



SOS-Convex Semialgebraic Programs and its Applications to Robust Optimization: A Tractable Class of Nonsmooth Convex Optimization

N. H. Chieu^{1,2} · J. W. Feng³ · W. Gao³ · G. Li¹ · D. Wu³

Received: 18 January 2017 / Accepted: 5 October 2017 / Published online: 14 November 2017
© Springer Science+Business Media B.V. 2017

Abstract In this paper, we introduce a new class of nonsmooth convex functions called SOS-convex semialgebraic functions extending the recently proposed notion of SOS-convex polynomials. This class of nonsmooth convex functions covers many common nonsmooth functions arising in the applications such as the Euclidean norm, the maximum eigenvalue function and the least squares functions with ℓ_1 -regularization or elastic net regularization used in statistics and compressed sensing. We show that, under commonly used strict feasibility conditions, the optimal value and an optimal solution of SOS-convex semialgebraic programs can be found by solving a single semidefinite programming problem (SDP). We achieve the results by using tools from semialgebraic geometry, convex-concave

Dedicated to the memory of Jon Borwein who was of great inspiration to us

Research was partially supported by a research grant from Australian Research Council. Research was also partially supported by the Vietnam National Foundation 465 for Science and Technology Development (NAFOSTED) under grant 101.01-2017.325.

✉ G. Li
g.li@unsw.edu.au

N. H. Chieu
chieunh@vinhuni.edu.vn

J. W. Feng
jinwen.feng@unsw.edu.au

W. Gao
w.gao@unsw.edu.au

D. Wu
di.wu@unsw.edu.au

¹ School of Mathematics Statistics, University of New South Wales, Sydney NSW 2052, Australia

² Department of Mathematics, Vinh University, Vinh, Nghe An 42118, Vietnam

³ School of Civil and Environmental Engineering, University of New South Wales, Sydney NSW 2052, Australia

minimax theorem and a recently established Jensen inequality type result for SOS-convex polynomials. As an application, we show that robust SOS-convex optimization problems under restricted spectrahedron data uncertainty enjoy exact SDP relaxations. This extends the existing exact SDP relaxation result for restricted ellipsoidal data uncertainty and answers an open question in the literature on how to recover a robust solution of uncertain SOS-convex polynomial programs from its semidefinite programming relaxation in this broader setting.

Keywords Nonsmooth optimization · Convex optimization · SOS-convex polynomial · Semidefinite program · Robust optimization

Mathematics Subject Classification (2010) 90C25 · 90C22 · 90C31

1 Introduction

Convex optimization is ubiquitous across science and engineering [3, 6]. It has found applications in a wide range of disciplines, such as automatic control systems, signal processing, electronic circuit design, data analysis, statistics (optimal design), and finance (see [3, 7] and the references therein). The key to the success in solving convex optimization problems is that convex functions exhibit a local to global phenomenon: every local minimizer is a global minimizer. Despite the great success of theoretical and algorithmic development and its wide application, we note that a convex optimization problem is, in general, NP-hard from the complexity point of view.

Recently, for convex polynomials, a new notion of sums-of-squares-convexity (SOS-convexity) [1, 15] has been proposed as a tractable sufficient condition for convexity based on semidefinite programming. The SOS-convex polynomials cover many commonly used convex polynomials such as convex quadratic functions and convex separable polynomials. An appealing feature of an SOS-convex polynomial is that deciding whether a polynomial is SOS-convex or not can be equivalently rewritten as a feasibility problem of a semidefinite programming problem (SDP) which can be validated efficiently. It has also been recently shown that for an SOS-convex optimization problem, its optimal value and optimal solution can be found by solving a single semidefinite programming problem [22] (see also [17, 18]). On the other hand, many modern applications of optimization to the area of statistics, machine learning, signal processing and image processing often result in structured nonsmooth convex optimization problems [7]. These optimization problems often take the following generic form $\min_{x \in \mathbb{R}^n} \{g(x) + h(x)\}$, where $g : \mathbb{R}^n \rightarrow \mathbb{R}$ is a convex quadratic function and $h : \mathbb{R}^n \rightarrow \mathbb{R}$ is a nonsmooth function. For example, in many signal processing applications g represents the quality of the recovered signal while h serves as a regularization which enforces prior knowledge of the form of the signal, such as simplicity/sparsity (in the sense that the solution has fewest nonzero entries). Some typical choices of the regularization function promoting the sparsity of the solution are the so-called ℓ_1 -norm and the weighted sum of ℓ_1 -norm and ℓ_2 -norm (referred as the elastic net regularization [26]), and is therefore nonsmooth. With these applications in mind, this then motivates the following natural and important question:

Is it possible to extend the SOS-convex polynomials and SOS-convex optimization problems to the nonsmooth setting which not only covers broad nonsmooth problems arising in common applications but also maintains the appealing feature of tractability (in terms of semidefinite programming)?

The purpose of this paper is to provide an affirmative answer for the above question. In particular, in this paper, we make the following contributions:

- (1) In Section 3, we identify a new class of nonsmooth convex functions which are referred to as SOS-convex semialgebraic functions (Definition 3.1). This class of nonsmooth convex functions covers not only convex functions which can be expressed as the maximum of finitely many SOS-convex polynomials (in particular, SOS-convex polynomials) but also many common nonsmooth functions arising in the applications such as the Euclidean norm, the maximum eigenvalue function (by identifying the space of symmetric matrices S^n as an Euclidean space with dimension $n(n + 1)/2$) and the least squares functions with ℓ_1 -regularizer or elastic net regularizer used in compressed sensing.
- (2) In Section 4, we show that, under a commonly used strict feasibility condition, the optimal value and an optimal solution of SOS-convex semialgebraic optimization problems can be found by solving a single semidefinite programming problem which extends the previous known result of SOS-convex polynomial optimization problems (Theorem 4.1 and Theorem 4.2). We achieve this by exploiting tools from semialgebraic geometry, convex-concave minimax theorem and a recently established Jensen inequality type result for SOS-convex polynomials.
- (3) In Section 5, as applications, we establish that robust SOS-convex optimization problems under restricted spectrahedron data uncertainty enjoy exact semidefinite programming relaxations under strict feasibility conditions. This extends the existing result for restricted ellipsoidal data uncertainty established in [19] and answers the open question left in [19] on how to recover a robust solution from the semidefinite programming relaxation in this broader setting.

2 Preliminaries

First of all, let us recall some notations and basic facts on sums-of-squares polynomial and semidefinite programming problems. Recall that S_n denotes the space of symmetric $(n \times n)$ matrices with the trace inner product and \succeq denotes the Löwner partial order of S^n , that is, for $M, N \in S^n$, $M \succeq N$ if and only if $(M - N)$ is positive semidefinite. Let $S^n_+ := \{M \in S_n \mid M \succeq 0\}$ be the closed convex cone of positive semidefinite symmetric $(n \times n)$ matrices. Note that for $M, N \in S^n_+$, denote the inner product as usual: $\langle M, N \rangle := \text{Tr} [MN]$, where $\text{Tr} [\cdot]$ refers to the trace operation. Note also that $M \succ 0$ means that M is positive definite. In the sequel, unless otherwise stated, the space \mathbb{R}^n is equipped with the Euclidean norm,

that is, $\|x\| := \left(\sum_{i=1}^n |x_i|^2\right)^{1/2}$ for all $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$. Consider a polynomial f with degree at most d where d is an even number. Let $\mathbb{R}_d[x_1, \dots, x_n]$ be the space consisting of all real polynomials on \mathbb{R}^n with degree at most d , and let $s(d, n) := \binom{n+d}{d}$ be the dimension of $\mathbb{R}_d[x_1, \dots, x_n]$. Write the canonical basis of $\mathbb{R}_d[x_1, \dots, x_n]$ by

$$x^{(d)} := (1, x_1, x_2, \dots, x_n, x_1^2, x_1x_2, \dots, x_2^2, \dots, x_n^2, \dots, x_1^d, \dots, x_n^d)^T$$

and let $x_\alpha^{(d)}$ be the α -th coordinate of $x^{(d)}$, $1 \leq \alpha \leq s(d, n)$. Then, we can write $f(x) = \sum_{\alpha=1}^{s(d,n)} f_\alpha x_\alpha^{(d)}$.

We say that a real polynomial f is sums-of-squares (cf. [21]) if there exist real polynomials $f_j, j = 1, \dots, q$, such that $f = \sum_{j=1}^q f_j^2$. The set consisting of all sum of squares

real polynomials in the variable x is denoted by $\Sigma^2[x]$. Moreover, the set consisting of all sum of squares real polynomials with degree at most d is denoted by $\Sigma_d^2[x]$. For a polynomial f , we use $\deg f$ to denote the degree of f . Let $l = d/2$. Then, f is a sum-of-squares polynomial if and only if there exists a positive semidefinite symmetric matrix $W \in S_+^{s(l,n)}$ such that

$$f(x) = (x^{(l)})^T W x^{(l)}, \tag{1}$$

where $x^{(l)} = (1, x_1, x_2, \dots, x_n, x_1^2, x_1x_2, \dots, x_2^2, \dots, x_n^2, \dots, x_1^l, \dots, x_n^l)^T$. For each $1 \leq \alpha \leq s(d, n)$, we denote $i(\alpha) = (i_1(\alpha), \dots, i_n(\alpha)) \in (\mathbb{N} \cup \{0\})^n$ to be the multi-index such that

$$x_\alpha^{(d)} = x^{i(\alpha)} := x_1^{i_1(\alpha)} \dots x_n^{i_n(\alpha)}.$$

Then, by comparing the coefficients in (1), we have the following linear matrix inequality characterization of a sum-of-squares polynomial.

Lemma 2.1 *Let d be an even number. For a polynomial f on \mathbb{R}^n with degree at most d , f is a sum-of-squares polynomial if and only if the following linear matrix inequality problem has a solution*

$$\begin{cases} W \in S_+^{s(l,n)} \\ f_\alpha = \sum_{1 \leq \beta, \gamma \leq s(l,n), i(\beta)+i(\gamma)=i(\alpha)} W_{\beta, \gamma}, \quad 1 \leq \alpha \leq s(d, n), \quad l = d/2. \end{cases}$$

We now recall the definition of SOS-convex polynomial. The notion of SOS-convex polynomial was first proposed in [15] and further developed in [1]. Here, for convenience of our discussion, we follow the definition used in [1].

Definition 2.1 (SOS-Convex Polynomials [15]) A real polynomial f on \mathbb{R}^n is called *SOS-convex* if the polynomial $F : (x, y) \mapsto f(x) - f(y) - \nabla f(y)^T(x - y)$ is a sums-of-squares polynomial on $\mathbb{R}^n \times \mathbb{R}^n$.

The significance of the class of SOS-convex polynomials is that checking whether a polynomial is SOS-convex is equivalent to solving a semidefinite programming problem (SDP) which can be done in polynomial time; while checking a polynomial is convex or not is, in general, an NP-hard problem [1, 15]. Moreover, another important fact is that, for SOS-convex polynomial program, an exact SDP relaxation holds under the usual strict feasibility condition. In contrast, solving a convex polynomial program, is again, in general, an NP hard problem [1].

Clearly, a SOS-convex polynomial is convex. However, the converse is not true, that is, there exists a convex polynomial which is not SOS-convex [1]. The sum of two SOS-convex polynomials and nonnegative scalar multiplication of an SOS-convex polynomial are still SOS convex polynomials. It is known that any convex quadratic function and any convex separable polynomial is an SOS-convex polynomial [17]. Moreover, an 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 which is non-quadratic and non-separable.

The following existence result for solutions of a convex polynomial optimization problem will also be useful for our later analysis.

Lemma 2.2 (Solution Existence of Convex Polynomial Programs [2, Theorem 3]) *Let f_0, f_1, \dots, f_m be convex polynomials on \mathbb{R}^n and let $C :=$*

$\{x \in \mathbb{R}^n : f_i(x) \leq 0, i = 1, \dots, m\}$ be nonempty. If $\inf_{x \in C} f_0(x) > -\infty$ then $\operatorname{argmin}_{x \in C} f_0(x) \neq \emptyset$.

3 SOS-Convex Semialgebraic Functions

We begin this section with introducing the notion of SOS-convex semialgebraic functions. The class of SOS-convex semialgebraic functions is a subclass of the class of locally Lipschitz nonsmooth convex functions, and includes SOS-convex polynomials.

Definition 3.1 (SOS-convex semialgebraic functions) We say $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is an SOS-convex semialgebraic function on \mathbb{R}^n if it admits a representation

$$f(x) = \max_{y \in \Omega} \{h_0(x) + \sum_{j=1}^m y_j h_j(x)\}, \quad m \in \mathbb{N}, \tag{2}$$

where

- (1) each $h_j, j = 0, 1, \dots, m$, is a polynomial, and, for each $y \in \Omega, h_0 + \sum_{j=1}^m y_j h_j$ is a SOS-convex polynomial on \mathbb{R}^n ;
- (2) Ω is a nonempty compact semidefinite program representable set given by

$$\Omega = \{y \in \mathbb{R}^m : \exists z \in \mathbb{R}^p \text{ such that } A_0 + \sum_{j=1}^m y_j A_j + \sum_{l=1}^p z_l B_l \succeq 0\}, \tag{3}$$

for some $p \in \mathbb{N}, A_j$ and $B_l, j = 0, 1, \dots, m, l = 1, \dots, p$, being $(t \times t)$ -symmetric matrices with some $t \in \mathbb{N}$.

Moreover, the maximum of the degree of the polynomial $h_j, j = 1, \dots, m$, is called the degree of the SOS-convex semialgebraic function f with respect to the representation (2).

Remark 3.1 (Discussion on the notion of SOS-convex semialgebraic functions) In classical algebraic geometry literature [5], a set C is called a semialgebraic set if it can be written as finite union of sets with the form $\{x \in \mathbb{R}^n : \phi_1(x) = 0, \dots, \phi_k(x) = 0, \psi_1(x) < 0, \dots, \psi_q(x) < 0\}$ for some real polynomials ψ_1, \dots, ψ_k and ψ_1, \dots, ψ_q . We also say a function f is semialgebraic if its graph $\{(x, f(x)) : x \in \mathbb{R}^n\}$ is a semialgebraic set [5, 24]. It is not hard to verify that any SOS-convex semialgebraic function is a semialgebraic function with additional SOS-convex structure. This provides the motivation of the name of this class of nonsmooth functions.

The class of SOS-convex semialgebraic functions contains many common nonsmooth convex functions. Below, we provide some typical examples.

Example 3.1 (Examples of SOS-convex semialgebraic functions)

- (1) Let $f(x) = \max_{1 \leq i \leq m} f_i(x)$ where each $f_i, i = 1, \dots, m$, is an SOS-convex polynomial. Note that $f(x) = \max_{y \in \Delta} g(x, y)$ where Δ is the simplex in \mathbb{R}^m given by $\Delta = \{y : y_i \geq 0, \sum_{i=1}^m y_i = 1\}$ and $g(x, y) = \sum_{i=1}^m y_i f_i(x)$. Then, we see that f is an SOS-convex semialgebraic function.

- (2) Let $f(x) = \|x\|$. Then, f is an SOS-convex semialgebraic function. To see this, we only need to note that

$$\|x\| = \max_{\|y\| \leq 1} \sum_{i=1}^n x_i y_i,$$

and the unit ball defined by $\|\cdot\|$ is a compact semidefinite program representable set. More generally, $f(x) = \|x\|_p := (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$ with $p = \frac{s}{s-1}$ and s being an even positive integer, is an SOS-convex semialgebraic function. To see this, we only need to note that

$$\|x\|_p = \max_{\|y\|_s \leq 1} \sum_{i=1}^n x_i y_i.$$

and the set $\{y \in \mathbb{R}^n : \|y\|_s \leq 1\} = \{y \in \mathbb{R}^n : \sum_{i=1}^n y_i^s \leq 1\}$ is described by an SOS-convex polynomial inequality (as s is even) and so, is a compact semidefinite program representable set [15].

- (3) Identify the space of $(n \times n)$ symmetric matrices S^n with the trace inner product $\text{Tr}(AB) = \sum_{ij} A_{ij} B_{ij}$ as $\mathbb{R}^{n(n+1)/2}$ with the usual inner product. Let $f : S^n \rightarrow \mathbb{R}$ be defined by $f(X) = \lambda_{\max}(X)$ where λ_{\max} is the maximum eigenvalue. Then, f is an SOS-convex semialgebraic function on S^n . To see this, we only need to notice that

$$\lambda_{\max}(X) = \max\{\text{Tr}(XY) : Y \in S^n, \text{Tr}(Y) = 1, Y \geq 0\}$$

and the set $\{Y \in S^n : \text{Tr}(Y) = 1, Y \geq 0\}$ is a compact semidefinite program representable set.

Next, we see that SOS-convex semialgebraic functions cover many least squares functions with regularization. To see this, we need the following simple proposition which shows that finite addition preserves SOS-convex semialgebraicity.

Proposition 3.1 *Let f_i be SOS-convex semialgebraic functions on \mathbb{R}^n , $i = 1, \dots, q$. Then, $\sum_{i=1}^q f_i$ is an SOS-convex semialgebraic function on \mathbb{R}^n .*

Proof To see the conclusion, it suffices to show the case where $q = 2$. We first show that $f_1 + f_2$ is an SOS-convex semialgebraic function. As $f_i, i = 1, 2$ are SOS-convex semialgebraic functions, $f_i(x) = \max_{y^i \in \Omega_i} \{h_0^i(x) + \sum_{j=1}^{m_i} y_j^i h_j^i(x)\}$, where $m_i \in \mathbb{N}$, $h_l^i, l = 0, 1, \dots, m$ are SOS-convex polynomials and Ω_i is a compact semidefinite program representable sets given by

$$\Omega_i = \{y^i \in \mathbb{R}^{m_i} : \exists z^i \in \mathbb{R}^{p_i} \text{ such that } A_0^i + \sum_{j=1}^{m_i} y_j^i A_j^i + \sum_{l=1}^{p_i} z_l^i B_l^i \geq 0\}.$$

Then,

$$f_1(x) + f_2(x) = \max_{(y^1, y^2) \in \Omega_1 \times \Omega_2} \left\{ h_0^1(x) + h_0^2(x) + \sum_{j=1}^{m_1} y_j^1 h_j^1(x) + \sum_{j=1}^{m_2} y_j^2 h_j^2(x) \right\}.$$

Note that $\Omega_1 \times \Omega_2$ is also a compact semidefinite program representable set. Thus, $f_1 + f_2$ is also an SOS-convex semialgebraic function. □

Example 3.2 (Further examples: least squares functions with regularizations) Let $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. From the preceding proposition, we see that the following functions which arises in sparse optimization are SOS-convex semialgebraic:

- (1) The least squares function with ℓ_1 -regularization $f(x) = \|Ax - b\|^2 + \mu\|x\|_1$ where $\mu > 0$. Note that since $\|x\|_1 := |x_1| + |x_2| + \dots + |x_n|$ and $|x_i| = \max\{x_i, -x_i\}$, $i = 1, 2, \dots, n$, it follows from Example 3.1 and Proposition 3.1 that $\|\cdot\|_1$ is an SOS-convex semialgebraic function, while the function $x \mapsto \|Ax - b\|^2$ is a convex quadratic function and thus is SOS-convex semialgebraic.
- (2) The least squares function with elastic net regularization [26] $f(x) = \|Ax - b\|^2 + \mu_1\|x\|_1 + \mu_2\|x\|^2$ where $\mu_1, \mu_2 > 0$.

4 Exact SDP Relaxation for SOS-Convex Semialgebraic Programs

In this section, we show that an SOS-convex semialgebraic program as the problem (P) below admits an exact SDP relaxation in the sense that the optimal value of the SDP relaxation problem equals the optimal value of the underlying SOS-convex semialgebraic program. Moreover, a solution for the SOS-convex semialgebraic program can be recovered from its SDP relaxation, under strict feasibility assumptions.

Consider the following SOS-convex semialgebraic program:

$$(P) \quad \inf f_0(x) \\ \text{s.t. } f_i(x) \leq 0, \quad i = 1, \dots, s,$$

where each $f_i, i = 0, 1, \dots, s$, is an SOS-convex semialgebraic function in the form

$$f_i(x) = \max_{y^i \in \Omega_i} \{h_0^i(x) + \sum_{j=1}^m y_j^i h_j^i(x)\}, \quad m \in \mathbb{N},$$

such that

- (1) each h_j^i is a polynomial with degree at most d , and for each $y^i = (y_1^i, \dots, y_m^i) \in \Omega_i$, the function $h_0^i + \sum_{j=1}^m y_j^i h_j^i$ is an SOS-convex polynomial on \mathbb{R}^n ;
- (2) $\Omega_i, i = 0, 1, \dots, s$, is a nonempty compact semidefinite program representable set given by

$$\Omega_i = \{(y_1^i, \dots, y_m^i) \in \mathbb{R}^m : \exists z^i = (z_1^i, \dots, z_{p_i}^i) \in \mathbb{R}^{p_i} \text{ such that } A_0^i + \sum_{j=1}^m y_j^i A_j^i + \sum_{l=1}^{p_i} z_l^i B_l^i \geq 0\},$$

for some $p_i \in \mathbb{N} \cup \{0\}$.

Without loss of generality, we always assume that d is an even number.

Inspired from [2, 20], we now introduce a relaxation problem for problem (P) as follows

$$\begin{aligned}
 (SDP) \quad & \sup_{\substack{\lambda_0^i \geq 0, (\lambda_1^i, \dots, \lambda_m^i) \in \mathbb{R}^m \\ z_l^i \in \mathbb{R}, \mu \in \mathbb{R}}} \left\{ \mu : h_0^0 + \sum_{j=1}^m \lambda_j^0 h_j^0 + \sum_{i=1}^s \left(\lambda_0^i h_0^i + \sum_{j=1}^m \lambda_j^i h_j^i \right) - \mu \in \Sigma_d^2[x], \right. \\
 & A_0^0 + \sum_{j=1}^m \lambda_j^0 A_j^0 + \sum_{l=1}^{p_0} z_l^0 B_l^0 \geq 0, \\
 & \left. \lambda_0^i A_0^i + \sum_{j=1}^m \lambda_j^i A_j^i + \sum_{l=1}^{p_i} z_l^i B_l^i \geq 0, i = 1, \dots, s \right\}.
 \end{aligned}$$

In the special case where the objective function can be expressed as a maximum of finitely many SOS-convex polynomials, the relaxation problem (SDP) collapses to the relaxation problem introduced in [20]. Moreover, direct application of Lemma 2.1 shows that (SDP) can be equivalently rewritten as the following semidefinite programming problem

$$\begin{aligned}
 & \sup_{\substack{\lambda_0^i \geq 0, (\lambda_1^i, \dots, \lambda_m^i) \in \mathbb{R}^m \\ z_l^i \in \mathbb{R}, \mu \in \mathbb{R}, W \in S^{s(d/2, n)}}} \left\{ \mu : (h_0^0)_1 + \sum_{j=1}^m \lambda_j^0 (h_j^0)_1 + \sum_{i=1}^s \left(\lambda_0^i (h_0^i)_1 + \sum_{j=1}^m \lambda_j^i (h_j^i)_1 \right) - \mu = W_{1,1}, \right. \\
 \text{s.t.} \quad & (h_0^0)_\alpha + \sum_{j=1}^m \lambda_j^0 (h_j^0)_\alpha + \sum_{i=1}^s \left(\lambda_0^i (h_0^i)_\alpha + \sum_{j=1}^m \lambda_j^i (h_j^i)_\alpha \right) \\
 & = \sum_{\substack{1 \leq \beta, \gamma \leq s(d/2, n) \\ i(\beta) + i(\gamma) = i(\alpha)}} W_{\beta, \gamma}, \quad 2 \leq \alpha \leq s(d, n) \\
 & f + \sum_{i=1}^m \left(\lambda_i^0 g_i^0 + \sum_{j=1}^s \lambda_j^i g_i^j \right) - \mu \in \Sigma_d^2, \\
 & W \geq 0, \quad A_0^0 + \sum_{j=1}^m \lambda_j^0 A_j^0 + \sum_{l=1}^{p_0} z_l^0 B_l^0 \geq 0, \\
 & \lambda_0^i A_0^i + \sum_{j=1}^m \lambda_j^i A_j^i + \sum_{l=1}^{p_i} z_l^i B_l^i \geq 0, i = 1, \dots, s.
 \end{aligned}$$

Next, we show that an exact SDP relaxation holds between (P) and (SDP) in the sense that their optimal values are the same. We start with a simple property for a bounded set which is described by linear matrix inequalities.

Lemma 4.1 *Let \mathcal{U} be a nonempty compact set with the form $\mathcal{U} = \{(u_1, \dots, u_m) \in \mathbb{R}^m : \exists z \in \mathbb{R}^p \text{ such that } A_0 + \sum_{j=1}^m u_j A_j + \sum_{l=1}^p z_l B_l \geq 0\}$ where $A_j, B_l \in S^q$. Let $(\lambda_0, \dots, \lambda_m) \in \mathbb{R}^{m+1}$ be such that $\lambda_0 A_0 + \sum_{j=1}^m \lambda_j A_j + \sum_{l=1}^p v_l B_l \geq 0$ for some $(v_1, \dots, v_p) \in \mathbb{R}^p$. Then, the following implication holds:*

$$\lambda_0 = 0 \implies \lambda_j = 0 \text{ for all } j = 1, \dots, m. \tag{4}$$

Proof Let $(\lambda_0, \dots, \lambda_m) \in \mathbb{R}^{m+1}$ be such that $\lambda_0 A_0 + \sum_{j=1}^m \lambda_j A_j + \sum_{l=1}^p v_l B_l \geq 0$ for some $(v_1, \dots, v_p) \in \mathbb{R}^p$. We proceed by the method of contradiction. Suppose that $\lambda_0 = 0$ and there exists $j_0 \in \{1, \dots, m\}$ with $\lambda_{j_0} \neq 0$. This means that

$$\sum_{j=1}^m \lambda_j A_j + \sum_{l=1}^p v_l B_l \geq 0 \text{ and } (\lambda_1, \dots, \lambda_m) \neq 0_{\mathbb{R}^m}.$$

Now take $\hat{u} = (\hat{u}_1, \dots, \hat{u}_m) \in \mathcal{U}$. Then, we have $A_0 + \sum_{j=1}^m \hat{u}_j A_j + \sum_{l=1}^p \hat{v}_l B_l \geq 0$ for some $(\hat{v}_1, \dots, \hat{v}_p) \in \mathbb{R}^p$, and so,

$$A_0 + \sum_{j=1}^m (\hat{u}_j + t\lambda_j) A_j + \sum_{l=1}^p (\hat{v}_l + tv_l) B_l \geq 0 \text{ for all } t \geq 0.$$

The latter implies that

$$(\hat{u}_1, \dots, \hat{u}_m) + t(\lambda_1, \dots, \lambda_m) \in \mathcal{U} \text{ for all } t \geq 0,$$

which contradicts the boundedness of \mathcal{U} . Thus, the conclusion follows. □

We are now ready to state and prove the first main result of this section, showing the exactness of the SDP relaxation for SOS-convex semialgebraic programs under a strict feasibility condition.

Theorem 4.1 (Exact SDP Relaxation for SOS-convex Semialgebraic Programs) *For problem (P), suppose the following strict feasibility condition holds: there exists $x_0 \in \mathbb{R}^n$ such that $f_i(x_0) < 0, i = 1, \dots, s$. Then, we have*

$$\text{val}(P) = \text{val}(\text{SDP}),$$

where $\text{val}(P)$ and $\text{val}(\text{SDP})$ are the optimal values of problems (P) and (SDP), respectively.

Proof We first justify that $\text{val}(P) \geq \text{val}(\text{SDP})$. Let $\lambda_0^i \geq 0, (\lambda_1^i, \dots, \lambda_m^i) \in \mathbb{R}^m, z_l^i \in \mathbb{R},$ and $\mu \in \mathbb{R},$ be feasible for (SDP). Then, we have

$$h_0^0 + \sum_{j=1}^m \lambda_j^0 h_j^0 + \sum_{i=1}^s \left(\lambda_0^i h_0^i + \sum_{j=1}^m \lambda_j^i h_j^i \right) - \mu \in \Sigma_d^2[x],$$

$$A_0^0 + \sum_{j=1}^m \lambda_j^0 A_j^0 + \sum_{l=1}^{p_0} z_l^0 B_l^0 \geq 0,$$

$$\lambda_0^i A_0^i + \sum_{j=1}^m \lambda_j^i A_j^i + \sum_{l=1}^{p_i} z_l^i B_l^i \geq 0, \quad i = 1, \dots, s.$$

Take any $x \in \mathbb{R}^n$ with $f_i(x) \leq 0, i = 1, \dots, s$. We want to show $f_0(x) \geq \mu$. For each $i = 1, \dots, s,$ pick $\bar{y}_i = (\bar{y}_1^i, \dots, \bar{y}_m^i) \in \Omega_i$. By the definition of $\Omega_i,$ there exist $\bar{z}^i \in \mathbb{R}^{p_i}$ such that $A_0^i + \sum_{j=1}^m \bar{y}_j^i A_j^i + \sum_{l=1}^{p_i} \bar{z}_l^i B_l^i \geq 0$. For each $j = 1, \dots, m$ and $i = 1, \dots, s,$ put

$$\tilde{y}_j^i := \begin{cases} \frac{\lambda_j^i}{\lambda_0^i} & \text{if } \lambda_0^i \neq 0, \\ \bar{y}_j^i & \text{if } \lambda_0^i = 0, \end{cases}$$

and

$$\tilde{z}_l^i := \begin{cases} \frac{z_l^i}{\lambda_0^i} & \text{if } \lambda_0^i \neq 0, \\ \bar{z}_l^i & \text{if } \lambda_0^i = 0. \end{cases}$$

Then, for each $i = 1, \dots, s$, we have

$$A_i^0 + \sum_{j=1}^m \tilde{y}_j^i A_j^i + \sum_{l=1}^{p_i} \tilde{z}_l^i B_l^i = \begin{cases} \frac{1}{\lambda_0^i} (\lambda_0^i A_0^i + \sum_{j=1}^m \lambda_j^i A_j^i + \sum_{l=1}^{p_i} z_l^i B_l^i) & \text{if } \lambda_0^i \neq 0, \\ A_0^i + \sum_{j=1}^m \bar{y}_j^i A_j^i + \sum_{l=1}^{p_i} \bar{z}_l^i B_l^i & \text{if } \lambda_0^i = 0, \end{cases}$$

which is always a positive semidefinite symmetric matrix. So, $\tilde{y}^i := (\tilde{y}_1^i, \dots, \tilde{y}_m^i) \in \Omega_i$ and hence

$$h_0^i(x) + \sum_{j=1}^m \tilde{y}_j^i h_j^i(x) \leq f_i(x) \leq 0 \quad \text{for all } i = 1, \dots, s. \tag{5}$$

Moreover, as the sets Ω_i are bounded, according to Lemma 4.1, for each $i = 1, \dots, s$, if $\lambda_0^i = 0$, then $\lambda_j^i = 0$ for all $j = 1, \dots, m$. This implies that

$$\begin{aligned} & h_0^0 + \sum_{j=1}^m \lambda_j^0 h_j^0 + \sum_{i=1}^s \lambda_0^i \left(h_0^i + \sum_{j=1}^m \tilde{y}_j^i h_j^i \right) - \mu \\ &= h_0^0 + \sum_{j=1}^m \lambda_j^0 h_j^0 + \sum_{i=1}^s \left(\lambda_0^i h_0^i + \sum_{j=1}^m (\lambda_0^i \tilde{y}_j^i) h_j^i \right) - \mu \\ &= h_0^0 + \sum_{j=1}^m \lambda_j^0 h_j^0 + \sum_{i=1}^s \left(\lambda_0^i h_0^i + \sum_{j=1}^m \lambda_j^i h_j^i \right) - \mu \in \Sigma_d^2[x]. \end{aligned}$$

So, noting that $(\lambda_1^0, \dots, \lambda_m^0) \in \Omega_0$ and $\lambda_0^i \geq 0$ for all $i = 1, \dots, s$, by (5) it holds that

$$\begin{aligned} f_0(x) &\geq h_0^0(x) + \sum_{j=1}^m \lambda_j^0 h_j^0(x) \\ &\geq h_0^0(x) + \sum_{j=1}^m \lambda_j^0 h_j^0(x) + \sum_{i=1}^s \lambda_0^i \left(h_0^i(x) + \sum_{j=1}^m \tilde{y}_j^i h_j^i(x) \right) \geq \mu. \end{aligned}$$

Therefore, $\text{val}(P) \geq \text{val}(SDP)$.

Next, we will justify that $\text{val}(P) \leq \text{val}(SDP)$. As $\text{val}(P) \geq \text{val}(SDP)$ always holds, it suffices to consider the case $\text{val}(P) > -\infty$. Noting that the feasible set of (P) is nonempty, we may assume that $\nu := \text{val}(P) \in \mathbb{R}$. Our assumptions guarantee that there exists x_0 such

that $f_i(x_0) < 0, i = 1, \dots, s$, and each f_i is a continuous convex function. So, the standard Lagrangian duality for convex programming problem shows that

$$\begin{aligned} v &:= \inf_{x \in \mathbb{R}^n} \{f_0(x) : f_i(x) \leq 0, i = 1, \dots, s\} \\ &= \max_{\lambda \in \mathbb{R}_+^s} \inf_{x \in \mathbb{R}^n} \{f_0(x) + \sum_{i=1}^s \lambda_i f_i(x)\} \\ &= \max_{\lambda \in \mathbb{R}_+^s} \inf_{x \in \mathbb{R}^n} \max_{y \in \prod_{i=0}^s \Omega_i} h_\lambda(x, y), \end{aligned} \tag{6}$$

where $\lambda := (\lambda_1, \dots, \lambda_s) \in \mathbb{R}^s, y := (y_1^0, \dots, y_m^0, \dots, y_1^s, \dots, y_m^s) \in \mathbb{R}^{m(s+1)}$, and

$$h_\lambda(x, y) := h_0^0(x) + \sum_{j=1}^m y_j^0 h_j^0(x) + \sum_{i=1}^s \lambda_i \left(h_0^i(x) + \sum_{j=1}^m y_j^i h_j^i(x) \right).$$

Note that $\prod_{i=0}^s \Omega_i$ is a convex compact set, and for any $\lambda \in \mathbb{R}_+^s$ the function $h_\lambda(x, y)$ is convex in x for each fixed y and is concave in y for each fixed x . Thus, for each $\lambda \in \mathbb{R}_+^s$, by the convex-concave minimax theorem we have

$$\inf_{x \in \mathbb{R}^n} \max_{y \in \prod_{i=0}^s \Omega_i} h_\lambda(x, y) = \max_{y \in \prod_{i=0}^s \Omega_i} \inf_{x \in \mathbb{R}^n} h_\lambda(x, y).$$

This together with (6) yields

$$\begin{aligned} v &= \max_{\lambda \in \mathbb{R}_+^s} \max_{y \in \prod_{i=0}^s \Omega_i} \inf_{x \in \mathbb{R}^n} h_\lambda(x, y) \\ &= \max_{\substack{y^i \in \Omega_i, 0 \leq i \leq s \\ \lambda_1 \geq 0, \dots, \lambda_s \geq 0}} \inf_{x \in \mathbb{R}^n} \{h_0^0(x) + \sum_{j=1}^m y_j^0 h_j^0(x) + \sum_{i=1}^s \lambda_i (h_0^i(x) + \sum_{j=1}^m y_j^i h_j^i(x))\}. \end{aligned}$$

In particular, the latter shows that there exist $(\tilde{y}_1^i, \dots, \tilde{y}_m^i) \in \Omega_i, 0 \leq i \leq s$, and $\tilde{\lambda}_1 \geq 0, \dots, \tilde{\lambda}_s \geq 0$, such that

$$\inf_{x \in \mathbb{R}^n} \{h_0^0(x) + \sum_{j=1}^m \tilde{y}_j^0 h_j^0(x) + \sum_{i=1}^s \tilde{\lambda}_i (h_0^i(x) + \sum_{j=1}^m \tilde{y}_j^i h_j^i(x))\} = v.$$

Denote $G(x) := h_0^0(x) + \sum_{j=1}^m \tilde{y}_j^0 h_j^0(x) + \sum_{i=1}^s \tilde{\lambda}_i (h_0^i(x) + \sum_{j=1}^m \tilde{y}_j^i h_j^i(x)) - v$. By Lemma 2.2, there exists $a \in \mathbb{R}^n$ such that $G(a) = \inf_{x \in \mathbb{R}^n} G(x) = 0$ (and so, $\nabla G(a) = 0$). As G is an SOS-convex polynomial, $H(x, y) := G(x) - G(y) - \nabla G(y)^T(x - y)$ is a sums-of-squares polynomial. Letting $y = a$, it follows that $G(x) = H(x, a)$ is also a sums-of-squares polynomial, that is,

$$h_0^0 + \sum_{j=1}^m \tilde{y}_j^0 h_j^0 + \sum_{i=1}^s \tilde{\lambda}_i (h_0^i + \sum_{j=1}^m \tilde{y}_j^i h_j^i) - v \in \Sigma_d^2[x]. \tag{7}$$

On the other hand, for each $i = 0, \dots, s$, since $(\tilde{y}_1^i, \dots, \tilde{y}_m^i) \in \Omega_i$, there exists $\tilde{z}^i = (\tilde{z}_1^i, \dots, \tilde{z}_{p_i}^i) \in \mathbb{R}^{p_i}$ such that

$$A_0^i + \sum_{j=1}^m \tilde{y}_j^i A_j^i + \sum_{l=1}^{p_i} \tilde{z}_l^i B_l^i \geq 0. \tag{8}$$

Now, let $\lambda_j^0 := \tilde{y}_j^0, j = 1, \dots, m, z_l^0 := \tilde{z}_l^0, l = 1, \dots, p_0, \lambda_0^i := \tilde{\lambda}_i$ and $\lambda_j^i := \tilde{\lambda}_i \tilde{y}_j^i$ and $z_l^i := \tilde{\lambda}_i \tilde{z}_l^i$ for each $j = 1, \dots, m, i = 1, \dots, s, l = 1, \dots, p_i$. From (7), (8) and $\tilde{\lambda}_i \geq 0$, we see that

$$h_0^0 + \sum_{j=1}^m \lambda_j^0 h_j^0 + \sum_{i=1}^s \left(\lambda_0^i h_0^i + \sum_{j=1}^m \lambda_j^i h_j^i \right) - \nu \in \Sigma_d^2[x],$$

$$A_0^0 + \sum_{j=1}^m \lambda_j^0 A_j^0 + \sum_{l=1}^{p_0} z_l^0 B_l^0 \geq 0,$$

and

$$\lambda_0^i A_0^i + \sum_{j=1}^m \lambda_j^i A_j^i + \sum_{l=1}^{p_i} z_l^i B_l^i \geq 0, i = 1, \dots, s.$$

This says that $\lambda_0^i \geq 0, (\lambda_1^i, \dots, \lambda_m^i) \in \mathbb{R}^m, z_l^i \in \mathbb{R}, \nu \in \mathbb{R}$ is feasible for (SDP). Thus,

$$\text{val}(P) = \nu \leq \text{val}(SDP),$$

and hence $\text{val}(P) = \text{val}(SDP)$. The proof is complete. □

Remark 4.1 (Special Cases: Min-max programs involving SOS-convex polynomials)

In the case where the objective function f_0 can be expressed as a finite maximum of SOS-convex polynomials and the constraint functions $f_i, i = 1, \dots, s$, are SOS-convex polynomials, Theorem 4.1 has established in [20, Theorem 3.1] for min-max programs.

In the preceding theorem, we see that the optimal value of a SOS-convex semialgebraic optimization problem (P) can be found by solving a single semidefinite programming problem, that is, its SDP relaxation problem (SDP). Next, we examine the important question that: how to recover an optimal solution of (P) from its SDP relaxation problem?

For a given $z = (z_\alpha) \in \mathbb{R}^{s(r,n)}$, we define a linear function $L_z : \mathbb{R}_r[x_1, \dots, x_n] \rightarrow \mathbb{R}$ by

$$L_z(u) = \sum_{\alpha=1}^{s(r,n)} u_\alpha z_\alpha \text{ with } u(x) = \sum_{\alpha=1}^{s(r,n)} u_\alpha x_\alpha^{(r)}. \tag{9}$$

For each $\alpha = 1, \dots, s(2r, n)$, define M_α to be the $(s(r, n) \times s(r, n))$ symmetric matrix such that

$$\text{Tr}(M_\alpha W) = \sum_{\substack{1 \leq \beta, \gamma \leq s(r,n) \\ i(\beta) + i(\gamma) = i(\alpha)}} W_{\beta, \gamma} \text{ for all } W \in S^{s(r,n)}.$$

Then, for $z = (z_\alpha) \in \mathbb{R}^{s(2r,n)}$, the moment matrix with respect to the sequence $z = (z_\alpha)$ with degree r is denoted by $\mathbf{M}_r(z)$, and is defined by

$$\mathbf{M}_r(z) = \sum_{1 \leq \alpha \leq s(2r,n)} z_\alpha M_\alpha.$$

As a simple illustration, let $r = 4$ and $n = 1$, for $z = (z_1, \dots, z_5)^T \in \mathbb{R}^{s(4,1)} = \mathbb{R}^5$,

$$L_z(u) = \sum_{i=1}^5 \alpha_i z_i, \text{ for all } u(x) = \alpha_1 + \alpha_2 x + \alpha_3 x^2 + \alpha_4 x^3 + \alpha_5 x^4.$$

Moreover, for $r = 1, n = 2$ and $z \in \mathbb{R}^{s(2,2)} = \mathbb{R}^6$

$$M_1(z) = \begin{pmatrix} z_1 & z_2 & z_3 \\ z_2 & z_4 & z_5 \\ z_3 & z_5 & z_6 \end{pmatrix}.$$

Recall that (SDP) can be equivalently rewritten as a semidefinite programming problem given as in (4). The Lagrangian dual of this semidefinite programming reformulation is

$$\begin{aligned} & \inf_{\substack{y=(y_\alpha) \in \mathbb{R}^{s(d,n)} \\ Z_i \geq 0}} \sup_{\substack{\lambda_0^i \geq 0, (\lambda_1^i, \dots, \lambda_m^i) \in \mathbb{R}^m \\ z_j^i \in \mathbb{R}, \mu \in \mathbb{R}, W \geq 0}} \left\{ \mu + y_1 \left((h_0^0)_1 + \sum_{j=1}^m \lambda_j^0 (h_j^0)_1 \right) \right. \\ & + \sum_{i=1}^s \left(\lambda_0^i (h_0^i)_1 + \sum_{j=1}^m \lambda_j^i (h_j^i)_1 \right) - \mu - W_{1,1} \Big) \\ & + \sum_{2 \leq \alpha \leq s(d,n)} y_\alpha \left((h_0^0)_\alpha + \sum_{j=1}^m \lambda_j^0 (h_j^0)_\alpha + \sum_{i=1}^s \left(\lambda_0^i (h_0^i)_\alpha + \sum_{j=1}^m \lambda_j^i (h_j^i)_\alpha \right) - \sum_{\substack{1 \leq \beta, \gamma \leq s(d/2,n) \\ i(\beta) + i(\gamma) = i(\alpha)}} W_{\beta, \gamma} \right) \\ & \left. + \text{Tr}(Z_0(A_0^0 + \sum_{j=1}^m \lambda_j^0 A_j^0 + \sum_{l=1}^{p_0} z_l^0 B_l^0)) + \sum_{i=1}^s \text{Tr}(Z_i(\lambda_0^i A_0^i + \sum_{j=1}^m \lambda_j^i A_j^i + \sum_{l=1}^{p_i} z_l^i B_l^i)) \right\}, \end{aligned}$$

which can be further simplified as

$$\begin{aligned} (SDP^*) \quad & \inf_{y=(y_\alpha) \in \mathbb{R}^{s(d,n)}, Z_i \geq 0} \sum_{1 \leq \alpha \leq s(d,n)} (h_0^0)_\alpha y_\alpha + \text{Tr}(Z_0 A_0^0) \\ & \text{s.t.} \quad \sum_{1 \leq \alpha \leq s(d,n)} (h_0^i)_\alpha y_\alpha + \text{Tr}(Z_i A_0^i) \leq 0, \quad i = 1, \dots, s, \\ & \quad \sum_{1 \leq \alpha \leq s(d,n)} (h_j^i)_\alpha y_\alpha + \text{Tr}(Z_i A_j^i) = 0, \quad i = 0, 1, \dots, s, \quad j = 1, \dots, m, \\ & \quad \text{Tr}(Z_i B_l^i) = 0, \quad i = 0, 1, \dots, s, \quad l = 1, \dots, p_i, \\ & \quad M_{\frac{d}{2}}(y) = \sum_{1 \leq \alpha \leq s(d,n)} y_\alpha M_\alpha \geq 0, \\ & \quad y_1 = 1. \end{aligned}$$

We note that the problem (SDP*) is also a semidefinite programming problem, and hence can be efficiently solved as well.

Next, we recall the following Jensen’s inequality for SOS-convex polynomial (cf [21]) which will play an important role in our later analysis.

Lemma 4.2 (Jensen’s inequality for SOS-convex polynomial [21, Theorem 5.13]) *Let f be an SOS-convex polynomial on \mathbb{R}^n with degree $2r$. Let $y \in \mathbb{R}^{s(2r,n)}$ with $y_1 = 1$ and $M_r(y) \geq 0$. Then, we have*

$$L_y(f) \geq f(L_y(X_1), \dots, L_y(X_n)),$$

where L_y is given as in (9) and X_i denotes the polynomial which maps a vector in \mathbb{R}^n to its i th coordinate.

The next main result of this section is the following theorem, providing a way to recover a solution of problem (P) from a solution of its SDP relaxation.

Theorem 4.2 (Recovery of the solution) For problem (P), suppose that the following strict feasibility conditions hold:

- (i) there exists $\bar{x} \in \mathbb{R}^n$ such that $f_i(\bar{x}) < 0$ for all $i = 1, \dots, s$;
- (ii) for each $i = 0, 1, \dots, s$, there exist $\bar{y}^i \in \mathbb{R}^m$ and $\bar{z}^i \in \mathbb{R}^{p_i}$ such that

$$A_0^i + \sum_{j=1}^m \bar{y}_j^i A_j^i + \sum_{l=1}^{p_i} \bar{z}_l^i B_l^i \succ 0.$$

Consider the SDP relaxation problem (SDP) and its Lagrangian dual problem (SDP*). Let $(y^*, Z_0^*, Z_1^*, \dots, Z_s^*)$ be an optimal solution for (SDP*) and let $x^* := (L_{y^*}(X_1), \dots, L_{y^*}(X_n))^T \in \mathbb{R}^n$ where X_i denotes the polynomial which maps a vector $x \in \mathbb{R}^n$ to its i th coordinate. Then, x^* is an optimal solution for (P).

Proof From condition (i), the exact SDP relaxation result (Theorem 4.1) gives us that $\text{val}(P) = \text{val}(\text{SDP})$. Note that (SDP) and (SDP*) are dual problems to each other. The usual weak duality for semidefinite programming implies that $\text{val}(\text{SDP}^*) \geq \text{val}(\text{SDP}) = \text{val}(P)$. Next, we establish that $\text{val}(\text{SDP}^*) = \text{val}(P)$, where $\text{val}(\text{SDP}^*)$ is the optimal value of problem (SDP*). To see this, let x be a feasible point of (P) and let $\rho := f_0(x)$. Then

$$f_0(x) = \max_{y^0 \in \Omega_0} \{h_0^0(x) + \sum_{j=1}^m y_j^0 h_j^0(x)\} = \rho$$

and

$$f_i(x) = \max_{y^i \in \Omega_i} \{h_0^i(x) + \sum_{j=1}^m y_j^i h_j^i(x)\} \leq 0, \quad i = 1, \dots, s,$$

where $\Omega_i, i = 0, 1, \dots, s$ are compact sets given by $\Omega_i = \{y^i := (y_1^i, \dots, y_m^i) \in \mathbb{R}^m : \exists z^i = (z_1^i, \dots, z_{p_i}^i) \in \mathbb{R}^{p_i}$ such that $A_0^i + \sum_{j=1}^m y_j^i A_j^i + \sum_{l=1}^{p_i} z_l^i B_l^i \geq 0\}$. This shows that

$$(y^0, z^0) \in \mathbb{R}^m \times \mathbb{R}^{p_0}, \quad A_0^0 + \sum_{j=1}^m y_j^0 A_j^0 + \sum_{l=1}^{p_0} z_l^0 B_l^0 \geq 0 \Rightarrow h_0^0(x) + \sum_{j=1}^m y_j^0 h_j^0(x) \leq \rho,$$

and

$$(y^i, z^i) \in \mathbb{R}^m \times \mathbb{R}^{p_i}, \quad A_0^i + \sum_{j=1}^m y_j^i A_j^i + \sum_{l=1}^{p_i} z_l^i B_l^i \geq 0 \Rightarrow h_0^i(x) + \sum_{j=1}^m y_j^i h_j^i(x) \leq 0, \quad i = 1, \dots, s.$$

It then follows from condition (ii) and the strong duality theorem for semidefinite programming that there exist $Z_i \geq 0, i = 0, 1, \dots, s$ such that

$$\left\{ \begin{array}{l} h_0^0(x) + \text{Tr}(Z_0 A_0^0) \leq \rho, \\ h_0^i(x) + \text{Tr}(Z_i A_0^i) \leq 0, \quad i = 1, \dots, s, \\ h_j^i(x) + \text{Tr}(Z_i A_j^i) = 0, \quad i = 0, 1, \dots, s, \quad j = 1, \dots, m, \\ \text{Tr}(Z_i B_l^i) = 0, \quad i = 0, 1, \dots, s, \quad l = 1, \dots, p_i, \end{array} \right.$$

Let $x^{(d)} = (1, x_1, x_2, \dots, x_n, x_1^2, x_1x_2, \dots, x_2^2, \dots, x_n^2, \dots, x_1^d, \dots, x_n^d)^T$. Then, $(x^{(d)}, Z_0, Z_1, \dots, Z_s)$ is feasible for (SDP^*) and

$$f_0(x) = \rho \geq h_0^0(x) + \text{Tr}(Z_0A_0^0) = \sum_{1 \leq \alpha \leq s(d,n)} (h_0^0)_\alpha x_\alpha^{(d)} + \text{Tr}(Z_0A_0^0).$$

This shows that $\text{val}(P) \geq \text{val}(SDP^*)$, and hence $\text{val}(P) = \text{val}(SDP^*)$. In particular, we have $\text{val}(P) = \text{val}(SDP) = \text{val}(SDP^*)$.

Now, let $(y^*, Z_0^*, Z_1^*, \dots, Z_s^*)$ be an optimal solution for (SDP^*) . Then, $Z_i^* \geq 0, i = 0, 1, \dots, s$, and

$$\begin{aligned} \sum_{1 \leq \alpha \leq s(d,n)} (h_0^i)_\alpha y_\alpha^* + \text{Tr}(Z_i^*A_0^i) &\leq 0, \quad i = 1, \dots, s, \\ \sum_{1 \leq \alpha \leq s(d,n)} (h_j^i)_\alpha y_\alpha^* + \text{Tr}(Z_i^*A_j^i) &= 0, \quad i = 0, 1, \dots, s, \quad j = 1, \dots, m, \\ \text{Tr}(Z_i^*B_l^i) &= 0, \quad i = 0, 1, \dots, s, \quad l = 1, \dots, p_i, \\ \mathbf{M}_{\frac{d}{2}}(y^*) &= \sum_{1 \leq \alpha \leq s(d,n)} y_\alpha^* M_\alpha \geq 0, \\ y_1^* &= 1. \end{aligned}$$

Note that for each $(y_1^0, \dots, y_m^0) \in \Omega_0$, one can find $z^0 = (z_1^0, \dots, z_{p_0}^0) \in \mathbb{R}^{p_0}$ such that

$$A_0^0 + \sum_{j=1}^m y_j^0 A_j^0 + \sum_{l=1}^{p_0} z_l^0 B_l^0 \geq 0.$$

So, for each $(y_1^0, \dots, y_m^0) \in \Omega_0$, it holds that

$$\begin{aligned} \sum_{1 \leq \alpha \leq s(d,n)} (h_0^0)_\alpha y_\alpha^* + \text{Tr}(Z_0^*A_0^0) &\geq \sum_{1 \leq \alpha \leq s(d,n)} (h_0^0)_\alpha y_\alpha^* - \text{Tr}(Z_0^* (\sum_{j=1}^m y_j^0 A_j^0 + \sum_{l=1}^{p_0} z_l^0 B_l^0)) \\ &= \sum_{1 \leq \alpha \leq s(d,n)} (h_0^0)_\alpha y_\alpha^* - \sum_{j=1}^m y_j^0 \text{Tr}(Z_0^*A_j^0) \\ &= \sum_{1 \leq \alpha \leq s(d,n)} (h_0^0)_\alpha y_\alpha^* + \sum_{j=1}^m y_j^0 \sum_{1 \leq \alpha \leq s(d,n)} (h_j^0)_\alpha y_\alpha^* \\ &= L_{y^*}(h_0^0 + \sum_{j=1}^m y_j^0 h_j^0). \end{aligned} \tag{10}$$

Since $h_0^0 + \sum_{j=1}^m y_j^0 h_j^0$ is SOS-convex, $\mathbf{M}_{\frac{d}{2}}(y^*) \geq 0$, and $y_1^* = 1$, by Lemma 4.2, we have

$$L_{y^*} \left(h_0^0 + \sum_{j=1}^m y_j^0 h_j^0 \right) \geq \left(h_0^0 + \sum_{j=1}^m y_j^0 h_j^0 \right) (L_{y^*}(X_1), \dots, L_{y^*}(X_n)) = h_0^0(x^*) + \sum_{j=1}^m y_j^0 h_j^0(x^*) \tag{11}$$

for every $(y_1^0, \dots, y_m^0) \in \Omega_0$. Taking supremum over all $(y_1^0, \dots, y_m^0) \in \Omega_0$ in (1) and using (2), it follows that

$$f_0(x^*) \leq \sum_{1 \leq \alpha \leq s(d,n)} (h_0^0)_\alpha y_\alpha^* + \text{Tr}(Z_0^*A_0^0).$$

Taking into account that $(y^*, Z_0^*, Z_1^*, \dots, Z_s^*)$ is an optimal solution for (SDP^*) , we get

$$\text{val}(SDP^*) = \sum_{1 \leq \alpha \leq s(d,n)} (h_0^0)_\alpha y_\alpha^* + \text{Tr}(Z_0^*A_0^0) \geq f_0(x^*).$$

We claim that x^* is feasible for (P). Granting this, we have

$$\text{val}(SDP^*) \geq f_0(x^*) \geq \text{val}(P) = \text{val}(SDP^*).$$

This forces that $f_0(x^*) = \text{val}(P)$, and so, x^* is an optimal solution for (P).

We now verify our claim. Take any $i = 1, \dots, s$ and $(y_1^i, \dots, y_m^i) \in \Omega_i$. Then one can find $z^i = (z_1^i, \dots, z_{p_i}^i) \in \mathbb{R}^{p_i}$ such that

$$A_0^i + \sum_{j=1}^m y_j^i A_j^i + \sum_{l=1}^{p_i} z_l^i B_l^i \geq 0.$$

Arguing as before, we arrive at

$$f_i(x^*) \leq \sum_{1 \leq \alpha \leq s(d,n)} (h_0^i)_\alpha y_\alpha^* + \text{Tr}(Z_i^* A_0^i) \leq 0.$$

This shows that x^* is feasible for (P). So, the conclusion follows. □

Finally, we illustrate how to find the optimal value and an optimal solution for an SOS-convex semialgebraic program by solving a single semidefinite programming problem.

Example 4.1 (Illustrative example) Consider the following simple 2-dimensional nonsmooth convex optimization problem:

$$\begin{aligned} (EP) \quad & \min x_1^4 - x_2 \\ & \text{s.t. } x_1^2 + x_2^2 + 2\|(x_1, x_2)\| - 1 \leq 0. \end{aligned}$$

Let

$$\Omega_1 = \{(y_1^1, y_2^1) : (y_1^1)^2 + (y_2^1)^2 \leq 1\} = \{(y_1^1, y_2^1) : \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + y_1^1 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + y_2^1 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \geq 0\}.$$

Let $h_0^1(x) = x_1^2 + x_2^2 - 1$ and $h_j^1(x) = 2x_j$, $j = 1, 2$. We first observe that, for each $(y_1^1, y_2^1) \in \Omega_1$, the function $h_0^1 + \sum_{j=1}^2 y_j^1 h_j^1$ is an SOS-convex polynomial. Denote $f_0(x) = x_1^4 - x_2$ and $f_1(x) = x_1^2 + x_2^2 + 2\|(x_1, x_2)\| - 1$. Then, $f_1(x) = \max_{(y_1^1, y_2^1) \in \Omega_1} \{h_0^1(x) + y_1^1 h_1^1(x) + y_2^1 h_2^1(x)\}$, and so, f_1 is an SOS-convex semialgebraic function. Obviously, f_0 is an SOS-convex polynomial and thus is also an SOS-convex semialgebraic function. This shows that (EP) is an SOS-convex semialgebraic program.

Let $x_0 = (0, 0)$. It can be verified that $f_1(x_0) = -1 < 0$. Thus, Theorem 4.1 implies that $\text{val}(EP) = \text{val}(ESDP)$ where (ESDP) is given by

$$\begin{aligned} (ESDP) \quad & \sup_{\lambda_0^1 \geq 0, \lambda_j^1 \in \mathbb{R}, \mu \in \mathbb{R}} \{ \mu : f_0 + \left(\lambda_0^1 h_0^1 + \sum_{j=1}^2 \lambda_j^1 h_j^1 \right) - \mu \in \Sigma_4^2[x], \\ & \lambda_0^1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \lambda_1^1 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + \lambda_2^1 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \geq 0 \}. \end{aligned}$$

Note that

$$\begin{aligned}
 & f_0 + \left(\lambda_0^1 h_1^0 + \sum_{j=1}^2 \lambda_j^1 h_j^1 \right) - \mu \in \Sigma_4^2[x] \\
 \Leftrightarrow & x_1^4 - x_2 + \lambda_0^1(x_1^2 + x_2^2 - 1) + 2\lambda_1^1 x_1 + 2\lambda_2^1 x_2 - \mu \\
 & = \left(1, x_1, x_2, x_1^2, x_1 x_2, x_2^2 \right) \begin{pmatrix} W_{11} & W_{12} & W_{13} & W_{14} & W_{15} & W_{16} \\ W_{12} & W_{22} & W_{23} & W_{24} & W_{25} & W_{26} \\ W_{13} & W_{23} & W_{33} & W_{34} & W_{35} & W_{36} \\ W_{14} & W_{24} & W_{34} & W_{44} & W_{45} & W_{46} \\ W_{15} & W_{25} & W_{35} & W_{45} & W_{55} & W_{56} \\ W_{16} & W_{26} & W_{36} & W_{46} & W_{56} & W_{66} \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1 x_2 \\ x_2^2 \end{pmatrix}, W = (W_{ij}) \in S_+^6, \\
 \Leftrightarrow & W_{11} = -\lambda_0^1 - \mu, W_{12} = \lambda_1^1, 2W_{14} + W_{22} = \lambda_0^1, \\
 & W_{33} + 2W_{16} = \lambda_0^1, 2W_{13} = -1 + 2\lambda_2^1, W_{44} = 1, \\
 & W_{23} = W_{24} = W_{34} = 0, W_{i5} = 0, i = 1, \dots, 5, W_{i6} = 0, i = 2, \dots, 6, \\
 & W = (W_{ij}) \in S_+^6.
 \end{aligned}$$

Thus, (ESDP) can be equivalently rewritten as the following semidefinite programming problem:

$$\begin{aligned}
 \sup_{\lambda_0^1 \geq 0, \lambda_j^1 \in \mathbb{R}, \mu \in \mathbb{R}, W \in S_+^6} \{ \mu : & W_{11} = -\lambda_0^1 - \mu, W_{12} = \lambda_1^1, 2W_{14} + W_{22} = \lambda_0^1, \\
 & W_{33} + 2W_{16} = \lambda_0^1, 2W_{13} = -1 + 2\lambda_2^1, W_{44} = 1, \\
 & W_{23} = W_{24} = W_{34} = 0, W_{i5} = 0, i = 1, \dots, 5, W_{i6} = 0, i = 2, \dots, 6, \\
 & W = (W_{ij}) \in S_+^6 \\
 & \lambda_0^1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \lambda_1^1 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + \lambda_2^1 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \geq 0 \}.
 \end{aligned}$$

Solving this semidefinite programming problem using CVX [10, 14], we obtain the optimal value $\text{val}(ESDP) = -0.414214 \approx 1 - \sqrt{2}$ and the dual variable $y^* = (y_1^*, \dots, y_{15}^*) \in \mathbb{R}^{15} = \mathbb{R}^{s(4,2)}$ with $y_1^* = 1, y_2^* = 0$ and $y_3^* = 0.414214 \approx \sqrt{2} - 1$. It can be verified that the conditions in Theorem 4.2 are satisfied. So, Theorem 4.2 implies that $x^* = (L_{y^*}(X_1), L_{y^*}(X_2)) = (y_2^*, y_3^*) = (0, \sqrt{2} - 1)$ is a solution for (EP).

Indeed, the optimality of $(0, \sqrt{2} - 1)$ for (EP) can be verified independently. To see this, note that for all (x_1, x_2) which is feasible for (EP), one has

$$x_1^2 + x_2^2 + 2\|(x_1, x_2)\| - 1 \leq 0.$$

In particular,

$$|x_2|^2 + 2|x_2| - 1 = x_2^2 + 2|x_2| - 1 \leq 0,$$

which implies that $|x_2| \leq \sqrt{2} - 1$. Thus, for all feasible point (x_1, x_2) for (EP), $x_1^4 - x_2 \geq -x_2 \geq -|x_2| \geq 1 - \sqrt{2}$, and so, $\text{val}(EP) \geq 1 - \sqrt{2}$. On the other hand, direct verification

shows that $(0, \sqrt{2} - 1)$ is feasible for (EP) with the object value $1 - \sqrt{2}$. So, $\text{val}(EP) = 1 - \sqrt{2}$ and $(0, \sqrt{2} - 1)$ is a solution of the problem (EP).

5 Applications to Robust Optimization

In this section, we show that how our results can be directly applied to the area of robust optimization [4, 8, 9] (for some recent development see [11–13, 16, 17, 23, 25]). Consider the following robust SOS-convex optimization problem under constraint data uncertainty

$$(RP) \min f(x)$$

$$\text{s.t. } g_i^{(0)}(x) + \sum_{j=1}^{t_i} u_i^{(j)} g_i^{(j)}(x) + \sum_{j=t_i+1}^s u_i^{(j)} g_i^{(j)}(x) \leq 0, \quad \forall u_i \in \mathcal{U}_i, \quad i = 1, \dots, m,$$

where $f, g_i^{(j)}, i = 1, \dots, m, j = 0, 1, \dots, t_i,$ are SOS-convex polynomials with degree $d, g_i^{(j)}, i = 1, \dots, m, j = t_i + 1, \dots, s,$ are affine functions, and u_i are uncertain parameters belonging to uncertainty sets $\mathcal{U}_i, i = 1, \dots, m.$ For simplicity, we only discuss the cases where the objective function does not affected by the the data uncertainty. Indeed, by introducing an additional variable, we can always assume the objective function is free of uncertainty (see [4, 16]).

In the case where \mathcal{U}_i is the so-called restricted ellipsoidal uncertainty set given by

$$\mathcal{U}_i^e = \{(u_i^1, \dots, u_i^{t_i}, u_i^{t_i+1}, \dots, u_i^s) : \|(u_i^1, \dots, u_i^{t_i})\| \leq 1, u_i^j \geq 0, j = 1, \dots, t_i\}$$

$$\|(u_i^{t_i+1}, \dots, u_i^s)\| \leq 1\},$$

this robust optimization problem was first examined in [13] in the special case of robust convex quadratic optimization problems, and then subsequently in [19] for general robust SOS-convex polynomial optimization problems. In particular, [19] showed that the optimal value of (RP) with $\mathcal{U}_i = \mathcal{U}_i^e$ can be found by solving a related semidefinite programming problem (SDP) and raised an open question that how to found an optimal solution of (RP) from the corresponding related SDP.

As we will see, as a simple application of the result in Section 4, we can extend the exact semidefinite programming relaxations result in [19] to a more general setting and answer the open question left in [19] on how to recover a robust solution from the semidefinite programming relaxation in this broader setting. To do this, we first introduce the notion of restricted spectrahedron data uncertainty set which is a convex compact set given by

$$\mathcal{U}_i^{res} = \{(u_i^{(1)}, \dots, u_i^{(t_i)}, u_i^{(t_i+1)}, \dots, u_i^{(s)}) \in \mathbb{R}^s : A_i^0 + \sum_{j=1}^s u_i^{(j)} A_i^j \geq 0,$$

$$(u_i^{(1)}, \dots, u_i^{(t_i)}) \in \mathbb{R}_+^{t_i}, (u_i^{(t_i+1)}, \dots, u_i^{(s)}) \in \mathbb{R}^{s-t_i}\}. \quad (12)$$

It is not hard to see that the restricted ellipsoidal uncertainty set is a special case of the restricted spectrahedron data uncertainty set as the norm constraint can be expressed as a linear matrix inequality.

5.1 Robust SOS-Convex Optimization with Restricted Spectrahedron Data Uncertainty

Now, let us consider the robust SOS-convex optimization problem with restricted spectrahedron data uncertainty set, that is, problem (RP) with $\mathcal{U}_i = \mathcal{U}_i^{res}$. We first see that this robust optimization problem can be regarded as an SOS-convex semialgebraic program.

Indeed, let $f_0(x) = f(x)$, $g_i(x, u_i) = g_i^{(0)}(x) + \sum_{j=1}^{t_i} u_i^{(j)} g_i^{(j)}(x) + \sum_{j=t_i+1}^s u_i^{(j)} g_i^{(j)}(x)$ and $f_i(x) = \max_{u_i \in \mathcal{U}_i^{res}} \{g_i(x, u_i)\}$, $i = 1, \dots, m$. From the construction of the restricted spectrahedron data uncertainty, for each $u_i \in \mathcal{U}_i^{res}$, $g_i(\cdot, u_i)$ is an SOS-convex polynomial. Moreover, each restricted spectrahedron data uncertainty set \mathcal{U}_i^{res} can be written as

$$\mathcal{U}_i^{res} = \{(u_i^{(1)}, \dots, u_i^{(t_i)}, u_i^{(t_i+1)}, \dots, u_i^{(s)}) \in \mathbb{R}^s : \tilde{A}_i^0 + \sum_{j=1}^s u_i^{(j)} \tilde{A}_i^j \geq 0\},$$

where

$$\tilde{A}_i^0 = \begin{pmatrix} 0_{t_i \times t_i} & 0 \\ 0 & A_i^0 \end{pmatrix}, \tilde{A}_i^j = \begin{pmatrix} \text{diag } e^j & 0 \\ 0 & A_i^j \end{pmatrix}, j = 1, \dots, t_i, \text{ and } \tilde{A}_i^j = \begin{pmatrix} 0_{t_i \times t_i} & 0 \\ 0 & A_i^j \end{pmatrix}, j = t_i + 1, \dots, s. \tag{13}$$

Here, $e^j \in \mathbb{R}^{t_i}$ denotes the vector whose j th element equals to one and 0 otherwise. Therefore, the robust optimization problem (RP) with $\mathcal{U}_i = \mathcal{U}_i^{res}$ can be cast as an SOS-convex semialgebraic program, and hence, one can associate this robust optimization problem a relaxation problem

$$\begin{aligned} (\widetilde{SDP}) \quad & \sup_{\lambda_i^0 \geq 0, \lambda_i^j \in \mathbb{R}, \mu \in \mathbb{R}} \{ \mu : f + \sum_{i=1}^m \left(\lambda_i^0 g_i^{(0)} + \sum_{j=1}^s \lambda_i^j g_i^{(j)} \right) - \mu \in \Sigma_d^2[x], \\ \text{s.t.} \quad & \lambda_i^0 A_i^0 + \sum_{j=1}^s \lambda_i^j A_i^j \geq 0, (\lambda_i^1, \dots, \lambda_i^{t_i}) \in \mathbb{R}_+^{t_i}, i = 1, \dots, m \}. \end{aligned}$$

Similar as in the previous section, we note that (\widetilde{SDP}) can be equivalently rewritten as the following semidefinite programming problem:

$$\begin{aligned} & \sup_{\lambda_i^0 \geq 0, \lambda_i^j \in \mathbb{R}, \mu \in \mathbb{R}, W \in S^{s(d/2, n)}} \{ \mu : f_1 + \sum_{i=1}^m \left(\lambda_i^0 (g_i^{(0)})_1 + \sum_{j=1}^s \lambda_i^j (g_i^{(j)})_1 \right) - \mu = W_{1,1}, \\ \text{s.t.} \quad & f_\alpha + \sum_{i=1}^m \left(\lambda_i^0 (g_i^{(0)})_\alpha + \sum_{j=1}^s \lambda_i^j (g_i^{(j)})_\alpha \right) = \sum_{\substack{1 \leq \beta, \gamma \leq s(d/2, n) \\ i(\beta) + i(\gamma) = i(\alpha)}} W_{\beta, \gamma}, \quad 2 \leq \alpha \leq s(d, n), \\ & \lambda_i^0 \tilde{A}_i^0 + \sum_{j=1}^s \lambda_i^j \tilde{A}_i^j \geq 0, W \geq 0, i = 1, \dots, m \}. \end{aligned} \tag{14}$$

The usual Lagrangian dual problem of the above semidefinite programming reformulation of (\widetilde{SDP}) is another semidefinite programming problem which can be expressed as

$$\begin{aligned}
 (\widetilde{SDP}^*) \quad & \inf_{y=(y_\alpha) \in \mathbb{R}^{s(d,n)}, Z_i \geq 0} \sum_{1 \leq \alpha \leq s(d,n)} f_\alpha y_\alpha \\
 \text{s.t.} \quad & \sum_{1 \leq \alpha \leq s(d,n)} (g_i^{(0)})_\alpha y_\alpha + \text{Tr}(Z_i \widetilde{A}_i^0) \leq 0, \quad i = 1, \dots, m, \\
 & \sum_{1 \leq \alpha \leq s(d,n)} (g_i^{(j)})_\alpha y_\alpha + \text{Tr}(Z_i \widetilde{A}_i^j) = 0, \quad i = 1, \dots, m, \quad j = 1, \dots, s, \\
 & \mathbf{M}_{\frac{d}{2}}(y) = \sum_{1 \leq \alpha \leq s(d,n)} y_\alpha M_\alpha \geq 0, \\
 & y_1 = 1.
 \end{aligned}$$

Now, we establish an exact SDP relaxation result for robust SOS-convex optimization problem with restricted spectrahedron data uncertainty set, and show how to recover a solution of the robust problem from its SDP relaxation problem. This extends the exact semidefinite programming relaxations result in [19] from restricted ellipsoidal data uncertainty to the more general setting of restricted spectrahedron data uncertainty, and answers the open question left in [19] on how to recover a robust solution from the semidefinite programming relaxation in this broader setting.

Proposition 5.1 *Consider (RP) under restricted spectrahedron data uncertainty, that is, $\mathcal{U}_i = \mathcal{U}_i^{res}$ where \mathcal{U}_i^{res} are convex compact sets given as in (12). Suppose that there exists $x_0 \in \mathbb{R}^n$ such that for all $i = 1, \dots, m$*

$$g_i^{(0)}(x_0) + \sum_{j=1}^s u_i^{(j)} g_i^{(j)}(x_0) < 0 \text{ for all } u_i \in \mathcal{U}_i^{res}.$$

Then, the following statements hold.

- (i) *It holds that $\text{val}(RP) = \text{val}(\widetilde{SDP})$.*
- (ii) *Suppose in addition that, for each $i = 1, \dots, m$, there exists $(\bar{u}_i^{(1)}, \dots, \bar{u}_i^{(s)}) \in \mathbb{R}^s$ such that $\bar{u}_i^{(j)} > 0$ for all $j = 1, \dots, t_i$, and*

$$A_i^0 + \sum_{j=1}^s \bar{u}_i^{(j)} A_i^j > 0. \tag{15}$$

Consider the SDP relaxation problem (\widetilde{SDP}) and its Lagrangian dual problem (\widetilde{SDP}^*) . Let (y^*, Z_i^*) be an optimal solution for (\widetilde{SDP}^*) and let $x^* := (L_{y^*}(X_1), \dots, L_{y^*}(X_n))^T \in \mathbb{R}^n$ where X_i denotes the polynomial which maps a vector to its i th coordinate. Then, x^* is a solution for (RP).

Proof of (i) Let $\tilde{A}_i^j, i = 1, \dots, m, j = 0, 1, \dots, s$ be defined as in (13). Then,

$$\mathcal{U}_i^{res} = \{(u_i^{(1)}, \dots, u_i^{(t_i)}, u_i^{(t_i+1)}, \dots, u_i^{(s)}) \in \mathbb{R}^s : \tilde{A}_i^0 + \sum_{j=1}^s u_i^{(j)} \tilde{A}_i^j \geq 0\}.$$

Now, as $g_i^{(j)}$, $j = 0, 1, \dots, t_i$, are SOS-convex polynomials and $g_i^{(j)}$, $j = t_i + 1, \dots, s$, are affine functions, $g_i^{(0)} + \sum_{j=1}^s u_i^{(j)} g_i^{(j)}$ is an SOS-convex polynomial for each $u_i \in \mathcal{U}_i^{res}$. Put

$$f_i(x) := \sup_{u_i} \{g_i^{(0)}(x) + \sum_{j=1}^s u_i^{(j)} g_i^{(j)}(x) : u_i := (u_i^{(1)}, \dots, u_i^{(s)}) \in \mathcal{U}_i^{res}\}.$$

We see that each f_i is an SOS-convex semialgebraic function, $i = 1, \dots, m$ and $f_i(x_0) < 0$ for all $i = 1, \dots, m$. Therefore, Theorem 4.1 implies that

$$\begin{aligned} \text{val}(RP) &= \sup_{\lambda_i^0 \geq 0, \lambda_i^j \in \mathbb{R}, \mu \in \mathbb{R}} \{ \mu : f + \sum_{i=1}^m \left(\lambda_i^0 g_i^{(0)} + \sum_{j=1}^s \lambda_i^j g_i^{(j)} \right) - \mu \in \Sigma_d^2, \\ &\quad \lambda_i^0 \tilde{A}_i^0 + \sum_{j=1}^s \lambda_i^j \tilde{A}_i^j \geq 0, i = 1, \dots, m \} \\ &= \sup_{\lambda_i^0 \geq 0, \lambda_i^j \in \mathbb{R}, \mu \in \mathbb{R}} \{ \mu : f + \sum_{i=1}^m \left(\lambda_i^0 g_i^{(0)} + \sum_{j=1}^s \lambda_i^j g_i^{(j)} \right) - \mu \in \Sigma_d^2, \\ &\quad \lambda_i^0 A_i^0 + \sum_{j=1}^s \lambda_i^j A_i^j \geq 0, (\lambda_i^1, \dots, \lambda_i^{t_i}) \in \mathbb{R}_+^{t_i}, i = 1, \dots, m \}. \end{aligned}$$

[Proof of (ii)] From the definition of \tilde{A}_i^j , $i = 1, \dots, m$, $j = 0, 1, \dots, s$, and our assumption (15), we see that, for each $i = 1, \dots, m$,

$$\tilde{A}_i^0 + \sum_{j=1}^s u_i^j \tilde{A}_i^j > 0.$$

Therefore, the assumption of Theorem 4.2 holds and the conclusion of (ii) follows from Theorem 4.2. □

References

1. Ahmadi, A.A., Parrilo, P.A.: A complete characterization of the gap between convexity and SOS-convexity. *SIAM J. Optim.* **23**, 811–833 (2013)
2. Belousov, E.G., Klatte, D.: A Frank-Wolfe type theorem for convex polynomial programs. *Comp. Optim. Appl.* **22**, 37–48 (2002)
3. Ben-Tal, A., Nemirovski, A.: Lectures on modern convex optimization. analysis, algorithms, and engineering applications. MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Programming Society (MPS), Philadelphia PA (2001)
4. Ben-Tal, A., El Ghaoui, L., Nemirovski, A.: Robust Optimization. Princeton U.P., Princeton (2009)
5. Bochnak, J., Coste, M., Roy, M.F.: Real Algebraic Geometry, vol. 36, p. x+430. Springer, Berlin (1998)
6. Borwein, J.M., Vanderwerff, J.D.: Convex Functions: Constructions, Characterizations and Counterexamples Encyclopedia of Mathematics and Its Applications, vol. 109. Cambridge University Press, Cambridge (2010)
7. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
8. Bertsimas, D., Sim, M.: The price of robustness. *Oper. Res.* **52**, 35–53 (2004)
9. Bertsimas, D., Brown, D.B., Caramanis, C.: Theory and applications of robust optimization. *SIAM Rev.* **53**(3), 464–501 (2011)
10. CVX Research, Inc.C.VX.: Matlab software for disciplined convex programming, version 2.0. <http://cvxr.com/cvx> (2011)

11. Dinh, D., Goberna, M.A., López, M.A., Volle, M.: A unifying approach to robust convex infinite optimization duality. to appear in *J. Optim. Theory & Appl.* <https://doi.org/10.1007/s10957-017-1136-x> (2017)
12. Goberna, M.A., Jeyakumar, V., Li, G., López, M.A.: Robust linear semi-infinite programming duality under uncertainty. *Math. Program.* **139**, 185–203 (2013)
13. Goldfarb, D., Iyengar, G.: Robust convex quadratically constrained programs. *Math. Program.* **97**, 495–515 (2003)
14. Grant, M., Boyd, S.: Graph implementations for nonsmooth convex programs, recent advances in learning and control. In: Blondel, V., Boyd, S., Kimura, H. (eds.) *Lecture notes in control and information sciences*, pp. 95–110. Springer (2008)
15. Helton, J.W., Nie, J.W.: Semidefinite representation of convex sets. *Math. Program.* **122**, 21–64 (2010)
16. Jeyakumar, V., Li, G.: Strong duality in robust convex programming: complete characterizations. *SIAM J. Optim.* **20**, 3384–3407 (2010)
17. Jeyakumar, V., Li, G.: A new class of alternative theorems for SOS-convex inequalities and robust optimization. *Appl. Anal.* **94**, 56–74 (2015)
18. Jeyakumar, V., Li, G.: Exact SDP relaxations for classes of nonlinear semidefinite programming problems. *Oper. Res. Letters* **40**, 529–536 (2012)
19. Jeyakumar, V., Li, G., Vicente-Pérez, J.: Robust SOS-convex polynomial optimization problems: exact SDP relaxations. *Optim. Letters* **9**, 1–18 (2015)
20. Jeyakumar, V., Vicente-Pérez, J.: Dual semidefinite programs without duality gaps for a class of convex minimax programs. *J. Optim. Theory Appl.* **162**, 735–753 (2014)
21. Lasserre, J.B.: *Moments, Positive Polynomials and their Applications*. Imperial College Press, London (2009)
22. Lasserre, J.B.: Convexity in semialgebraic geometry and polynomial optimization. *SIAM J. Optim.* **19**, 1995–2014 (2008)
23. Jeyakumar, V., Li, G.Y., Lee, G.M.: A robust von Neumann minimax theorem for zero-sum games under bounded payoff uncertainty. *Oper. Res. Letters* **39**(2), 109–114 (2011)
24. Lee, G.M., Pham, T.S.: Stability and genericity for semialgebraic compact programs. *J. Optim. Theory Appl.* **169**, 473–495 (2016)
25. Li, G., Jeyakumar, V., Lee, G.M.: Robust conjugate duality for convex optimization under uncertainty with application to data classification. *Nonlinear Anal.* **74**(6), 2327–2341 (2011)
26. Zou, H., Hastie, T.: Regularization and variable selection via the elastic net. *J. R. Statist. Soc. B* **67**, 310–320 (2005)