Exact SDP Relaxations for Classes of Nonlinear Semidefinite Programming Problems*

V. Jeyakumar† and G. Li‡

Revised Version:August 31, 2012

Abstract

An exact semidefinite linear programming (SDP) relaxation of a nonlinear semidefinite programming problem is a highly desirable feature because a semidefinite linear programming problem can efficiently be solved. This paper addresses the basic issue of which nonlinear semidefinite programming problems possess exact SDP relaxations under a constraint qualification. We do this by establishing exact SDP relaxations for classes of nonlinear semidefinite programming problems with SOS-convex polynomials. These classes include SOS-convex semidefinite programming problems and fractional semidefinite programming problems with SOS-convex polynomials. The class of SOS-convex polynomials contains, in particular, convex quadratic functions and separable convex polynomials. Consequently, we also derive numerically checkable conditions, completely characterizing minimizers of these classes of semidefinite programming problems. We finally present a constraint qualification, which is, in a sense, the weakest condition guaranteeing these checkable optimality conditions. The SOS-convexity of polynomials is a sufficient condition for convexity and it can be checked by semidefinite programming whereas deciding convexity is generally NP-hard.

Key words. Polynomial optimization, convex semi-definite programming, SOS-convex polynomials, sum of squares polynomials, fractional programs.

---

*Authors are grateful to the referee for the comments and suggestions which have contributed to the final preparation of the paper. Research was partially supported by a grant from the Australian Research Council.

†Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia. E-mail: v.jeyakumar@unsw.edu.au

‡Department of Mathematics, University of New South Wales, Sydney 2052, Australia. E-mail: g.li@unsw.edu.au

1 Introduction

Polynomial optimization problems are often solved by approximating them by a hierarchy of semidefinite linear programming (SDP) relaxation problems. Recent research has shown that the convergence of such an approximating scheme can even be finite for some classes of problems. Lasserre [17, 19] has indeed shown that a hierarchy of SDP relaxations to convex polynomial optimization problems with convex polynomial inequality constraints has, under some assumptions, finite convergence [19, 16]. In particular, such a SDP relaxation is *exact* for a convex quadratic polynomial programming problem under the Slater constraint qualification (Theorem 5.2, [20]). For other examples, see [21, 16]. Thus, for such a convex programming problem, the optimal values of the original problem and its sum of squares relaxation problem are equal and the relaxation problem, which is representable as a semidefinite linear programming problem, attains its optimum (see Theorem 3.4 Lasserre [19]). It is a highly desirable feature because SDP relaxation problems can efficiently be solved [6, 14, 26].

Unfortunately, however, such a SDP relaxation may not be exact for a convex (polynomial) optimization problem with *semidefinite constraints* (see Example 3.1), even under the Slater condition. Convex polynomial optimization model problems with semidefinite constraints arise in many areas of applications and they frequently appear in robust optimization [5]. In fact, robust counterpart of a convex quadratic programming problem with second-order cone constraint under (unstructured) norm-bounded uncertainty can equivalently be written as a convex quadratic optimization problem with semidefinite constraints (Theorem 6.3.2, [5]). This raises the very basic issue of which *nonlinear semidefinite programming problems* possess exact SDP relaxations under a constraint qualification.

It is known that the class of SOS-convex polynomials [1, 9] enjoys various forms of sum of squares representations which often lead to exact SDP relaxations for S0S-convex programming problems [17, 18, 23]. This provides us with the motivation for examining exact SDP relaxations for classes of nonlinear semidefinite programming problems with SOS-convex polynomials [1, 9].

The class of SOS-convex polynomials is an important subclass of convex polynomials that contains the classes of separable convex polynomials and convex quadratic functions. For other examples of SOS-convex polynomials, see [9, 10]. The significance of SOS-convexity is that one can easily check whether a polynomial is SOS-convex or not by solving a single semi-definite programming problem [1, 9].

We make the following contributions to semidefinite programming and polynomial optimization.

First, we establish exactness, under a closed cone constraint qualification, between a SOS-convex semidefinite optimization problem and its SDP relaxation problem. We do this by way employing Fenchel's duality theorem [7, 22, 25] and invoking an important link between SOS-convex polynomials and sum of squares polynomials [10]. The constraint qualification is guaranteed by the strict feasibility condition. The importance of SOS-convexity for our result is also illustrated by a numerical example.

Second, we show that a class of nonlinear fractional semidefinite programming problems with SOS-convex polynomials enjoys exact SDP relaxations. This class includes fractional quadratic programs with semidefinite constraints. We derive this result by reformulating the fractional programming problem as an equivalent SOS-convex programming problem and by applying the exact SDP relaxation result of SOS-convex programming problems.

Third, we provide conditions in terms of sum of squares polynomials characterizing optimality of semidefinite programming problems with SOS-convex polynomials. These conditions yield numerically checkable characterizations for a feasible point to be a minimizer of these problems. For such a checkable characterization, we provide a weakest condition guaranteeing the characterization condition.

The outline of the paper is as follows. Section 2 provides preliminaries on convex and SOS-convex polynomials. Section 3 presents exact relaxation results for SOS-convex programming problems. Section 4 establishes exact SDP relaxations for fractional semidefinite programming problems. Section 5 provides conditions characterizing optimality of nonlinear semidefinite programming problems.

2 Preliminaries on Convexity and SOS-Convexity

In this Section, we present some definitions and preliminaries that will be used later in the paper. For a set $A \subset \mathbb{R}^n$, the closure of A will be denoted by $\operatorname{cl} A$ and the cone generated by A will be denoted by $\operatorname{cone} A := \{td \ : \ t \geq 0, \ a \in A\}$. The indicator function δ_A is defined as $\delta_A(x) = 0$ if $x \in A$ and $\delta_A(x) = +\infty$ if $x \notin A$. For $u, x \in \mathbb{R}^n$, the inner product is given by $\langle u, x\rangle := u^T x$. The support function of A, σ_A, is defined by $\sigma_A(u) = \sup_{x\in A} u^T x$. The (convex) normal cone of A at a point $x \in \mathbb{R}^n$ is defined as

$$N_A(x) = \begin{cases} \{y \in \mathbb{R}^n : y^T(a-x) \leq 0 \text{ for all } a \in A\} & \text{if } x \in A, \\ \emptyset, & \text{otherwise.} \end{cases}$$

For a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty, +\infty\}$, the conjugate function of f is denoted by $f^* : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty, +\infty\}$ and is defined by $f^*(v) = \sup\{v^T x - f(x) \mid x \in \mathbb{R}^m\}$. The function f is said to be *proper* if f does not take on the value $-\infty$ and $\operatorname{dom} f \neq \emptyset$, where the *domain* of f, $\operatorname{dom} f$, is given by $\operatorname{dom} f = \{x \in \mathbb{R}^n \mid f(x) < +\infty\}$. The *epigraph of* f, $\operatorname{epi} f$, is defined by $\operatorname{epi} f = \{(x, r) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq r\}$.

Let S^m be the space of symmetric $(m\times m)$ matrices with the trace inner product. Note that, for $M, N \in S^m$, the trace inner product, (M, N), is defined by $(M, N) = \text{Tr}\,[MN]$, where $\text{Tr}\,[.]$ refers to the trace operation. The Löwner partial order $\succeq$ of S^m is given by $M, N \in S^m$, $M \succeq N$ if and only if $(M - N)$ is positive semidefinite. Let $S^m_+ := \{M \in S^m \mid M \succeq 0\}$ be the closed convex cone of positive semidefinite $(m \times m)$ matrices. The notation $M \succ 0$ means that M is positive definite. For $F_i \in S^m$, $i = 1, 2, \ldots, n$, the linear operator $\hat{F} : \mathbb{R}^n \to S^m$ is defined by $\hat{F}(x) = \sum_{i=1}^n x_i F_i$. Then the adjoint operator $\hat{F}^* : S^m \to \mathbb{R}^n$ is given by $(\hat{F}^*(Z))_i = (\text{Tr}\,[F_i Z])$. Let

$$D := \left\{ \begin{pmatrix} -\hat{F}^*(Z) \\ Tr[ZF_0] + r \end{pmatrix} \ \mid Z \in S^m_+, \ r \geq 0 \right\}.$$

Then D is a convex cone and it can be expressed as $D = B^*(S^m_+) + \{0\} \times \mathbb{R}_+$, where B^* denotes the adjoint of the linear map $B : \mathbb{R}^n \times \mathbb{R} \to S^m$, given by $B(x, \beta) = -\hat{F}(x) + \beta F_0$. The convex cone D is, in general, not closed. However, if the Slater condition that there exists $x^* \in \mathbb{R}^n$ such that $F_0 + \sum_{i=1}^n F_i x_i^* \succ 0$ holds then D is closed.

On the other hand, for $F^{-1}(S^m_+) := \{x : F_0 + \sum_{i=1}^n x_i F_i \succeq 0\}$, it is easy to verify that $\{-\hat{F}^*(Z) : Z \in S^m_+, \mathrm{Tr}[ZF(x)] = 0\} \subset N_{F^{-1}(S^m_+)}(x)$. The equality

$$N_{F^{-1}(S^m_+)}(x) = \{-\hat{F}^*(Z) : Z \in S^m_+, \mathrm{Tr}[ZF(x)] = 0\} \tag{2.1}$$

holds at each x whenever the convex cone D is closed. The equality (2.1) is known as the *normal cone condition* at x. For details, see [12, 13].

We say that a real polynomial f is sum of squares [8, 15, 16, 23] if there exist real polynomials f_j, $j = 1, \ldots, r$, such that $f = \sum_{j=1}^r f_j^2$. The set consisting of all sum of squares real polynomial is denoted by Σ^2. Similarly, The set consisting of all sum of squares real polynomial with degree at most d is denoted by Σ^2_d. Also, we say a matrix polynomial $F \in \mathbb{R}[x]^{n\times n}$ is a SOS matrix polynomial if $F(x) = H(x)H(x)^T$ where $H(x) \in \mathbb{R}[x]^{n\times r}$ is a matrix polynomial for some $r \in \mathbb{N}$. We now introduce the notion of a SOS-convex polynomial.

Definition 2.1. (SOS-Convexity) *[1, 10] A real polynomial f on $\mathbb{R}^n$ is called SOS-convex if the Hessian matrix function $H : x \mapsto \nabla^2 f(x)$ is a SOS matrix polynomial.*

Clearly, a SOS-convex polynomial is convex. However, the converse is not true. Thus, there exists a convex polynomial which is not SOS-convex [1]. So, the class of SOS-convex polynomials is a proper subclass of convex polynomials. It is known that any convex quadratic function and any convex separable polynomial is a SOS-convex polynomial. Moreover, a SOS-convex polynomial can be non-quadratic and non-separable. For instance, $f(x) = x_1^8 + x_1^2 + x_1x_2 + x_2^2$ is a SOS-convex polynomial (see [9]) which is non-quadratic and non-separable.

We end this section with the following two useful results. The first one provides an interesting link between a SOS-convex polynomial and a sum of squares polynomial while the second one gives us the existence of the optimal solution for convex polynomial optimization problems.

Lemma 2.1. *[10] Let f be a real polynomial on $\mathbb{R}^n$ such that $f(x^*) = 0$ and $\nabla f(x^*) = 0$ for some $x^* \in \mathbb{R}^n$. If f is SOS-convex then f is a sum of squares polynomial.*

Lemma 2.2. *[3] Let $f_0, f_1, \ldots, f_m$ be convex polynomials on $\mathbb{R}^n$. Let $C := \bigcap_{j=1}^m \{x \in \mathbb{R}^n : f_j(x) \le 0\}$. Suppose that* $\inf_{x\in C} f_0(x) > -\infty$*. Then,* $\mathrm{argmin}_{x\in C} f_0(x) \neq \emptyset$.

3 SOS-Convex Semidefinite Programming

Consider the SOS-convex semidefinite optimization model problem

$$(P) \qquad \inf_{x\in\mathbb{R}^n} \; f(x)$$
$$\text{s.t.} \quad F_0 + \sum_{i=1}^n x_i F_i \succeq 0,$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a SOS-convex polynomial and $F_i \in S^m$, for $i = 0, \ldots, n$. The model problem (P) covers convex quadratic programming problems with convex quadratic constraints and the semidefinite separable convex minimization problems.

As we shall see later in this Section, an SOS-convex programming problem [17, 18, 23], where we minimize an SOS-convex function over SOS-convex inequalities, can be solved using (P), whereas an SOS-convex semidefinite programming problem (P) can not, in general, be represented by an SOS-convex programming problem.

For (P), the associated SDP relaxation problem is given by

$$(D) \max_{\mu \in \mathbb{R}, Z \in S^m_+} \{\mu : f - \langle \hat{F}^*(Z), \cdot \rangle - \text{Tr } [ZF_0] - \mu \in \Sigma^2_d\},$$

where $d \in \mathbb{N}$ is the smallest even number such that $d \geq \deg f$ and Σ^2_d is the set of sum of squares polynomials with degree no larger than d. Recall that $\hat{F}^* : S^m \to \mathbb{R}^n$ is given by $(\hat{F}^*(Z))_i = (\text{Tr } [F_i Z])$ and $\langle \hat{F}^*(Z), x \rangle = \sum_{i=1}^n \text{Tr } [ZF_i] x_i$, for each $x \in \mathbb{R}^n$. Although it is known that sum of squares relaxation problems are representable semidefinite linear programming problems [17, 18, 24], for the sake of self containment of the paper, we show in the Appendix how our relaxation problem can be represented by a semidefinite linear programming problem.

We also note that the dual of our relaxation problem (D) can be interpreted as a first moment-relaxation [18, 19, 20, 11] of (P) in a generalized Lasserre hierarchy for polynomial optimization problems with linear matrix inequality constraints.

We say that the optimization problem (P) exhibits an exact SDP relaxation whenever the optimal values of the original problem (P) and its relaxation are equal and the relaxation problem, which is representable as a semidefinite linear programming problem, attains its optimum, that is,

$$\inf_{x \in \mathbb{R}^n} \{f(x) : F_0 + \sum_{i=1}^n x_i F_i \succeq 0\} = \max_{\mu \in \mathbb{R}, Z \in S^m_+} \{\mu : f - \langle \hat{F}^*(Z), \cdot \rangle - \text{Tr } [ZF_0] - \mu \in \Sigma^2_d\}.$$

We first begin with an example showing that this exact SDP relaxation may fail for a convex semidefinite programming problem.

Example 3.1. (Failure of Exact SDP Relaxation for Convex SDP) *Let f be a convex homogeneous polynomial with degree at least 2 in $\mathbb{R}^n$ which is not a sum of squares polynomial (see [4, 16], for the existence of such polynomials). Consider the following convex semidefinite programming problem*

$$\begin{array}{ll} \min_{x \in \mathbb{R}^n} & f(x) \\ s.t. & \begin{pmatrix} I_n & 0 \\ 0 & 1 \end{pmatrix} + \sum_{i=1}^n x_i \begin{pmatrix} 0 & e_i \\ e_i^T & 0 \end{pmatrix} \succeq 0 \end{array}$$

where $F_0 = \begin{pmatrix} I_n & 0 \\ 0 & 1 \end{pmatrix}$, $F_i = \begin{pmatrix} 0 & e_i \\ e_i^T & 0 \end{pmatrix}$, $i = 1, \ldots, n$, I_n is an $n \times n$ identity matrix and $e_i = (0, 0, \ldots, 1, \ldots, 0)^T$ is the vector in $\mathbb{R}^n$ whose ith component is one and all the other components are all zero. It is easy to see that the strict feasibility condition is satisfied.

We now show that our SDP relaxation is not exact. To see this, as f is a convex homogeneous polynomial with degree at least 2 (which is necessarily nonnegative), we first note that $\inf_{x\in\mathbb{R}^n}\{f(x) : F_0 + \sum_{i=1}^n x_iF_i \succeq 0\} = 0$. *The claim will follow if we show that for any* $Z = \begin{pmatrix} A & a \\ a^T & r \end{pmatrix} \succeq 0$, $f - \langle \hat{F}^*(Z), \cdot\rangle - \mathrm{Tr}\ [ZF_0] - 0 \notin \Sigma_d^2$. *Otherwise, there exists a sum of squares polynomial* σ *with degree at most* d *such that*

$$f(x) - (2a^Tx + \mathrm{Tr}\ [A] + r) = \sigma(x), \quad \text{for all } x \in \mathbb{R}^n. \tag{3.2}$$

Note that $f(0) = 0$ *and* $\mathrm{Tr}\ [A] + r \geq 0$ *as* $Z \succeq 0$. *Letting* $x = 0$ *in (3.2), we see that*

$$0 \geq -(\mathrm{Tr}\ [A] + r) = f(0) - (\mathrm{Tr}\ [A] + r) = \sigma(0) \geq 0.$$

This forces that $\mathrm{Tr}\ [A] + r = 0$, *and hence* $f(x) - 2a^Tx = \sigma(x)$, *for all* $x \in \mathbb{R}^n$. *In particular, we have* $f(x) \geq 2a^Tx$. *As* f *is a convex homogeneous polynomial with degree* m *and* $m \geq 2$, *this implies that* $a = 0$. *(Indeed, if* $a \neq 0$, *then there exists* x_0 *such that* $a^Tx_0 > 0$. *Then, we have* $t^m f(x_0) = f(tx_0) \geq 2a^T(tx_0) = 2ta^Tx_0$, *for all* $t > 0$. *So,* $\frac{f(x_0)}{2a^Tx_0} \geq \frac{1}{t^{m-1}}$, *for all* $t > 0$. *This is a contradiction as* $\frac{1}{t^{m-1}} \to \infty$ *as* $t \to 0$.*) Hence,* $f = \sigma$ *and so* f *is a sum of squares polynomial. This contradicts our construction of* f. *Therefore, the relaxation is not exact.*

In the following, we show, under a closed cone constraint qualification, that the SDP relaxation of a convex semidefinite programming problem is exact whenever the objective function is a SOS-convex polynomial.

Theorem 3.1. (Exact SDP Relaxation) *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be a SOS-convex real polynomial. Let* $K := \{x \in \mathbb{R}^n : F_0 + \sum_{i=1}^n x_iF_i \succeq 0\} \neq \emptyset$, *where* $F_i \in S^m$, *for* $i = 0, 1, \ldots, n$. *Suppose that the convex cone* D *is closed. Then,*

$$\inf_{x\in\mathbb{R}^n}\{f(x) : F_0 + \sum_{i=1}^n x_iF_i \succeq 0\} = \max_{\mu\in\mathbb{R}, Z\in S_+^m}\{\mu : f - \langle \hat{F}^*(Z), \cdot\rangle - \mathrm{Tr}\ [ZF_0] - \mu \in \Sigma_d^2\}.$$

Proof. Clearly, the inequality,

$$\inf_{x\in\mathbb{R}^n}\{f(x) : F_0 + \sum_{i=1}^n x_iF_i \succeq 0\} \geq \max_{\mu\in\mathbb{R}, Z\in S_+^m}\{\mu : f - \langle \hat{F}^*(Z), \cdot\rangle - \mathrm{Tr}\ [ZF_0] - \mu \in \Sigma_d^2\}.$$

always holds by construction. The conclusion will follow if we establish the reverse inequality. To see this, without loss of generality, we may assume that $\inf_{x\in K} f(x) > -\infty$, otherwise the required equality trivially follows. As $K \neq \emptyset$, we see that $\beta := \inf_{x\in K} f(x) \in \mathbb{R}$. Now, by the Fenchel duality theorem (see [7, 22, 25])

$$\beta = \inf_{x\in K} f(x) = \inf_{x\in\mathbb{R}^n}\{f(x) + \delta_K(x)\} = \max_{x^*\in\mathbb{R}^n}\{-f^*(x^*) - \delta_K^*(-x^*)\}.$$

Then we can find $x^* \in \mathbb{R}^n$ such that $-f^*(x^*) - \delta_K^*(-x^*) = \beta$. So, $(-x^*, -f^*(x^*) - \beta) \in \mathrm{epi}\delta_K^*$.

We claim that epi$\delta_K^* = \text{cl}(D)$. Granting this, we obtain that epi$\delta_K^* = D$, because of the assumption that D is closed. So, we can find $Z \in S_+^m$ and $r \geq 0$ such that $x_i^* = \text{Tr }[ZF_i]$ and $-f^*(x^*) - \beta = \text{Tr }[ZF_0] + r$. Then, for each $x \in \mathbb{R}^n$, $\sum_{i=1}^n \text{Tr }[ZF_i]x_i - f(x) = \sum_{i=1}^n x_i^* x_i - f(x) = x^{*T}x - f(x) \leq f^*(x^*)$ and $f^*(x^*) = -\beta - \text{Tr }[ZF_0] - r$. Then, $\sum_{i=1}^n \text{Tr }[ZF_i]x_i - f(x) \leq -\beta - \text{Tr }[ZF_0] - r$. Thus, for each $x \in \mathbb{R}^n$,

$$f(x) - \text{Tr }[Z(F_0 + \sum_{i=1}^n F_i x_i)] - \beta \geq r \geq 0.$$

Let $\sigma_0(x) := f(x) - (\text{Tr }[ZF_0] + \sum_{i=1}^m \text{Tr }[ZF_i]x_i) - \beta$, for each $x \in \mathbb{R}^n$. Then, σ_0 is a SOS-convex polynomial, because f is a SOS-convex polynomial, any affine function is a SOS-convex polynomial and the sum of two SOS-convex polynomial is still SOS-convex. So, $\min_{x\in\mathbb{R}^n} \sigma_0(x)$ is attained by Lemma 2.2 since σ_0 is a nonnegative convex polynomial. Let $\sigma_0(x^*) = \min_{x\in\mathbb{R}^n} \sigma_0(x) \geq 0$. Define $\widehat{\sigma_0}(x) := \sigma_0(x) - \sigma_0(x^*)$. Then, $\widehat{\sigma_0}$ is SOS-convex polynomial, $\widehat{\sigma_0}(x^*) = 0$ and $\nabla\widehat{\sigma_0}(x^*) = 0$. So, by Lemma 2.1, $\widehat{\sigma_0}$ is a sum of squares polynomial. This shows that $\sigma_0(x) = \widehat{\sigma_0}(x) + \sigma_0(x^*)$ is a sum of squares polynomial as $\sigma_0(x^*)$ is nonnegative. That is, $\sigma_0(x) = f(x) - \beta - \langle \hat{F}^*(Z), x\rangle - \text{Tr }[ZF_0]$ is a sum of squares polynomial and σ_0 has degree at most deg$f$. On the other hand, any sum of squares polynomial must have even degree. So, deg$\sigma_0 \leq d$, $f - \langle \hat{F}^*(Z), \cdot\rangle - \text{Tr }[ZF_0] - \beta \in \Sigma_d^2$ and $(\beta, Z) \in \mathbb{R} \times S_+^n$. Hence,

$$\inf_{x\in\mathbb{R}^n}\{f(x) : F_0 + \sum_{i=1}^n x_i F_i \succeq 0\} = \beta \leq \max_{\mu\in\mathbb{R}, Z\in S_+^m}\{\mu : f - \langle \hat{F}^*(Z), \cdot\rangle - \text{Tr }[ZF_0] - \mu \in \Sigma_d^2\}.$$

and the conclusion follows.

Finally, we establish the claim that epi$\delta_K^* = \text{cl}(D)$. It is easy to check that $D \subseteq$ epiδ_K^*. So, $\text{cl}(D) \subseteq$ epiδ_K^*. To see the reverse inclusion, we assume that there exists $(u, \alpha) \in$ epi$\delta_K^* \backslash \text{cl}(D)$. Then, the separation theorem gives us that there exists $(v, s) \neq (0, 0)$ such that

$$v^T(-\hat{F}^*(Z)) + s(Tr[ZF_0] + r) \leq 0 < v^T u + s\alpha, \quad \text{for all } Z \in S_+^m,\ r \geq 0. \tag{3.3}$$

Letting $Z = 0$ in (3.3), we see that $sr \leq 0$, for every $r \geq 0$ and in particular, $s \leq 0$, by taking $r = 1$. If $s = 0$, then $v^T(-\hat{F}^*(Z)) \leq 0 < v^T u$, for all $Z \in S_+^m$. This implies that $\text{Tr}[(\sum_{i=1}^n v_i F_i)Z] \geq 0$, for all $Z \in S_+^m$. So, $\sum_{i=1}^n v_i F_i \succeq 0$. Fix $a \in K$. Then $a + tv \in K$ for all $t > 0$. Thus, $(a + tv)^T u \leq \sup_{x\in K} u^T x = \delta_K^*(u) \leq \alpha$ for all $t > 0$. Letting $t \to \infty$, we see that $v^T u \leq 0$. This is a contradiction as $v^T u > 0$. So, $s < 0$. Now, let $w = v/(-s)$. Then, we obtain from (3.3) that

$$w^T(-\hat{F}^*(Z)) - (\text{Tr}[ZF_0] + r) \leq 0 < w^T u - \alpha, \quad \text{for all } Z \in S_+^m,\ r \geq 0.$$

So, in particular, $\text{Tr}[(-F_0 - \sum_{i=1}^n w_i F_i)Z] = w^T(-\hat{F}^*(Z)) - \text{Tr}[ZF_0] \leq 0$, for all $Z \in S_+^m$ which implies that $F_0 + \sum_{i=1}^n w_i F_i \succeq 0$. Hence, $w = (w_1, \ldots, w_n)^T \in K$ and so, $w^T u \leq \sup_{x\in K} u^T x = \delta_K^*(u) \leq \alpha$. This is impossible as $w^T u - \alpha > 0$. Therefore, epi$\delta_K^* = \text{cl}(D)$. This completes the proof. □

Remark 3.1. (Importance of SOS-convexity in Theorem 3.1) *It is worth noting that SOS-convexity can not be dropped in Theorem 3.1. To see this observe that any convex homogeneous polynomial with degree at least 2 which is SOS-convex must be a sum of squares polynomial (cf. [2, Lemma 3.2] or [10, Lemma 8]). In Example 3.1, the objective function is not a SOS-convex polynomial and we have shown that the conclusion of Theorem 3.1 fails.*

Now, we give a simple numerical example of a SOS-convex semi-definite programming problem, verifying our exact relaxation result. We illustrate that the sum of squares relaxation problem is in fact represented by a SDP problem. We also observe that the problem cannot be expressed as a standard SOS-convex programming problem with SOS-convex inequality constraints.

Example 3.2. (Verifying Exactness of Relaxation) *Consider the following SOS-convex semidefinite programming problem:*

$$(EP_1) \quad \inf_{x\in\mathbb{R}^2} \quad x_1^4$$
$$s.t. \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + x_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \succeq 0.$$

It can be easily verified that, in this case, the optimal value of (EP_1) is zero but the infimum of (EP_1) is not attained.

We now verify the equality between the optimal values between (EP_1) and its sum of squares relaxation, which can be written as

$$\max_{\mu\in\mathbb{R}, Z\in S^2_+} \{\mu : f - \langle \hat{F}^*(Z), \cdot\rangle - \mathrm{Tr}\,[ZF_0] - \mu \in \Sigma^2_d\}$$
$$= \max_{\mu\in\mathbb{R}, z_1,z_2,z_3\in\mathbb{R}} \{\mu : x_1^4 - 2z_2 + z_1x_1 + z_3x_2 - \mu \in \Sigma^2_4, \begin{pmatrix} z_1 & z_2 \\ z_2 & z_3 \end{pmatrix} \succeq 0\}.$$

Let (μ, z_1, z_2, z_3) be any feasible point for the relaxation problem. Note that $x_1^4 - 2z_2 + z_1x_1 + z_3x_2 - \mu \in \Sigma^2_4$ implies that $z_3 = 0$. This together with the fact that $\begin{pmatrix} z_1 & z_2 \\ z_2 & z_3 \end{pmatrix} \succeq 0$ implies that $z_3 = 0, z_2 = 0$ and $z_1 \geq 0$. So, the relaxation problem can be simplified as

$$\max_{\mu\in\mathbb{R}, Z\in S^2_+} \{\mu : f - \langle \hat{F}^*(Z), \cdot\rangle - \mathrm{Tr}\,[ZF_0] - \mu \in \Sigma^2_d\}$$
$$= \max_{\mu\in\mathbb{R}, z_1\in\mathbb{R}} \{\mu : x_1^4 + z_1x_1 - \mu \in \Sigma^2_4, z_1 \geq 0\}.$$

Clearly, the representation $x_1^4 + z_1x_1 - \mu \in \Sigma^2_4$ is equivalent to the condition that there exists

$$W = \begin{pmatrix} W_{11} & W_{12} & W_{13} \\ W_{12} & W_{22} & W_{23} \\ W_{13} & W_{23} & W_{33} \end{pmatrix} \succeq 0$$

such that

$$x_1^4 + z_1 x_1 - \mu = \begin{pmatrix} 1 \\ x_1 \\ x_1^2 \end{pmatrix}^T \begin{pmatrix} W_{11} & W_{12} & W_{13} \\ W_{12} & W_{22} & W_{23} \\ W_{13} & W_{23} & W_{33} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ x_1^2 \end{pmatrix},$$

which is further equivalent to $W_{33} = 1$*,* $W_{23} = 0$*,* $2W_{13} + W_{22} = 0$*,* $2W_{12} = z_1$ *and* $W_{11} = -\mu$*. As* $W \succeq 0$*, we have* $W_{13} = W_{22} = 0$ *and so,* $z_1 = W_{12} = 0$*. Thus, the problem can be equivalently represented by the simple linear semi-definite programming problem:*

$$\max_{\mu \in \mathbb{R}} \{\mu : \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \mu \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \succeq 0\}.$$

Thus, the optimal value of the relaxation problem is zero. So, the equality between the optimal values of (EP_1) *and its sum of squares relaxation is verified.*

Note that the feasible set

$$K := \{x \in \mathbb{R}^2 : \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + x_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \succeq 0\}$$

of the problem (EP_1) *cannot be written as* $\{x : g_i(x) \le 0, i = 1, \dots, l\}$ *where* g_i*,* $i = 1, \dots, l$ *are convex polynomials. To see this, we proceed by the method of contradiction and suppose that* $K = \{x : g_i(x) \le 0, i = 1, \dots, l\}$ *for some convex polynomials* g_i*,* $i = 1, \dots, l$*. Then Lemma 2.2 shows that* $f(x_1, x_2) = x_1^4$ *must attains its minimum over the set* K *which is impossible.*

Remark 3.2. (SOS-convex Minimization over Semidefinite Representable sets) *Consider minimizing a SOS-convex polynomial* f *over a convex feasible set* K*. Suppose that the feasible set* K *is semidefinite representable, that is, there exist* $l \in \mathbb{N}$ *and* $G_j \in S^m$*,* $j = 1, \dots, l$*, such that*

$$K = \{x \in \mathbb{R}^n : \exists y = (y_1, \dots, y_l) \in \mathbb{R}^l \text{ such that } F_0 + \sum_{i=1}^n x_i F_i + \sum_{j=1}^l y_j G_j \succeq 0\}.$$

Then, the minimization problem $\min_{x \in K} f(x)$ *is equivalent to the following SOS-convex optimization problem with linear matrix inequality constraint*

$$(P') \qquad \inf_{x \in \mathbb{R}^n, y \in \mathbb{R}^l} \quad f(x)$$
$$\text{s.t.} \qquad F_0 + \sum_{i=1}^n x_i F_i + \sum_{j=1}^l y_j G_j \succeq 0,$$

So, the exact relaxation theorem (i.e. Theorem 3.1) provides us with a way of finding the optimal value of (P') *by solving the corresponding sum of squares relaxation problem.*

In particular, an SOS-convex programming problem, where the feasible set, $K = \{x : g_k(x) \le 0, k = 1, \dots, p\}$*, and each* g_k *is a SOS-convex polynomial, can also be solved in this way by solving its sum of squares relaxation problem because* K *is semi-definite representable [10].*

In passing, we note that the exact SDP relaxation might not be the "best" way to solve the original problem (P). However, it shows that the problem (P) can be solved by solving a standard semidefinite program. For other approaches for solving the nonlinear semidefinite linear programming problems such as (P), see [27] and other references therein.

4 Fractional Semidefinite Programming

In this Section, we show that, our SDP relaxation is exact for a class of nonlinear semidefinite fractional programming model problems.

Theorem 4.1. (SDP Relaxations for Fractional Programs) *Let $f : \mathbb{R}^n \to \mathbb{R}$ be SOS-convex polynomial; let $c \in \mathbb{R}^n$, $d \in \mathbb{R}$ and $F_i \in S^m$, for $i = 1, 2, \ldots, n$. Suppose that $c^T x + d > 0$ for all x with $F_0 + \sum_{i=1}^n x_i F_i \succeq 0$. If the convex cone D is closed then*

$$\inf\{\frac{f(x)}{c^T x + d} : F_0 + \sum_{i=1}^n x_i F_i \succeq 0\} = \max_{r\in\mathbb{R}, Z\in S^m_+} \{r : f - \langle \hat{F}^*(Z) + rc, \cdot\rangle - \text{Tr }[ZF_0] - rd \in \Sigma^2_p\}.$$

where p is the smallest even number such that $p \geq \deg f$.

Proof. For any x with $F_0 + \sum_{i=1}^n x_i F_i \succeq 0$ and for any $r \in \mathbb{R}, Z \in S^m_+$ with $f - \langle \hat{F}^*(Z) + rc, \cdot\rangle - \text{Tr }[ZF_0] - rd \in \Sigma^2_p$,

$$\frac{f(x)}{c^T x + d} \geq \frac{r(c^T x + d) + \text{Tr }[Z(F_0 + \sum_{i=1}^n x_i F_i)]}{c^T x + d} \geq r,$$

because, for each $x \in \mathbb{R}^n$, $f(x) - \sum_{i=1}^n \text{Tr }[ZF_i] x_i - rc^T x - \text{Tr }[ZF_0] - rd = \sigma_0(x) \geq 0$, for some $\sigma_0 \in \sum_p^2$. So,

$$\inf\{\frac{f(x)}{c^T x + d} : F_0 + \sum_{i=1}^n x_i F_i \succeq 0\} \geq \max_{r\in\mathbb{R}, Z\in S^m_+} \{r : f - \langle \hat{F}^*(Z) + rc, \cdot\rangle - \text{Tr }[ZF_0] - rd \in \Sigma^2_p\}.$$

To see the converse inequality, let $r_0 = \inf\{\frac{f(x)}{c^T x+d} : F_0 + \sum_{i=1}^n x_i F_i \succeq 0\}$. Without loss of generality, we may assume that $r_0 > -\infty$. As $\{x : F_0 + \sum_{i=1}^n x_i F_i \succeq 0\} \neq \emptyset$, we see that $r_0 \in \mathbb{R}$. Then, we have

$$F_0 + \sum_{i=1}^n x_i F_i \succeq 0 \;\Rightarrow\; \frac{f(x)}{c^T x + d} \geq r_0.$$

This is equivalent to

$$F_0 + \sum_{i=1}^n x_i F_i \succeq 0 \;\Rightarrow\; f(x) - r_0(c^T x + d) \geq 0.$$

Thus, $\inf_{x\in K}\{f(x) - r_0(c^Tx+d)\} \geq 0$. Now, it follows from Theorem 3.1 that there exist $\beta \in \mathbb{R}$ and $Z \in S^m_+$ such that $0 \leq \inf_{x\in K}\{f(x) - r_0(c^Tx+d)\} = \beta$ and, for each $x \in \mathbb{R}^n$,

$$f(x) - r_0(c^Tx+d) - \mathrm{Tr}[Z(F_0 + \sum_{i=1}^n x_iF_i)] - \beta = \sigma_0(x) \text{ for all } x \in \mathbb{R}^n,$$

for some $\sigma_0 \in \sum^2_p$. As $\beta \geq 0$,

$$f(x) - r_0(c^Tx+d) - \mathrm{Tr}[Z(F_0 + \sum_{i=1}^n x_iF_i)] = \beta + \sigma_0(x) = \sigma_1(x),$$

where $\sigma_1 \in \sum^2_p$. So, we have

$$\inf\{\frac{f(x)}{c^Tx+d} : F_0 + \sum_{i=1}^n x_iF_i \succeq 0\} \leq \max_{r\in\mathbb{R}, Z\in S^m_+}\{r : f - \langle \hat{F}^*(Z) + rc, \cdot\rangle - \mathrm{Tr}\ [ZF_0] - rd \in \Sigma^2_p\}.$$

Therefore, the conclusion follows. □

As a corollary, we derive an explicit exact SDP relaxation for a fractional quadratic programming problem.

Corollary 4.1. (SDP Relaxation for Fractional Quadratic Programs) *Let $A \in S^n_+$, $b, c \in \mathbb{R}^n$, $d, e \in \mathbb{R}$ and let $F_i \in S^m$, for $i = 1, 2, \ldots, n$. Suppose that the convex cone D is closed and $c^Tx + d > 0$ for all x with $F_0 + \sum_{i=1}^n x_iF_i \succeq 0$. Then,*

$$\begin{aligned}
&\inf\{\frac{\frac{1}{2}x^TAx + b^Tx + e}{c^Tx+d} : F_0 + \sum_{i=1}^n x_iF_i \succeq 0\} \\
= &\max_{r,\gamma\in\mathbb{R}, u\in\mathbb{R}^n}\{r : \begin{pmatrix} Z & 0 & 0 \\ 0 & A & b - rc - u \\ 0 & (b-rc-u)^T & 2(e - rd - \gamma)\end{pmatrix} \succeq 0, u = \hat{F}^*(Z), \gamma = \mathrm{Tr}[ZF_0]\}.
\end{aligned}$$

Proof. We apply the preceding theorem with $d = 2$. So,

$$\begin{aligned}
&\inf\{\frac{\frac{1}{2}x^TAx + b^Tx + e}{c^Tx+d} : F_0 + \sum_{i=1}^n x_iF_i \succeq 0\} \\
= &\max_{r\in\mathbb{R}, Z\in S^m_+}\{r : \frac{1}{2}x^TAx + b^Tx + e - r(c^Tx+d) - \mathrm{Tr}\ [ZF(x)] \in \Sigma^2_2\} \\
= &\max_{r\in\mathbb{R}, Z\in S^m_+}\{r : \begin{pmatrix} A & b - rc - \hat{F}^*(Z) \\ (b - rc - \hat{F}^*(Z))^T & 2(e - rd - \mathrm{Tr}[ZF_0])\end{pmatrix} \succeq 0\} \\
= &\max_{r,\gamma\in\mathbb{R}, u\in\mathbb{R}^n}\{r : \begin{pmatrix} Z & 0 & 0 \\ 0 & A & b - rc - u \\ 0 & (b-rc-u)^T & 2(e - rd - \gamma)\end{pmatrix} \succeq 0, u = \hat{F}^*(Z), \gamma = \mathrm{Tr}[ZF_0]\}
\end{aligned}$$

□

5 Sum of Squares Characterizations of Optimality

In this section, we provide conditions in terms of sum of squares polynomials characterizing optimality of semidefinite programming problems with SOS-convex polynomials. These conditions yield numerically checkable characterizations for a feasible point to be a minimizer of these problems. We consider again the model problem.

$$(P) \qquad \inf_{x\in\mathbb{R}^n} \; f(x)$$
$$\text{s.t.} \quad F_0 + \sum_{i=1}^{n} x_i F_i \succeq 0,$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a SOS-convex function and $F_i \in S^m$, for $i = 0, \ldots, n$. Recall that $F : \mathbb{R}^n \to S^m$ is defined by $F(x) := F_0 + \sum_{i=1}^{n} x_i F_i$ for each $x \in \mathbb{R}^m$ and, for each $Z \in S^m$, $\mathrm{Tr}[ZF(x)] := \mathrm{Tr}[Z(F_0 + \sum_{i=1}^{n} F_i x_i)]$.

Theorem 5.1. (Strong Saddle-Point Conditions) *For the problem (P), let f be SOS-convex. Let the Lagrangian function $L : \mathbb{R}^n \times S^m \to \mathbb{R}$ of (P) be defined by $L(x, Z) = f(x) - \mathrm{Tr}[ZF(x)]$. Suppose that the convex cone D is closed. Then, the following statements are equivalent.*

(i) *The point x^* is a minimizer of (P).*

(ii) $(\exists Z^* \in S^m_+)$ $L(\cdot, Z^*) - L(x^*, Z^*) \in \Sigma^2$ *and* $L(x^*, Z) \leq L(x^*, Z^*)$, $\forall Z \in S^m_+$.

Proof. [(ii) ⇒ (i)] If (ii) holds, then, for all $Z \in S^m_+$

$$\mathrm{Tr}[ZF(x^*)] - \mathrm{Tr}[Z^* F(x^*)] = L(x^*, Z^*) - L(x^*, Z) \geq 0.$$

and so,

$$\mathrm{Tr}[Z^* F(x^*)] \leq 0 \text{ and } F(x^*) \in S^m_+.$$

Hence, x^* is feasible for (P). Moreover, for each $x \in \mathbb{R}^n$

$$\begin{aligned} f(x) - f(x^*) - \mathrm{Tr}[Z^* F(x)] &\geq f(x) - \mathrm{Tr}[Z^* F(x)] - f(x^*) + \mathrm{Tr}[Z^* F(x^*)] \\ &= L(x, Z^*) - L(x^*, Z^*) = \sigma_0(x) \geq 0. \end{aligned}$$

So, for each feasible point x of (P), $f(x) \geq f(x^*)$, and hence, x^* is a global minimizer of (P).

[(i) ⇒ (ii)] If x^* is a minimizer of (P), then by Theorem 3.1,

$$f(x^*) = \inf(P) = \max_{\mu\in\mathbb{R}, Z\in S^m_+} \{\mu : f - \langle \hat{F}^*(Z), \cdot\rangle - \mathrm{Tr}\ [ZF_0] - \mu \in \Sigma^2_d\}.$$

So, there exist $\mu^* \in \mathbb{R}$ and $Z^* \in S^m_+$ such that $f(x^*) = \mu^*$ and, for each $x \in \mathbb{R}^n$, $f(x) - \langle \hat{F}^*(Z^*), x\rangle - \mathrm{Tr}\ [Z^* F_0] - \mu^* = \sigma_0(x)$ for some $\sigma_0 \in \Sigma^2_d$. This give us that, for each $x \in \mathbb{R}^n$, $f(x) - f(x^*) - \mathrm{Tr}\ [Z^* F(x)] = \sigma_0(x)$. Letting $x = x^*$, we have $-\mathrm{Tr}\ [Z^* F(x^*)] = \sigma_0(x^*) \geq 0$.

As x^* is a minimizer (and so, is feasible), Tr $[Z^*F(x^*)] \geq 0$. Then, Tr $[Z^*F(x^*)] = 0$. So, for all $Z \in S_+^m$,

$$L(x^*, Z^*) - L(x^*, Z) = \text{Tr}[ZF(x^*)] - \text{Tr}[Z^*F(x^*)] = \text{Tr}[ZF(x^*)] \geq 0.$$

Moreover, for each $x \in \mathbb{R}^n$,

$$L(x, Z^*) - L(x^*, Z^*) = f(x) - \text{Tr}[Z^*F(x)] - f(x^*) + \text{Tr}[Z^*F(x^*)] = f(x) - f(x^*) - \text{Tr}[Z^*F(x)] = \sigma_0(x).$$

So, (ii) holds. □

The following Corollary provides an easily numerically checkable optimality condition characterizing a minimizer.

Corollary 5.1. (Numerically Checkable Characterization of Optimality) *For the problem (P), let x^* be a feasible point. Assume that the cone D is closed. Then, the following statements are equivalent*

(i) *The point x^* is a minimizer of (P).*

(ii) $(\exists Z \in S_+^m)$ $(\exists \sigma_0 \in \Sigma^2)$ $(\forall x \in \mathbb{R}^n)$ $f(x) - f(x^*) - \text{Tr}\ [ZF(x)] = \sigma_0(x)$.

Proof. [(ii) ⇒ (i)] If there exist $Z \in S_+^m$ and $\sigma_0 \in \Sigma^2$ be such that

$$f(x) - f(x^*) - \text{Tr}\ [ZF(x)] = \sigma_0(x).$$

Then, for each feasible point x,

$$f(x) - f(x^*) = \text{Tr}\ [ZF(x)] + \sigma_0(x) \geq 0,$$

because $F(x) \succeq 0$, for each feasible point x, and so, Tr $[ZF(x)] \geq 0$.

[(i) ⇒ (ii)] Suppose that x^* is a minimizer of (P). Then, it follows from Theorem 3.1 that there exists $Z^* \in S_+^m$ such that $L(\cdot, Z^*) - L(x^*, Z^*) \in \Sigma^2$ and $L(x^*, Z) \leq L(x^*, Z^*)$, $\forall Z \in S_+^m$. This shows that $\text{Tr}[Z^*F(x^*)] = 0$ and, for each $x \in \mathbb{R}^n$), $f(x) - f(x^*) - \text{Tr}\ [Z^*F(x)] = \sigma_0(x)$, for some $\sigma_0 \in \Sigma^2$. □

Consider the following semi-definite fractional programming problem

$$(P_1) \quad \min \quad \frac{f(x)}{g(x)}$$
$$\text{s.t.} \quad F_0 + \sum_{i=1}^n x_i F_i \succeq 0,$$

where $f, -g$ are SOS-convex polynomials on $\mathbb{R}^n$ such that $g(x) > 0$ and $f(x) \geq 0$ for all x satisfies $F_0 + \sum_{i=1}^n x_i F_i \succeq 0$.

Corollary 5.2. *For (P_1), assume that x^* is a feasible point. Suppose that the convex cone D is closed. Then, the following statements are equivalent.*

(i) *The point* x^* *is a minimizer of* (P_1).

(ii) $(\exists Z \in S^m_+)\ (\exists \sigma_0 \in \Sigma^2)\ (\forall x \in \mathbb{R}^n)\ f(x) - \frac{f(x^*)}{g(x^*)} g(x) - \text{Tr}\ [ZF(x)] = \sigma_0(x).$

Proof. If x^* is a minimizer of (P_1), then x^* is a solution of

$$(P_1') \quad \min \quad f(x) - \frac{f(x^*)}{g(x^*)} g(x)$$
$$\text{s.t.} \quad F_0 + \sum_{i=1}^{n} x_i F_i \succeq 0.$$

Then, (ii) follows from Theorem 5.1. The converse implication easily follows.

□

In the following Theorem, we show that the normal cone condition (2.1) is the weakest condition guaranteeing the characterization of optimality of a SOS-convex semidefinite programming problem in Corollary 4.1.

Theorem 5.2. (Weakest Condition for Checkable Characterization) *The following assertions are equivalent.*

(i) $N_{F^{-1}(S^m_+)}(x) = \{-\hat{F}^*(Z) : Z \in S^m_+, \text{Tr}[ZF(x)] = 0\}, \quad \forall x \in F^{-1}(S^m_+)$

(ii) *For each SOS-convex function* $f : \mathbb{R}^n \to \mathbb{R}$ *and for each minimizer* x^* *of* f *over* $F^{-1}(S^m_+)$,

$$(\exists Z \in S^m_+)\ (\exists \sigma_0 \in \Sigma^2)\ (\forall x \in \mathbb{R}^n)\ f(x) - f(x^*) - \text{Tr}\ [ZF(x)] = \sigma_0(x).$$

Proof. [(i) ⇒ (ii)] Let f be a SOS-convex function and let x^* be a minimizer of f over $F^{-1}(S^m_+)$. Then,

$$F_0 + \sum_{i=1}^{n} x_i F_i \succeq 0 \ \Rightarrow\ f(x) - f(x^*) \geq 0.$$

So, $-\nabla f(x^*) \in N_{F^{-1}(S^m_+)}(x^*)$. Thus, the normal condition implies that

$$-\nabla f(x^*) \in \{-\hat{F}^*(Z) : Z \in S^m_+, \text{Tr}[ZF(x^*)] = 0\}$$

and so, there exists $Z \in S^m_+$ with $\text{Tr}[ZF(x^*)] = 0$ such that $\nabla f(x^*) = \hat{F}^*(Z)$. Now, consider the function $\sigma_0(x) = f(x) - f(x^*) - \text{Tr}\ [ZF(x)]$. Clearly, σ_0 is convex function with $\sigma_0(x^*) = 0$ and $\nabla \sigma_0(x^*) = \nabla f(x^*) - \hat{F}^*(Z) = 0$. This implies that x^* is a global minimizer of σ_0, and hence $\sigma_0(x) \geq \sigma_0(x^*) = 0$ for all $x \in \mathbb{R}^n$. Then, by similar arguments as in Theorem 3.1, we can show that σ_0 is a sum of squares polynomial.

[(ii) ⇒ (i)] We proceed by the method of contradiction and suppose that (i) fails. Noting that $\{-\hat{F}^*(Z) : Z \in S^m_+, \text{Tr}[ZF(x)] = 0\} \subset N_{F^{-1}(S^m_+)}(x)$ for all $x \in F^{-1}(S^m_+)$ always holds, we can find $x^* \in F^{-1}(S^m_+)$ such that

$$a \in N_{F^{-1}(S^m_+)}(x^*) \backslash \{-\hat{F}^*(Z) : Z \in S^m_+, \text{Tr}[ZF(x^*)] = 0\}.$$

Now, consider the affine function $f(x) = -a^T x$. As $a \in N_{F^{-1}(S_+^m)}(x^*)$, $a^T(x - x^*) \leq 0$ for all $x \in F^{-1}(S_+^m)$. It follows that x^* is a global minimizer of f over $F^{-1}(S_+^m)$. So, by (ii), there exist $Z \in S_+^m$ and $\sigma_0 \in \Sigma^2$ such that for all $x \in \mathbb{R}^n$

$$f(x) - f(x^*) - \text{Tr } [ZF(x)] = \sigma_0(x).$$

Note that $f(x) - f(x^*) - \text{Tr } [ZF(x)]$ is an affine function and any affine function $c^T x + \alpha$ is sum of squares if and only if $c = 0$ and $\alpha \geq 0$. So, we see that $a = -\hat{F}^*(Z)$ and $\text{Tr } [ZF(x^*)] \leq 0$. As $x^* \in F^{-1}(S_+^m)$ and $Z \in S_+^m$, $\text{Tr } [ZF(x^*)] \geq 0$. So, $\text{Tr } [ZF(x^*)] = 0$. However, this contradicts the fact that $a \notin \{-\hat{F}^*(Z) : Z \in S_+^m, \text{Tr}[ZF(x^*)] = 0\}$. □

6 Appendix: SDP Representations of Relaxations

In this Section, for the sake of completeness, we show how our relaxation problem can be represented by a semidefinite linear programming problem.

Let $k \in \mathbb{N}$. Let $P_k(\mathbb{R}^n)$ be the space consisting of all real polynomials on $\mathbb{R}^n$ with degree k and let $C(k, n)$ be the dimension of $P_k(\mathbb{R}^n)$. Write the basis of $P_k(\mathbb{R}^n)$ by

$$z^{(k)} := [1, x_1, x_2, \ldots, x_n, x_1^2, x_1 x_2, \ldots, x_2^2, \ldots, x_n^2, \ldots, x_1^k, \ldots, x_n^k]^T$$

and $z_\alpha^{(k)}$ be the α-th coordinate of $z^{(k)}$, $1 \leq \alpha \leq C(k, n)$. Let $g(x) = -\text{Tr } [ZF(x)]$. Note that f is of degree d. We can write

$$f = \sum_{1 \leq \alpha \leq C(d,n)} f_\alpha z_\alpha^{(d)} \text{ and } g = \sum_{1 \leq \alpha \leq C(d,n)} g_\alpha z_\alpha^{(d)}.$$

where $g_\alpha = -\text{Tr } [ZF_{\alpha-1}(x)]$ for $1 \leq \alpha \leq n + 1$ and $g_\alpha = 0$ for any $\alpha > n + 1$. Let $2l$ be the smallest even number which larger or equal to d. Then condition "$f + g - \mu$ is a SOS-polynomial" is equivalent to finding a matrix $W \in S_+^{C(l,n)}$ such that

$$-\mu + \sum_{1 \leq \alpha \leq C(d,n)} (f_\alpha + g_\alpha) z_\alpha^{(d)} = z^{(l)} W z^{(l)},$$

which is in turn equivalent to finding matrices $W \in S_+^{C(l,n)}$ and $Z \in S_+^n$ such that

$$\begin{cases} -\mu + f_1 + g_1 = W_{1,1} \\ f_\alpha + g_\alpha = \sum_{1 \leq \beta, \gamma \leq C(l,n), \beta + \gamma = \alpha} W_{\beta,\gamma} \quad (2 \leq \alpha \leq C(d, n)). \\ g_\alpha = -\text{Tr } [ZF_{\alpha-1}(x)], \ 1 \leq \alpha \leq n + 1 \\ g_\alpha = 0, \ \alpha > n + 1. \end{cases}$$

Therefore, the relaxation problem

$$(D) \max_{\mu \in \mathbb{R}, Z \in S_+^m} \{\mu : f - \text{Tr } [ZF(x)] - \mu \in \Sigma_d^2\},$$

is equivalent to the following semi-definite programming problem

$$\max_{\mu\in\mathbb{R},\, W\in S_+^{C(l,n)},\, Z\in S_+^n} \quad \mu$$

$$\begin{aligned} & -\mu + f_1 + g_1 = W_{1,1} \\ & f_\alpha + g_\alpha = \sum_{1\le\beta,\gamma\le C(l,n),\beta+\gamma=\alpha} W_{\beta,\gamma} \quad (2 \le \alpha \le C(d,n)). \\ & g_\alpha = -\mathrm{Tr}\,[ZF_{\alpha-1}(x)],\ 1\le\alpha\le n+1 \\ & g_\alpha = 0,\ \alpha > n+1. \end{aligned}$$

References

[1] A.A. Ahmadi and P.A. Parrilo. A convex polynomial that is not SOS-convex. *Math. Program.* (2011), DOI: 10.1007/s10107-011-0457-z.

[2] A.A. Ahmadi and P.A. Parrilo. A complete characterization of the gap between convexity and SOS-convexity, preprint (2011), arXiv:1111.4587v1 [math.OC] .

[3] E. G. Belousov and D. Klatte, A Frank-Wolfe type theorem for convex polynomial programs, *Comp. Optim. & Appl.,* **22** (2002), 37-48.

[4] G. Blekherman, Nonnegative polynomials and sums of squares, arXiv:1010.3465, (to appear in the *Journal of the AMS* (2012).

[5] A. Ben-Tal, L. El Ghaoui and A. Nemirovski, Robust Optimization, Princeton Series in Applied Mathematics, 2009.

[6] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge (2004).

[7] R.S. Burachik, and V. Jeyakumar, A new geometric condition for Fenchel's duality in infinite dimensional spaces, *Math. Program.*, **104** (2005), 229-233.

[8] H. V. Ha and T. S. Pham, Representation of positive polynomials and optimization on noncompact semialgebraic sets, *SIAM J. Optim*, **20** (2010), 3082-3103.

[9] J.W. Helton and J.W. Nie, Structured Semidefinite Representation of Some Convex Sets , *Proceedings of 47th IEEE Conference on Decision and Control*, 4797-4800, Cancun, Mexico, Dec. 9-11, 2008.

[10] J.W. Helton and J.W. Nie, Semidefinite representation of convex sets *Math. Program.*, **122** (2010), 21-64.

[11] D. Henrion and J. B. Lasserre, Convergenet relaxations of polynomial matrix inequalities and static output feedback, *IEEE Trans. Automat. Cont.*, **51** (2006), 192-202.

[12] V. Jeyakumar, A note on strong duality in convex semidefinite optimization: necessary and sufficient conditions, *Optim. Lett.* **2** (2008), 15-25.

[13] V. Jeyakumar, and M. Nealon, Complete characterizations of optimality for convex semidefinite programming, *Canadian Math. Soc. Conf. Proc*, **27** (2000), 165-173.

[14] E. de Klerk, Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications, Vol. 65, Applied Optimization Series, Kluwer Academic Publishers, 2002.

[15] E. de. Klerk, M. Laurent and P.Parrilo, *On the equivalence of algebraic approaches to the minimization of forms on the simplex*. In D. Henrion & A. Garulli (Eds.), Positive Polynomials in Control. Berlin: Springer Verlag. (LNCIS, 312), 2005.

[16] E. de. Klerk and M. Laurent, On the Lasserre hierarchy of semidefinite programming relaxations of convex polynomial optimization problems. *SIAM J. Optim.* **21** (2011), 824-832,

[17] J.B. Lasserre, *Moments, Positive Polynomials and Their Applications*, Imperial College Press, 2009.

[18] J.B. Lasserre, Representation of nonnegative convex polynomials, *Arch. Math.* **91** (2008), no. 2, 126-130.

[19] J.B. Lasserre, Convexity in semialgebraic geometry and polynomial optimization, **19** *SIAM J. Optim.*, (2009), 1995-2014.

[20] J. B. Lasserre, Global optimization with polynomials and the problem of moments, *SIAM J. Optim.*, **11** (2000/01), 796-817.

[21] J. B. Lasserre, Convex sets with semidefinite representation, *Math. Program*, Ser A, **120** (2009), 457-477.

[22] G. Li, and K.F. Ng, On extension of Fenchel duality and its application, *SIAM J. Optim.*, **19** (2008), 1489-1509.

[23] M. Laurent, Sums of squares, moment matrices and optimization over polynomials. Emerging applications of algebraic geometry, 157-270, *IMA Vol. Math. Appl.*, 149, Springer, New York, 2009.

[24] P. A. Parrilo, Semidefinite programming relaxations for semialgebraic problems, *Math. Program,* **96** (2003), 293-320.

[25] R. T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, Princeton, N. J, 1970.

[26] L. Vandenberghe and S. Boyd, Semidefinite programming, *SIAM Rev.*, **38** (1996), 49-95.

[27] H. Yamashita and H. Yabe, Local and superlinear convergence of a primal-dual interior point method for nonlinear semidefinite programming, *Math. Program*, Ser. A, **132** (2012), 1–30.