Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 16, No. 3, pp. 395–418, 2020

SEMIDEFINITE PROGRAMMING RELAXATIONS FOR LINEAR SEMI-INFINITE POLYNOMIAL PROGRAMMING

Feng Guo* and Xiaoxia Sun†

Abstract: This paper studies a class of so-called linear semi-infinite polynomial programming (LSIPP) problems. It is a subclass of linear semi-infinite programming problems whose constraint functions are polynomials in parameters and index sets are basic semialgebraic sets. We present a hierarchy of semidefinite programming (SDP) relaxations for LSIPP problems. Convergence rate analysis of the SDP relaxations is established based on some existing results. We show how to verify the compactness of feasible sets of LSIPP problems. In the end, we extend the SDP relaxation method to more general semi-infinite programming problems.

Key words: *linear semi-infinite programming, semidefinite programming relaxations, sum of squares, polynomial optimization*

Mathematics Subject Classification: *65K05, 90C22*

---

1 introduction

We consider the following *linear semi-infinite polynomial programming* (LSIPP) problem

$$(P) \qquad \begin{cases} p^* := \inf_{x\in\mathbb{R}^m} \ c^T x \\ \qquad\qquad \text{s.t. } a(y)^T x + b(y) \geq 0, \quad \forall y \in S \subseteq \mathbb{R}^n, \end{cases} \tag{1.1}$$

where $c \in \mathbb{R}^m$, $b(Y) \in \mathbb{R}[Y] := \mathbb{R}[Y_1, \dots, Y_n]$ the polynomial ring in Y over the real field, $a(Y) = (a_1(Y), \dots, a_m(Y))^T \in \mathbb{R}[Y]^m$, and the *index set* S is a basic semialgebraic set defined by

$$S := \{y \in \mathbb{R}^n \mid g_1(y) \geq 0, \dots, g_s(y) \geq 0\}, \tag{1.2}$$

where $g_j(Y) \in \mathbb{R}[Y]$, $j = 1, \dots, s$. Lowercase letters (e.g. x, y, w) are hereinafter used for denoting points in a space while uppercase letters (e.g. X, Y, W) for the corresponding variables. In this paper, we assume that (1.1) is feasible and bounded from below, i.e., $-\infty < p^* < \infty$. Note that the problem (1.1) is NP-hard. Indeed, it is obvious that the problem of minimizing a polynomial $f(Y) \in \mathbb{R}[Y]$ over S can be regarded as a special LSIPP

*Feng Guo is supported by the Chinese National Natural Science Foundation under grants 11401074, 11571350.

†Xiaoxia Sun is supported by the Chinese National Natural Science Foundation under grant 11801064, the Foundation of Liaoning Education Committee under grant LN2017QN043.

problem (see Example 3.20). As is well known, the polynomial optimization problem is NP-hard even when $n > 1$, $f(Y)$ is a nonconvex quadratic polynomial and $g_j(Y)$'s are linear [36]. Hence, a general LSIPP problem cannot be expected to be solved in polynomial time unless P=NP.

LSIPP can be seen as a special branch of *linear semi-infinite programming* (LSIP), or more general, of *semi-infinite programming* (SIP), in which the involved functions are not necessarily polynomials. Numerically, SIP problems can be solved by different approaches including, for instance, discretization methods, local reduction methods, exchange methods, simplex-like methods and so on. See the surveys [10, 11, 28] and the references therein for details. One of main difficulties in numerical treatment of general SIP problems is that the feasibility test of a point $\bar{u} \in \mathbb{R}^m$ is equivalent to *globally* solving the problem of minimizing the constraint function with fixed $\bar{u}$ over the index set, which is called the lower level subproblem. Typically, when solving SIP problems by existing methods in the literature, the main difficulty lies in solving the nonlinear lower level subproblems at each iteration.

LSIPP, as a special subclass of SIP, has many applications like minimax problems, functional approximation problems. However, to the best of our knowledge, few of the numerical methods mentioned above are specially designed by exploiting features of polynomial optimization problems. Parpas and Rustem [37] proposed a discretization-like method to solve minimax polynomial optimization problems, which can be reformulated as semi-infinite polynomial programming (SIPP) problems. Using polynomial approximation and an appropriate hierarchy of *semidefinite programming* (SDP) relaxations, Lasserre presented an algorithm to solve the generalized SIPP problems in [24]. Based on an exchange scheme, an SDP relaxation method for solving SIPP problems was proposed in [44]. By using representations of nonnegative polynomials in the univariate case, an SDP method was given in [46] for LSIPP problems (1.1) with S being closed intervals.

In this paper, we propose a hierarchy of SDP relaxations for LSIPP (1.1). The dual problem of LSIPP is a special case of the generalized moment problems (Section 2.2), which has been well investigated, see [2–4, 19, 20, 23, 34] and the references therein. Lasserre [23] proposed an SDP relaxation method for generalized moment problems based on Putinar's Positivstellensatz [39]. Although the SDP relaxations presented in this paper can be seen as the dual of Lasserre's relaxations for GPM, they are of their independent interest because of the following desirable features they enjoy. First, some (approximate) minimizers of (1.1) can be extracted by these SDP relaxations, which is very useful in some applications, like functional approximation problems (Example 3.8); Second, convergence rate of these SDP relaxations can be estimated (Section 3.2) by using the complexity analysis of Putinar's Positivstellensatz in [35]; Third, these SDP relaxations can be easily extended to more general semi-infinite programming problems (Section 4), like problems of the form (1.1) with semi-algebraic functions, or with s.o.s-convex objectives. As the feasible set of (1.1) is assumed to be compact in the convergence rate estimation, we also show that the compactness can be verified by computing a positive lower bound of the infima of several LSIPP problems. It can be done by the proposed SDP relaxations if the finite convergence happens (in particular, if S is a closed and bounded interval (Section 3.3)).

This paper is organized as follows. We introduce some notation and preliminaries in Section 2. SDP relaxations of LSIPP problems and the convergence rate analysis is given in Section 3, where we also discuss how to verify the compactness of feasible sets of LSIPP problems. In Section 4, we extend the SDP relaxation method to more general semi-infinite programming problems. Some conclusions are made in Section 5.

2 Notation and Preliminaries

Here is some notation used in this paper. The symbol $\mathbb{N}$ (resp., $\mathbb{R}$) denotes the set of nonnegative integers (resp., real numbers). For any $t \in \mathbb{R}$, $\lceil t \rceil$ denotes the smallest integer that is not smaller than t. For $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$, $\|y\|_2$ denotes the standard Euclidean norm of y. For $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, $\|\alpha\|_1 := \alpha_1 + \cdots + \alpha_n$. For $k \in \mathbb{N}$, denote $\mathbb{N}^n_k = \{\alpha \in \mathbb{N}^n \mid \|\alpha\|_1 \le k\}$. For $y \in \mathbb{R}^n$ and $\alpha \in \mathbb{N}^n$, y^α denotes $y_1^{\alpha_1} \cdots y_n^{\alpha_n}$. $\mathbb{R}[Y] = \mathbb{R}[Y_1, \cdots, Y_n]$ denotes the ring of polynomials in $(Y_1, \cdots, Y_n)$ with real coefficients. For $k \in \mathbb{N}$, denote by $\mathbb{R}[Y]_k$ the set of polynomials in $\mathbb{R}[Y]$ of total degree up to k. For a symmetric matrix Z, $Z \succeq 0 (\succ 0)$ means that Z is positive semidefinite (definite). For $m \in \mathbb{N}$, $\mathbb{R}^{m\times m}$ denotes the set of $m \times m$ real matrices and $\mathbb{S}^m_+ \subset \mathbb{R}^{m\times m}$ denotes its subset of positive semidefinite matrices. For two symmetric matrices A, B of the same size, $\langle A, B\rangle$ denotes the inner product of A and B.

2.1 Sums of squares and moments

We recall some background about *sums of squares* (s.o.s) of polynomials and the dual theory of *moment matrices*. For any $f(Y) \in \mathbb{R}[Y]_k$, let $\mathbf{f}$ denote its column vector of coefficients in the canonical monomial basis of $\mathbb{R}[Y]_k$. A polynomial $f(Y) \in \mathbb{R}[Y]$ is said to be a sum of squares of polynomials if it can be written as $f(Y) = \sum_{i=1}^t f_i(Y)^2$ for some $f_1(Y), \ldots, f_t(Y) \in \mathbb{R}[Y]$. The symbol $\Sigma^2[Y]$ denotes the set of polynomials that are s.o.s.

Let $G := \{g_1, \ldots, g_s\}$ be the set of polynomials that define the semialgebraic set S (1.2). We denote by

$$\mathcal{Q}(G) := \left\{ \sum_{j=0}^{s} \sigma_j g_j \;\middle|\; g_0 = 1,\ \sigma_j \in \Sigma^2[Y], j = 0, 1, \ldots, s \right\} \tag{2.1}$$

the *quadratic module* generated by G and denote by

$$\mathcal{Q}_k(G) := \left\{ \sum_{j=0}^{s} \sigma_j g_j \;\middle|\; g_0 = 1,\ \sigma_j \in \Sigma^2[Y], \deg(\sigma_j g_j) \le 2k, j = 0, 1, \ldots, s \right\} \tag{2.2}$$

its k-th *quadratic module*. It is clear that if $f \in \mathcal{Q}(G)$, then $f(y) \ge 0$ for any $y \in S$. However, the converse is not necessarily true, see Example 3.18. Note that checking whether $f \in \mathcal{Q}_k(G)$ for a fixed $k \in \mathbb{N}$ is an SDP feasibility problem [21].

For $k \in \mathbb{N}$, denote $s(k) := \binom{n+k}{n}$. Consider a finite sequence of real numbers $z := (z_\alpha)_{\alpha \in \mathbb{N}^n_{2k}} \in \mathbb{R}^{s(2k)}$ whose elements are indexed by n-tuples $\alpha \in \mathbb{N}^n_{2k}$. z is called a *truncated moment sequence* up to order $2k$ if there exists a Borel measure μ on $\mathbb{R}^n$ such that

$$z_\alpha = \int Y^\alpha \mathrm{d}\mu(y),\ \forall \alpha \in \mathbb{N}^n_{2k}. \tag{2.3}$$

In this case, we say that z has a *representing measure* μ. The associated k-th *moment matrix* is the matrix $M_k(z)$ indexed by $\mathbb{N}^n_k$, with (α, β)-th entry $z_{\alpha+\beta}$ for $\alpha, \beta \in \mathbb{N}^n_k$. Given a polynomial $f(Y) = \sum_\alpha f_\alpha Y^\alpha$, for $k \ge d_f := \lceil \deg(f)/2 \rceil$, the $(k - d_f)$-th *localizing moment matrix* $M_{k-d_f}(fz)$ is defined as the moment matrix of the *shifted vector* $((fz)_\alpha)_{\alpha \in \mathbb{N}^n_{2(k-d_f)}}$ with $(fz)_\alpha = \sum_\beta f_\beta z_{\alpha+\beta}$. $\mathscr{M}_{2k}$ denotes the space of all sequences $z = (z_\alpha)_{\alpha \in \mathbb{N}^n_{2k}} \in \mathbb{R}^{s(2k)}$ with order at most $2k$. For any $z \in \mathscr{M}_{2k}$, the corresponding Riesz functional $\mathscr{L}_z$ on $\mathbb{R}[Y]_{2k}$

is defined by

$$\mathscr{L}_z\left(\sum_\alpha q_\alpha Y_1^{\alpha_1}\cdots Y_n^{\alpha_n}\right) := \sum_\alpha q_\alpha z_\alpha, \quad \forall q(Y)\in\mathbb{R}[Y]_{2k}. \tag{2.4}$$

From the definition of the localizing moment matrix $M_{k-d_f}(fz)$, it is easy to check that

$$\mathbf{q}^T M_{k-d_f}(fz)\mathbf{q} = \mathscr{L}_z(f(Y)q(Y)^2), \quad \forall q(Y)\in\mathbb{R}[Y]_{k-d_f}. \tag{2.5}$$

Let $d_j := \lceil \deg(g_j)/2\rceil$ for each $j=1,\dots,s$. For any $v\in S$, let $\zeta_{2k,v} := [v^\alpha]_{\alpha\in\mathbb{N}^n_{2k}}$ be the *Zeta vector* of v up to degree $2k$, i.e.,

$$\zeta_{2k,v} = [1 \quad v_1 \quad \cdots \quad v_n \quad v_1^2 \quad v_1v_2 \quad \cdots \quad v_n^{2k}]. \tag{2.6}$$

Then, $M_k(\zeta_{2k,v})\succeq 0$ and $M_{k-d_j}(g_j\zeta_{2k,v})\succeq 0$ for $j=1,\dots,s$. In fact, let $g_0=1$, then for each $j=0,1,\dots,s$,

$$\mathbf{q}^T M_{k-d_j}(g_j\zeta_{2k,v})\mathbf{q} = \mathscr{L}_{\zeta_{2k,v}}(g_j(Y)q(Y)^2) = g_j(v)q(v)^2 \geq 0,\ \forall q(Y)\in\mathbb{R}[Y]_{k-d_j}. \tag{2.7}$$

Definition 2.1. We say that $\mathcal{Q}(G)$ is *Archimedean* if there exists $\psi\in\mathcal{Q}(G)$ such that the inequality $\psi(y)\geq 0$ defines a compact set in $\mathbb{R}^n$.

Note that the Archimedean property implies that S is compact but the converse is not necessarily true. However, for any compact set S we can always force the associated quadratic module to be Archimedean by adding a redundant constraint $M-\|y\|_2^2\geq 0$ in the description of S for sufficiently large M.

Theorem 2.2 ([39, Putinar's Positivstellensatz])**.** *Suppose that $\mathcal{Q}(G)$ is Archimedean.*

(i) *If a polynomial $f\in\mathbb{R}[Y]$ is positive on S, then $f\in\mathcal{Q}_k(G)$ for some $k\in\mathbb{N}$;*

(ii) *If $M_k(z)\succeq 0$ and $M_k(g_jz)\succeq 0$ for all $j=1,\dots,s$, and all $k=0,1,\dots,$ then $z=(z_\alpha)_{\alpha\in\mathbb{N}^n}\in\mathbb{R}^{\mathbb{N}^n}$ has a representing measure μ supported by S.*

2.2 Dual problems and GPM

The Lagrangian dual problem [17, 28, 41] of (1.1) is

$$\begin{cases} d^* := \sup\limits_{\mu\in M^+(S)} & -\displaystyle\int_S b(y)\mathrm{d}\mu(y) \\ \qquad\qquad \text{s.t.} & \displaystyle\int_S a_i(y)\mathrm{d}\mu(y) = c_i,\ i=1,\dots,m, \end{cases} \tag{2.8}$$

where $M^+(S)$ is the space of all nonnegative bounded regular Borel measure supported by S. The dual problem (2.8) is in fact a special case of the so-called generalized problems of moments (GPM), which is to maximize a linear function over a linear section of the moment cone. We refer the interested readers to [2, 3, 20] and the references therein for various methodologies and applications of GPM problems. For numerical treatment of GPM problems, see [4, 19] for some geometric approaches and [23, 34] for SDP relaxation methods for GPM problems with polynomial data.

Now we introduce the main idea of the SDP relaxation method for (2.8) proposed by Lasserre in [23]. Assume that S is compact, by Putinar's Positivstellensatz (part *(ii)* of

Theorem 2.2), a sequence $z = (z_\alpha)_{\alpha\in\mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}^n}$ has a representing measure μ supported by S if

$$M_k(z) \succeq 0, \quad M_k(g_j z) \succeq 0, \ j = 1, \dots, s, \ k = 0, 1, \dots. \tag{2.9}$$

Define

$$\begin{aligned} &d_j := \lceil \deg(g_j)/2 \rceil, \ d_S := \max\{1, d_1, \dots, d_s\}, \\ &d_P := \max\{d_S, \lceil \deg(a_1)/2 \rceil, \cdots, \lceil \deg(a_m)/2 \rceil, \lceil \deg(b)/2 \rceil\}. \end{aligned} \tag{2.10}$$

Let $k \geq d_P$ and $z = (z_\alpha)_{\alpha\in\mathbb{N}^n_{2k}} \in \mathbb{R}^{s(2k)}$, the k-th semidefinite relaxation of (2.8) is

$$\begin{cases} p_k^{\text{mom}} := \sup\limits_{z\in\mathbb{R}^{s(2k)}} \ -\sum\limits_{\alpha\in\mathbb{N}^n_{2k}} b_\alpha z_\alpha \\ \qquad\qquad \text{s.t.} \sum\limits_{\alpha\in\mathbb{N}^n_{2k}} a_{i,\alpha} z_\alpha = c_i, \quad i = 1, \dots, m, \\ \qquad\qquad\qquad M_k(z) \succeq 0, \ M_{k-d_j}(g_j z) \succeq 0, \ j = 1, \dots, s. \end{cases} \tag{2.11}$$

Under certain assumptions, Lasserre proved [23] that p_k^{mom} decreasingly converges to p^*. The SDP relaxations (2.11) can be easily implemented and solved by the software GloptiPoly [16] developed by Henrion, Lasserre and Löfberg.

Condition 2.1. An optimizer z^* of the k-th SDP relaxation (2.11) satisfies the *flat extension condition* when

$$\text{rank} M_{k-d_S}(z^*) = \text{rank} M_k(z^*). \tag{2.12}$$

Based on [7, Theorem 1.1], Lasserre [23, Theorem 2] showed that the finite convergence of (2.11) happens at order k if the flat extension condition holds.

3 SDP Relaxations of LSIPP

In this section, we present a hierarchy of SDP relaxations for LSIPP problems. These SDP relaxations can be seen as the dual of Lasserre's relaxations for GPM and enjoy several desirable features. For example, (approximate) mininizers can be extracted and the convergence rate can be estimated by using some existing results. We shall also see in Section 4 that these SDP relaxations can be easily extended to more general semi-infinite programming problems.

3.1 SDP relaxations of LSIPP problems

We assume that S in (1.1) is compact. For a given feasible point $x \in \mathbb{R}^m$ of the LSIPP problem (1.1), the constraint requires that the polynomial $a(Y)^T x + b(Y) \in \mathbb{R}[Y]$ is nonnegative on S. Since every polynomial in the quadratic module $\mathcal{Q}(G)$ of S generated by G is nonnegative on S, we can relax the problem (1.1) as follows

$$p^{\text{sos}} := \inf_{x\in\mathbb{R}^m} \quad c^T x \quad \text{s.t. } a(Y)^T x + b(Y) \in \mathcal{Q}(G). \tag{3.1}$$

Clearly, any feasible point of (3.1) is also feasible for (1.1). Hence, we have $p^{\text{sos}} \geq p^*$.

Definition 3.1. We say that the Slater condition holds for the problem (1.1) if there exists $\bar{x} \in \mathbb{R}^m$ such that $a(y)^T \bar{x} + b(y) > 0$ for all $y \in S$.

Theorem 3.2. *If $\mathcal{Q}(G)$ is Archimedean and the Slater condition holds for the LSIPP problem (1.1), then $p^{\text{sos}} = p^*$.*

Proof. Fix an $\varepsilon > 0$ and a feasible $\bar{x} \in \mathbb{R}^m$ of (1.1) such that $a(y)^T\bar{x} + b(y) > 0$ for all $y \in S$. We next show that $p^{\text{sos}} - p^* < \varepsilon$. By Putinar's Positivstellensatz, $\bar{x}$ is a feasible point of (3.1) and thus we can assume that $c \neq 0$ without loss of generality. If $c^T\bar{x} - p^* < \varepsilon$, then $p^{\text{sos}} - p^* \leq c^T\bar{x} - p^* < \varepsilon$ and we are done. Hence, we assume that $c^T\bar{x} - p^* \geq \varepsilon$ in the following. Then we can fix another feasible point $x' \in \mathbb{R}^m$ of (1.1) such that $c^T\bar{x} > c^Tx'$ and $c^Tx' - p^* < \varepsilon/2$. Let

$$\delta := \frac{\varepsilon}{2c^T(\bar{x} - x')} > 0 \quad \text{and} \quad \hat{x} := (1-\delta)x' + \delta\bar{x}. \tag{3.2}$$

Then we have $0 < \delta < 1$ and hence

$$a(y)^T\hat{x} + b(y) = (1-\delta)[a(y)^Tx' + b(y)] + \delta[a(y)^T\bar{x} + b(y)] > 0, \quad \forall y \in S. \tag{3.3}$$

Since $\mathcal{Q}(G)$ is Archimedean, $a(Y)^T\hat{x} + b(Y) \in \mathcal{Q}(G)$ by Putinar's Positivstellensatz. That is, $\hat{x}$ is feasible for both (1.1) and (3.1). We have

$$\begin{aligned} p^{\text{sos}} - p^* &\leq c^T\hat{x} - p^* \\ &= (1-\delta)c^Tx' + \delta c^T\bar{x} - p^* \\ &= (c^Tx' - p^*) + \delta c^T(\bar{x} - x') \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \end{aligned} \tag{3.4}$$

which means that $p^{\text{sos}} \leq p^*$ since $\varepsilon > 0$ is arbitrary. As $p^{\text{sos}} \geq p^*$, we can conclude that $p^{\text{sos}} = p^*$. □

Note that we do not require that p^* is attainable in the above proof. For $k \geq d_P$, replacing $\mathcal{Q}(G)$ in (3.1) by its k-th truncation $\mathcal{Q}_k(G)$, we obtain

$$\begin{cases} p_k^{\text{sos}} := \inf\limits_{x \in \mathbb{R}^m} \ c^Tx \\ \qquad \text{s.t. } a(Y)^Tx + b(Y) = \sum\limits_{j=0}^{s} \sigma_j(Y)g_j(Y), \\ \qquad\qquad g_0 = 1, \sigma_j \in \Sigma^2[Y], \deg(\sigma_j g_j) \leq 2k, \ j = 0, \ldots, s. \end{cases} \tag{3.5}$$

Now we reformulate (3.5) as an SDP problem. For any $t \in \mathbb{N}$, let $m_t(Y)$ be the column vector consisting of all the monomials in Y of degree up to t. Recall that $s(t) = \binom{n+t}{n}$ which is the dimension of $m_t(Y)$. For each $j = 0, 1, \ldots, s$, there exists a positive semidefinite matrix $Z_j \in \mathbb{R}^{s(k-d_j)\times s(k-d_j)}$ such that

$$\sigma_j(Y) = m_{k-d_j}(Y)^T \cdot Z_j \cdot m_{k-d_j}(Y). \tag{3.6}$$

For each $\alpha \in \mathbb{N}^n_{2k}$, we can find a symmetric matrix $C_{j,\alpha} \in \mathbb{R}^{s(k-d_j)\times s(k-d_j)}$ such that the coefficient of $\sigma_j g_j$ equals $\langle Z_j, C_{j,\alpha}\rangle$ for each $j = 0, 1, \ldots, s$. Let

$$b(Y) = \sum_{\alpha \in \mathbb{N}^n_{2k}} b_\alpha Y^\alpha \quad \text{and} \quad a_i(Y) = \sum_{\alpha \in \mathbb{N}^n_{2k}} a_{i,\alpha} Y^\alpha, \ i = 1, \ldots, m. \tag{3.7}$$

Then (3.5) can be written as the following SDP problem

$$\begin{cases} p_k^{\text{sos}} = \inf\limits_{Z_j \succeq 0, x \in \mathbb{R}^m} \ c^Tx \\ \qquad \text{s.t. } \sum\limits_{i=1}^{m} x_i a_{i,\alpha} + b_\alpha = \sum\limits_{j=0}^{s} \langle Z_j, C_{j,\alpha}\rangle, \ \forall \alpha \in \mathbb{N}^n_{2k}. \end{cases} \tag{3.8}$$

It follows that

Theorem 3.3. *If $\mathcal{Q}(G)$ is Archimedean and the Slater condition holds for the LSIPP problem* (1.1)*, then $p_k^{\rm sos}$ decreasingly converges to p^* as $k \to \infty$.*

Proof. For any $\varepsilon > 0$, let $\hat{x}$ be defined as in the proof of Theorem 3.2. We have $a(Y)^T\hat{x} + b(Y) \in \mathcal{Q}_k(G)$ for some $k \in \mathbb{N}$ and then $p_k^{\rm sos} - p^* \leq c^T\hat{x} - p^* < \varepsilon$. Since ε is arbitrary, $p_k^{\rm sos}$ decreasingly converges to p^* as $k \to \infty$. □

The Lagrangian dual problem of (3.5) is eactly the SDP relaxation (2.11) derived by Lasserre in [23]. By the 'weak duality', we have $p_k^{\rm mom} \leq p_k^{\rm sos}$. Consequently, we can reprove the convergence of (2.11).

Theorem 3.4. *If $\mathcal{Q}(G)$ is Archimedean and the Slater condition holds for the LSIPP problem* (1.1)*, then $p_k^{\rm mom}$ decreasingly converges to p^* as $k \to \infty$.*

Proof. Since S is compact and the Slater condition holds for (1.1), $p^* = d^*$ and d^* is attainable (c.f. [5]). It is clear that $p_k^{\rm mom} \geq d^* = p^*$ for each $k \geq d_P$. Then the conclusion follows from Theorem 3.3 and the 'weak duality'. □

For any feasible point $x \in \mathbb{R}^m$ of (1.1), the *active index set* of x is

$$\{y \in S \mid a(y)^Tx + b(y) = 0\}. \tag{3.9}$$

Consider the flat extension condition (Condition 2.1). If it happens, then $p_k^{\rm mom} = p^*$ and by [7, Theorem 1.1], z^* has a unique r-atomic measure supported by S, i.e., there exist r positive real numbers $\lambda_1, \ldots, \lambda_r$ and r distinct points $v_1, \ldots, v_r \in S$ such that

$$z^* = \lambda_1\zeta_{2k,v_1} + \cdots + \lambda_r\zeta_{2k,v_r}, \tag{3.10}$$

where ζ_{2k,v_i} is the Zeta vector of v_i up to degree $2k$.

Proposition 3.5. *Suppose that $\mathcal{Q}(G)$ is Archimedean and the Slater condition holds for the LSIPP problem* (1.1)*. Then, $v_1, \ldots, v_r$ in* (3.10) *belong to the active index set of each minimizer x^* of* (1.1)*.*

Proof. As

$$p^* = c^Tx^* = \sum_{i=1}^{r} \lambda_i a(v_i)^Tx^* \geq -\sum_{i=1}^{r} \lambda_i b(v_i) = p_k^{\rm mom} = p^* \tag{3.11}$$

for any minimizer x^* of (1.1), the conclusion follows. □

The extraction procedure of the points v_i's can be found in [15] and has been implemented in GloptiPoly.

Remark 3.6. Note that the flat extension condition is only a sufficient condition which means that it might not hold when the finite convergence of (2.11) happens. A weaker stopping criterion called *flat truncation condition* was proposed by Nie in [32] for SDP relaxations of polynomial optimziaiton problems. It can also be used as a sufficient condition to certify the finite convergence of (2.11). Precisely, if an optimizer z^* of the k-th SDP relaxation (2.11) satisfies

$$\mathrm{rank} M_{t-d_S}(z^*) = \mathrm{rank} M_t(z^*) \tag{3.12}$$

for some integer $t \in [d_P, k]$, then $p_k^{\rm mom} = p^*$ and the points $v_1, \ldots, v_r$ can also be extracted. See [32] for details.

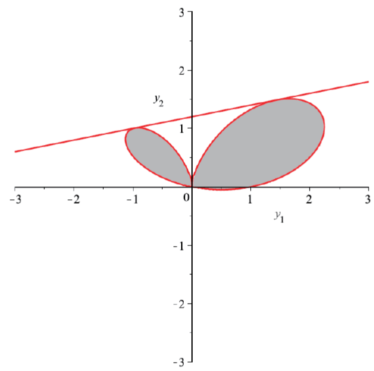

Figure 1: The semialgebraic set S (gray) in Example 3.7 and the line $x_1^* y_1 + x_2^* - y_2 = 0$ (red).

Compared with existing numerical approaches for LSIP problems, the SDP relaxations (3.5) and (2.11) are applicable for LSIPP problems with index sets being arbitrary basic semialgebraic sets, not necessarily box-shaped.

Example 3.7. Consider the following problem

$$\begin{cases} \inf\limits_{x\in\mathbb{R}^2} & x_2 \\ \text{s.t.} & x_1 y_1 + x_2 - y_2 \geq 0, \ \forall y \in S, \end{cases} \tag{3.13}$$

where

$$S := \{y \in \mathbb{R}^2 \mid (y_1 + 5y_2)y_1^2 - (y_1^2 + y_2^2)^2 \geq 0\} \tag{3.14}$$

which is the gray region in Figure 1. Clearly, it is equivalent to the bilevel problem

$$\min_{x_1\in\mathbb{R}} \max_{y\in S} \ y_2 - x_1 y_1. \tag{3.15}$$

By replacing the lower level maximality condition by the KKT condition, it is easy to check that the minimizer is $x^* = (\frac{1}{5}, \frac{125}{104})$ and its active index set consists of

$$\left(\frac{625}{2704} + \frac{1875}{2704}\sqrt{3}, \ \frac{3375}{2704} + \frac{375}{2704}\sqrt{3}\right) \approx (1.4322, 1.4884),$$
$$\left(\frac{625}{2704} - \frac{1875}{2704}\sqrt{3}, \ \frac{3375}{2704} - \frac{375}{2704}\sqrt{3}\right) \approx (-0.9699, 1.0079).$$

Thus, the optimum is $\frac{125}{104} \approx 1.2019$. Using GloptiPoly, we get $p_2^{\text{mom}} = 1.2982$ and $p_3^{\text{mom}} = 1.2019$. The flat extension condition holds at the order $k = 3$. We can extract the active index set $\{(1.4321, 1.4883), (-0.9699, 1.0079)\}$.

Although the optimal value p^* of (1.1) can be approximated by solving the dual problem (2.8) with the SDP relaxations (2.11) given in [23], the hierarchy of SDP relaxations (3.5) of (1.1) itself is of independent interest. For example, we can solve the relaxation (3.5) and extract the optimal solution $(x^{(k)}, Z_0^{(k)}, \ldots, Z_s^{(k)})$ (if it exists) by the software YALMIP [27]. As p_k^{sos} may not be attainable, let $(\tilde{x}^{(k)}, \tilde{Z}_0^{(k)}, \ldots, \tilde{Z}_s^{(k)})$ be an $\frac{\varepsilon}{2}$-optimal solution of (3.5). Since $\tilde{x}^{(k)}$ is feasible for (1.1) and $p_k^{\text{sos}} - p^* < \frac{\varepsilon}{2}$ for some $k \in \mathbb{N}$, a subsequence of $\{\tilde{x}^{(k)}\}_{k\in\mathbb{N}}$ converges to an ε-optimal solution of (1.1) if the feasible set of (1.1) is bounded.

Example 3.8. Consider the following problem

$$\begin{cases} \min\limits_{x_0, x_{i_1,i_2} \in \mathbb{R}} \; x_0 \\ \quad\text{s.t.} \; \left| \sum\limits_{i_1=0}^{t} \sum\limits_{i_2=0}^{t} x_{i_1,i_2} y_1^{i_1} y_2^{i_2} - b(y) \right| \le x_0, \quad \forall y \in [-1,1]^2, \end{cases} \tag{3.16}$$

which is to approximate the function $b(Y)$ from the spans of $Y_1^{i_1} Y_2^{i_2}$ in some sense. Hence, it is more useful to give the minimizers x^* which are the corresponding optimal coefficients for the basis functions $Y_1^{i_1} Y_2^{i_2}$ in the approximations. Here, we consider two cases [45]:

$$\text{(i): } t = 2,\; b(Y) = \frac{1}{Y_1 + 2Y_2 + 4}; \quad \text{and (ii): } t = 2,\; b(Y) = \sqrt{Y_1 + 2Y_2 + 4}. \tag{3.17}$$

While $b(Y)$ in (ii) is a semialgebraic function, (3.5) still works by adding some lifted variables, see Section 4.1. Solving (3.5) with YALMIP, the obtained coefficients are listed below

$$\begin{aligned} \text{(i)} &: (0.2341, -0.0468, -0.1507, 0.1203, -0.0706, -0.1292, 0.0837, 0.0136, 0.0927), \\ \text{(ii)} &: (2.0043, 0.2494, 0.5125, -0.0747, 0.0194, 0.0350, -0.0133, -0.0178, -0.0708), \end{aligned} \tag{3.18}$$

in the order $(x^*_{0,0}, x^*_{1,0}, x^*_{0,1}, x^*_{1,1}, x^*_{2,1}, x^*_{1,2}, x^*_{2,2}, x^*_{2,0}, x^*_{0,2})$.

At the end of this part, let us briefly introduce the SDP relaxation method for general SIPP problems given in [44] which can also be used to solve (1.1). The approach in [44] is based on the following exchange scheme. At its k-th iteration, we solve the linear programming problem

$$\min_{x \in \mathbb{R}^m} \quad c^T x \quad \text{s.t. } a(y)^T x + b(y) \ge 0,\; y \in Y_k, \tag{3.19}$$

where $Y_k \subset S$ is a finite set and generated at the last iteration. Next, choose a minimizer $x^{(k)}$ of (3.19) and globally solve the polynomial optimizaion problem

$$\min_{y \in S} \quad a(y)^T x^{(k)} + b(y) \tag{3.20}$$

by Lasserre's SDP relaxation method [21]. Extract the set $\mathcal{S}_k$ of global minimizers of (3.20) and let $Y_{k+1} = Y_k \cup \mathcal{S}_k$, then go to next iteration. To guarantee the convergence, the feasible set of (1.1) need to be compact and $\mathcal{S}_k \neq \emptyset$ for each k. However, if the flat extension condition or flat trucation condition is not satisfied when solving (3.20) by Lasserre's relaxation (when $\mathcal{S}_k$ is infinite, for example), it is hard to obtain the set $\mathcal{S}_k$ without which the iteration goes into dead loop.

3.2 Convergence rate analysis

Denote by $\mathcal{F}$ and $\mathcal{F}_k$ the feasible sets of (1.1) and (3.5), respectively. In this subsection, we assume that $\mathcal{F}$ and S are compact. For the simplicity in the convergence rate analysis of the SDP relaxations (3.5), we consider the following assumption which holds possibly after some rescaling.

Assumption 3.1. It holds that $\mathcal{F} \subseteq (-1,1)^m$ and $S \subseteq (-1,1)^n$ for (1.1).

Let

$$\omega := \max\{\deg(a_1), \ldots, \deg(a_m), \deg(b)\}. \tag{3.21}$$

For a polynomial $h(Y) = \sum_\alpha h_\alpha Y^\alpha \in \mathbb{R}[Y]$, define the norm

$$\|h\| := \max_\alpha \frac{|h_\alpha|}{\binom{\|\alpha\|_1}{\alpha}}.$$

Recalling the proof of Theorem 3.2, we have

Theorem 3.9. *Suppose that Assumption 3.1 holds, $\mathcal{Q}(G)$ is Archimedean and $\bar{x}$ is a Slater point of* (1.1)*. Let $r^*_{\bar{x}} := \min_{y\in S} a(y)^T\bar{x} + b(y) > 0$, then there exists $\gamma > 0$ depending on g_i's in* (1.2) *such that for any $\varepsilon > 0$, it holds that $0 \le p_k^{\text{sos}} - p^* \le \varepsilon$ whenever*

$$k \ge \gamma \exp\left[\left(\omega^2 n^\omega \frac{\sum_{i=1}^m \|a_i\| + \|b\|}{\kappa(\varepsilon) r_{\bar{x}^*}}\right)^\gamma\right], \tag{3.22}$$

where $\kappa(\varepsilon) := \min\{1, \frac{\varepsilon}{4\sqrt{m}\|c\|_2}\}$.

Proof. If $c^T\bar{x} - p^* \le \varepsilon$, let $x^{(\varepsilon)} = \bar{x}$; otherwise, let $x^{(\varepsilon)} = \hat{x}$ as defined in (3.2). As $\mathcal{F} \subseteq (-1,1)^m$, we have $\delta \ge \frac{\varepsilon}{4\sqrt{m}\|c\|_2}$ in (3.2). In either case, it holds from (3.3) that $a(y)^T x^{(\varepsilon)} + b(y) \ge \kappa(\varepsilon) r_{\bar{x}^*} > 0$ for any $y \in S$. By [35, Theorem 6], there exists $\gamma > 0$ depending on g_i's such that $a(Y)^T x^{(\varepsilon)} + b(Y) \in \mathcal{Q}_{\bar{k}}(G)$ where

$$\bar{k} := \gamma \exp\left[\left(\omega^2 n^\omega \frac{\|a(Y)^T x^{(\varepsilon)} + b(Y)\|}{\min_{y\in S} a(y)^T x^{(\varepsilon)} + b(y)}\right)^\gamma\right]. \tag{3.23}$$

Clearly,

$$\bar{k} \le \gamma \exp\left[\left(\omega^2 n^\omega \frac{\sum_{i=1}^m \|a_i\| + \|b\|}{\kappa(\varepsilon) r_{\bar{x}^*}}\right)^\gamma\right] \le k. \tag{3.24}$$

Hence, $a(Y)^T x^{(\varepsilon)} + b(Y) \in \mathcal{Q}_k(G)$ and $p_k^{\text{sos}} - p^* \le c^T x^{(\varepsilon)} - p^* \le \varepsilon$ by (3.4). □

For any $\varepsilon > 0$, compared with (1.1), consider the problem

$$(P_\varepsilon) \qquad \begin{cases} p^*_\varepsilon := \inf_{x\in\mathbb{R}^m} c^T x \\ \qquad\quad \text{s.t. } a(y)^T x + b(y) \ge \varepsilon, \quad \forall y \in S. \end{cases} \tag{3.25}$$

Obviously, $p^* \le p^*_\varepsilon$ for any $\varepsilon > 0$. Moreover, by the stability of optimal values of linear semi-infinite programming problems (c.f. [12, Theorem 5.1.5]), it follows that

Lemma 3.10. *If Assumption 3.1 and the Slater condition hold for* (1.1)*, then there exist scalars $\bar{\varepsilon} > 0$ and $L > 0$ such that $p^*_\varepsilon - p^* \le L\varepsilon$ for any $\varepsilon \le \bar{\varepsilon}$.*

Proof. Since $\mathcal{F}$ is compact, the optimal solution set of (1.1) is non-empty and compact. As S is compact, a Slater point of (1.1) is also a strong Slater point. Then, the conclusion follows by the Lipschitz continuity of the optimal value function of (1.1) (see, [12, Theorem 5.1.5]). □

For any $\varepsilon > 0$, denote the feasible set of (3.25) by

$$\mathcal{F}_\varepsilon := \{x \in \mathbb{R}^m \mid a(y)^T x + b(y) \ge \varepsilon, \quad \forall\, y \in S\}. \tag{3.26}$$

Lemma 3.11. *Suppose that Assumption 3.1 holds and $\mathcal{Q}(G)$ is Archimedean. Then, there exists some $\gamma > 0$ depending on g_i's in* (1.2) *such that for all integers $k > \gamma \exp((2\omega^2 n^\omega)^\gamma)$, we have $\mathcal{F}_\varepsilon \subseteq \mathcal{F}_k$ whenever*

$$\varepsilon \geq \varepsilon_k := \frac{6\omega^3 n^{2\omega}(\sum_{i=1}^m \|a_i\| + \|b\|)}{\sqrt[\gamma]{\log \frac{k}{\gamma}}}. \tag{3.27}$$

Proof. Fix a point $u \in \mathcal{F}_\varepsilon$. Let $r_u^* := \min_{y \in S} a(y)^T u + b(y)$. Then, by [35, Theorem 8], there exists some $\gamma > 0$ depending on g_i's such that for all $k > \gamma \exp((2\omega^2 n^\omega)^\gamma)$, it holds that

$$a(Y)^T u + b(Y) - r_u^* + \frac{6\omega^3 n^{2\omega}\|a(Y)^T u + b(Y)\|}{\sqrt[\gamma]{\log \frac{k}{\gamma}}} \in \mathcal{Q}_k(G). \tag{3.28}$$

As $u \in \mathcal{F}_\varepsilon$, we have $r_u^* \geq \varepsilon$. Since $\mathcal{F}_\varepsilon \subseteq \mathcal{F}$, the assumption $\mathcal{F} \subseteq (-1,1)^m$ implies that $\|a(Y)^T u + b(Y)\| \leq \sum_{i=1}^m \|a_i(Y)\| + \|b(Y)\|$. Consequently,

$$r_u^* - \frac{6\omega^3 n^{2\omega}\|a(Y)^T u + b(Y)\|}{\sqrt[\gamma]{\log \frac{k}{\gamma}}} \geq \varepsilon - \varepsilon_k \geq 0. \tag{3.29}$$

Hence, we have $a(Y)^T u + b(Y) \in \mathcal{Q}_k(G)$ and $u \in \mathcal{F}_k$. □

Theorem 3.12. *Suppose that Assumption 3.1, the Slater condition hold for* (1.1) *and $\mathcal{Q}(G)$ is Archimedean. Then, there exist some $\gamma > 0$ depending on g_i's in* (1.2) *and scalars $\bar{\varepsilon} > 0$, $L > 0$ such that for all integers*

$$k > \max\left\{\gamma \exp((2\omega^2 n^\omega)^\gamma),\ \gamma \exp\left[\frac{6\omega^3 n^{2\omega}(\sum_{i=1}^m \|a_i\| + \|b\|)}{\bar{\varepsilon}}\right]^\gamma\right\}, \tag{3.30}$$

it holds that

$$0 \leq p_k^{\text{sos}} - p^* \leq L \frac{6\omega^3 n^{2\omega}(\sum_{i=1}^m \|a_i\| + \|b\|)}{\sqrt[\gamma]{\log \frac{k}{\gamma}}}. \tag{3.31}$$

Proof. Note that all assumptions in Lemma 3.10 and 3.11 hold. Then, there exist $\gamma > 0$ depending on g_i's, $\bar{\varepsilon} > 0$ and $L > 0$ as described in the conclusions of Lemma 3.10 and 3.11. Recall ε_k defined in (3.27). By Lemma 3.11, it holds that $\mathcal{F}_{\varepsilon_k} \subseteq \mathcal{F}_k$ which implies $p_{\varepsilon_k}^* \geq p_k^{\text{sos}}$. Moreover, it is easy to check that $\varepsilon_k \leq \bar{\varepsilon}$ and hence $p_{\varepsilon_k}^* - p^* \leq L\varepsilon_k$ by Lemma 3.10. Consequently, $p_k^{\text{sos}} - p^* \leq L\varepsilon_k$. □

3.3 On compactness of $\mathcal{F}$

In the last subsection, we assume that the feasible set $\mathcal{F}$ of (1.1) is compact in order to estimate the convergence rate of the SDP relaxations (3.5). In the following, we show that the compactness of $\mathcal{F}$ can be determined by solving some LSIPP problems, which can be done by the SDP relaxations (3.5) in some cases. Denote by $\mathbf{0}$ the vector of all zeros in $\mathbb{R}^n$. In this subsection, without loss of generality,we assume that

Assumption 3.2. $\mathbf{0} \in \mathcal{F}$, or equivalently, $b(y) \geq 0$ for all $y \in S$.

Denote by $0^+\mathcal{F} \subset \mathbb{R}^m$ the *recession cone* of $\mathcal{F}$, i.e., $u \in 0^+\mathcal{F}$ if and only if $x + tu \in \mathcal{F}$ for all $x \in \mathcal{F}$ and $t \geq 0$. As $\mathcal{F}$ is closed and convex, $\mathcal{F}$ is compact if and only if $0^+\mathcal{F} = \{\mathbf{0}\}$ by [40, Theorem 8.4].

Consider the minimax problem

$$r_a^* := \max_{\|u\|_2=1} \min_{y\in S} \ a(y)^T u \tag{3.32}$$

Proposition 3.13. *Suppose that Assumption 3.2 holds for* (1.1). *Then, its feasible set* $\mathcal{F}$ *is compact if and only if* $r_a^* < 0$.

Proof. As $\mathbf{0} \in \mathcal{F}$, by [40, Theorem 8.3], a vector $u \in 0^+\mathcal{F}$ if and only if $tu \in \mathcal{F}$ for all $t \geq 0$. That is, $(a(y)^T u) \cdot t + b(y) \geq 0$ for all $y \in S$ and $t \geq 0$, which is true if and only if $a(y)^T u \geq 0$ for all $y \in S$ since $b(y) \geq 0$ on S. Note that since S is compact, $\min_{y\in S}\ a(y)^T u$ is continuous in u and then r_a^* is attainable. Therefore, if $\mathcal{F}$ is compact, then $u \notin 0^+\mathcal{F}$ for any nonzero $u \in \mathbb{R}^n$ and hence $r_a^* < 0$. Conversely, assume that $r_a^* < 0$. If $u \in 0^+\mathcal{F}$ for some nonzero $u \in \mathbb{R}^n$, then we have $r_a^* \geq 0$, a contradiction. Then, $0^+\mathcal{F} = \{\mathbf{0}\}$ and hence $\mathcal{F}$ is compact. $\square$

It is clear that the minimax problem (3.32) is equivalent to the following problem

$$\begin{cases} r_a^* = \displaystyle\sup_{\|u\|_2=1,\lambda\in\mathbb{R}} \ \lambda \\ \qquad\qquad \text{s.t. } a(y)^T u - \lambda \geq 0, \quad \forall\, y \in S. \end{cases} \tag{3.33}$$

By some rescalings, we can reformulate the problem (3.33) as the following LSIPP problems of the form (1.1):

$$(P_i^+) \qquad \begin{cases} r_{a,i}^+ := \displaystyle\inf_{u_i=1,\lambda\in\mathbb{R}} \ -\lambda \\ \qquad\qquad \text{s.t. } a(y)^T u - \lambda \geq 0, \quad \forall y \in S, \end{cases} \tag{3.34}$$

and

$$(P_i^-) \qquad \begin{cases} r_{a,i}^- := \displaystyle\inf_{u_i=-1,\lambda\in\mathbb{R}} \ -\lambda \\ \qquad\qquad \text{s.t. } a(y)^T u - \lambda \geq 0, \quad \forall y \in S, \end{cases} \tag{3.35}$$

for $i = 1, \ldots, m$. Then,

Corollary 3.14. *Suppose that Assumption 3.2 holds for* (1.1). *Then, its feasible set* $\mathcal{F}$ *is compact if and only if* $\min\{r_{a,i}^+,\ r_{a,i}^-,\ i = 1, \ldots, m\} > 0$.

Consequently, the compactness of $\mathcal{F}$ can be verified by a positive lower bound of $\min\{r_{a,i}^+,\ r_{a,i}^-,\ i = 1, \ldots, m\}$ which can be obtained by solving (P_i^+)'s and (P_i^-)'s using, for instance, discretization methods.

Note that the SDP relaxations (3.5) and (2.11) produce *upper* bounds of p^* of (1.1). The compactness of $\mathcal{F}$ can also be verified by these SDP relaxations of (P_i^+)'s and (P_i^-)'s if finite convergence happens for each problem, which can be detected by the flat extension condition (or the weaker flat truncation condition in Remark 3.6). In particular, when S is a closed and bounded interval, the SDP relaxations (3.5) and (2.11) of the smallest order are exact for (1.1) by the representation result of nonnegative polynomials in the univariate case. This result has been investigated in [46]. Precisely, without loss of generality, we can assume that $S = [-1, 1]$. Let

$$[-1, 1] = \{y_1 \in \mathbb{R} \mid g_1(y_1) \geq 0\}, \quad \text{where} \quad g_1(Y_1) = 1 - Y_1^2.$$

Recall the well-known result

Theorem 3.15 (c.f. [26,38]). *Let $h \in \mathbb{R}[Y_1]$ and $h \geq 0$ on $[-1,1]$, then $h = \sigma + \sigma_1(1 - Y_1^2)$ where $\sigma, \sigma_1 \in \Sigma^2[Y_1]$ and $\deg(\sigma)$, $\deg(\sigma_1(1-Y_1^2)) \leq 2\lceil \deg(h)/2 \rceil$.*

It follows that $p^{\text{sos}}_{k_P} = p^{\text{mom}}_{k_P} = p^*$ holds for (3.5) and (2.11) in this case [46]. Therefore, the compactness of $\mathcal{F}$ can always be verified by SDP relaxations (3.5) and (2.11) of (P_i^+)'s and (P_i^-)'s when S is a closed and bounded interval.

Example 3.7 revisited. Consider the feasible set $\mathcal{F}$ of (3.13) in Example 3.7, which is clearly noncompact. Note that $\mathbf{0} \notin \mathcal{F}$ and we have proved that the minimizer is $(\frac{1}{5}, \frac{125}{104})$. Let $w_1 = x_1 - \frac{1}{5}$, $w_2 = x_2 - \frac{125}{104}$ and move the set $\mathcal{F}$ to

$$\left\{ w \in \mathbb{R}^2 \mid \left(w_1 + \frac{1}{5}\right) y_1 + \left(w_2 + \frac{125}{104}\right) - y_2 \geq 0,\ \forall\ y \in S \right\}, \tag{3.36}$$

which contains $\mathbf{0}$. Consider the LSIPP problem

$$(P_1^+) \qquad \begin{cases} r^+_{a,1} := \inf\limits_{u_2, \lambda \in \mathbb{R}} \ -\lambda \\ \qquad\qquad \text{s.t. } y_1 + u_2 - \lambda \geq 0, \quad \forall y \in S. \end{cases} \tag{3.37}$$

It is easy to check that $y_1 \geq -2$ for all $y \in S$. Then, $(u_2 = N+2, \lambda = N)$ is feasible for (P_1^+) for all $N \in \mathbb{N}$. Hence, we have $r^+_{a,1} = -\infty$ and therefore $\mathcal{F}$ is noncompact by Corollary 3.14.

Example 3.17. Consider the ellipse

$$\mathcal{F} := \{(x_1, x_2) \in \mathbb{R}^2 \mid 2x_1^2 + x_2^2 + 2x_1x_2 + 2x_1 \leq 0\} \tag{3.38}$$

which can be represented by

$$\{(x_1, x_2) \in \mathbb{R}^2 \mid a(y_1)^T x + b(y_1) \geq 0,\ \forall y_1 \in S\} \tag{3.39}$$

where

$$a(Y_1) = (-Y_1^4 - 2Y_1^3 + 3Y_1^2 + 2Y_1 - 1, -2Y_1(Y_1^2 - 1))^T, \quad b(Y_1) = 2Y_1^2, \tag{3.40}$$

and $S = [-1, 1]$ (see [11]). Clearly, $\mathcal{F}$ is compact and $\mathbf{0} \in \mathcal{F}$. As $S = [-1,1]$, all problems (P_i^+)'s and (P_i^-)'s can be solved by the SDP relaxations (3.5) and (2.11) of order $d_P = 2$. Using GloptiPoly, we first solve the SDP relaxation (2.11) of

$$(P_1^+) \qquad \begin{cases} r^+_{a,1} := \inf\limits_{u_2, \lambda \in \mathbb{R}} \ -\lambda \\ \qquad\qquad \text{s.t. } -y_1^4 - 2y_1^3 + 3y_1^2 + 2y_1 - 1 - 2y_1(y_1^2 - 1)u_2 - \lambda \geq 0, \\ \qquad\qquad\quad \forall y_1 \in S. \end{cases} \tag{3.41}$$

As the infeasibility of the SDP problem is detected by the SDP solver SeDuMi [42] called by GloptiPoly, we have $r^+_{a,1} = +\infty$. We continue to solve (P_1^-), (P_2^+) and (P_2^-). The results solved by GloptiPoly are $r^-_{a,1} = 1$, $r^+_{a,2} = 0.5491$ and $r^-_{a,2} = 0.7698$, which imply the compactness of $\mathcal{F}$ by Corollary 3.14.

Note that the index set S is required to be compact to guarantee the convergence of the SDP relaxations (3.5) and (2.11). To end this section, we consider two examples to illustrate how to deal with the case when S is noncompact by the homogenization technique and its applications in polynomial optimzation problems.

Example 3.18. Consider the LSIPP problem

$$p^* := \inf_{x\in\mathbb{R}} \ -\frac{x}{2} \quad \text{s.t. } (1-3y_2)x + 3y_1 \geq 0,\ \forall y \in S, \tag{3.42}$$

where

$$S := \{y \in \mathbb{R}^2 \mid y_1 \geq 0, y_1^2 - y_2^3 \geq 0\}. \tag{3.43}$$

Since $(0,0) \in S$, a feasible x must be nonnegative. Clearly, $x = 0$ is a feasible point. $x > 0$ is feasible if and only if

$$0 \geq \max_{y\in S} \left\{ y_2 - \frac{1}{3} - \frac{y_1}{x} \right\} = \max_{y\in S} \left\{ y_1^{\frac{2}{3}} - \frac{1}{3} - \frac{y_1}{x} \right\}. \tag{3.44}$$

The latter maximum is attained at $\frac{8x^3}{27}$ with optimal value $\frac{4x^2}{27} - \frac{1}{3}$. Thus, the feasible set of (3.42) is $[0, \frac{3}{2}]$ and the minimizer is $x^* = \frac{3}{2}$.

Obviously, $\mathcal{Q}(G)$ is not Archimedean. For any $k \in \mathbb{N}$, we know from [13, Example 2.10] that $(1-3Y_2)x + 3Y_1 \in \mathcal{Q}_k(G)$ if and only if $x = 0$, i.e., $p_k^{\text{sos}} = 0$ for each $k \geq d_P$. Now we show that $p_k^{\text{mom}} = p_k^{\text{sos}}$ for each $k \geq d_P$. In fact, for the SDP relaxation (2.11) of the problem (3.42), let μ be a probability measure with uniform distribution in the following subset of S:

$$S_1 := \{(y_1, y_2) \in \mathbb{R}^2 \mid 1 \leq y_1 \leq 2,\ 0 \leq y_2 \leq 1\} \tag{3.45}$$

and $z^{(\mu)}$ be the truncated moment sequence with representing measure μ up to order $2k$. It can be verified that $z^{(\mu)}$ is a feasible point of (2.11) and its corresponding truncated moment matrix and localizing moment matrices are positive definite since S_1 has nonempty interior. Then $p_k^{\text{mom}} = p_k^{\text{sos}}$ follows by the conic duality theorem. Hence, both SDP relaxations (3.5) and (2.11) do not converge to the optimum.

Now let us see how to solve this issue by homogenization. We first homogenize the defining polynomials of S by new variable y_0 and define the following bounded set

$$\widetilde{S}_> := \{\tilde{y} = (y_0, y_1, y_2) \in \mathbb{R}^3 \mid y_1 \geq 0,\ y_0 y_1^2 - y_2^3 \geq 0,\ y_0 > 0,\ \|\tilde{y}\|_2^2 = 1\}. \tag{3.46}$$

Then, we homogenize the constraint polynomial of (3.42) with respect to Y and consider the problem

$$\inf_{x\in\mathbb{R}} \ -\frac{x}{2} \quad \text{s.t. } (y_0 - 3y_2)x + 3y_1 \geq 0,\ \forall \tilde{y} = (y_0, y_1, y_2) \in \mathsf{closure}(\widetilde{S}_>), \tag{3.47}$$

which is equivalent to (3.42) by [44, Proposition 4.2]. However, the set $\mathsf{closure}(\widetilde{S}_>)$ is not in the form of basic semialgebraic sets. Hence, we define the following compact set

$$\widetilde{S} := \{(y_0, y_1, y_2) \in \mathbb{R}^3 \mid y_1 \geq 0,\ y_0 y_1^2 - y_2^3 \geq 0,\ y_0 \geq 0,\ \|\tilde{y}\|_2^2 = 1\}. \tag{3.48}$$

We say S is *closed at* ∞ [33] if $\widetilde{S} = \mathsf{closure}(\widetilde{S}_>)$, in which case (3.42) is equivalent to

$$\inf_{x\in\mathbb{R}} \ -\frac{x}{2} \quad \text{s.t. } (y_0 - 3y_2)x + 3y_1 \geq 0,\ \forall \tilde{y} = (y_0, y_1, y_2) \in \widetilde{S}. \tag{3.49}$$

Note that S is indeed closed at ∞. In fact, for every $(0, v_1, v_2) \in \widetilde{S} \backslash \widetilde{S}_>$, let

$$v^{(\varepsilon)} := \left(\varepsilon,\ v_1,\ \sqrt[3]{\varepsilon v_1^2 + v_2^3} \right). \tag{3.50}$$

Then $\{v^{(\varepsilon)}/\|v^{(\varepsilon)}\|_2\}_{\varepsilon>0} \subseteq \widetilde{S}_>$ and $\lim_{\varepsilon\to 0} v^{(\varepsilon)}/\|v^{(\varepsilon)}\|_2 = (0, v_1, v_2)$. Hence, we have $\widetilde{S}\backslash\widetilde{S}_> \subseteq$ $\mathsf{closure}(\widetilde{S}_>)$ and so S is closed at ∞. Clearly, the quadratic module associated with $\widetilde{S}$ is Archimedean and $\bar{x} = 1$ is a Slater point of (3.49). With GloptiPoly, we solve the SDP relaxations (2.11) of (3.49) and get the following numerical results: $p_2^{\rm mom} = -1.2124 \times 10^{-8}$ and $p_3^{\rm mom} = -0.7500$. The flat extension condition is satisfied for $k = 3$ and we obtain the certified optimum -0.7500. By Proposition 3.5, the extracted numerical active index set of the minimizer $x^* = 3/2$ is $(0.5773, 0.5774, 0.5774)$ which corresponds to $(1, 1) \in S$.

Remark 3.19. Note that not every set S of the form (1.2) is closed at ∞ even when it is compact [31, Example 5.2]. However, it is shown in [44, Theorem 4.10] that the closedness at ∞ is a *generic* property.

Example 3.20. Consider the following polynomial optimization problem

$$\begin{cases} \inf\limits_{y\in\mathbb{R}^2} \; f(y) := y_1^2 + y_2^2 \\ \text{s.t.} \;\; y \in S := \{y \in \mathbb{R}^2 \mid g_1(y) \geq 0, \; g_2(y) \geq 0, \; g_3(y) \geq 0\}, \end{cases} \tag{3.51}$$

where

$$g_1(Y) = Y_2^2 - 1, \; g_2(Y) = Y_1^2 - Y_1Y_2 - 1, \; g_3(Y) = Y_1^2 + Y_1Y_2 - 1. \tag{3.52}$$

It was shown in [8, 29, 33] that the global minimizers and global minimum are

$$\left(\pm\frac{1+\sqrt{5}}{2}, \pm 1\right) \approx (\pm 1.618, \pm 1) \quad \text{and} \quad 2 + \frac{(1+\sqrt{5})}{2} \approx 3.618. \tag{3.53}$$

Because S is noncompact, the classic Lasserre's SDP relaxations [21] of (3.51) can only provide lower bounds 2 no matter how large the order is (c.f. [8]).

Clearly, any polynomial optimization problem of the form (3.51) can be equivalently reformulated to the following LSIPP problem

$$f^* = \sup_{x\in\mathbb{R}} \; x \quad \text{s.t.} \;\; f(y) - x \geq 0, \; \forall y \in S. \tag{3.54}$$

As S is noncompact, we use the homogenization technique in Example 3.18 to convert this LSIPP problem to

$$\tilde{f}^* := \sup_{x\in\mathbb{R}} \; x \quad \text{s.t.} \;\; f^h(\tilde{y}) - x y_0^{\deg(f)} \geq 0, \; \forall \tilde{y} \in \widetilde{S}, \tag{3.55}$$

where f^h is the homogenization of f and $\widetilde{S}$ is defined as in Example 3.18. Suppose that $f^* > -\infty$, then the Slater condition holds for (3.55) if and only if

$$\hat{f}(y) > 0, \quad \forall\, y \in \widehat{S} := \{y \in \mathbb{R}^n \mid \hat{g}_1(y) \geq 0, \; \ldots, \; \hat{g}_s(y) \geq 0, \; \|y\|_2^2 = 1\}, \tag{3.56}$$

where $\hat{f}$ and $\hat{g}_i$'s are the homogeneous parts of f and g_i's of the highest degree. Moreover, if the condition (3.56) holds for (3.55), it is easy to see that any feasible point of (3.54) is also feasible for (3.55). Thus, $\tilde{f}^* = f^*$ and we can compute them by the SDP relaxations (3.5) and (2.11).

Obviously, the condition (3.56) holds for (3.51). We compute the relaxations (2.11) of (3.55) with GloptiPoly. For $k = 3$, the flat extension condition is satisfied and we get the numerically certified optimum $f_3^{\rm mom} = 3.6180$. The extracted active index set is $\{(0.4653, \pm 0.7529, \pm 0.4653)\}$ which corresponds to the set of global minimizers $(\pm 1.6181, \pm 1)$.

Remark 3.21. (i) By [30, Theorem 5.1 and 5.3], the condition (3.56) holds if and only if f is *stably bounded from below* on S, i.e., f remains bounded from below on S for all sufficiently small perturbations of the coefficients of $f, g_1, \ldots, g_s$. Therefore, we give an SDP relaxation method in Example 3.20 for solving the class of polynomial optimization problems whose objective polynomials are stably bounded from below on noncompact feasible sets; (ii) Note that the stably boundedness from below of f on S is irrelevant to the closedness at ∞ of S. For example, the set $\{y \in \mathbb{R}^2 \mid y_2 \geq y_1^2\}$ is not closed at ∞ but Y_2^2 is stably bounded from below on it; the set S in Example 3.18 is closed at ∞ but Y_1 is not stably bounded from below on it.

4 Some Extensions

In this section, we discuss some extensions of the SDP relaxations (3.5) for (1.1) in Section 3 to more general semi-infinite programming problems.

4.1 LSIP with semi-algebraic functions

Inspired by Lasserre and Putinar's work [25], we would like to point out that the SDP relaxation method proposed in this paper is applicable to a more general subclass of LSIP problems. Denote by $\mathcal{X} \subseteq \mathbb{R}^m$ a convex polyhedron defined by finitely many linear inequalities in the variables X. Denote by $\mathcal{A}$ the algebra consisting of functions generated by finitely many of the dyadic operations $\{+, -, /, \vee, \wedge\}$ and monadic operations $\{|\cdot|, (\cdot)^{1/p}, p \in \mathbb{N}\}$ on polynomials in $\mathbb{R}[Y]$, where $f \vee g := \max[f, g]$ and $f \wedge g := \min[f, g]$ for $f, g \in \mathbb{R}[Y]$. For example,

$$\sqrt{|f(Y)| + g(Y)^2} \wedge \left(\frac{1}{g(Y)} \vee f(Y)\right) \in \mathcal{A}. \tag{4.1}$$

Note that every function in $\mathcal{A}$ has a *lifted basic semi-algebraic representation* [25, Definition 1]. Then, the SDP relaxations (3.5) and (2.11) can be extended for more general LSIP problems of the form

$$\begin{cases} p^* := \inf_{x \in \mathcal{X}} \ c^T x \\ \qquad\quad \text{s.t.} \ a^l(y)^T x + b_l(y) \geq 0, \quad \forall y \in S \text{ and } l = 1, \ldots, t, \end{cases} \tag{4.2}$$

where $c \in \mathbb{R}^m$, $a^l(Y) \in \mathcal{A}^m$, $b_l(Y) \in \mathcal{A}$, $l = 1, \ldots, t$ and

$$S := \{y \in \mathbb{R}^n \mid g_1(y) \geq 0, \ldots, g_s(y) \geq 0\}, \tag{4.3}$$

where $g_j(Y) \in \mathcal{A}$, $j = 1, \ldots, s$. In fact, as shown in [25], the nonnegativity test of semi-algebraic functions in $\mathcal{A}$ on the set (4.3) can be reduced to an equivalent polynomial funcation case in a *lifted* space by adding some new variables. For instance, with $f, h, g_1, g_2 \in \mathbb{R}[Y_1]$,

$$\sqrt{f(y_1)} - 1/h(y_1) \geq 0 \quad \text{on} \quad \{y_1 \in \mathbb{R} \mid |g_1(y_1)| g_2(y_1) \geq 1\} \tag{4.4}$$

can be written as $y_2 - y_3 \geq 0$ on

$$\{y \in \mathbb{R}^4 \mid f(y_1) = y_2^2,\ y_2 \geq 0,\ h(y_1)y_3 = 1,\ y_4 g_2(y_1) \geq 1,\ g_1(y_1)^2 = y_4^2,\ y_4 \geq 0\}. \tag{4.5}$$

Consequently, the extension to $\mathcal{A}$ of Putinar's Positivstellensatz ([25, Theorem 2]) provides us representations of each nonnegativity constraint in (4.3) via s.o.s and the dual theory of moments. Notice that the constraint $x \in \mathcal{X}$ is linear in X. Hence, SDP relaxations as (3.5) and its dual (2.11) can be similarly derived for (4.2) by lifting the parameter space. Moreover, the convergence results and stopping criterion, as Theorem 3.3, 3.4 and Proposition 3.5, can also be analogously established. As might be expected, additional parameters in the lifted space can cause more computational burden in resulting SDP problems. However, as pointed out in [25], the *running intersection property* holds true for these lifted parameters. Hence, like for polynomial optimization problems [22,43], some sparse SDP relaxations for (4.2) can be explored to reduce the computational cost.

Example 4.1. Consider the one-sided L_1 approximation problem

$$\begin{cases} \min\limits_{x\in\mathbb{R}^n} \ \sum\limits_{i=1}^{n} \dfrac{x_i}{i} \\ \quad \text{s.t.} \ \sum\limits_{i=1}^{n} y^{i-1}x_i - b(y) \geq 0, \quad \forall y \in [0,1]. \end{cases} \tag{4.6}$$

Here, we approximate two (semi-algebraic) functions [9] on $[0,1]$:

$$\text{(i): } b(y) = \frac{1}{2-y}, \ n = 8; \quad \text{and} \quad \text{(ii): } b(y) = -\frac{1}{1+y^2}, \ n = 10. \tag{4.7}$$

Clearly, in order to convert this problem into LSIPP, we can add lifted variable z such that $(2-y)z = 1$ for case (i) and $(1+y^2)z = -1$ for case (ii). Then, we solve the SDP relaxations (3.5) with order $k = 4$ for (i) and $k = 5$ for (ii) by YALMIP. The obtained coefficients x_i's are listed below

$$\begin{aligned} &\text{(i): } (0.5000, 0.2501, 0.1227, 0.0787, -0.0258, 0.1226, -0.0967, 0.0484), \\ &\text{(ii): } (-1.0000, -0.0000, 1.0016, -0.0202, -0.8566, -0.6123, 2.6222, -2.6059, \\ &\qquad 1.1881, -0.2168) \end{aligned} \tag{4.8}$$

We show the accuracy of the computed optimal approximations (denoted by f) of $b(Y)$ in Figure 2.

Example 4.2. In (x_1, x_2)-plane, consider the intersection area $\mathcal{F}$ of $x_2 \geq 0$, $1 - x_1^2 - x_2^2 \geq 0$ and $x_1 + 1 - x_2^2 \geq 0$. Then, $\mathcal{F}$ can also be seen as the interstion of $x_2 \geq 0$, the half planes defined by the lines tangent to $1 - x_1^2 - x_2^2 = 0$ in the first quadrant and to $x_1 + 1 - x_2^2 = 0$ in the second quadrant, as shown in Figure 3. Therefore, it is easy to check that

$$\mathcal{F} = \{(x_1, x_2) \in \mathbb{R}^2 \mid a(y_1)^T x + b(y_1) \geq 0, \ \forall \ y_1 \in [-1, 1]\} \cap \mathcal{X}, \tag{4.9}$$

where $\mathcal{X} = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 \geq 0\}$,

$$\begin{aligned} a_1(y_1) &= -\min[y_1, 0] - \max[y_1, 0]y_1, \\ a_2(y_1) &= 2\min[y_1, 0]\sqrt{y_1 + 1} - \max[y_1, 0]\sqrt{1 - y_1^2}, \\ b(y_1) &= -\min[y_1, 0](2 + y_1) + \max[y_1, 0]. \end{aligned} \tag{4.10}$$

Here, the equations $a(y_1)^T X + b(y_1) = 0$ for $y_1 \in [-1, 1]$, in fact, represent the tangent lines mentioned above. Consider the LSIP problem

$$\begin{cases} p^* := \min\limits_{x\in\mathcal{X}} \ c^T x \\ \qquad\quad \text{s.t. } a(y_1)^T x + b(y_1) \geq 0, \ \ \forall y \in [-1, 1], \end{cases} \tag{4.11}$$

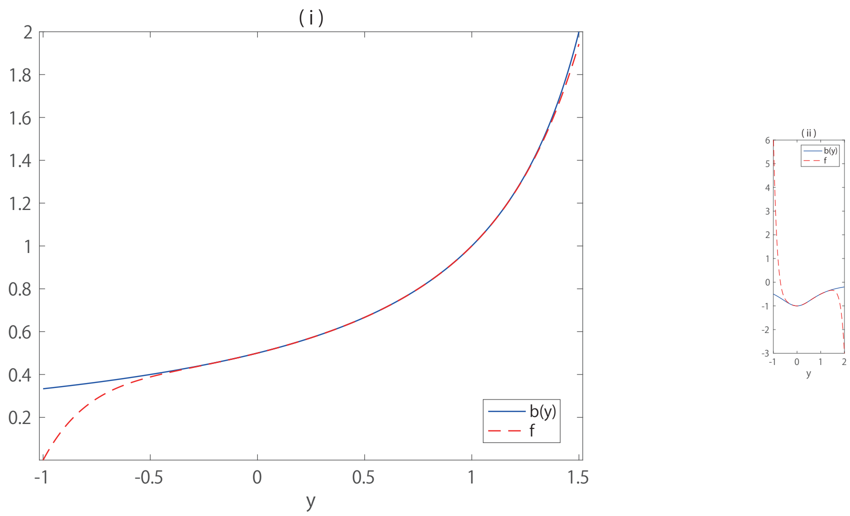

Figure 2: Pictures for the univariate approximation problems in Example 4.1.

in two cases: (i) $c = (1, -1)$; (ii) $c = (-1, -1)$. We can verify that the minima and minimizers are: (i) $p^* = -\frac{5}{4}$, $x^* = (-\frac{3}{4}, \frac{1}{2})$; (ii) $p^* = -\sqrt{2}$, $x^* = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$. Now, we first convert this LSIP problem into LSIPP by the lifting method and then solve it by the SDP relaxations (3.5). As $2 \cdot \min[y_1, 0] = y_1 - |y_1|$ and $2 \cdot \max[y_1, 0] = y_1 + |y_1|$, we can write

$$\mathcal{F} = \{(x_1, x_2) \in \mathbb{R}^2 \mid \tilde{a}(y)^T x + \tilde{b}(y) \geq 0, \ \forall \ y \in S\} \cap \mathcal{X}, \tag{4.12}$$

where

$$\begin{aligned} \tilde{a}_1(y) &= -y_1 + y_2 - y_1^2 - y_1 y_2, \\ \tilde{a}_2(y) &= 2y_1 y_3 - 2y_2 y_3 - y_1 y_4 - y_2 y_4, \\ \tilde{b}(y) &= -(2 + y_1)(y_1 - y_2) + y_1 + y_2, \end{aligned} \tag{4.13}$$

and

$$S = \{y \in \mathbb{R}^4 \mid -1 \leq y_1 \leq 1, y_2^2 = y_1^2, y_2 \geq 0, 1 + y_1 = y_3^2, y_3 \geq 0, 1 - y_1^2 = y_4^2, y_4 \geq 0\}. \tag{4.14}$$

Now we can use the SDP relaxations (3.5) to solve the obtained LSIPP problems. The approximate minima and minimizers are: (i) $p_5^{\text{sos}} = -1.2496$, $\tilde{x}^{(5)} = (-0.7575, 0.4921)$; (ii) $p_5^{\text{sos}} = -1.3936$, $\tilde{x}^{(5)} = (0.6544, 0.7392)$.

4.2 S.O.S-convex objectives

Next, by the exact SDP relaxations for classes of nonlinear SDP problems proposed in [18], we extend the SDP relaxation method in Section 3 to the following semi-infinite programming problem

$$\begin{cases} h^* := \inf\limits_{x \in \mathbb{R}^m} \ h(x) \\ \qquad\qquad \text{s.t. } a(y)^T x + b(y) \geq 0, \quad \forall y \in S, \end{cases} \tag{4.15}$$

where $a(Y)$, $b(Y)$, S are defined as in (1.1) and $h(X) \in \mathbb{R}[X]$ is s.o.s-convex polynomial. Recall that

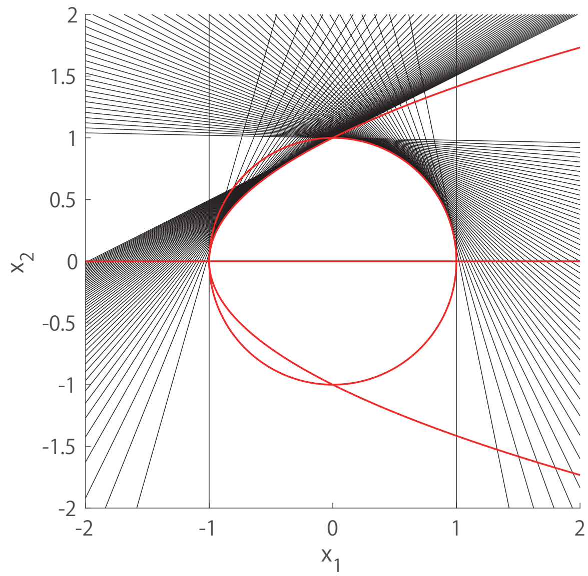

Figure 3: The feasible set $\mathcal{F}$ in Example 4.2.

Definition 4.3 ([14])**.** A polynomial $h \in \mathbb{R}[Y]$ is s.o.s-convex if its Hessian $\nabla^2 h$ is a s.o.s, i.e., there are an integer r and a matrix polynomial $H \in \mathbb{R}[Y]^{r\times n}$ such that $\nabla^2 h(Y) = H(Y)^T H(Y)$.

While checking the convexity of a polynomial is generally NP-hard [1], s.o.s-convexity can be checked numerically by solving an SDP, see [14].

We first relax (4.15) with arbitrary convex polynomial objective function as

$$h^{\text{sos}} := \inf_{x\in\mathbb{R}^m} \quad h(x) \quad \text{s.t. } a(Y)^T x + b(Y) \in \mathcal{Q}(G). \tag{4.16}$$

Theorem 4.4. *If $h(X)$ is convex, $\mathcal{Q}(G)$ is Archimedean and the Slater condition holds for* (4.15)*, then $h^{\text{sos}} = h^*$.*

Proof. As $h(X)$ is convex, replacing the function $c^T X$ in the proof of Theorem 3.2 by $h(X)$, all arguments in the proof of Theorem 3.2 are still valid. In particular, (3.4) become

$$\begin{aligned} p^{\text{sos}} - p^* &\le h(\hat{x}) - p^* \\ &\le (1-\delta)h(x') + \delta h(\bar{x}) - p^* \\ &= (h(x') - p^*) + \delta(h(\bar{x}) - h(x')) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \end{aligned} \tag{4.17}$$

due to the convexity of h. Then the conclusion follows. □

Recall that $\mathcal{F}_k$ denotes the feasible set of (3.5). For $k \ge d_P$, replacing $\mathcal{Q}(G)$ in (4.16) by its k-th truncation $\mathcal{Q}_k(G)$, we get

$$h_k^{\text{sos}} := \inf \quad h(x) \quad \text{s.t. } x \in \mathcal{F}_k. \tag{4.18}$$

Consequently, it follows from the proof of Theorem 3.3 that

Theorem 4.5. *If $h(X)$ is convex, $\mathcal{Q}(G)$ is Archimedean and the Slater condition holds for* (4.15)*, then h_k^{sos} decreasingly converges to h^* as $k \to \infty$.*

Moreover, if $h(X)$ is s.o.s-convex, we point out that for each $k \geq d_P$, (4.18) is equivalent to a *single* SDP problem under certain conditions as shown in [18]. In fact, it is easy to see that there exist some integers l, t and real $t \times t$ symmetric matrices $\{A_i\}_{i=0}^m$ and $\{B_j\}_{j=1}^l$ such that $\mathcal{F}_k$ is identical with

$$\left\{ x \in \mathbb{R}^m \;\middle|\; \exists w \in \mathbb{R}^l, \text{ s.t. } A_0 + \sum_{i=1}^m A_i x_i + \sum_{j=1}^l B_j w_j \succeq 0 \right\}, \tag{4.19}$$

where w_j's correspond to the entries of Z_j's in (3.8). Thus, (4.18) becomes

$$\begin{cases} h_k^{\text{sos}} = \inf\limits_{x \in \mathbb{R}^m, w \in \mathbb{R}^l} & h(x) \\ & \text{s.t. } A_0 + \sum\limits_{i=1}^m A_i x_i + \sum\limits_{j=1}^l B_j w_j \succeq 0. \end{cases} \tag{4.20}$$

Let $d \in \mathbb{N}$ is the *smallest* even number such that $d \geq \deg h$. Denote the variables $W = (W_1, \ldots, W_l)$ and let $\Sigma_d^2[X, W]$ be the set of sums of squares of polynomials in $\mathbb{R}[X, W]$ of degree up to d. Consider the dual problem of (4.20)

$$\begin{cases} \lambda_k := \sup\limits_{\lambda \in \mathbb{R}, V \in \mathbb{S}_+^m} \lambda \\ \qquad \text{s.t. } h(X) - \sum\limits_{i=1}^m \langle A_i, V \rangle \cdot X_i - \sum\limits_{j=1}^l \langle B_j, V \rangle \cdot W_j - \langle A_0, V \rangle - \lambda \\ \qquad \in \Sigma_d^2[X, W], \end{cases} \tag{4.21}$$

which can be reduced to an SDP problem as shown in Section 3. Clearly, $h_k^{\text{sos}} \geq \lambda_k$. We have $h_k^{\text{sos}} = \lambda_k$ under certain conditions (c.f. [18, Theorem 3.1]). Therefore, an SDP relaxation method is drived for (4.15) with s.o.s-convex objectives.

Example 4.6. [6] Consider the followin SIP problem

$$\begin{cases} h^* = \inf\limits_{x \in \mathbb{R}^3} x_1^2 + x_2^2 + x_3^2 \\ \qquad \text{s.t. } -(y_1 + y_2^2 + 1)x_1 - (y_1 y_2 - y_2^2)x_2 - (y_1 y_2 + y_2^2 + y_2)x_3 - 1 \geq 0, \\ \qquad \forall y \in [0,1]^2. \end{cases} \tag{4.22}$$

The exact minimum and minimizer are $h^* = 1$ and $x^* = (-1, 0, 0)$ [6]. Clearly, the objective function is s.o.s-convex. Solving the SDP problem (4.21), we obtain $\lambda_1 = 1.0000$ with approximate minimizer $(-1.0000, 2.3149 \times 10^{-6}, -1.0410 \times 10^{-5})$.

Example 4.7. Consider the following SIP problem

$$\begin{cases} h^* = \inf\limits_{x \in \mathbb{R}^2} h(x) = \frac{5}{8}x_1^2 - 4x_1 + \frac{3}{4}x_1 x_2 - 4x_2 + \frac{5}{8}x_2^2 + 8 \\ \qquad \text{s.t. } 1 - y_1 x_1 - y_2 x_2 \geq 0, \ \forall y \in \{y \in \mathbb{R}^2 \mid y_1^2 + y_2^2 = 1\}. \end{cases} \tag{4.23}$$

Clearly, $h(X)$ is s.o.s-convex. It is easy to see that the feasible set is the closed unit disk around the origin. If we let $f(X) = (X_1 - 2)^2 + \frac{(X_2-2)^2}{4}$, then $h(X) = f(\frac{(X_1-2)+(X_2-2)}{\sqrt{2}} +$

$2, \frac{-(X_1-2)+(X_2-2)}{\sqrt{2}}+2)$. Geometrically, for any $r>0$, the curve $h(X)=r$ can be obtained by rotating the ellipse $f(X)=r$ around $(2,2)$ by $45°$ counterclockwise. Therefore, the minimizer is $x^*=(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2})$ with the minimum $h^*=9-4\sqrt{2}\approx 3.3431$. Solving the SDP problem (4.21), we obtain $\lambda_1=3.3432$ with approximate minimizer $(0.7071, 0.7071)$.

5 Conclusion

In this paper, a hierarchy of SDP relaxations for LSIPP problems is presented. It can be seen as the dual of Lasserre's relaxations for GPM problems and enjoys several desirable features. Some (approximate) minimizers of LSIPP problems can be extracted using these SDP relaxations, which is useful in some applications. Convergence rate of these SDP relaxations is estimated using some existing results. We also extend this SDP relaxation method to more general semi-infinite programming problems.

Acknowledgments:

The authors are very grateful for the comments of two anonymous referees which helped to improve the presentation. The authors would like to thank Miguel A. Goberna for helpful comments on the Lipschitz continuity of the optimal value function of (1.1) used in Lemma 3.10.

References

[1] A.A. Ahmadi, A. Olshevsky, P.A. Parrilo, and J.N. Tsitsiklis, NP-hardness of deciding convexity of quartic polynomials and related problems. *Math. Program.* 137 (2013) 453–476.

[2] N.I. Akhiezer, *The Classical Moment Problem*, Hafner, New York, NY, 1965.

[3] N.I. Akhiezer and M.G. Krein, *Some Questions in the Theory of Moment*, American Mathematical Society, Providence, RI, 1962.

[4] G.A. Anastassiou, *Moments in Probability Theory and Approximation Theory*, Pitman Research Notes in Mathematics Series. Longman Scientific & Technical, 1993.

[5] A. Charnes, W. W. Cooper and K. Kortanek. On representations of semi-infinite programs which have no duality gaps, *Manag. Sci.* 12 (1965) 113–121.

[6] I.D. Coope and G.A. Watson, A projected lagrangian algorithm for semi-infinite programming, *Math. Program.* 32 (1985) 337–356.

[7] R.E. Curto and L.A. Fialkow, Truncated K-moment problems in several variables, *J. Operator Theory* 54 (2005) 189–226.

[8] J. Demmel, J. Nie and V. Powers, Representations of positive polynomials on noncompact semialgebraic sets via KKT ideals, *J. Pure Appl. Algebra* 209 (2007) 189 – 200.

[9] K. Glashoff and S.A. Gustafson, *Linear Optimization and Approximation.* Springer-Verlag, Berlin, 1983.

[10] M.Á. Goberna, Linear semi-infinite optimization: Recent advances, in: *Continuous Optimization*, volume 99 of Applied Optimization, V. Jeyakumar and A. Rubinov (eds), Springer, US, 2005, pp. 3–22.

[11] M.Á. Goberna and M.A. López, *Linear Semi-Infinite Optimization*, John Wiley & Sons, Chichester, 1998.

[12] M.A. Goberna and M.A. López, *Post-Optimal Analysis in Linear Semi-Infinite Optimization*, SpringerBriefs in Optimization, Springer, New York, NY, 2014.

[13] F. Guo, C. Wang and L. Zhi, Semidefinite representations of noncompact convex sets. *SIAM J. Optim.* 25 (2015) 377–395.

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

[15] D. Henrion and J.B. Lasserre, Detecting global optimality and extracting solutions in gloptipoly. in: *Positive Polynomials in Control*, D. Henrion and A. Garulli (eds), Springer, Berlin, Heidelberg, 2005, pp. 293–310.

[16] D. Henrion, J.B. Lasserre and J. Löfberg, GloptiPoly 3: moments, optimization and semidefinite programming. *Optim. Methods Softw.* 24 (2009) 761–779.

[17] R. Hettich and K.O. Kortanek, Semi-infinite programming: Theory, methods, and applications. *SIAM Rev.* 35 (1993) 380–429.

[18] V. Jeyakumar and G. Li, Exact SDP relaxations for classes of nonlinear semidefinite programming problems, *Oper. Res. Lett.* 40 (2012) 529–536.

[19] J.H.B. Kemperman, Geometry of the moment problem, in: *Moments in Mathematics*, Proceedings of Symposia in Applied Mathematics 37, H.J. Landau (edi.), American Mathematical Society, Providence, 1980, pp. 16–53.

[20] H.J. Landau, *Moments in Mathematics*, Proceedings of Symposia in Applied Mathematics 37, American Mathematical Society, Providence, 1980.

[21] J.B. Lasserre, Global optimization with polynomials and the problem of moments, *SIAM J. Optim.* 11 (2001) 796–817.

[22] J.B. Lasserre, Convergent SDP-relaxations in polynomial optimization with sparsity, *SIAM J. Optim.* 17 (2006) 822–843.

[23] J.B. Lasserre, A semidefinite programming approach to the generalized problem of moments. *Math. Program.* 112 (2008) 65–92.

[24] J.B. Lasserre, An algorithm for semi-infinite polynomial optimization, *TOP* 20 (2012) 119–129.

[25] J.B. Lasserre and M. Putinar, Positivity and optimization for semi-algebraic functions, *SIAM J. Optim.* 20 (2010) 3364–3383.

[26] M. Laurent, Sums of squares, moment matrices and optimization over polynomials. in: *Emerging Applications of Algebraic Geometry*, M. Putinar and S. Sullivant (eds), Springer, New York, NY, 2009, pp.157–270.

[27] J. Löfberg, YALMIP : a toolbox for modeling and optimization in MATLAB. in: *2004 IEEE International Conference on Robotics and Automation (IEEE Cat. No.04CH37508)*, 2004, pp. 284–289

[28] M. López and G. Still, Semi-infinite programming. *European J. Oper. Res.* 180 (2007) 491–518, 2007.

[29] Z.-Q. Luo, N. D. Sidiropoulos, P. Tseng and S. Zhang, Approximation bounds for quadratic optimization with homogeneous quadratic constraints, *SIAM J. Optim.* 18 (2007) 1–28.

[30] M. Marshall,, Optimization of polynomial functions. *Canad. Math. Bull.* 46 (2003) 575–587.

[31] J. Nie, Discriminants and nonnegative polynomials, *J. Symbolic Comput.* 47 (2012) 167–191.

[32] J. Nie, Certifying convergence of Lasserre's hierarchy via flat truncation, *Math. Program. Ser. A* 142 (2013) 485–510.

[33] J. Nie, An exact Jacobian SDP relaxation for polynomial optimization, *Math. Program. Ser. A* 137 (2013) 225–255.

[34] J. Nie, Linear optimization with cones of moments and nonnegative polynomials, *Math. Program.* 153 (2015) 247–274.

[35] J. Nie and M. Schweighofer, On the complexity of Putinar's positivstellensatz, *J. Complexity* 23 (2007) 135 – 150.

[36] P.M. Pardalos and S.A. Vavasis, Quadratic programming with one negative eigenvalue is NP-hard, *J. Global Optim.* 1 (1991) 15–22.

[37] P. Parpas and B. Rustem, An algorithm for the global optimization of a class of continuous minimax problems, *J. Optim. Theory Appl.* 141 (2009) 461–473.

[38] V. Powers and B. Reznick, Polynomials that are positive on an interval, *Trans. Amer. Math. Soc.* 352 (2000) 4677–4692.

[39] M. Putinar, Positive polynomials on compact semi-algebraic sets, *Indiana Univ. Math. J.* 42 (1993) 969–984.

[40] R.T. Rockafellar, *Convex Analysis.* Princeton Landmarks in Mathematics and Physics. Princeton University Press, Princeton, 1996.

[41] A. Shapiro, Semi-infinite programming, duality, discretization and optimality conditions, *Optimization* 58 (1999) 133–161.

[42] J.F. Sturm, Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones, *Optim. Methods Softw.* 11 (1999) 625–653.

[43] H. Waki, S. Kim, M. Kojima and M. Muramatsu, Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity, *SIAM J. Optim.* 17 (2006) 218–242.

[44] L. Wang and F. Guo, Semidefinite relaxations for semi-infinite polynomial programming, *Comput. Optim. Appl.* 58 (2013) 133–159.

[45] G.A. Watson, A multiple exchange algorithm for multivariate chebyshev approximation, *SIAM J. Numer. Anal.* 12 (1975) 46–52.

[46] Y. Xu, W. Sun and L. Qi, On solving a class of linear semi-infinite programming by SDP method, *Optimization* 64 (2015) 603–616.

Manuscript received 19 July 2018
revised 8 February 2019
accepted for publication 31 October 2019

Feng Guo
School of Mathematical Sciences
Dalian University of Technology
Dalian, 116024, China
E-mail address: fguo@dlut.edu.cn

Xiaoxia Sun
School of Mathematics
Dongbei University of Finance and Economics
Dalian, 116025, China
E-mail address: xiaoxiasun@dufe.edu.cn