sums of squares of polynomials to linear matrix inequalities

The lemma below is well-known ([3, 10, 13]). We provide a proof for the reader to see how we reduce the computation of unknown polynomials $g^1(\boldsymbol{x}), g^2(\boldsymbol{x}), \ldots, g^r(\boldsymbol{x})$ satisfying the identity (2) without knowing r.

Lemma 2.1. *A polynomial $f(\boldsymbol{x})$ of the form (1) is a sum of squares of a finite number of polynomials with a common support $\mathcal{G}$ if and only if there exists a $\boldsymbol{V} \in \mathcal{S}^{\mathcal{G}}_+$ such that*

$$f(\boldsymbol{x}) = \boldsymbol{u}(\boldsymbol{x}, \mathcal{G})^T\boldsymbol{V}\boldsymbol{u}(\boldsymbol{x}, \mathcal{G}) = \sum_{\boldsymbol{\alpha}\in\mathcal{G}}\sum_{\boldsymbol{\beta}\in\mathcal{G}} V_{\alpha\beta}\boldsymbol{x}^{\boldsymbol{\alpha}+\boldsymbol{\beta}}. \tag{7}$$

Proof: Suppose that $f(\boldsymbol{x})$ is a sum of squares of a finite number of polynomials $g^1(\boldsymbol{x}), g^2(\boldsymbol{x}), \ldots, g^r(\boldsymbol{x})$ in (2) represented as (6) for some $\boldsymbol{v}^i = (v^i_\alpha : \boldsymbol{\alpha} \in \mathcal{G}) \in \mathbb{R}^{\mathcal{G}}$ $(i = 1, 2, \ldots, r)$. Substituting $g^i(\boldsymbol{x})$ of the form (6) into (2), we see that

$$\begin{aligned} f(\boldsymbol{x}) &= \sum_{i=1}^{r} g^i(\boldsymbol{x})^2 \\ &= \sum_{i=1}^{r} \left((\boldsymbol{v}^i)^T \boldsymbol{u}(\boldsymbol{x}, \mathcal{G})\right)^2 \\ &= \boldsymbol{u}(\boldsymbol{x}, \mathcal{G})^T \left(\sum_{i=1}^{r} \boldsymbol{v}^i (\boldsymbol{v}^i)^T \right) \boldsymbol{u}(\boldsymbol{x}, \mathcal{G}) \\ &= \boldsymbol{u}(\boldsymbol{x}, \mathcal{G})^T \boldsymbol{V} \boldsymbol{u}(\boldsymbol{x}, \mathcal{G}), \end{aligned}$$

where $\boldsymbol{V} = \sum_{i=1}^{r} \boldsymbol{v}^i (\boldsymbol{v}^i)^T \in \mathcal{S}^{\mathcal{G}}_+$. Thus, we have shown that if $f(\boldsymbol{x})$ is a sum of squares of finite number of polynomials with a common support $\mathcal{G}$, then there exists a $\boldsymbol{V} \in \mathcal{S}^{\mathcal{G}}_+$ satisfying (7). The converse is true since each $\boldsymbol{V} \in \mathcal{S}^{\mathcal{G}}_+$ ca be factorized as $\boldsymbol{V} = \sum_{i=1}^{r} \boldsymbol{v}^i (\boldsymbol{v}^i)^T$ for some $\boldsymbol{v}^i = (v^i_\alpha : \boldsymbol{\alpha} \in \mathcal{G}) \in \mathbb{R}^{\mathcal{G}}$ $(i = 1, 2, \ldots, r)$ and some positive integer r. ∎

Now, we convert the problem of finding $\boldsymbol{V} \in \mathcal{S}^{\mathcal{G}}_+$ satisfying (7) to an LMI (linear matrix inequality) [2]. Substituting $f(\boldsymbol{x}) = \sum_{\boldsymbol{\gamma} \in \mathcal{F}} c_\gamma \boldsymbol{x}^{\boldsymbol{\gamma}}$ into (7), we see that

$$\sum_{\boldsymbol{\gamma} \in \mathcal{F}} c_\gamma \boldsymbol{x}^{\boldsymbol{\gamma}} = \sum_{\boldsymbol{\alpha} \in \mathcal{G}} \sum_{\boldsymbol{\beta} \in \mathcal{G}} V_{\alpha\beta} \boldsymbol{x}^{\boldsymbol{\alpha}+\boldsymbol{\beta}} = \sum_{(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{C}} V_{\alpha\beta} \boldsymbol{x}^{\boldsymbol{\alpha}+\boldsymbol{\beta}} + \sum_{(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{C}, \boldsymbol{\alpha} \neq \boldsymbol{\beta}} V_{\alpha\beta} \boldsymbol{x}^{\boldsymbol{\alpha}+\boldsymbol{\beta}}.$$

Here $\mathcal{C} = \{(\boldsymbol{\alpha}, \boldsymbol{\beta}) : \boldsymbol{\alpha} \in \mathcal{G}, \boldsymbol{\beta} \in \mathcal{G}, \boldsymbol{\alpha} \preceq_\ell \boldsymbol{\beta}\}$, and $\boldsymbol{\alpha} \preceq_\ell \boldsymbol{\beta}$ denotes that $\boldsymbol{\alpha}$ is lexicographically smaller than or equal to $\boldsymbol{\beta}$. By comparing the coefficients of the monomials $\boldsymbol{x}^{\boldsymbol{\gamma}}$ $(\boldsymbol{\gamma} \in \mathcal{G}+\mathcal{G})$, we see that $\boldsymbol{V} \in \mathcal{S}^{\mathcal{G}}_+$ satisfies this identity for every $\boldsymbol{x} \in \mathbb{R}^n$ if and only if it satisfies a linear system of equations

$$\left.\begin{aligned} c_\gamma &= \sum_{(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{C}, \boldsymbol{\gamma} = \boldsymbol{\alpha}+\boldsymbol{\beta}} V_{\alpha\beta} + \sum_{(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{C}, \boldsymbol{\gamma} = \boldsymbol{\alpha}+\boldsymbol{\beta}, \boldsymbol{\alpha} \neq \boldsymbol{\beta}} V_{\alpha\beta} \quad \text{for every } \boldsymbol{\gamma} \in \mathcal{F}, \\ 0 &= \sum_{(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{C}, \boldsymbol{\gamma} = \boldsymbol{\alpha}+\boldsymbol{\beta}} V_{\alpha\beta} + \sum_{(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{C}, \boldsymbol{\gamma} = \boldsymbol{\alpha}+\boldsymbol{\beta}, \boldsymbol{\alpha} \neq \boldsymbol{\beta}} V_{\alpha\beta} \\ &\qquad\qquad \text{for every } \boldsymbol{\gamma} \in (\mathcal{G}+\mathcal{G}) \setminus \mathcal{F}. \end{aligned}\right\} \tag{8}$$

Let σ and τ denote the cardinalities of the sets $\mathcal{G}$ and $\mathcal{G}+\mathcal{G}$, respectively. It is convenient to introduce $\sigma(\sigma+1)/2$ constant matrices $\boldsymbol{E}_{\alpha\beta} \in \mathcal{S}^{\mathcal{G}}$ with 1 in $(\boldsymbol{\alpha}, \boldsymbol{\beta})$th and $(\boldsymbol{\beta}, \boldsymbol{\alpha})$th positions and 0 elsewhere for $\boldsymbol{\alpha} \preceq_\ell \boldsymbol{\beta}$. Then, the positive semidefinite condition $\boldsymbol{V} \in \mathcal{S}^{\mathcal{G}}_+$ can be rewritten as

$$\boldsymbol{V} = \sum_{(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{C}} V_{\alpha\beta} \boldsymbol{E}_{\alpha\beta} \in \mathcal{S}^{\mathcal{G}}_+. \tag{9}$$

Thus (8) and (9) forms an LMI (linear matrix inequality) in the variables $V_{\alpha\beta}$ $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathcal{C}$. We mention here that there are many software packages [17, 18, 19], based on interior-point methods for semidefinite programs, to solve LMIs.

3 Representation of a nonnegative polynomial as a sum of squares

Throughout this section, we assume that a polynomial $f(\boldsymbol{x})$ given in (1) with a support $\mathcal{F} \subset \mathbb{Z}^n_+$ and real coefficients $c_\alpha \neq 0$ $(\boldsymbol{\alpha} \in \mathcal{F})$ is a sum of squares of a finite number of unknown polynomials $g^i(\boldsymbol{x})$ $(i = 1, 2, \ldots, r)$ given in (4) with a common support $\mathcal{G} \subset \mathbb{Z}^n_+$ and coefficients v^i_α $(\boldsymbol{\alpha} \in \mathcal{G},\ i = 1, 2, \ldots, r)$, *i.e.*, the identity (2) holds.

Lemma 3.1. *Assume that*

$$\left.\begin{array}{l} \emptyset \neq \mathcal{H} \subset \mathcal{G},\ \emptyset \neq \mathcal{B} \subset \mathcal{G},\ \mathcal{G} = \mathcal{H} \cup \mathcal{B}, \\ (\mathcal{B}+\mathcal{B}) \cap \mathcal{F} = \emptyset \ \ \textit{and}\ \ (\mathcal{B}+\mathcal{B}) \cap (\mathcal{G}+\mathcal{H}) = \emptyset. \end{array}\right\} \tag{10}$$

Then,

$$v^i_\alpha = 0 \ \ \textit{for every}\ \boldsymbol{\alpha} \in \mathcal{B}\ (i = 1, 2, \ldots, r)\ \textit{and}\ \ f(\boldsymbol{x}) = \sum_{i=1}^{r} \left(\sum_{\boldsymbol{\alpha} \in \mathcal{H}} v^i_\alpha \boldsymbol{x}^{\boldsymbol{\alpha}} \right)^2 .$$

(Note that the condition (10) implies $\mathcal{B} \cap \mathcal{H} = \emptyset$.)

Proof: First we observe that

$$\begin{aligned} f(\boldsymbol{x}) &= \sum_{\boldsymbol{\alpha} \in \mathcal{F}} c_\alpha \boldsymbol{x}^{\boldsymbol{\alpha}} \\ &= \sum_{i=1}^{r} \left(\sum_{\boldsymbol{\alpha} \in \mathcal{B} \cup \mathcal{H}} v^i_\alpha \boldsymbol{x}^{\boldsymbol{\alpha}} \right)^2 \\ &= \sum_{i=1}^{r} \left(\sum_{\boldsymbol{\alpha} \in \mathcal{B}} v^i_\alpha \boldsymbol{x}^{\boldsymbol{\alpha}} + \sum_{\boldsymbol{\alpha} \in \mathcal{H}} v^i_\alpha \boldsymbol{x}^{\boldsymbol{\alpha}} \right)^2 \\ &= \sum_{i=1}^{r} \left(\left(\sum_{\boldsymbol{\alpha} \in \mathcal{B}} v^i_\alpha \boldsymbol{x}^{\boldsymbol{\alpha}} \right)^2 + 2 \sum_{\boldsymbol{\alpha} \in \mathcal{B}} \sum_{\boldsymbol{\beta} \in \mathcal{H}} v^i_\alpha v^i_\beta \boldsymbol{x}^{\boldsymbol{\alpha}+\boldsymbol{\beta}} + \left(\sum_{\boldsymbol{\beta} \in \mathcal{H}} v^i_\beta \boldsymbol{x}^{\boldsymbol{\beta}} \right)^2 \right). \end{aligned}$$

Letting

$$\begin{aligned} \varphi_1(\boldsymbol{x}) &= \sum_{i=1}^{r} \left(\sum_{\boldsymbol{\alpha} \in \mathcal{B}} v^i_\alpha \boldsymbol{x}^{\boldsymbol{\alpha}} \right)^2 \quad \text{and} \\ \varphi_2(\boldsymbol{x}) &= 2 \sum_{i=1}^{r} \sum_{\boldsymbol{\alpha} \in \mathcal{B}} \sum_{\boldsymbol{\beta} \in \mathcal{H}} v^i_\alpha v^i_\beta \boldsymbol{x}^{\boldsymbol{\alpha}+\boldsymbol{\beta}} + \sum_{i=1}^{r} \left(\sum_{\boldsymbol{\beta} \in \mathcal{H}} v^i_\beta \boldsymbol{x}^{\boldsymbol{\beta}} \right)^2 , \end{aligned}$$

we thus obtain the identity $f(\boldsymbol{x}) = \varphi_1(\boldsymbol{x}) + \varphi_2(\boldsymbol{x})$. Note that the supports of the polynomials $f(\boldsymbol{x})$, $\varphi_1(\boldsymbol{x})$ and $\varphi_2(\boldsymbol{x})$ are $\mathcal{F}$, $\mathcal{B}+\mathcal{B}$ and $\mathcal{G}+\mathcal{H}$ for which the relation (10) holds. Hence the identify above implies that $\varphi_1(\boldsymbol{x}) = \sum_{i=1}^{r} \left(\sum_{\boldsymbol{\alpha} \in \mathcal{B}} v^i_\alpha \boldsymbol{x}^{\boldsymbol{\alpha}} \right)^2 = 0$. Therefore, the desired result follows. ∎

By restricting $\mathcal{B}$ to a singleton in $\mathcal{G}$, we have the following corollary.

Corollary 3.2. ((2) of Proposition 3.7 of [3]). *Assume that*

$$\hat{\boldsymbol{\alpha}} \in \mathcal{G},\ \mathcal{G}\backslash\{\hat{\boldsymbol{\alpha}}\} \neq \emptyset,\ 2\hat{\boldsymbol{\alpha}} \notin \mathcal{F}^e \quad \textit{and}\ 2\hat{\boldsymbol{\alpha}} \notin (\mathcal{G} + \mathcal{G}\backslash\{\hat{\boldsymbol{\alpha}}\})\,. \tag{11}$$

Then,

$$v^i_{\hat{\alpha}} = 0\ (i = 1, 2, \ldots, r)\ \textit{and}\ \ f(\boldsymbol{x}) = \sum_{i=1}^{r} \left(\sum_{\boldsymbol{\alpha} \in \mathcal{G}\backslash\{\hat{\boldsymbol{\alpha}}\}} v^i_{\alpha} \boldsymbol{x}^{\boldsymbol{\alpha}} \right)^2 .$$

As mentioned in the Introduction, we know that the support $\mathcal{F}$ of $f(\boldsymbol{x})$ is contained in $\mathrm{co}(\mathcal{F}^e)$, the convex hull of the set $\mathcal{F}^e$ of integer vectors $\boldsymbol{\alpha} \in \mathcal{F}$ with coordinates α_k $(k = 1, 2, \ldots, n)$ of even nonnegative integers, and we can take $\mathcal{G}^0$ defined by (3) for an effective common support of unknown polynomials $g^i(\boldsymbol{x})$ $(i = 1, 2, \ldots, r)$. Let $\mathcal{G} = \mathcal{G}^0$. By applying Lemma 3.1, we can recursively define a family Γ of common supports of unknown polynomials $g^1(\boldsymbol{x}), g^2(\boldsymbol{x}), \ldots, g^r(\boldsymbol{x})$ for which (2) holds as the following steps:

(i) $\mathcal{G}^0 \in \Gamma$.

(ii) if $\mathcal{G} \in \Gamma$ and there exist $\mathcal{B}$ and $\mathcal{H}$ satisfying (10), then $\mathcal{H} \in \Gamma$.

We know that $\frac{1}{2}\mathcal{F}^e \subset \mathcal{G}^0$ and that any $\hat{\boldsymbol{\alpha}} \in \frac{1}{2}\mathcal{F}^e$ can not satisfy (11). Hence

$$\frac{1}{2}\mathcal{F}^e \subset \mathcal{G} \ \text{ and } \#\mathcal{F}^e \leq \#\mathcal{G} \ \text{ for every } \mathcal{G} \in \Gamma. \tag{12}$$

We show that the family Γ is closed under intersection: if $\mathcal{G} \in \Gamma$ and $\mathcal{G}' \in \Gamma$ then $\mathcal{G} \cap \mathcal{G}' \in \Gamma$. This property of Γ guarantees the existence of the smallest support $\mathcal{G}^*$ in Γ, which is unique in the sense that $\mathcal{G}^* \subset \mathcal{G}$ for every $\mathcal{G} \in \Gamma$. We use $\mathcal{G}^*$ in practical computation of unknown polynomials $g^1(\boldsymbol{x}), g^2(\boldsymbol{x}), \ldots, g^r(\boldsymbol{x})$. Numerical methods for $\mathcal{G}^*$ are described in Section 4.

In order to prove that the family Γ is closed under intersection, we need the following lemma.

Lemma 3.3. *Let $\mathcal{F}$, $\mathcal{B}$, $\mathcal{H}$ and $\mathcal{G}$ be finite subsets of $\mathbb{Z}^n_+$ satisfying (10) of Lemma 3.1. Assume that $\mathcal{B}$ contains at least two integer vectors. Then, there exists an $\hat{\boldsymbol{\alpha}} \in \mathcal{B}$ satisfying (11).*

Proof: Let $\hat{\boldsymbol{\alpha}}$ be a vertex of the convex hull of $\mathcal{B}$. Since $\hat{\boldsymbol{\alpha}} \in \mathcal{B} \subset \mathcal{G}$, the first relation $\hat{\boldsymbol{\alpha}} \in \mathcal{G}$ and the second relation $\mathcal{G}\backslash\{\hat{\boldsymbol{\alpha}}\} \neq \emptyset$ in (11) are apparent. The relation $2\hat{\boldsymbol{\alpha}} \notin \mathcal{F}^e$ in (11) follows from the relation $(\mathcal{B} + \mathcal{B}) \cap \mathcal{F} = \emptyset$ in (10). To show the relation $2\hat{\boldsymbol{\alpha}} \notin (\mathcal{G} + \mathcal{G}\backslash\{\hat{\boldsymbol{\alpha}}\})$ in (11), it suffices to show

$$2\hat{\boldsymbol{\alpha}} = \boldsymbol{\alpha} + \boldsymbol{\beta},\ \boldsymbol{\alpha} \in \mathcal{G},\ \boldsymbol{\beta} \in \mathcal{G} \text{ implies } \boldsymbol{\alpha} = \boldsymbol{\beta} = \hat{\boldsymbol{\alpha}}. \tag{13}$$

This is because (13) indicates that there exist no $\boldsymbol{\alpha} \in \mathcal{G}$ and $\boldsymbol{\beta} \in \mathcal{G}\backslash\{\hat{\boldsymbol{\alpha}}\}$ such that $2\hat{\boldsymbol{\alpha}} = \boldsymbol{\alpha} + \boldsymbol{\beta}$. We assume that

$$\hat{\boldsymbol{\alpha}} + \hat{\boldsymbol{\alpha}} = \boldsymbol{\alpha} + \boldsymbol{\beta},\ \boldsymbol{\alpha} \in \mathcal{G} \ \text{ and } \boldsymbol{\beta} \in \mathcal{G}, \tag{14}$$

and derive $\boldsymbol{\alpha} = \boldsymbol{\beta} = \hat{\boldsymbol{\alpha}}$. Since $\mathcal{G} = \mathcal{B} \cup \mathcal{H}$ and $\mathcal{B} \cap \mathcal{H} = \emptyset$, it suffices to consider the following three cases

(a) $\boldsymbol{\alpha} \in \mathcal{B}$ and $\boldsymbol{\beta} \in \mathcal{H}$.

(b) $\boldsymbol{\alpha} \in \mathcal{H}$ and $\boldsymbol{\beta} \in \mathcal{H}$.

(c) $\boldsymbol{\alpha} \in \mathcal{B}$ and $\boldsymbol{\beta} \in \mathcal{B}$.

Cases (a) and (b) can not occur because either of them contradicts the last relation of (10). In case (c), (14) implies that $\boldsymbol{\alpha} = \boldsymbol{\beta} = \hat{\boldsymbol{\alpha}}$ since $\hat{\boldsymbol{\alpha}}$ is a vertex of the convex hull of $\mathcal{B}$. Hence, we have shown (11). ∎

Suppose that both (10) and (11) hold. Let $\mathcal{G}_1 = \mathcal{G}\backslash\{\hat{\boldsymbol{\alpha}}\}$, $\mathcal{B}_1 = \mathcal{B}\backslash\{\hat{\boldsymbol{\alpha}}\}$ and $\mathcal{H}_1 = \mathcal{H}$. Then,

$$\left.\begin{array}{l} \emptyset \neq \mathcal{H}_1 \subset \mathcal{G}_1,\ \emptyset \neq \mathcal{B}_1 \subset \mathcal{G}_1 \text{ and } \mathcal{H}_1 \cup \mathcal{B}_1 = \mathcal{G}_1, \\ (\mathcal{B}_1 + \mathcal{B}_1) \cap \mathcal{F} = \emptyset \ \text{ and } \ (\mathcal{B}_1 + \mathcal{B}_1) \cap (\mathcal{G}_1 + \mathcal{H}_1) = \emptyset. \end{array}\right\}$$

We can apply Lemma 3.3 again to the quadruple $\mathcal{F}$, $\mathcal{B}_1$, $\mathcal{H}_1$ and $\mathcal{G}_1$ whenever $\mathcal{B}_1$ and $\mathcal{H}_1$ are nonempty. Thus Lemma 3.3 not only shows a close relation of Corollary 3.2 with Lemma 3.1, but also ensures that we can replace the condition (ii) by a simpler condition

(ii)' if $\mathcal{G} \in \Gamma$ and there exists an $\hat{\boldsymbol{\alpha}} \in \mathcal{G}$ satisfying (11), then $\mathcal{G}\backslash\{\hat{\boldsymbol{\alpha}}\} \in \Gamma$.

in the definition of the family Γ above. The latter fact plays an essential role in the proof of the theorem below and also in a numerical method given in Section 4 for computing the smallest set $\mathcal{G}^*$ in the family Γ.

Now, we are ready to show the main theorem of this section.

Theorem 3.4. *The family Γ is closed under intersection: if two $\mathcal{G}$ and $\mathcal{G}'$ lie in the family Γ, then so does their intersection $\mathcal{G} \cap \mathcal{G}'$.*

Proof: Suppose that $\mathcal{G},\ \mathcal{G}' \in \Gamma$. We want to show that $\mathcal{G} \cap \mathcal{G}' \in \Gamma$. If $\mathcal{G}' \subset \mathcal{G}$ then $\mathcal{G}' \cap \mathcal{G} = \mathcal{G}' \in \Gamma$. So we assume that $\mathcal{G}' \not\subset \mathcal{G}$. In view of Lemma 3.3 and the discussion above, we can construct $\mathcal{G}$ from $\mathcal{G}^0$ by eliminating an integer vector of $\mathcal{G}^{p-1} \in \Gamma$ to generate $\mathcal{G}^p \in \Gamma$ $(p = 1, 2, \ldots, q)$ successively such that

$$\left.\begin{array}{l} \mathcal{G}^q = \mathcal{G},\ \mathcal{G}^{p-1}\backslash\{\boldsymbol{\alpha}^{p-1}\} = \mathcal{G}^p \ \text{ for some } \boldsymbol{\alpha}^{p-1} \in \mathcal{G}^{p-1}\ (p = 1, 2, \ldots, q), \\ 2\boldsymbol{\alpha}^{p-1} \notin \mathcal{F} \ \text{ and } 2\boldsymbol{\alpha}^{p-1} \notin \left(\mathcal{G}^p + \mathcal{G}^{p-1}\right) \end{array}\right\} \tag{15}$$

Let $\mathcal{H}^p = \mathcal{G}' \cap \mathcal{G}^p$ $(p = 1, 2, \ldots, q)$. Then

$$\mathcal{H}^p = \mathcal{G}' \cap \mathcal{G}^p = \mathcal{G}' \cap \left(\mathcal{G}^{p-1}\backslash\{\boldsymbol{\alpha}^{p-1}\}\right) = \mathcal{H}^{p-1}\backslash\{\boldsymbol{\alpha}^{p-1}\}\ (p = 1, 2, \ldots, q).$$

Hence we see that

$$\mathcal{H}^{p-1} = \left\{ \begin{array}{ll} \mathcal{H}^p & \text{if } \boldsymbol{\alpha}^{p-1} \notin \mathcal{H}^{p-1}, \\ \mathcal{H}^p \cup \{\boldsymbol{\alpha}^{p-1}\} & \text{if } \boldsymbol{\alpha}^{p-1} \in \mathcal{H}^{p-1} \end{array} \right. \tag{16}$$

Since $\mathcal{G} = \mathcal{G}^q$, it suffices to show $\mathcal{H}^p \in \Gamma$ $(p = 0, 1, 2, \ldots, q)$ by induction. When $p = 0$, $\mathcal{H}^0 = \mathcal{G}'$; hence $\mathcal{H}^0 \in \Gamma$. We assume that $\mathcal{H}^j \in \Gamma$ $(j = 0, 1, \ldots, p-1)$ for some $p \leq q$ and then show that $\mathcal{H}^p \in \Gamma$. If $\mathcal{H}^{p-1} = \mathcal{H}^p$ then $\mathcal{H}^p \in \Gamma$ since $\mathcal{H}^{p-1} \in \Gamma$ by induction. Suppose that $\mathcal{H}^{p-1} \neq \mathcal{H}^p$. By (16), we see that $\mathcal{H}^p \cup \{\boldsymbol{\alpha}^{p-1}\} = \mathcal{H}^{p-1}$. On the other hand, we see from the relations on the last line of (15), $\mathcal{H}^{p-1} \subset \mathcal{G}^{p-1}$ and $\mathcal{H}^p \subset \mathcal{G}^p$ that

$$2\boldsymbol{\alpha}^{p-1} \notin \mathcal{F} \ \text{ and } 2\boldsymbol{\alpha}^{p-1} \notin \left(\mathcal{H}^p + \mathcal{H}^{p-1}\right)$$

Therefore, we conclude by the definition of Γ that $\mathcal{H}^p \in \Gamma$. ∎

In the example (5), if we define subsets $\mathcal{G}^1,\ \mathcal{G}^2$ and $\mathcal{G}^3$ of $\mathcal{G}^0$ recursively by

$$\begin{aligned}
\boldsymbol{\alpha}^0 &= \begin{pmatrix} 3 \\ 3 \end{pmatrix},\ \mathcal{G}^1 = \mathcal{G}^0 \backslash \{\boldsymbol{\alpha}^0\}, \\
\boldsymbol{\alpha}^1 &= \begin{pmatrix} 2 \\ 2 \end{pmatrix},\ \mathcal{G}^2 = \mathcal{G}^1 \backslash \{\boldsymbol{\alpha}^1\}, \\
\boldsymbol{\alpha}^2 &= \begin{pmatrix} 1 \\ 1 \end{pmatrix},\ \mathcal{G}^3 = \mathcal{G}^2 \backslash \{\boldsymbol{\alpha}^2\}.
\end{aligned}$$

Then, $\mathcal{G}^p \in \Gamma$ $(p = 0, 1, 2, 3)$ and $\mathcal{G}^* = \mathcal{G}^3$ is the smallest set in Γ.

4 Numerical methods for computing the smallest set $\mathcal{G}^*$ in Γ

Computation of the smallest set $\mathcal{G}^*$ in the class Γ is achieved by solving a combinatorial counting problem. The problem can be divided into two phases. The first phase generates the initial set $\mathcal{G}^0$ given by (3), and the second phase eliminates redundant integer vectors from $\mathcal{G}^0$ to obtain the smallest set $\mathcal{G}^*$ in Γ.

4.1 Phase 1

The aim of the first phase is to compute the initial set $\mathcal{G}^0$. One way for $\mathcal{G}^0$ is as follows:

(i) Generate a very fast rough outer approximation $\mathcal{A}$ of $\mathcal{G}^0$; $\mathcal{G}^0 \subset \mathcal{E} = \{\boldsymbol{\alpha}^1, \boldsymbol{\alpha}^2, \ldots, \boldsymbol{\alpha}^m\} \subset \mathbb{Z}_+^n$.

(ii) Describe the polytope $\frac{1}{2}\mathrm{co}(\mathcal{F}^e)$ in terms of its facet inequalities

$$\boldsymbol{a}_j^T \boldsymbol{\alpha} \leq b_j\ (j = 1, 2, \ldots, \ell).$$

(iii) Check whether each $\boldsymbol{\alpha}$ in the outer approximation $\mathcal{E}$ of $\mathcal{G}^0$ satisfies the system of linear inequalities above; $\boldsymbol{\alpha}$ lies in $\mathcal{G}^0$ if and only if it satisfies the facet inequalities.

SOSTOOLS [11] employs this method with MATLAB's function convhulln to obtain the facet inequality description of the polytope $\frac{1}{2}\mathrm{co}(\mathcal{F}^e)$ as given in (ii). The software cdd+ [4] can also be used for (ii). The method described above, however, may not be efficient for the cases that the dimension n and/or the number of elements in $\mathcal{F}^e$ becomes larger because computing all facet inequalities of the polytope $\frac{1}{2}\mathrm{co}(\mathcal{F}^e)$ requires much work. This can be observed by the numerical results showing that the number of facets is much larger than the number of integer points in a polytope in Section 6. We also should mention that the goal of phase 1 is not the facet inequalities but to obtain $\mathcal{G}^0$.

Barvinok [1] presented a numerical method to find the integer points in a rational polytope described as a set of linear inequalities with integer coefficients or the convex hull of a finite set of integer points. However, the software LattE [9] implementing his numerical

method accepts only a polytope represented by linear inequalities as input. Therefore, another way to compute $\mathcal{G}^0$ is that we first implement the step (ii), which can be done with the software cdd+ by [4]. Then, use the software LattE to enumerate the integer points in $\mathcal{G}^0$.

Instead of computing all facet inequalities that may require a great deal of computational time, we develop a simple enumeration method to generate $\mathcal{G}^0$. Let us define the lower and upper bounds of each coordinate of the integer points in the polytope as follows.

$$\begin{aligned}
\boldsymbol{\ell}^0 &= (\ell_1^0, \ell_2^0, \ldots, \ell_n^0)^T, \quad \ell_k^0 = \min\left\{\alpha_k : \boldsymbol{\alpha} \in \frac{1}{2}\mathrm{co}(\mathcal{F}^e)\right\} \ (k = 1, 2, \ldots, n), \\
\boldsymbol{u}^0 &= (u_1^0, u_2^0, \ldots, u_n^0)^T, \quad u_k^0 = \max\left\{\alpha_k : \boldsymbol{\alpha} \in \frac{1}{2}\mathrm{co}(\mathcal{F}^e)\right\} \ (k = 1, 2, \ldots, n), \\
\mathcal{E}^0 &= \left\{\boldsymbol{\alpha} \in \mathbb{Z}_+^n : \boldsymbol{\ell}^0 \leq \boldsymbol{\alpha} \leq \boldsymbol{u}^0\right\}.
\end{aligned}$$

We observe that $\mathcal{G}^0 = \frac{1}{2}\mathrm{co}(\mathcal{F}^e) \cap \mathbb{Z}_+^n \subset \mathcal{E}^0$. The set $\mathcal{E}^0$ defined above serves as an initial outer approximation of $\mathcal{G}^0$. The size of $\mathcal{E}^0$ can be reduced using the fact that the lower and upper bounds of a higher dimensional coordinate (*e.g.* $k \leq n$) of integer vectors depend on lower dimensional coordinates (*e.g.* 1 to $k-1$) whose values are set to appropriate integers. If the first coordinate of the all integer vectors is fixed as an integer value between the lower and upper bounds ℓ_1^0 and u_1^0, new lower and upper bounds of all the other coordinates can be computed. Then, use a different value between ℓ_1^0 and u_1^0 to fix the value of the first coordinate of the integer vectors and compute the lower and upper bounds of the other coordinates. We continue this process until all the values in ℓ_1^0 and u_1^0 are consumed. Obviously, the size of the resulting set by these lower and upper bounds is a smaller than $\mathcal{E}^0$. This process can be applied to fix the second to $(n-1)$th coordinates with some value and obtain lower and upper bounds of the coordinates that are not fixed. If all the possible integer values in the first to $(n-1)$th coordinates are fixed to compute the lower and upper bounds of the nth coordinate, the integer vectors with the nth coordinate between the lower and upper bounds yield $\mathcal{G}^0$.

4.2 Phase 2

Now we focus our attention on how to generate $\mathcal{G}^*$ from $\mathcal{G}^0$. Let $p = 0$. We construct a digraph (directed graph) having the nodes corresponding to all the integer vectors of $\mathcal{F}^e$ and $\mathcal{G}^p$. We denote the nodes by $\{\nu(\boldsymbol{\gamma}, \mathcal{F}^e) : \boldsymbol{\gamma} \in \mathcal{F}^e\} \cup \{\nu(\boldsymbol{\alpha}, \mathcal{G}^p) : \boldsymbol{\alpha} \in \mathcal{G}^p\}$. Note that an $\boldsymbol{\alpha} \in \mathbb{Z}_+^n$ may lie in both $\mathcal{G}^p$ and $\mathcal{F}^e$, but two nodes $\nu(\boldsymbol{\alpha}, \mathcal{G}^p)$ and $\nu(\boldsymbol{\alpha}, \mathcal{F}^e)$ need to be distinguished. We attach two types of (direct) edges to some pairs of the nodes as follows:

(a) a node $\nu(\boldsymbol{\alpha}, \mathcal{G}^p)$ and a node $\nu(\boldsymbol{\gamma}, \mathcal{F}^e)$ have an edge $(\nu(\hat{\boldsymbol{\alpha}}, \mathcal{G}^p), \nu(\boldsymbol{\gamma}, \mathcal{F}^e))$ if and only if $2\hat{\boldsymbol{\alpha}} = \boldsymbol{\gamma}$.

(b) distinct two nodes $\nu(\hat{\boldsymbol{\alpha}}, \mathcal{G}^p)$ and $\nu(\boldsymbol{\alpha}, \mathcal{G}^p)$ have an edge $(\nu(\hat{\boldsymbol{\alpha}}, \mathcal{G}^p), \nu(\boldsymbol{\alpha}, \mathcal{G}^p))$ if and only if $2\hat{\boldsymbol{\alpha}} = \boldsymbol{\alpha} + \boldsymbol{\beta}$ for some $\boldsymbol{\beta} \in \mathcal{G}^p$.

Then we can recognize each $\hat{\boldsymbol{\alpha}} \in \mathcal{G}^p$ satisfying $2\hat{\boldsymbol{\alpha}} \notin \mathcal{F}^e$ and $2\hat{\boldsymbol{\alpha}} \notin \mathcal{G}^p + (\mathcal{G}^p \backslash \{\hat{\boldsymbol{\alpha}}\})$ as a node $\nu(\hat{\boldsymbol{\alpha}}, \mathcal{G}^p)$ with no outgoing edges.

If there does not exist such a node then $\mathcal{G}^* = \mathcal{G}^p$ is the smallest set in Γ. Otherwise let $\nu(\boldsymbol{\alpha}^p, \mathcal{G}^p)$ be such a node $\nu(\hat{\boldsymbol{\alpha}}, \mathcal{G}^p)$ and $\mathcal{G}^{p+1} = \mathcal{G}^p \backslash \{\hat{\boldsymbol{\alpha}}\}$. Then eliminate the node $\nu(\boldsymbol{\alpha}^p, \mathcal{G}^p)$ and all edges connected to the node $\nu(\boldsymbol{\alpha}^p, \mathcal{G}^p)$. Then the resulting graph has the nodes corresponding to all the integer vectors of $\mathcal{F}^e$ and $\mathcal{G}^{p+1}$ and two types of edges characterized as (a) and (b) above with $p = p + 1$. Therefore, replacing p by $p + 1$, we can continue this procedure until we obtain the digraph that does not have any node with no outgoing edges or the smallest set $\mathcal{G}^* = \mathcal{G}^p$ in Γ.

5 Some properties on $\mathcal{G}^0$ and $\mathcal{G}^*$

We present some basic properties of $\mathcal{G}^0$ and $\mathcal{G}^*$. These properties give an idea of the types of test problems to be generated for numerical experiments. We begin with a monotonicity property. Let $\mathcal{F}_1$ and $\mathcal{F}_2$ be the supports of two different polynomials in $\boldsymbol{x} \in \mathbb{R}^n$ to be represented with sums of squares. Suppose that $\mathcal{F}_1 \subset \mathcal{F}_2$. Then $\mathcal{F}_1^e \subset \mathcal{F}_2^e$, $\mathcal{G}_1^0 \subset \mathcal{G}_2^0$ and $\mathcal{G}_1^* \subset \mathcal{G}_2^*$. The first two inclusion relations follow directly from the definition. The last one follows from the facts

- $\mathcal{G}_1^0 \subset \mathcal{G}_2^0$,
- if $\hat{\boldsymbol{\alpha}} \in \mathcal{G}_1' \subset \mathcal{G}_2'$, $\mathcal{G}_1' \backslash \{\hat{\boldsymbol{\alpha}}\} \neq \emptyset$, and if the triplet $\hat{\boldsymbol{\alpha}}$, $\mathcal{F}^e = \mathcal{F}_2^e$ and $\mathcal{G} = \mathcal{G}_2'$ satisfy (11), then so do the triplet $\hat{\boldsymbol{\alpha}}$, $\mathcal{F}^e = \mathcal{F}_1^e$ and $\mathcal{G} = \mathcal{G}_1'$.

Next we focus on the fact that for given an $\mathcal{F} \subset \mathbb{Z}_+^n$, the family Γ is completely determined by $\mathcal{F}^e \subset 2\mathbb{Z}_+^n$; any integer vector in $\mathcal{F} \backslash \mathcal{F}^e$ is irrelevant for constructing the family Γ. (Recall that the family Γ is characterized by (i) and (ii)' in Section 3.) Hence, if $\mathcal{F}_1^e = \mathcal{F}_2^e$ in the discussion above, we know that $\mathcal{G}_1^* = \mathcal{G}_2^*$. It is sufficient to consider $\mathcal{F}^e \subset 2\mathbb{Z}_+^n$ as an input instead of $\mathcal{F} \subset \mathbb{Z}_+^n$ for the proposed method to calculate $\mathcal{G}^0$ and $\mathcal{G}^*$. On the other hand, if $\mathcal{F} \subset 2\mathbb{Z}_+^n$, then $\mathcal{F}^e = \mathcal{F}$ and the polynomial $\sum_{\boldsymbol{\alpha} \in \mathcal{F}} \boldsymbol{x}^{\boldsymbol{\alpha}}$ forms a sum of squares. This means that any $\mathcal{F} \subset 2\mathbb{Z}_+^n$ can be a candidate for the supports of polynomials to be represented as a sum of squares. Therefore, "a randomly generated finite subset $\mathcal{F}$ of $2\mathbb{Z}_+^n$" gives a fair sample for numerical experiments on computing $\mathcal{G}^0$ and $\mathcal{G}^*$.

Now, we mention a special case where it is not possible to eliminate any integer vector of $\mathcal{G}^0$ for the construction of $\mathcal{G}^*$.

Proposition 5.1. *Let ρ be a positive integer, and let $\mathcal{F} \subset 2\mathbb{Z}_+^n$ be such that*

$$\left\{\mathbf{0}, 2\rho\boldsymbol{e}^1, 2\rho\boldsymbol{e}^2, \ldots, 2\rho\boldsymbol{e}^n\right\} \subset \mathcal{F} \subset co\left(\left\{\mathbf{0}, 2\rho\boldsymbol{e}^1, 2\rho\boldsymbol{e}^2, \ldots, 2\rho\boldsymbol{e}^n\right\}\right).$$

Then, $\mathcal{G}^* = \mathcal{G}^0 = \left\{\boldsymbol{\alpha} \in \mathbb{Z}_+^n : \boldsymbol{e}^T\boldsymbol{\alpha} \leq \rho\right\}$. Here $\boldsymbol{e}^k \in \mathbb{R}^n$ denotes the kth unit coordinate vector with 1 in the kth component and 0 elsewhere.

Proof: First we observe that

$$\begin{aligned}
&\left\{\mathbf{0}, 2\rho\boldsymbol{e}^1, 2\rho\boldsymbol{e}^2, \ldots, 2\rho\boldsymbol{e}^n\right\} \subset \mathcal{F}^e \subset \mathrm{co}\left(\left\{\mathbf{0}, 2\rho\boldsymbol{e}^1, 2\rho\boldsymbol{e}^2, \ldots, 2\rho\boldsymbol{e}^n\right\}\right), \\
&\mathcal{G}^0 = \left(\frac{1}{2}\mathrm{co}\left(\mathcal{F}^e\right)\right) \cap \mathbb{Z}_+^n \\
&\quad = \left(\mathrm{co}\left(\left\{\mathbf{0}, \rho\boldsymbol{e}^1, \rho\boldsymbol{e}^2, \ldots, \rho\boldsymbol{e}^n\right\}\right)\right) \cap \mathbb{Z}_+^n \\
&\quad = \left\{\boldsymbol{\alpha} \in \mathbb{Z}_+^n : \boldsymbol{e}^T\boldsymbol{\alpha} \leq \rho\right\},
\end{aligned}$$

where $\boldsymbol{e}$ denotes the n-dimensional column vector of ones. To prove $\mathcal{G}^* = \mathcal{G}^0$, it suffices to show that any $\hat{\boldsymbol{\alpha}} \in \mathcal{G}^0$ does not satisfy the relations $2\hat{\boldsymbol{\alpha}} \notin \mathcal{F}^e$ and $2\hat{\boldsymbol{\alpha}} \notin \left(\mathcal{G}^0 + \mathcal{G}^0 \backslash \{\hat{\boldsymbol{\alpha}}\}\right)$ simultaneously in (11), or equivalently, that any $\hat{\boldsymbol{\alpha}} \in \mathcal{G}^0$ satisfies $2\hat{\boldsymbol{\alpha}} \in \mathcal{F}^e$ or

$$2\hat{\boldsymbol{\alpha}} \in \left(\mathcal{G}^0 + \mathcal{G}^0 \backslash \{\hat{\boldsymbol{\alpha}}\}\right). \tag{17}$$

Suppose that $\hat{\boldsymbol{\alpha}} \in \mathcal{G}^0$. We obviously see that every vertex $\hat{\boldsymbol{\alpha}}$ of $\mathcal{G}^0$satisfies $2\hat{\boldsymbol{\alpha}} \in \mathcal{F}^e$. So we assume that $\hat{\boldsymbol{\alpha}} \in \mathcal{G}^0$ is not a vertex of $\mathcal{G}^0$. Then $\rho \geq 2$, and we have two cases: $0 < \boldsymbol{e}^T\hat{\boldsymbol{\alpha}} < \rho$ and $\boldsymbol{e}^T\hat{\boldsymbol{\alpha}} = \rho$. In the former case, there is a j such that $0 < \hat{\alpha}_j < \rho$. Hence we see that

$$2\hat{\boldsymbol{\alpha}} = \left(\hat{\boldsymbol{\alpha}} - \boldsymbol{e}^j\right) + \left(\hat{\boldsymbol{\alpha}} + \boldsymbol{e}^j\right),\ \left(\hat{\boldsymbol{\alpha}} - \boldsymbol{e}^j\right) \in \mathcal{G}^0 \backslash \{\hat{\boldsymbol{\alpha}}\} \text{ and } \left(\hat{\boldsymbol{\alpha}} + \boldsymbol{e}^j\right) \in \mathcal{G}^0 \backslash \{\hat{\boldsymbol{\alpha}}\}.$$

Thus (17) holds. In the later case, there are two different indices j and k such that $0 < \hat{\alpha}_j < \rho$ and $0 < \hat{\alpha}_k < \rho$. Hence we see that

$$\begin{aligned} &2\hat{\boldsymbol{\alpha}} = \left(\hat{\boldsymbol{\alpha}} - \boldsymbol{e}^j + \boldsymbol{e}^k\right) + \left(\hat{\boldsymbol{\alpha}} + \boldsymbol{e}^j - \boldsymbol{e}^k\right), \\ &\left(\hat{\boldsymbol{\alpha}} - \boldsymbol{e}^j + \boldsymbol{e}^k\right) \in \mathcal{G}^0 \backslash \{\hat{\boldsymbol{\alpha}}\} \text{ and } \left(\hat{\boldsymbol{\alpha}} + \boldsymbol{e}^j - \boldsymbol{e}^k\right) \in \mathcal{G}^0 \backslash \{\hat{\boldsymbol{\alpha}}\}, \end{aligned}$$

which imply (17). ∎

We obtain the following Corollary from the proposition above and the monotonicity argument at the beginning of this section.

Corollary 5.2. *Let ρ be a positive integer and J be a nonempty subset of $\{1, 2, \ldots, n\}$. Assume that $\mathcal{F} \subset 2\mathbb{Z}^n_+$ satisfies $\{\boldsymbol{0},\ 2\rho\boldsymbol{e}^j\ (j \in J)\} \subset co\,(\mathcal{F})$. Then $co\,(\{\boldsymbol{0},\ \rho\boldsymbol{e}^j\ (j \in J)\}) \cap \mathbb{Z}^n_+ \subset \mathcal{G}^*$.*

Proof: It follows from the assumption that $\{\boldsymbol{0},\ 2\rho_j\boldsymbol{e}^j\ (j \in J)\} \subset \mathcal{F}$ for some positive integer $\rho_j \geq \rho\ (j \in J)$. Hence

$$\mathrm{co}\left(\left\{\boldsymbol{0},\ \rho\boldsymbol{e}^j\ (j \in J)\right\}\right) \cap \mathbb{Z}^n_+ \subset \mathrm{co}\left(\left\{\boldsymbol{0},\ \rho_j\boldsymbol{e}^j\ (j \in J)\right\}\right) \cap \mathbb{Z}^n_+ \subset \mathcal{G}^0.$$

Applying the same argument as in the proof of Proposition 5.1 to every

$$\hat{\boldsymbol{\alpha}} \in \mathrm{co}\left(\left\{\boldsymbol{0},\ \rho\boldsymbol{e}^j\ (j \in J)\right\}\right) \cap \mathbb{Z}^n_+,$$

we obtain $\mathrm{co}\,(\{\boldsymbol{0},\ \rho\boldsymbol{e}^j\ (j \in J)\}) \cap \mathbb{Z}^n_+ \subset \mathcal{G}^*$. ∎

We mention a difference between the smallest support $\mathcal{G}^*$ in Γ and a support of a minimal representation over all possible representation of sums of square of a polynomial. Given a polynomial $f(\boldsymbol{x})$ with a support $\mathcal{F}$, it is true that $\mathcal{G}^0$ includes any support of a polynomial that can be used for some sum of squares representation of $f(\boldsymbol{x})$, and we can eliminate an $\boldsymbol{\alpha} \in \mathcal{G}^0$ only if any sum of squares representation of $f(\boldsymbol{x})$ never uses $\boldsymbol{\alpha}$. In general, $\mathcal{G}^*$ does not necessarily provide a minimal representation over all sums of square representations of $f(\boldsymbol{x})$, but it still covers any support of a polynomial that can appear in some sum of squares representation of $f(\boldsymbol{x})$. To illustrate this, let us consider a simple example

$$f(\boldsymbol{x}) = 2 + 2x_1^4 + 2x_2^4 \text{ for every } \boldsymbol{x} = (x_1, x_2)^T.$$

We may regard $\mathcal{G}^\sharp = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right\}$ as the support for a minimal sum of squares representation of $f(\boldsymbol{x})$. We may expect that $\mathcal{G}^* = \mathcal{G}^\sharp$. But this is not true. In fact, we see by Proposition 5.1 that

$$\mathcal{G}^0 = \mathcal{G}^* = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right\}.$$

We also see that $f(\boldsymbol{x})$ is represented as another sum of squares

$$f(\boldsymbol{x}) = (x_1^2 - 1)^2 + 2x_1^2 + (x_2^2 - 1)^2 + 2x_2^2 + (x_1^2 - x_2^2)^2 + 2(x_1x_2)^2.$$

That is, $f(\boldsymbol{x})$ can be represented as a sum of squares of polynomials whose supports use all integer vectors of $\mathcal{G}^*$.

6 Preliminary numerical experiments

Phases 1 and 2 described in Section 4 were implemented with a MATLAB program. The main focus of the numerical experiments in this section is to observe how the number of integer vectors changes from $\mathcal{G}^0$ to $\mathcal{G}^*$ with some randomly generated test problems. We are also interested in how effectively the procedure by phase 1 and 2 reduces the sizes of $\mathcal{G}^0$ and $\mathcal{G}^*$ as the number of integer vectors in $\mathcal{F}^e \subset 2\mathbb{Z}^n_+$ increases. The numerical results reported here, however, should be regarded as preliminary for the following reasons:

- The current algorithms for phases 1 and 2 are primitive, and the MATLAB program can handle only small size $\mathcal{F}$'s. For the largest case wtih 15 dimension and 61 elements shown in Table 3, more than 5,000 cpu seconds were consumed to compute $\mathcal{G}^0$ and $\mathcal{G}^*$.
- The tested problems consist of a small number of artificially generated examples, and they are not samples from practical problems.

We used a MATLAB function convhulln to compute the number of the facets of $\frac{1}{2}\mathrm{co}(\mathcal{F}^e)$. It should be noted that the number of facets does not affect the implementation of the method because the method does not need any information on the facets. The number of facets is included for illustration and comparison to the method described as (i), (ii) and (iii) in the beginning of Section 4.1.

As mentioned in the previous section, we may regard a randomly generated finite subset $\mathcal{F}$ of $2\mathbb{Z}^n_+$ as a reasonable sample for numerical experiments on computing $\mathcal{G}^0$ and $\mathcal{G}^*$. We fixed the dimension $n = 10$, and generated $\mathcal{H}_{pq} \subset 2\mathbb{Z}^{10}_+$ $(p = 1, 2, 3, 4, q = 1, 2, 3, 4, 5)$ such that $\#\mathcal{H}_{pq} = 10$ and each integer vector of $\mathcal{H}_{pq}$ was chosen randomly from $2\{\boldsymbol{\alpha} \in \mathbb{Z}^{10}_+ : \boldsymbol{\alpha} \leq 4\boldsymbol{e}\}$.

For Corollary 5.2, we generated $\mathcal{H}_{0q} \in 2\mathbb{Z}^{10}_+$ $(q = 1, 2, 3, 4, 5)$ such that each $\mathcal{H}_{0q}$ consists of 10 coordinate vectors $2\rho_j\boldsymbol{e}^j$ with a randomly generated number ρ_j in $\{1, 2, 3, 4\}$ $(j = 1, 2, \ldots, 10)$.

Numerical results are presented for

$$\mathcal{F}_{pq} \;=\; (\{\boldsymbol{0}\}) \cup (\cup_{k=1}^{p} \mathcal{H}_{kq}) \;\; (p = 2, 3, 4, \;\; q = 1, 2, 3, 4, 5)$$

in Table 1, and for

$$\widetilde{\mathcal{F}}_{pq} \;=\; (\{\mathbf{0}\}) \cup (\cup_{k=0}^{p} \mathcal{H}_{kq}) \;\; (p = 0, 2, 3, \;\; q = 1, 2, 3, 4, 5).$$

in Table 2. In Tables 1 and 2, $\#\mathcal{F}^e_{pq}$ ($\#\widetilde{\mathcal{F}}pq^e$, $\#\mathcal{G}^0$, $\#\widetilde{\mathcal{G}}^0$, $\#\mathcal{G}^*$, $\#\widetilde{\mathcal{G}}^*$, respectively) denotes the number of integer vectors in $\mathcal{F}^e_{pq}$ ($\widetilde{\mathcal{F}}^e_{pq}$, $\mathcal{G}^0$, $\widetilde{\mathcal{G}}^0$, $\mathcal{G}^*$, $\widetilde{\mathcal{G}}^*$, respectively), and #facets the number of the facets of the polytope $\frac{1}{2}\mathrm{co}\big(\mathcal{F}^e_{pq}\big)$ in Tables 1 and the polytope $\frac{1}{2}\mathrm{co}\Big(\widetilde{\mathcal{F}}^e_{pq}\Big)$ in Table 2. We measured cpu time in seconds for each of the computations $\mathcal{G}^0_{pq}$ ($q = 1, 2, 3, 4, 5$) ($\mathcal{G}^*_{pq}$ ($q = 1, 2, 3, 4, 5$), $\widetilde{\mathcal{G}}^0_{pq}$ ($q = 1, 2, 3, 4, 5$) or $\widetilde{\mathcal{G}}^*_{pq}$ ($q = 1, 2, 3, 4, 5$)), then show the minimum and maximum cpu time of the measured cpu times in Table 1 and Table 2, denoted by min.cpu and max.cpu, respectively.

By construction, we notice that

$$\begin{aligned}
&\mathbf{0} \in \mathcal{F}_{2q} \subset \mathcal{F}_{3q} \subset \mathcal{F}_{4q} \subset 2\{\boldsymbol{\alpha} \in \mathbb{Z}^n_+ : \boldsymbol{\alpha} \leq 4\boldsymbol{e}\} \;\; (q = 1, 2, 3, 4, 5),\\
&\mathcal{H}_{0q} = \widetilde{\mathcal{F}}_{0q} \subset \widetilde{\mathcal{F}}_{2q} \subset \widetilde{\mathcal{F}}_{3q} \subset 2\{\boldsymbol{\alpha} \in \mathbb{Z}^n_+ : \boldsymbol{\alpha} \leq 4\boldsymbol{e}\} \;\; (q = 1, 2, 3, 4, 5),\\
&\mathcal{F}_{2q} \subset \widetilde{\mathcal{F}}_{2q}, \;\; \mathcal{F}_{3q} \subset \widetilde{\mathcal{F}}_{3q} \;\; (q = 1, 2, 3, 4, 5).
\end{aligned}$$

Table 1 displays the results for three cases (a) $\mathcal{F}_{2q}$, (b) $\mathcal{F}_{3q}$ and (c) $\mathcal{F}_{4q}$, where $\mathcal{F}_{pq}$ consists of $\mathbf{0} \in \mathbb{Z}^{10}_+$, and $10p$ randomly generated integer vectors from $2\{\boldsymbol{\alpha} \in \mathbb{Z}^{10}_+ : \boldsymbol{\alpha} \leq 4\boldsymbol{e}\}$. In all cases, the proposed method is shown to work very effectively. We also observe that a very large number of facets of $\frac{1}{2}\mathrm{co}\big(\mathcal{F}^e_{pq}\big)$ is attained, much larger than the number vertices of $\mathcal{F}_{pq}$, $1 + 10p$ ($p = 2, 3, 4, q = 1, 2, 3, 4, 5$).

When $\mathcal{F}$ involves $\mathbf{0}$ and 10 coordinate vectors $2\mu_j\boldsymbol{e}^j$ with $\mu_j \in \{1, 2, 3, 4\}$ ($j = 1, 2, \ldots, 10$), as stated in Corollary 5.2, the effectiveness of the method for eliminating integer vectors of $\mathcal{G}_0$ is expected to decrease. We confirm this in Table 2. In case (c), we have $\widetilde{\mathcal{G}}^0_{0q} = \widetilde{\mathcal{G}}^*_{0q}$ ($q = 1, 2, 3, 4, 5$). For example, $\widetilde{\mathcal{F}}_{03}$ is given by

$$\widetilde{\mathcal{F}}_{03} = \mathcal{H}_{03} = 2\left\{\mathbf{0}, 4\boldsymbol{e}^1, 3\boldsymbol{e}^2, 4\boldsymbol{e}^3, 3\boldsymbol{e}^4, 3\boldsymbol{e}^5, 3\boldsymbol{e}^6, \boldsymbol{e}^7, \boldsymbol{e}^8, 4\boldsymbol{e}^9, 4\boldsymbol{e}^{10}\right\}.$$

From Corollary 5.2, we know that $\widetilde{\mathcal{G}}^*_{03}$ contains the set $\mathrm{co}\,(\{\mathbf{0}, 3\boldsymbol{e}^j \;\; (j = 1, 2, 3, 4, 5, 6, 9, 10)\}) \cap \mathbb{Z}^{10}_+$, which consists of 165 integer vectors. As more randomly generated integer vectors from $2\{\boldsymbol{\alpha} \in \mathbb{Z}^{10}_+ : \boldsymbol{\alpha} \leq 4\boldsymbol{e}\}$ are added to $\widetilde{\mathcal{F}}_{0q}$, the effectiveness of the proposed method increases as shown in cases (e) and (f) of Table 2. If we compare (b) of Table 1 and (e) of Table 2, we notice that both $\mathcal{F}_{3q}$ and $\widetilde{\mathcal{F}}_{2q}$ consist of 31 integer vectors but the method favors the former case; $\mathcal{G}^*_{3q}$ is much smaller than $\widetilde{\mathcal{G}}^*_{2q}$ ($q = 1, 2, 3, 4, 5$). This difference could be seen because $\widetilde{\mathcal{F}}_{2q}$ contained the coordinate vectors $2\mu_j\boldsymbol{e}^j$ with $\mu_j \in 2\{1, 2, 3, 4\}$ ($j = 1, 2, \ldots, 10$). We also observe a similar difference between (c) of Table 1 and (f) of Table 2. If we compare (a) of Table 1 with (e) of Table 2, and (b) of Table 1 with (f) of Table 2, then we notice that adding coordinate vectors $2\mu_j\boldsymbol{e}^j$ with $\mu_j \in 2\{1, 2, 3, 4\}$ ($j = 1, 2, \ldots, 10$) makes $\mathcal{G}^0$ larger but the method still works effectively to reduce $\mathcal{G}^0$ in cases (e) and (f). In cases (e) and (f) of Table 2, the number of facets of $\frac{1}{2}\mathrm{co}\Big(\widetilde{\mathcal{F}}^e_{pq}\Big)$ becomes very large, when compared with the number integer points of $\widetilde{\mathcal{F}}_{pq}$, $10p + 1$ ($p = 2, 3, q = 1, 2, 3, 4, 5$).

	(a) $\mathcal{F}_{2q}$, $\#\mathcal{F}^e_{2q} = 21$			(b) $\mathcal{F}_{3q}$, $\#\mathcal{F}^e_{3q} = 31$			(c) $\mathcal{F}_{4q}$, $\#\mathcal{F}^e_{4q} = 41$		
q	$\#\mathcal{G}^0_{2q}$	$\#\mathcal{G}^*_{2q}$	#facets	$\#\mathcal{G}^0_{3q}$	$\#\mathcal{G}^*_{3q}$	#facets	$\#\mathcal{G}^0_{4q}$	$\#\mathcal{G}^*_{4q}$	#facets
1	46	21	2916	156	31	20,950	526	42	55,876
2	37	22	2996	167	32	16,833	468	42	52,581
3	41	21	2850	151	31	18,880	544	42	57,954
4	38	22	3180	201	32	18,947	644	43	50,853
5	38	23	2831	135	35	19,741	354	45	59,543
min.cpu	53.5	0.28		159.0	11.9		349.3	218.6	
max.cpu	71.9	0.53		227.7	39.1		625.3	1370.0	

Table 1: $\mathbf{0} \in \mathcal{F}_{pq} \subset 2\{\boldsymbol{\alpha} \in \mathbb{Z}^{10}_+ : \boldsymbol{\alpha}_i \leq 4\boldsymbol{e}\}$

.

	(d) $\widetilde{\mathcal{F}}_{0q}$, $\#\widetilde{\mathcal{F}}^e_{0q} = 11$			(e) $\widetilde{\mathcal{F}}_{2q}$, $\#\widetilde{\mathcal{F}}^e_{2q} = 31$			(f) $\widetilde{\mathcal{F}}_{3q}$, $\#\widetilde{\mathcal{F}}^e_{3q} = 41$		
q	$\#\widetilde{\mathcal{G}}^0_{0q}$	$\#\widetilde{\mathcal{G}}^*_{0q}$	#facets	$\#\widetilde{\mathcal{G}}^0_{2q}$	$\#\widetilde{\mathcal{G}}^*_{2q}$	#facets	$\#\widetilde{\mathcal{G}}^0_{3q}$	$\#\widetilde{\mathcal{G}}^*_{3q}$	#facets
1	63	63	11	182	83	18,257	424	94	59,876
2	96	96	11	254	116	20,000	627	215	50,191
3	202	202	11	437	237	18,212	802	259	53,435
4	158	158	11	372	204	18,310	883	216	56,728
5	73	73	11	164	105	15,240	381	127	51,339
min.cpu	14.3	0.57		134.3	19.1		359.6	211.4	
max.cpu	47.5	12.69		324.8	224.3		1718.0	3234.4	

Table 2: $\mathbf{0},\ \rho_j \boldsymbol{e}^j \in \widetilde{\mathcal{F}}_{pq} \subset 2\{\boldsymbol{\alpha} \in \mathbb{Z}^{10}_+ : \boldsymbol{\alpha} \leq 4\boldsymbol{e}\}\ (j = 1, 2, \ldots, 10)$

.

To observe how $\#\mathcal{G}^0$, $\#\mathcal{G}^*$ and #facets change as n increases to 11, 12, 13, 14 and 15 for case (f) of Table 2, we took $\widehat{\mathcal{F}_n}$ consisting of $1 + 4n$ integer vectors

$$\mathbf{0} \in \mathbb{R}^n,\ 2\mu_i \boldsymbol{e}^j\ (j = 1, 2, \ldots, n),\ 2\boldsymbol{\alpha}_r\ (r = 1, 2, \ldots, 3n)$$

where μ_j was chosen randomly from $\{1, 2, 3, 4\}$ and $\boldsymbol{\alpha}_r$ randomly from $2\{\boldsymbol{\alpha} \in \mathbb{Z}^n_+ : \boldsymbol{\alpha} \leq 4\boldsymbol{e}\}$. The numerical results are shown in Table 3. The proposed method works effectively in this case, and that the number of facets of $\frac{1}{2}\mathrm{co}\,(\mathcal{F}^e)$ increases rapidly as n increases; for $n = 14,\ 15$, it was not possible to obtain the number of facets using the MATLAB function convhulln as a result of out of memory in MATLAB.

7 Concluding remarks

We have shown theoretical properties on supports for sum of squares representations of a polynomial in Section 3 and used the properties to propose a numerical method to compute the smallest support among the class Γ of supports for sums of square representations of a given polynomial. Numerical experiments in Section 6 exhibit the effectiveness of the proposed method.

	(f)' $\widehat{\mathcal{F}}_n$, $\#\widetilde{\mathcal{F}}_n^e = 1+4n$				
n	$\#\widehat{\mathcal{G}}_n^0$	(cpu time)	$\#\widehat{\mathcal{G}}_n^*$	(cpu time)	#facets
11	656	(1128.7)	130	(1656.6)	165,236
12	616	(1260.3)	160	(1358.8)	563,955
13	582	(1489.8)	163	(1174.2)	2,059,342
14	715	(2615.5)	167	(2433.0)	-
15	668	(3321.2)	186	(2022.8)	-

Table 3: $\mathbf{0},\ \rho_j \boldsymbol{e}^j \in \widehat{\mathcal{F}}_n \subset 2\{\boldsymbol{\alpha} \in \mathbb{Z}_+^n : \boldsymbol{\alpha} \leq 4\boldsymbol{e}\}\ (j = 1, 2, \ldots, n)$

.

The test problems used in Section 6 do not represent a wide range of polynomials arising in applications. They are randomly generated polynomials. As mentioned in Section 4, we have used a simple enumeration method for phase 1 in the numerical experiments. Moreover, the MATLAB implementation of phases 1 and 2 is not efficient to solve large scale problems. Developing more efficient codes and extensive numerical tests for various polynomials are subjects of future study.

In this paper, we have assumed that a given polynomial $f(\boldsymbol{x})$ with a support $\mathcal{F} \subset \mathbb{Z}_+^n$ has a sum of square representation and derived $\mathcal{G}^*$ as the smallest support in the class Γ of supports for sums of square representations of $f(\boldsymbol{x})$. It should be noted that the numerical procedure to obtain $\mathcal{G}^*$ can be used to determine whether $f(\boldsymbol{x})$ has a sum of squares representation even when it is not known in advance.

Separability is one of the important issues related to sparse polynomials. A polynomial $f(\boldsymbol{x})$ is called separable if it can be written as $f(\boldsymbol{x}) = \sum_{k=1}^m f_k(\boldsymbol{x}_k)$, where

$$\boldsymbol{x}_k \in \mathbb{R}^{n_k}\ (k = 1, 2, \ldots, m),\ \sum_{k=1}^m n_k = n,\ \boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m).$$

In this case, if $f(\boldsymbol{x})$ is a sum of squares of polynomials in the joint vector variable $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m)$, we can think of a conjecture that it can be represented as a sum of squares of polynomials each of which is a polynomial in exactly one of the vector variables $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m$. We have confirmed that this conjecture is true when $f(\boldsymbol{x})$ attains 0 for some $\boldsymbol{x}$ as its minimum. With the conjecture, the problem of finding the smallest support for a separable polynomial is divided into m subproblems with smaller sizes. Then, the computational method discussed in Section 4 can be applied to each of m subproblems. Therefore, separable polynomials can be handled efficiently in search for the smallest support.

Acknowledgments: The authors would like to thank Prof. A. Barvinok for providing information on numerical methods for enumerating integer vectors in a convex hull, Prof. J.A. De Loera, Prof. K. Fukuda, and Dr. Pablo Parillo for helpful comments on their software packages.

References

[1] A. Barvinok, J. E. Pommersheim, "An algorithmic theory of lattice points in polyhedra," New perspectives in algebraic combinatorics (Berkeley, CA, 1996–97), 91–147, *Math. Sci. Res. Inst. Publ.*, **38,** Cambridge Univ. Press, Cambridge, 1999.

[2] S. Boyd, L. E. Ghaoui, E. Feron and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, (SIAM, Philadelphia, 1994).

[3] 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.

[4] K. Fukuda, "cdd and cddplus homepage", http://www.cs.mcgill.ca/∼fukuda/soft/cdd_home/cdd.html, Computer Science, McGill University, 3480 University, Montreal , Quebec, Canada H3A 2A7.

[5] K. Gatermann and P. A. Parrilo, "Symmetry groups, semidefinite programs and sums of squares", Working paper, Konrad-Zuse-Zentrum fur Informationstechnik, Takustr. 7, D-14195, Berlin, Germany, 2003.

[6] 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.

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

[8] J. B. Lasserre, "An Explicit Equivalent Positive Semidefinite Program for 0-1 Nonlinear Programs", 2002. To appear in *SIAM Journal on Optimization*.

[9] J. A. De Loera, R. Hemmecke, J. Tauzer and R. Yoshida, LattE, http://www.math.ucdavis.edu/ latte, University of California at San Diago.

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

[11] 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.

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

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

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

[15] B. Reznick, "Some concrete aspects of Hilbert's 17th problem", *In Contemporary Mathematics*, **253** (2000) 251-272.

[16] N.Z. Shor, " Class of global minimization bounds of polynomial functions", *Cybernetics*, **23** (1987) 731-734.

[17] J. F. Strum, "SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones", *Optimization Methods and Software*, **11 & 12** (1999) 625-653.

[18] M. J. Todd, K. C. Toh and R. H. Tütüncü, "SDPT3 – a MATLAB software package for semidefinite programming, version 1.3," *Optimization Methods and Software*, **11 & 12** (1999) 545-581.

[19] M. Yamashita, K. Fujisawa and M. Kojima, "Implementation and Evaluation of SDPA 6.0 (SemiDefinite Programming Algorithm 6.0)", September 2002. To appear in *Optimization Methods and Software*.