CERTIFYING CONVERGENCE OF LASSERRE'S HIERARCHY VIA FLAT TRUNCATION

JIAWANG NIE

ABSTRACT. Consider the optimization problem of minimizing a polynomial function subject to polynomial constraints. A typical approach for solving it globally is applying Lasserre's hierarchy of semidefinite relaxations, based on either Putinar's or Schmüdgen's Positivstellensatz. A practical question in applications is: how to certify its convergence and get minimizers? In this paper, we propose *flat truncation* as a certificate for this purpose. Assume the set of global minimizers is nonempty and finite. Our main results are: i) Putinar type Lasserre's hierarchy has finite convergence if and only if flat truncation holds, under some generic assumptions; the same conclusion holds for the Schmüdgen type one under weaker assumptions. ii) Flat truncation is asymptotically satisfied for Putinar type Lasserre's hierarchy if the archimedean condition holds; the same conclusion holds for the Schmüdgen type one if the feasible set is compact. iii) We show that flat truncation can be used as a certificate to check exactness of standard SOS relaxations and Jacobian SDP relaxations.

1. INTRODUCTION

Given polynomials $f, g_1, \ldots, g_m$, consider the optimization problem

$$(1.1) \qquad \begin{cases} \min\limits_{x \in \mathbb{R}^n} & f(x) \\ \text{s.t.} & g_1(x) \geq 0, \ldots, g_m(x) \geq 0. \end{cases}$$

Let K be its feasible set, and f_{min} be its global minimum value. The k-th Lasserre's relaxation [6] for solving (1.1) is (k is also called the *relaxation order*)

$$(1.2) \qquad \max \quad \gamma \quad \text{s.t.} \ f - \gamma \in Q_k(g).$$

Here, the set $Q_k(g)$ denotes the k-th truncated quadratic module generated by the tuple $g := (g_1, \ldots, g_m)$ (for convenience, denote $g_0 = 1$ and say a polynomial is SOS if it is a sum of squares of polynomials with real coefficients):

$$Q_k(g) := \left\{ \sum_{i=0}^{m} g_i \sigma_i \ \middle| \ \sigma_i \text{ is SOS, } \deg(g_i \sigma_i) \leq 2k \text{ for every } i \right\}.$$

The relaxation (1.2) is equivalent to a semidefinite program (SDP), and thus could be solved efficiently by numerical methods like interior point algorithms. Let f_k be the optimal value of (1.2) for a given order k. Clearly, every $f_k \leq f_{min}$ and the sequence $\{f_k\}$ is monotonically increasing. Under the archimedean condition (i.e., there exists $\phi \in Q_\ell(g)$ for some ℓ such that the inequality $\phi(x) \geq 0$ defines a compact set in x), Lasserre proved the convergence $f_k \to f_{min}$ as $k \to \infty$ [6]. An estimation of its convergence rate is given in [14]. The sequence of (1.2) as $k \to \infty$ is called

Lasserre's hierarchy in the literature. As demonstrated by extensive numerical experiments, it occurs quite a lot that $f_k = f_{min}$ for a finite order k in applications. If this happens, we say Lasserre's hierarchy has *finite convergence.* This raises a very practical question: since f_{min} is typically unknown, how do we certify its finite convergence if it happens? If it is certified, how do we get minimizers? A frequently used sufficient condition for this purpose is *flat extension* introduced by Curto and Fialkow (cf. [3]), but it is a strong condition that might not be satisfied, i.e., it is not necessary. To the author's best knowledge, there is very little work on proving a certificate for checking finite convergence of Lasserre's hierarchy, and the question is almost completely open. The motivation of this paper is to address this issue. Our main result is that a more suitable condition called *flat truncation* could generically serve as such a certificate.

1.1. **Background.** Typically, to extract a global minimizer, one needs to consider the dual optimization problem of (1.2) whose description uses localizing matrices. Define degree integers

$$(1.3) \qquad d_i = \lceil \deg(g_i)/2 \rceil, \quad d_g = \max\{1, d_1, \ldots, d_m\}, \quad d_f = \lceil \deg(f)/2 \rceil.$$

(Here $\lceil a \rceil$ denotes the smallest integer that is greater than or equal to a.) Let y be a sequence indexed by $\alpha := (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ ($\mathbb{N}$ is the set of nonnegative integers) with $|\alpha| := \alpha_1 + \cdots + \alpha_n \leq 2k$, i.e., y is a *truncated moment sequence (tms)* of degree $2k$. Denote by $\mathscr{M}_{2k}$ the space of all tms whose degrees are $2k$. A tms $y \in \mathscr{M}_{2k}$ defines a Riesz functional $\mathscr{L}_y$ on $\mathbb{R}[x]_{2k}$ (the space of real polynomials in $x := (x_1, \ldots, x_n)$ with degrees at most $2k$, and denote $\mathbb{R}[x] := \sum_i \mathbb{R}[x]_i$) as

$$\mathscr{L}_y \left(\sum_\alpha p_\alpha x_1^{\alpha_1} \cdots x_n^{\alpha_n} \right) := \sum_\alpha p_\alpha y_\alpha.$$

For convenience, denote $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ and

$$\langle p, y \rangle := \mathscr{L}_y(p).$$

The k-th *localizing matrix* $L_h^{(k)}(y)$ generated by a polynomial h and a tms $y \in \mathscr{M}_{2k}$ is a symmetric matrix satisfying (denote $d_h := \lceil \deg(h)/2 \rceil$)

$$p^T L_h^{(k)}(y)q := \mathscr{L}_y(hpq) \quad \forall\, p, q \in \mathbb{R}[x]_{k-d_h}.$$

(Here, for convenience, we still use p to denote the coefficient vector of the polynomial p.) When $h = 1$, $L_h^{(k)}(y)$ is called a *moment matrix* and is denoted as

$$M_k(y) := L_1^{(k)}(y).$$

The columns and rows of $L_h^{(k)}(y)$, as well as $M_k(y)$, are indexed by integral vectors $\alpha \in \mathbb{N}^n$ with $|\alpha| \leq k - d_h$.

The dual optimization problem of (1.2) is (cf. [6, 9])

$$(1.4) \qquad \begin{cases} \min\limits_{y \in \mathscr{M}_{2k}} & \langle f, y \rangle \\ \text{s.t.} & L_{g_i}^{(k)}(y) \succeq 0 \, (i = 0, 1, \ldots, m), \, \langle 1, y \rangle = 1. \end{cases}$$

In the above, $X \succeq 0$ means a matrix X is positive semidefinite. Let f_k^* be the optimal value of (1.4) for order k. By weak duality, $f_k^* \leq f_k$ for every k. If K has nonempty interior, then (1.4) has an interior point, (1.2) achieves its optimal value and $f_k^* = f_k$, i.e., there is no duality gap between (1.2) and (1.4) (cf. [6]). Clearly,

every $f_k^* \leq f_{min}$, and the sequence $\{f_k^*\}$ is also monotonically increasing. We refer to Lasserre's book [9] and Laurent's survey [11] for related work in this area.

Suppose y^* is an optimizer of (1.4). If the *flat extension* condition (cf. [3])

$$(1.5) \qquad \operatorname{rank} M_{k-d_g}(y^*) \quad = \quad \operatorname{rank} M_k(y^*)$$

holds (d_g is from (1.3)), then we can extract $r = \operatorname{rank} M_k(y^*)$ global optimizers for (1.1). By Theorem 1.1 of Curto and Fialkow [3], if y^* is feasible for (1.4) and (1.5) is satisfied, then y^* admits a unique r-atomic measure supported on K, i.e., there exist $(\lambda_1, \ldots, \lambda_r) > 0$ and r distinct points $v_1, \ldots, v_r \in K$ such that

$$(1.6) \qquad y^* \quad = \quad \lambda_1 [v_1]_{2k} + \cdots + \lambda_r [v_r]_{2k}.$$

Here, for $x \in \mathbb{R}^n$, $[x]_{2k}$ denotes the vector defined as

$$[x]_{2k} := \begin{bmatrix} 1 & x_1 & \cdots & x_n & x_1^2 & x_1 x_2 & \cdots & x_n^{2k} \end{bmatrix}^T.$$

In (1.6), the constraint $\langle 1, y^* \rangle = 1$ implies $\lambda_1 + \cdots + \lambda_r = 1$. If (1.5) holds, then $f_k = f_{min}$ (suppose there is no duality gap between (1.2) and (1.4)), all $v_1, \ldots, v_r$ are global minimizers of (1.1), and they could typically be obtained by solving some SVD and eigenvalue problems, as shown by Henrion and Lasserre [5]. Generally, (1.5) is a sufficient but not necessary condition for checking finite convergence of Lasserre's hierarchy.

For a tms $z \in \mathscr{M}_{2t}$, we say z is *flat with respect to g* if z is feasible in (1.4) for $k = t$ and satisfies the condition $\operatorname{rank} M_{t-d_g}(z) = \operatorname{rank} M_t(z)$ (cf. [3]). If the tuple g is clear in the context, we just simply say z is flat.

1.2. **Flat truncation.** To get a minimizer of (1.1) from an optimizer y^* of (1.4), the flat extension condition (1.5) would be weakened. For instance, if y^* has a *flat truncation*, i.e., there exists an integer $t \in [\max\{d_f, d_g\}, k]$ such that

$$(1.7) \qquad \operatorname{rank} M_{t-d_g}(y^*) \quad = \quad \operatorname{rank} M_t(y^*),$$

then we could still extract $r = \operatorname{rank} M_t(y^*)$ minimizers for (1.1). Let $z = y^*|_{2t}$ (denote $y^*|_{2t} := (y_\alpha^*)_{|\alpha| \leq 2t}$) be the truncation. Then, the tms z is flat, and there exist r distinct points $u_1, \ldots, u_r \in K$ and scalars $\lambda_1, \ldots, \lambda_r$ such that

$$z = \lambda_1 [u_1]_{2t} + \cdots + \lambda_r [u_r]_{2t},$$

$$\lambda_1 > 0, \ldots, \lambda_r > 0, \quad \lambda_1 + \cdots + \lambda_r = 1.$$

Clearly, $\langle f, z \rangle = \langle f, y^* \rangle = f_k^* \leq f_{min}$ (note $2t \geq \deg(f)$) and

$$\lambda_1 f(u_1) + \cdots + \lambda_r f(u_r) = \langle f, z \rangle = \langle f, y^* \rangle \leq f_{min}.$$

Since every $f(u_i) \geq f_{min}$, each u_i must be a global minimizer of (1.1). Hence, if an optimizer of (1.4) has a flat truncation, then $f_k^* = f_{min}$, and $f_k = f_{min}$ if there is no duality gap between (1.2) and (1.4).

The flat extension condition (1.5) and flat truncation condition (1.7) are different in checking convergence of Lasserre's hierarchy. We show the difference in the following example.

Example 1.1. Consider the univariate cubic optimization problem

$$\min_{x \in \mathbb{R}} \quad x^3 \quad \text{s.t.} \quad x \geq 0, 1 - x \geq 0.$$

The feasible set is compact and has interior points, and the archimedean condition holds. The global minimum $f_{min} = 0$ and the origin is the unique global minimizer. It could be easily seen that $f_k^* = f_k = 0$ for all $k \geq 2$. So, its Lasserre's hierarchy

has finite convergence. Both (1.2) and (1.4) achieve their optimal values. However, the flat extension condition (1.5) is violated for almost all optimizers of (1.4), while the flat truncation condition (1.7) is satisfied for all the optimizers. To see this, let y^* be an optimizer of (1.4), then $y_3^* = 0$. It could be easily shown that the constraints of (1.4) implies that $y_\alpha^* = 0$ for all $|\alpha| \le 2k - 1$. (For instance, for $k = 2$, its constraints are

$$M_2(y) = \begin{bmatrix} 1 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{bmatrix} \succeq 0, \qquad L_x^{(2)}(y) = \begin{bmatrix} y_1 & y_2 \\ y_2 & y_3 \end{bmatrix} \succeq 0,$$

$$L_{1-x}^{(3)}(y) = \begin{bmatrix} 1 - y_1 & y_1 - y_2 \\ y_1 - y_2 & y_2 - y_3 \end{bmatrix} \succeq 0,$$

and $y_3^* = 0$ implies $y_1^* = y_2^* = 0$.) Thus, for each $k \ge 2$, every optimizer of (1.4) is in the form

$$y^*(\epsilon) := (1, 0, \dots, 0, \epsilon)$$

with a paramter $\epsilon \ge 0$. If (1.4) is solved by interior-point methods, then typically a value $\epsilon > 0$ is returned. For instance, for $k = 2$, a numerical value 0.7908 of ϵ is returned when (1.4) is solved by the SDP package `SeDuMi` [24]. Clearly, (1.5) is violated for $y^*(\epsilon)$ with $\epsilon > 0$, while (1.7) is always satisfied for $y^*(\epsilon)$ with every $\epsilon \ge 0$ and $1 \le t \le k-1$. This shows that flat truncation is a more suitable condition than flat extension in certifying convergence of Lasserre's hierachy.   □

The rank condition (1.7) was used in [9, 11] as a sufficient condition to verify exactness of Lasserre's relaxations in polynomial optimization and generalized problems of moments. When the feasible set K is defined by polynomial equalities $h_1(x) = \cdots = h_s(x) = 0$, if the equations $h_i(x) = 0$ have finitely many zeros, then Lasserre's hierarchy has finite convergence (cf. [10, §3] or [11, §6.5]), and for k big enough every tms y that is feasible for the resulting (1.4) has a flat truncation (cf. [7, Prop. 4.6] or [11, Theorem 6.20]). For general polynomial optimization problems, there are no such results in the existing literature.

As we have seen earlier, for an optimizer of (1.4) to have a flat truncation, a necessary condition is $f_k = f_{min}$ (suppose (1.2) and (1.4) has no duality gap). That is, flat truncation is a sufficient condition for Lasserre's hierarchy to have finite convergence. Thus, one is wondering whether flat truncation is also necessary: if Lasserre's hierarchy converges in finitely many steps, does *every* [1] optimizer of (1.4) have a flat truncation? If so, flat truncation can be used as a certificate for checking its finite convergence and some minimizers of (1.1) can also be obtained. This issue was addressed very little in the literature.

Another important issue in applications is how to certify asymptotic convergence of Lasserre's hierarchy. Under the archimedean condition, we know $f_k \to f_{min}$ as $k \to \infty$ (cf. [6]), but the convergence to the set of global minimizers of (1.1) is not known very well. When (1.1) has a unique global minimizer u^*, Schweighofer [23] showed that the subvector consisting of linear moments of a nearly optimizer of (1.4) converges to u^* as $k \to \infty$. However, for the more general case that (1.1) has more than one global minimizer, there are no such results in the existing literature. It is possible that $f_k = f_k^* < f_{min}$ for every k, i.e., Lasserre's hierarchy may have no finite convergence, as implied by Scheiderer's work [20]. For instance, for the

---

[1] If $f_k = f_{min}$ and (1.1) has a minimizer, say x^*, then (1.4) always has an optimizer that is flat, e.g., $[x^*]_{2k}$. Thus, it is more sensible to ask whether every optimizer has a flat truncatioin.

problem of minimizing the Motzkin polynomial over the unit ball, the Lasserre's hierarchy does not converge within finitely many steps (cf. [16, Example 5.3]). In such cases, one should not expect any minimizer of (1.4) to have a flat truncation. However, how about the limit points of truncations of minimizers of (1.4) as $k \to \infty$? Is every such limit point flat, or does it have a flat truncation? There is very little work on this issue.

1.3. **Contributions.** This paper focuses on proving a certificate for checking convergence of Lasserre's hierarchy. We assume (1.1) has a nonempty set S of global minimizers and its cardinality $|S|$ is finite. Our main result is that flat truncation can generically serve as such a certificate.

First, we study how to certify finite convergence. For Lasserre's hierarchy of (1.2) (it is also called a Putinar type one), we show that: it has finite convergence if and only if every minimizer of (1.4) has a flat truncation when k is sufficiently large, under some generic conditions (e.g., Assumption 2.1). For Schmüdgen type Lasserre's hierarchy, which is a refined version of (1.2) by using cross products of g_j (cf. (2.6)), we show the same conclusion holds under some slightly weaker conditions. This will be shown in Section 2.

Second, we study how to certify asymptotical convergence. Let $\{y^{(k)}\}_{k=1}^{\infty}$ be a sequence of asymptotically optimal solutions of (1.4). We prove that: under the archimedean condition, for any fixed order $t \geq \max\{d_f, d_g + |S| - 1\}$, the truncated sequence $\{y^{(k)}|_{2t}\}_{k=1}^{\infty}$ is bounded, and its every limit point is flat, i.e., flat truncation is asymptotically satisfied for both Lasserre's hierarchies. The same conclusion holds for the Schmüdgen type one when K is compact (the archimedean condition is not required). This will be shown in Section 3.

Third, we show that flat truncation can be used as a certificate to check exactness of standard SOS relaxations and Jacobian SDP relaxations. This will be shown in Section 4.

Last, we make some discussions about flat truncation in Section 5.

2. CERTIFYING FINITE CONVERGENCE

As we have seen in §1.2, flat truncation is a sufficient condition for Lasserre's hierarchy of (1.2) to converge in finitely many steps. In this section, we show that flat truncation is also a necessary condition in the generic case. Thus, it could generically serve as a certificate for checking finite convergence of (1.2). Lasserre's hierarchy of (1.2) is also called a Putinar type one, since it uses Putinar's Positivstellensatz [19] in representing positive polynomials. A refining of (1.2) is Schmüdgen type Lasserre's hierarchy, which uses cross products of the constraining polynomials. Similarly, flat truncation could also be generically used as a certificate for checking finite convergence of the Schmüdgen type one. We present the results in two separate subsections.

2.1. **Putinar Type Lasserre's relaxation.** The quadratic module generated by the tuple g is

$$Q(g) \quad := \quad \bigcup_{k=1}^{\infty} Q_k(g).$$

The archimedean condition requires that there exists $\phi \in Q(g)$ such that the inequality $\phi(x) \geq 0$ defines a compact set in x. Note that the archimedean condition

implies the feasible set K is compact. The convergence for Lasserre's hierarchy of (1.2) is based on Putinar's Positivstellensatz [19]: if a polynomial p is positive on K and the archimedean condition holds, then $p \in Q(g)$. To certify finite convergence of Putinar type Lasserre's hierarchy, we need the following assumption on $f, g_1, \ldots, g_m$.

Assumption 2.1. *There exists* $\rho \in Q(g)$ *such that for every* $J \subseteq \{1, \ldots, m\}$ *and*

$$\mathcal{V}_J := \{x \in \mathbb{R}^n : g_j(x) = 0 \ (\forall j \in J), \ rank \, G_J(x) \leq |J|\},$$

the intersection $\mathcal{V}_J \cap \mathcal{M} \cap \mathcal{P}$ *is finite. Here, denote* $J = \{j_1, \ldots, j_l\}$,

$$G_J(x) := \begin{bmatrix} \nabla f(x) & \nabla g_{j_1}(x) & \cdots & \nabla g_{j_l}(x) \end{bmatrix},$$

and $\mathcal{M} := \{x \in \mathbb{R}^n : f(x) = f_{min}\}$, $\mathcal{P} := \{x \in \mathbb{R}^n : \rho(x) \geq 0\}$.

Assumption 2.1 requires that for every J the optimization problem

$$\min \quad f(x) \quad \text{s.t.} \quad g_j(x) = 0 \ (j \in J)$$

has finitely many critical points lying on $\mathcal{M} \cap \mathcal{P}$ (if u is a critical point of the above, then $rank \, G_J(u) \leq |J|$, cf. [15, §2]). Let S be the set of global minimizers of (1.1) and

$$\mathcal{V} := \cup_J \mathcal{V}_J.$$

Clearly, $S \subseteq \mathcal{V}$, and Assumption 2.1 implies that S is finite.

We would like to remark that Assumption 2.1 is *generically* true. (Here, we say a property is generically true if it holds in the space of data except a set of Lebesgue measure zero, cf. [15]). As shown by Proposition 2.1 of [15], if f and every g_j have generic coefficients, then $\mathcal{V}_J$ is finite for every index set J, and Assumption 2.1 is satisfied by simply choosing $\rho = 0$.

Assumption 2.1 also holds in the following cases:

- Suppose S is finite, $\mathcal{V} \cap \mathcal{M}$ is infinite, but $(\mathcal{V} \cap \mathcal{M}) \backslash K$ is contained in a compact set T not intersecting K. Then, there exists $\delta > 0$ such that

  $$dist(x, K) \geq \delta \ \forall \, x \in T, \qquad dist(x, K) = 0 \ \forall \, x \in K.$$

  (Here $dist(x, K) = \min_{u \in K} \|x - u\|_2$ and $\| \cdot \|_2$ denotes the standard 2-norm.) The function $\tau(x) := dist(x, K) - \delta/2$ is continuous in x. Then, by the compactness of K and T, there exists a polynomial η that is an approximation of τ and satisfies

  $$\eta(x) \geq \delta/4 \ \forall \, x \in T, \qquad \eta(x) \leq -\delta/4 \ \forall \, x \in K.$$

  Clearly, $-\eta$ is positive on K. Assume the archimedean condition holds for g, then $-\eta \in Q(g)$ by Putinar's Positivstellensatz. Assumption 2.1 is satisfied by choosing $\rho = -\eta$, because $\mathcal{V} \cap \mathcal{M} \cap \mathcal{P} = S$ is finite.

- Suppose $\mathcal{V} \cap \mathcal{M}$ is unbounded but, except finitely many points, lies on a real algebraic variety

  $$\{x \in \mathbb{R}^n : w_1(x) = \cdots = w_r(x) = 0\}$$

  not intersecting K. Suppose the archimedean condition holds for g, then K is compact and there exists $\epsilon > 0$ such that the polynomial $w := w_1^2 + \cdots + w_r^2 - \epsilon$ is positive on K. By Putinar's Positivstellensatz, $w \in Q(g)$. Assumption 2.1 is satisfied by choosing $\rho = w$.

Our main result of this subsection is the following theorem.

Theorem 2.2. *Suppose Assumption 2.1 holds, the set S of global minimizers of (1.1) is nonempty, and for k big enough the optimal value of (1.2) is achievable and there is no duality gap between (1.2) and (1.4). Then, Lasserre's hierarchy of (1.2) has finite convergence if and only if every minimizer of (1.4) has a flat truncation for k sufficiently large.*

Remark 2.3. In Theorem 2.2, 1) if K has nonempty interior, then for every order k (1.2) achieves its optimum, and it has no duality gap; 2) if finite convergence occurs, then (1.4) always has a minimizer for k big enough (e.g., $[x^*]_{2k}$ is one for any $x^* \in S$), and there is no duality gap; 3) when k is big enough, for every minimizer y^* of (1.4), $y^*|_{2k-2}$ is always flat, as shown in the proof later; 4) we do not need to assume K is compact or the archimedean condition holds.

Remark 2.4. As pointed out in Laurent's survey [11, §6.6], we could get the following properties about flat truncation. 1) In Theorem 2.2, Assumption 2.1 implies S is finite. If rank $M_k(y^*)$ is maximum over the optimizers of (1.4), then for any flat truncation of y^*, say, $y^*|_{2t}$, rank $M_t(y^*) = |S|$. This means that all the minimizers of (1.1) could be extracted from $y^*|_{2t}$. 2) When S is infinite and rank $M_k(y^*)$ is maximum over the optimizers of (1.4), then y^* could not have a flat truncation. 3) When (1.2) and (1.4) are solved by primal-dual interior-point algorithms, a minimizer y^* near the analytic center of the face of optimizers of (1.4) is usually returned and rank $M_k(y^*)$ is typically maximum. Therefore, if the conditions of Theorem 2.2 are satisfied, by solving a sufficient high order Lasserre's relaxation via interior point methods, then we could typically find all minimizers of (1.1) if S is finite, and flat truncation could not be satisfied if S is infinite.

To prove Theorem 2.2, we need some properties about the kernels of moment and localizing matrices. (Given a polynomial p, we also denote its coefficient vector by p, for convenience of notations.) For a localizing matrix $L_h^{(k)}(y)$, recall that if $\deg(hp^2) \leq 2k$ then

$$p^T \cdot L_h^{(k)}(y) \cdot p \quad = \quad \mathscr{L}_y(hp^2).$$

If $L_h^{(k)}(y) \cdot p = 0$, we say $p \in \ker L_h^{(k)}(y)$. When $M_k(y) = L_1^{(k)}(y)$ is a moment matrix, we similarly say $p \in \ker M_k(y)$ if $M_k(y) \cdot p = 0$.

Lemma 2.5. *Let $y \in \mathscr{M}_{2k}, h \in \mathbb{R}[x]$ be such that $L_h^{(k)}(y) \succeq 0$.*

i) *([7, 11]) Suppose $M_k(y) \succeq 0$. Let $p, q \in \mathbb{R}[x]$. If $\deg(pq) \leq k - 1$ and $q \in \ker M_k(y)$, then $pq \in \ker M_k(y)$. If $q^\ell \in \ker M_k(y)$ and $2\lceil \ell/2 \rceil \deg(q) \leq k - 1$, then $q \in \ker M_k(y)$.*

ii) *([4]) Let s be an SOS polynomial with $\deg(hs) \leq 2k$. Then, $\langle hs, y \rangle \geq 0$. If $\langle hs, y \rangle = 0$, then for any $\phi \in \mathbb{R}[x]_{2\ell}$ with $\deg(hs) + 2\ell \leq 2k - 2$ we have $\langle hs\phi, y \rangle = 0$.*

iii) *Let $\{p_j\}_{j=1}^\infty \subset \mathbb{R}[x]$ be a sequence such that each $\deg(hp_j^2) \leq 2k$ and*

$$L_h^{(k)}(y) \cdot p_j \to 0 \quad as \quad j \to \infty.$$

If $q \in \mathbb{R}[x]$ and every $\deg(p_j q) \leq k - \lceil \deg(h)/2 \rceil - 1$, then

$$L_h^{(k)}(y) \cdot (p_j q) \to 0 \quad as \quad j \to \infty.$$

iv) *Let $\{s_j\}_{j=1}^\infty$ be a sequence of SOS polynomials such that*

$$\deg(hs_j) < 2k \ (\forall j), \quad and \quad \langle hs_j, y \rangle \to 0 \ as \ j \to \infty.$$

If $\phi \in \mathbb{R}[x]_{2\ell}$ and every $\deg(hs_j) + 2\ell \leq 2k - 2$, then

$$\langle hs_j\phi, y \rangle \to 0 \quad as \quad j \to \infty.$$

Proof. i) The first part is basically from [11, Lemma 5.7] or [10, Lemma 21]. For the second part, if ℓ is even, the result follows Lemma 3.9 of [7]; if ℓ is odd, then $q^{\ell+1} \in \ker M_k(y)$ from the first part, and the result is still true.

ii) Write $s = \sum_i s_i^2$, then

$$\langle hs, y \rangle = \sum_i \langle hs_i^2, y \rangle = \sum_i s_i^T L_h^{(k)}(y) s_i \geq 0.$$

Now, assume $\langle hs, y \rangle = 0$. One could always write $\phi = \sum_j p_j q_j$ with all $p_j, q_j \in \mathbb{R}[x]_\ell$. Then, we have

$$\langle hs, y \rangle = \sum_i s_i^T L_h^{(k)}(y) s_i, \quad \langle hs\phi, y \rangle = \sum_{i,j} (s_i p_j)^T L_h^{(k)}(y)(s_i q_j).$$

Since $L_h^{(k)}(y) \succeq 0$, $\langle hs, y \rangle = 0$ implies every $L_h^{(k)}(y)s_i = 0$. Then, by Lemma 2.3 of [4], every $L_h^{(k)}(y)(s_i q_j) = 0$ and $(s_i p_j)^T L_h^{(k)}(y)(s_i q_j) = 0$. So, $\langle hs\phi, y \rangle = 0$.

iii) Let $r = k - \lceil \deg(h)/2 \rceil$. By a simple induction on $\deg(q)$, it suffices to prove the lemma for $q = x_i$ ($1 \leq i \leq n$). Note that every $\deg(p_j) + 1 \leq r - 1$. Let

$$u^{(j)} := L_h^{(k)}(y) \cdot p_j, \qquad v^{(j)} := L_h^{(k)}(y) \cdot (p_j \cdot x_i).$$

They are indexed by integral vectors $\alpha \in \mathbb{N}^n$, and could also be expressed as

$$u^{(j)} = \mathscr{L}_y\big(p_j \cdot h \cdot \mathfrak{m}_r\big), \quad v^{(j)} = \mathscr{L}_y\big(x_i \cdot p_j \cdot h \cdot \mathfrak{m}_r\big).$$

(Here, $\mathfrak{m}_r$ is the vector of monomials of degrees $\leq r$.) Then $u^{(j)} \to 0$ implies

$$\lim_{j\to\infty} \mathscr{L}_y\big(p_j \cdot h \cdot x^\beta\big) = 0 \quad \text{if} \quad |\beta| \leq r.$$

The α-th entry of $v^{(j)}$ is

$$v_\alpha^{(j)} := \mathscr{L}_y\big(p_j \cdot h \cdot x^\alpha \cdot x_i\big).$$

If $\deg(x^\alpha \cdot x_i) \leq r$, then $v_\alpha^{(j)} = u_{\tilde{\alpha}}^{(j)} \to 0$ ($\tilde{\alpha}$ is the exponent of $x^\alpha \cdot x_i$). So,

$$\lim_{j\to\infty} v_\alpha^{(j)} = 0 \quad \text{if} \quad |\alpha| \leq r - 1.$$

This implies that (note $L_h^{(k-1)}(y)$ is a leading principal submatrix of $L_h^{(k)}(y)$)

$$\lim_{j\to\infty} \Big(L_h^{(k-1)}(y)\Big) \cdot (p_j \cdot x_i) = \lim_{j\to\infty} \mathscr{L}_y\big(x_i \cdot p_j \cdot h \cdot \mathfrak{m}_{r-1}\big) = 0.$$

Since $L_h^{(k-1)}(y)$ is symmetric and positive semidefinite, there exist orthogonal vectors $a_1, \ldots, a_t$ such that $L_h^{(k-1)}(y) = a_1 a_1^T + \cdots + a_t a_t^T$. So, we have

$$a_\ell^T (p_j \cdot x_i) \to 0 \quad \text{as } j \to \infty, \quad \ell = 1, \ldots, t.$$

By the above, it holds that

$$\lim_{j\to\infty} (p_j \cdot x_i)^T \cdot \Big(L_h^{(k-1)}(y)\Big) \cdot (p_j \cdot x_i) = \lim_{j\to\infty} \sum_{\ell=1}^t \big(a_\ell^T (p_j \cdot x_i)\big)^2 = 0.$$

When $|\alpha| > r - 1$, the coefficient of x^α in the polynomial $p_j \cdot x_i$ is zero. So

$$(p_j \cdot x_i)^T \cdot \Big(L_h^{(k)}(y)\Big) \cdot (p_j \cdot x_i) = (p_j \cdot x_i)^T \cdot \Big(L_h^{(k-1)}(y)\Big) \cdot (p_j \cdot x_i).$$

Hence, we also have

$$\lim_{j \to \infty} (p_j \cdot x_i)^T \cdot \left(L_h^{(k)}(y) \right) \cdot (p_j \cdot x_i) = 0.$$

Since $L_h^{(k)}(y) \succeq 0$, the above implies that

$$\lim_{j \to \infty} \left(L_h^{(k)}(y) \right) \cdot (p_j \cdot x_i) = 0.$$

iv) Write $s_j = \sum_t p_{j,t}^2$ (the length of summation is at most $\binom{n+k}{k}$, since $\deg(s_j) \leq 2k$). By item ii), $\langle h s_j, y \rangle \geq \langle h p_{j,t}^2, y \rangle \geq 0$. So, $\langle h s_j, y \rangle \to 0$ implies that for each t

$$\langle h p_{j,t}^2, y \rangle = (p_{j,t})^T L_h^{(k)}(y) p_{j,t} \to 0 \quad \text{as } j \to \infty.$$

Since $L_h^{(k)}(y) \succeq 0$, we have $L_h^{(k)}(y) p_{j,t} \to 0$ as $j \to \infty$. Like in item ii), one could write $\phi = \sum_i u_i v_i$ with all $u_i, v_i \in \mathbb{R}[x]_\ell$. By item iii), it holds that for all t, i

$$L_h^{(k)}(y)(p_{j,t} u_i) \to 0, \ L_h^{(k)}(y)(p_{j,t} v_i) \to 0 \quad \text{as } j \to \infty.$$

So, we also have

$$\left(L_h^{(k)}(y) \right)^{\frac{1}{2}} (p_{j,t} u_i) \to 0, \ \left(L_h^{(k)}(y) \right)^{\frac{1}{2}} (p_{j,t} v_i) \to 0 \quad \text{as } j \to \infty.$$

(The $X^{\frac{1}{2}}$ denotes the standard matrix square root of a symmetric positive semidefinite matrix X.) The above implies that

$$(p_{j,t} u_i)^T L_h^{(k)}(y)(p_{j,t} v_i) = (p_{j,t} u_i)^T \left(L_h^{(k)}(y) \right)^{\frac{1}{2}} \cdot \left(L_h^{(k)}(y) \right)^{\frac{1}{2}} (p_{j,t} v_i) \to 0$$

and

$$\langle h s_j \phi, y \rangle = \sum_{t,i} \langle h p_{j,t}^2 u_i v_i, y \rangle = \sum_{t,i} (p_{j,t} u_i)^T L_h^{(k)}(y)(p_{j,t} v_i) \to 0,$$

as $j \to \infty$. □

Proof of Theorem 2.2   The sufficiency of flat truncation was observed in §1.2, if there is no duality gap between (1.2) and (1.4). We only need to prove its necessity. Suppose $f_{k_0} = f_{min}$ for k_0 big enough. Since (1.2) has a maximizer when its order is big enough, we could assume (1.2) achieves its optimum f_{min} for order k_0 (otherwise increase k_0), i.e., $f - f_{min} \in Q_{k_0}(g)$. So, there exist SOS polynomials $s_0, s_1, \ldots, s_m$ such that every $\deg(g_i s_i) \leq 2k_0$ and

$$(2.1) \qquad\qquad f - f_{min} = s_0 + g_1 s_1 + \cdots + g_m s_m.$$

Let y^* be an arbitrary minimizer of (1.4) (it always has one when $k \geq k_0$, e.g., $[x^*]_{2k}$, for any $x^* \in S$). Clearly, $\langle f, y^* \rangle = f_{\min}$. Let C be the semialgebraic set defined as (ρ is from Assumption 2.1)

$$(2.2) \qquad C = \{x \in \mathbb{R}^n : s_0(x) = g_1(x)s_1(x) = \cdots = g_m(x)s_m(x) = 0, \rho(x) \geq 0\}.$$

We complete the proof in three steps.

Step 1   We show that C is a finite set. In the identity (2.1), differentiating its both sides in x results in (note $g_0 = 1$)

$$\nabla f = \sum_{i=0}^{m} \left(s_i \cdot \nabla g_i + g_i \nabla s_i \right).$$

Choose an arbitrary $u \in C$, then clearly $u \in \mathcal{M} \cap \mathcal{P}$. Let $J(u) = \{i : g_i(u) = 0\}$. Note that for every $i \notin J(u)$, $s_i(u) = 0$, and it implies $\nabla s_i(u) = 0$ (because s_i is SOS and u is a minimizer of s_i). So, by the above, it holds that

$$\nabla f(u) = \sum_{i \in J(u)} s_i(u) \cdot \nabla g_i(u).$$

Hence, $u \in \mathcal{V}_{J(u)} \cap \mathcal{M} \cap \mathcal{P}$. Since there are at most 2^m active index sets like $J(u)$, by Assumption 2.1, C must be finite.

Step 2 We show that every generator of the vanishing ideal of C belongs to the kernel of $M_k(y^*)$ for k big enough. Since C is finite, its vanishing ideal

$$I(C) := \{p \in \mathbb{R}[x] : p(u) = 0 \quad \forall u \in C\}$$

is zero dimensional. Let $\{h_1, \ldots, h_r\}$ be a Grobner basis of $I(C)$ with respect to a total degree ordering. Clearly, each h_i vanishes on C. By Real Nullstellensatz (cf. Corollary 4.4.3 of [1]), there exist $\ell \in \mathbb{N}$, $\phi_1, \ldots, \phi_m \in \mathbb{R}[x]$, and $\varphi \in Q(\rho)$ (the quadratic module generated by the single polynomial ρ, which is also equal to the prepordering generated by ρ) such that

$$h_i^{2\ell} + \varphi + g_1 s_1 \phi_1 + \cdots + g_m s_m \phi_m = 0.$$

Applying the Riesz functional $\mathscr{L}_{y^*}$ to the above (suppose $2k$ is bigger than the degrees of all the products there), we get

$$(2.3) \qquad \langle h_i^{2\ell}, y^* \rangle + \langle \varphi, y^* \rangle + \sum_{j=1}^{m} \langle g_j s_j \phi_j, y^* \rangle = 0.$$

Applying $\mathscr{L}_{y^*}$ to (2.1) results in (note $\langle f, y^* \rangle = f_{min}$)

$$(2.4) \qquad 0 = \langle f - f_{min}, y^* \rangle = \sum_{j=0}^{m} \langle g_j s_j, y^* \rangle.$$

Since each s_j is SOS, every $\langle g_j s_j, y^* \rangle \geq 0$, by item ii) of Lemma 2.5. Thus, from the above, we know every $\langle g_j s_j, y^* \rangle = 0$. Again, by item ii) of Lemma 2.5, if $2k > 2 + \deg(g_j s_j \phi_j)$, then every $\langle g_j s_j \phi_j, y^* \rangle = 0$. So, from (2.3), we can get

$$(2.5) \qquad \langle h_i^{2\ell}, y^* \rangle + \langle \varphi, y^* \rangle = 0.$$

Since $h_i^{2\ell}$ is SOS, we similarly have $\langle h_i^{2\ell}, y^* \rangle \geq 0$. Since $Q(\rho) \subset Q(g)$, $\varphi \in Q(g)$ and one could write $\varphi = \sum_{j=0}^{m} g_j \sigma_j$ with each σ_j being SOS. Hence,

$$\langle \varphi, y^* \rangle = \langle g_0 \sigma_0, y^* \rangle + \langle g_1 \sigma_1, y^* \rangle + \langle g_m \sigma_m, y^* \rangle \geq 0,$$

by item ii) of Lemma 2.5. Then (2.5) implies $\langle h_i^{2\ell}, y^* \rangle = 0$, i.e., $h_i^{\ell} \in \ker M_k(y^*)$. By item i) of Lemma 2.5 and $M_k(y^*) = L_{g_0}^{(k)}(y^*) \succeq 0$, if k is big enough, we get

$$h_i \in \ker M_k(y^*).$$

Step 3 It's enough to show that $y^*|_{2k-2}$ is flat. Since C is finite, the quotient space $\mathbb{R}[x]/I(C)$ is finitely dimensional. Let $\{b_1, \ldots, b_L\}$ be a standard basis of $\mathbb{R}[x]/I(C)$. Then, for every $\alpha \in \mathbb{N}^n$, we can write

$$x^\alpha = \eta(\alpha) + \sum_{i=1}^{r} \theta_i h_i, \quad \deg(\theta_i h_i) \leq |\alpha|, \quad \eta(\alpha) \in \text{span}\{b_1, \ldots, b_L\}.$$

Because every $h_i \in \ker M_k(y^*)$, we know

$$\theta_i h_i \in \ker M_k(y^*) \quad \text{if } |\alpha| \leq k-1,$$

by item i) of Lemma 2.5. Thus,

$$x^\alpha - \eta(\alpha) \in \ker M_k(y^*) \quad \text{if } |\alpha| \leq k-1.$$

Set $d_b := \max_j \deg(b_j)$. Then, every α-th column $(d_b + 1 \leq |\alpha| \leq k-1)$ of $M_k(y^*)$ is a linear combination of β-th columns of $M_k(y^*)$ with $|\beta| \leq d_b$, so

$$\operatorname{rank} M_{d_b}(y^*) = \operatorname{rank} M_t(y^*), \quad t = d_b + 1, \ldots, k-1.$$

Hence, if $k - 1 - d_g \geq d_b$, then

$$\operatorname{rank} M_{k-1-d_g}(y^*) = \operatorname{rank} M_{k-1}(y^*).$$

That is, $y^*|_{2k-2}$ is flat and y^* has a flat truncation, when k is big enough. $\qquad\square$

2.2. Schmüdgen type Lasserre's relaxation.

Now we consider Schmüdgen type Lasserre's hierarchy, which refines (1.2) as:

$$(2.6) \qquad\qquad \max \quad \gamma \quad \text{s.t. } f - \gamma \in Pr_k(g).$$

The above $Pr_k(g)$ denotes the k-th truncated preordering generated by the tuple g (denote $g_\nu := g_1^{\nu_1} \cdots g_m^{\nu_m}$):

$$Pr_k(g) := \left\{ \sum_{\nu \in \{0,1\}^m} g_\nu \sigma_\nu \,\middle|\, \sigma_\nu \text{ is SOS, } \deg(g_\nu \sigma_\nu) \leq 2k \text{ for every } \nu \right\}.$$

The dual optimization problem of (2.6) is (cf. [6, 9])

$$(2.7) \qquad\qquad \begin{cases} \min\limits_{y \in \mathscr{M}_{2k}} & \langle f, y \rangle \\ \text{s.t.} & L_{g_\nu}^{(k)}(y) \succeq 0 \, (\nu \in \{0,1\}^m), \, \langle 1, y \rangle = 1. \end{cases}$$

Let $\widetilde{f}_k$ and $\widetilde{f}_k^*$ be the optimal values of (2.6) and (2.7) respectively, for a given order k. By weak duality, $\widetilde{f}_k \leq \widetilde{f}_k^*$ for every k. If K has nonempty interior, then $\widetilde{f}_k = \widetilde{f}_k^*$, i.e., there is no duality gap. Clearly, every $\widetilde{f}_k^* \leq f_{min}$. Both sequences $\{\widetilde{f}_k\}$ and $\{\widetilde{f}_k^*\}$ are monotonically increasing. The relaxation (2.6) is stronger than (1.2), because $Q_k(g) \subseteq Pr_k(g)$, so we have $f_k \leq \widetilde{f}_k$ for every k. By Schmüdgen's Positivstellensaz (if K is compact and a polynomial p is positive on K, then $p \in Pr_\ell(g)$ for some ℓ, cf. [21]), $\widetilde{f}_k$ converges to f_{min}, and so does $\widetilde{f}_k^*$, when K is compact. If $\widetilde{f}_{k_1} = f_{min}$ for some order k_1, we say Schmüdgen type Lasserre's hierarchy has *finite convergence*.

Suppose y^* is an optimizer of (2.7). If y^* has a flat truncation, say $y^*|_{2t}$, then, as shown in §1.2, one could not only extract $\operatorname{rank} M_t(y^*)$ global optimizers of (1.1) from y^*, but also get a certificate for $\widetilde{f}_k = f_{min}$ if there is no duality gap between (2.6) and (2.7). So, flat truncation is also a sufficient condition for Schmüdgen type Lasserre's hierarchy to have finite convergence. Is it also a necessary condition? If so, flat truncation could serve as a certificate for checking finite convergence of the hierarchy of (2.6). Like for the Putinar type one, this is also generically true. A similar result like Theorem 2.2 holds, and weaker conditions are required.

Theorem 2.6. *Suppose the set S of global minimizers of (1.1) is nonempty and finite, and there is no duality gap between (2.6) and (2.7) for k big enough. Then, Schmüdgen type Lasserre's hierarchy of (2.6) has finite convergence if and only if every minimizer of (2.7) has a flat truncation for k sufficiently large.*

Remark 2.7. The comments in Remarks 2.3 and 2.4 for Theorem 2.2 are all applicable to Theorem 2.6. Here, we point out some differences: 1) Theorem 2.6 does not require the optimum of (2.6) to be achievable. 2) Assumption 2.1 is slightly stronger than that S is finite, although they are both generically true. 3) If it occurs that $Q_k(g) = Pr_k(g)$ (e.g., this is the case when (1.1) has equality constraints and/or a single inequality constraint), then in Theorem 2.2 Assumption 2.1 could be replaced by $|S| < \infty$ and (1.2) is not required to achieve its optimum.

Proof of Theorem 2.6   Like in the proof of Theorem 2.2, we only need to prove the necessity of flat truncation. Suppose $\widehat{f}_{k_1} = f_{min}$ for some order k_1, then for every $\epsilon > 0$ we have $f - f_{min} + \epsilon \in Pr_{k_1}(g)$. Write

$$(2.8) \qquad\qquad f - f_{min} + \epsilon = \sum_{\nu \in \{0,1\}^m} \sigma_\nu^\epsilon \cdot g_\nu$$

for some SOS polynomials σ_ν^ϵ with $\deg(\sigma_\nu^\epsilon g_\nu) \le 2k_1$. Note that as $\epsilon \to 0$ some coefficients of σ_ν^ϵ might go to infinity while its degree is bounded. Let y^* be an arbitrary minimizer of (2.7) (if $k \ge k_1$, (2.7) always has one, e.g., $[x^*]_{2k}$, for any $x^* \in S$). Then $\mathscr{L}_{y^*}(f) = \widehat{f}_k = f_{min}$. Applying $\mathscr{L}_{y^*}$ to (2.8), we get

$$(2.9) \qquad\qquad \epsilon = \sum_{\nu \in \{0,1\}^m} \langle \sigma_\nu^\epsilon g_\nu, y^* \rangle.$$

Since every σ_ν^ϵ is SOS, by item ii) of Lemma 2.5, $\langle \sigma_\nu^\epsilon g_\nu, y^* \rangle \ge 0$ and

$$(2.10) \qquad\qquad \lim_{\epsilon \to 0} \langle \sigma_\nu^\epsilon g_\nu, y^* \rangle = 0.$$

We complete the proof in three steps.

Step 1   The set S is finite, so its vanishing ideal

$$I(S) \quad = \quad \{p \in \mathbb{R}[x] : p(u) = 0 \quad \forall u \in S\}$$

is zero dimensional. Let $\{h_1, \ldots, h_r\}$ be a Grobner basis of $I(S)$ with respect to a total degree ordering. Clearly, each h_i vanishes on

$$S = \{x \in \mathbb{R}^n : -(f(x) - f_{min}) = 0, g_1(x) \ge 0, \ldots, g_m(x) \ge 0\}.$$

By Real Nullstellensatz (cf. Corollary 4.4.3 of [1]), there exist $\ell \in \mathbb{N}$, $\varphi \in \mathbb{R}[x]$ and SOS polynomials ϕ_ν ($\nu \in \{0,1\}^m$) such that

$$h_i^{2\ell} - (f - f_{min})\varphi + \sum_{\nu \in \{0,1\}^m} \phi_\nu g_\nu = 0.$$

Applying $\mathscr{L}_{y^*}$ to the above (suppose $2k$ is bigger than the degrees of all the above products) results in

$$(2.11) \qquad \langle h_i^{2\ell}, y^* \rangle + \sum_{\nu \in \{0,1\}^m} \langle \phi_\nu g_\nu, y^* \rangle = \langle (f - f_{min})\varphi, y^* \rangle.$$

By (2.8), for every $\epsilon > 0$, we get

$$(2.12) \qquad \langle (f - f_{min} + \epsilon)\varphi, y^* \rangle = \sum_{\nu \in \{0,1\}^m} \langle g_\nu \sigma_\nu^\epsilon \varphi, y^* \rangle.$$

Step 2 By (2.10) and item iv) of Lemma 2.5, we can get

$$\lim_{\epsilon \to 0} \langle (g_\nu \sigma_\nu^\epsilon \varphi, y^* \rangle = 0$$

for k big enough. Hence, from (2.12) and the above, it holds that

$$\langle (f - f_{min})\varphi, y^* \rangle = \lim_{\epsilon \to 0} \langle (f - f_{min} + \epsilon)\varphi, y^* \rangle = 0.$$

So, (2.11) results in the equality

$$\langle h_i^{2\ell}, y^* \rangle + \sum_{\nu \in \{0,1\}^m} \langle \phi_\nu g_\nu, y^* \rangle = 0.$$

Since $h_i^{2\ell}$ and every ϕ_ν are SOS, each $\langle \phi_\nu g_\nu, y^* \rangle \geq 0$ and $\langle h_i^{2\ell}, y^* \rangle \geq 0$, by item ii) of Lemma 2.5. Hence, $\langle h_i^{2\ell}, y^* \rangle = 0$, i.e., $h_i^\ell \in \ker M_k(y^*)$. Again, by item i) of Lemma 2.5 and $M_k(y^*) = L_{g_0}^{(k)}(y^*) \succeq 0$, if k is big enough, then

$$h_i \in \ker M_k(y^*).$$

Step 3 Like Step 3 in the proof of Theorem 2.2, we would prove $y^*|_{2k-2}$ is flat by repeating the same argument there, and omit it here for cleanness. $\square$

3. Asymptotical Convergence

In this section, we consider the case that Lasserre's hierarchy of (1.2) has asymptotic but not finite convergence. Under the archimedean condition, Lasserre proved $f_k \to f_{min}$ as $k \to \infty$. Since it is possible that $f_k = f_k^* < f_{min}$ for every k, we should not expect flat truncation holds in such a case. When (1.1) has a unique global minimizer u^*, Schweighofer [23] proved: the subvector consisting of linear moments (indexed by $\alpha \in \mathbb{N}^n$ with $|\alpha| = 1$) of an almost optimizer $y^{(k)}$ of (1.4) for order k converges to u^* as $k \to \infty$. When (1.1) has more than one minimizer, do we have a similar convergence result? To the author's best knowledge, there is little work on this question. This section is addressing this issue. Generally, we need to use a higher order truncation of $y^{(k)}$, and consider its limit points. The main result of this section is that every limit point of a truncation of $y^{(k)}$ is flat if (1.1) has finitely many global minimizers. In other words, flat truncation is asymptotically satisfied when Lasserre's hierarchy has asymptotic convergence.

We assume the archimedean condition holds for the tuple g: there exist $R > 0$ and $k_0 \in \mathbb{N}$ such that

$$(3.1) \qquad\qquad R - \|x\|_2^2 \in Q_{k_0}(g).$$

For a Borel set $T \subseteq \mathbb{R}^n$, denote by $\mathtt{Prob}(T)$ the set of all probability measures supported on T. Let S be the set of global minimizers of (1.1). For each integer $t > 0$, denote

$$(3.2) \qquad\qquad F_{2t} := \left\{ \int_S [x]_{2t} d\mu \; : \; \mu \in \mathtt{Prob}(S) \right\}.$$

Proposition 3.1. *Let S and F_{2t} be defined as above. If $0 < |S| < \infty$, then for any integer $t \geq \max\{d_f, d_g + |S| - 1\}$, every tms $z \in F_{2t}$ is flat.*

Proof. Let $\ell := |S|$ and write $S = \{(a_{i,1}, \ldots, a_{i,n}) : i = 1, \ldots, \ell\}$, and I be the ideal generated by the following polynomial equations

$$(x_{i_1} - a_{1,i_1})(x_{i_2} - a_{2,i_2}) \cdots (x_{i_\ell} - a_{l,i_n}) = 0, \quad \forall i_1, \ldots, i_\ell \in \{1, \ldots, n\}.$$

Clearly, the zero set of the above equations is S. For each $\alpha \in \mathbb{N}^n$ with $|\alpha| = \ell$, the x^α is a leading monomial in one of the above defining polynomials, so there exists $p_\alpha \in \mathbb{R}[x]_{\ell-1}$ such that

$$(3.3) \qquad\qquad x^\alpha - p_\alpha \equiv 0 \mod I.$$

Choose an arbitrary $z \in F_{2t}$, then there exists $\mu \in \texttt{Prob}(S)$ satisfying

$$z = \int_S [x]_{2t} d\mu, \qquad M_t(z) = \int_S [x]_t [x]_t^T d\mu.$$

By a simple induction on $|\alpha|$, one could show that for every $|\alpha| \in [\ell, t]$, there exists $p_\alpha \in \mathbb{R}[x]_{\ell-1}$ satisfying (3.3). Clearly, $x^\alpha - p_\alpha$ vanishes on S and

$$x^\alpha - p_\alpha \in \ker M_t(z).$$

This means that every α-th ($|\alpha| \geq \ell$) column of $M_t(z)$ is a linear combination of its β-th columns ($|\beta| \leq \ell - 1$). So, it holds that

$$\operatorname{rank} M_{\ell-1}(z) = \operatorname{rank} M_r(z), \qquad \text{for } r = \ell, \ell+1, \ldots, t.$$

Thus, for all $t \geq d_g + \ell - 1$, we have

$$\operatorname{rank} M_{t-d_g}(z) = \operatorname{rank} M_t(z).$$

Clearly, every $L_{g_i}^{(t)}(z) \succeq 0$, and hence z is flat by the definition.   □

Denote by $\mathscr{M}_\infty$ the space of all full moment sequences $w = (w_\alpha)$ indexed by vectors $\alpha \in \mathbb{N}^n$. For all $w, z \in \mathscr{M}_\infty$, define

$$\langle w, z \rangle = \sum_\alpha w_\alpha z_\alpha, \quad \|w\|_2 = \sqrt{\langle w, w \rangle}.$$

This induces a Hilbert space

$$\mathscr{M}_\infty^2 = \{ w \in \mathscr{M}_\infty : \|w\|_2 < \infty \}.$$

If we think of $\mathscr{M}_\infty^2$ as a Banach space, then it is self-dual. Clearly, (1.1) is equivalent to the optimization problem:

$$(3.4) \qquad \min_{y \in \mathscr{M}_\infty} \langle f, y \rangle \quad \text{s.t.} \quad y \in \left\{ \int_K [x]_\infty d\mu : \mu \in \texttt{Prob}(K) \right\}.$$

The set of its optimizers is precisely $\texttt{Prob}(S)$.

Lemma 3.2. *Suppose (3.1) holds. If y is feasible for (1.4) and $k \geq k_1 \geq k_0$, then*

$$(3.5) \qquad\qquad \| y|_{2(k_1-k_0)} \|_2^2 \leq 1 + R + \cdots + R^{k_1-k_0}.$$

Proof. By (3.1), there exist SOS polynomials $s_0, s_1, \ldots, s_m$ such that

$$R - \|x\|_2^2 = \sum_{i=0}^m g_i s_i, \quad \text{each} \quad \deg(g_i s_i) \leq 2k_0.$$

(Here $\|x\|_2 := \sqrt{x_1^2 + \cdots + x_n^2}$.) Thus, for every $j = 1, \ldots, k_1 - k_0$, we have

$$R \cdot \|x\|_2^{2j-2} - \|x\|_2^{2j} = \sum_{i=0}^m g_i \cdot s_i \cdot \|x\|_2^{2j-2}.$$

By item ii) of Lemma 2.5, applying $\mathscr{L}_y$ to the above gives

$$R \mathscr{L}_y(\|x\|_2^{2j-2}) - \mathscr{L}_y(\|x\|_2^{2j}) \geq 0, \quad j = 1, \ldots, k_1 - k_0.$$

Therefore, it holds that

$$\mathscr{L}_y(\|x\|_2^2) \leq R, \quad \mathscr{L}_y(\|x\|_2^4) \leq R^2, \ldots, \quad \mathscr{L}_y(\|x\|_2^{2j}) \leq R^j.$$

The moment matrix $M_k(y) = L_{g_0}^{(k)}(y) \succeq 0$ implies its submatrix $M_{k_1-k_0}(y) \succeq 0$. In the below, we denote by $\|A\|_F$ the Frobenius norm of a matrix A, i.e., $\|A\|_F = \sqrt{Trace(A^T A)}$. Recall that $\|A\|_F \leq Trace(A)$ if $A \succeq 0$. Clearly,

$$\sum_{|\alpha| \leq 2(k_1-k_0)} |y_\alpha|^2 \leq \|M_{k_1-k_0}(y)\|_F^2 \leq (Trace(M_{k_1-k_0}(y)))^2,$$

$$Trace(M_{k_1-k_0}(y)) = \sum_{j=0}^{k_1-k_0} \sum_{|\alpha|=2j} y_\alpha \leq \sum_{j=0}^{k_1-k_0} \mathscr{L}_y(\|x\|_2^{2j}) \leq \sum_{j=0}^{k_1-k_0} R^j.$$

So, the inequality (3.5) is true. $\qquad\square$

It is possible that (1.4) might not have a minimizer, i.e., the optimal value f_k^* of (1.4) may not be achievable (cf. [23, Example 4.8]). Thus, we consider an almost optimizer of (1.4). Let $\{y^{(k)}\}$ be a sequence such that $y^{(k)}$ is feasible for (1.4) with order k. We say $\{y^{(k)}\}$ is *asymptotically optimal* if

$$\lim_{k\to\infty} \langle f, y^{(k)} \rangle = \lim_{k\to\infty} f_k^*.$$

(A different notion *nearly optimality* was used in [23].) Note that if (3.1) holds and $\{y^{(k)}\}$ is asymptotically optimal, then $\langle f, y^{(k)} \rangle \to f_{min}$ as $k \to \infty$.

Theorem 3.3. *Assume the archimedean condition (3.1) holds and the set S of global minimizers of (1.1) is nonempty and finite. Let $\{y^{(k)}\}$ be asymptotically optimal for (1.4). Then, for every $t \geq \max\{d_f, d_g + |S| - 1\}$, the truncated sequence $\{y^{(k)}|_{2t}\}$ is bounded, and its every limit point belongs to F_{2t} and is flat.*

Proof. If we replace $k_1 - k_0$ by t in (3.5) of Lemma 3.2, then the sequence $\{y^{(k)}|_{2t}\}$ is clearly bounded. One could generally assume $R < 1$, because otherwise we can scale x in (3.1) so that $R < 1$ and Lasserre's hierarchy remains equivalent.

Let v be an arbitrary limit point of $\{y^{(k)}|_{2t}\}$. One could generally assume $y^{(k)}|_{2t} \to v$. We need to prove $v \in F_{2t}$. Suppose otherwise $v \notin F_{2t}$. For every k, define

$$z^{(k)} := y^{(k)}|_{2(k-k_0)}.$$

Each $z^{(k)}$ could be treated as a vector in $\mathscr{M}_\infty^2$ by adding zero entries to the tailing. So, for $k > k_0$, the inequality (3.5) implies

$$\|z^{(k)}\|_2^2 \leq 1 + R + R^2 + R^3 + \cdots = 1/(1-R).$$

The sequence $\{z^{(k)}\}$ is bounded in $\mathscr{M}_\infty^2$. By Alaoglu's Theorem (cf. [2, Theorem V.3.1] or [9, Theorem C.18]), it has a subsequence $\{z^{(k_j)}\}$ that is convergent in the weak-$*$ topology. That is, there exists $z^* \in \mathscr{M}_\infty^2$ such that

$$\langle c, z^{(k_j)} \rangle \to \langle c, z^* \rangle \quad \text{as } j \to \infty$$

for every $c \in \mathscr{M}_\infty^2$. If we choose c as $\langle c, w \rangle = w_\alpha$ for each α, then one could get

(3.6) $$y^{(k_j)}|_\alpha = z^{(k_j)}|_\alpha \to z^*|_\alpha, \qquad z^*|_{2t} = v.$$

Note that every $L_{g_i}^{(r)}(z^{(k_j)}) \succeq 0$ if $k_j \geq 2r$. By (3.6), it holds that for all r

$$L_{g_i}^{(r)}(z^*) \succeq 0, \quad i = 0, 1, \ldots, m.$$

Hence, $z^* \in \mathscr{M}_\infty^2$ is a full moment sequence such that the corresponding localizing matrices of all orders are positive semidefinite. By Lemma 3.2 of Putinar [19] and (3.1), z^* admits a probability measure (note $\mathscr{L}_{z^*}(1) = 1$) supported on K, and it is feasible for (3.4). The polynomial f defines a continuous linear functional acting on $\mathscr{M}_\infty^2$ as $\langle f, w \rangle$. By the weak-$*$ convergence of $z^{(k_j)} \to z^*$, it holds that

$$\langle f, z^* \rangle = \lim_{j \to \infty} \langle f, z^{(k_j)} \rangle = \lim_{j \to \infty} \langle f, y^{(k_j)} \rangle = \lim_{j \to \infty} f^*_{k_j} = f_{min}.$$

The above last two equalities are because (3.1) holds and the sequence $\{y^{(k)}\}$ is asymptotically optimal. This means that z^* is also a minimizer of (3.4). Hence, $v = z^*|_{2t} \in F_{2t}$. But, this contradicts the earlier assertion that $v \notin F_{2t}$. Therefore, every limit point of $\{y^{(k)}|_{2t}\}$ belongs to F_{2t}. By Proposition 3.1, we know every $v \in F_{2t}$ is flat if $t \geq \max\{d_f, d_g + |S| - 1\}$.   □

In Theorem 3.3, for a fixed t, it is possible that the truncation $y^{(k)}|_{2t}$ is not flat for every k. But, for every $\epsilon > 0$, if k is big enough, there exists a $z \in F_{2t}$ such that $\|y^{(k)}|_{2t} - z\|_2 \leq \epsilon$. In other words, flat truncation is asymptotically satisfied when S is finite. Theorem 3.3 also implies that the distance between $y^{(k)}|_{2t}$ and F_{2t} tends to zero as $k \to \infty$. Therefore, in numerical computations, flat truncation has a good chance to be satisfied, even if Lasserre's hierarchy does not have finite convergence.

Remark 3.4. A similar version of Theorem 3.3 also holds for Schmüdgen type Lasserre's hierarchy. This is because (2.7) can be thought of as a Putinar type one applied to the 2^m constraints $g_\nu(x) \geq 0 \, (\nu \in \{0, 1\}^m)$. The archimedean condition (3.1) for the corresponding constraints is automatically satisfied for compact K, by Schmüdgen's Positivstellensatz. Hence, when K is compact and S is finite, flat truncation is also asymptotically satisfied for the hierarchy of (2.7).

4. Some Applications

In this section, we consider two interesting cases of semidefinite relaxations for solving polynomial optimization, for which flat truncation could serve as a certificate to check their exactness and be used to get minimizers.

4.1. **Standard SOS relaxations.** For a polynomial $f \in \mathbb{R}[x]$, consider the unconstrained polynomial optimization problem

$$(4.1) \qquad \min_{x \in \mathbb{R}^n} \quad f(x).$$

To have a finite minimum f_{min}, assume f has an even degree $2d$. The standard SOS relaxation (cf. [6, 17]) for solving (4.1) is

$$(4.2) \qquad \max \quad \gamma \quad \text{s.t.} \quad f - \gamma \quad \text{is SOS.}$$

Its dual optimization problem is

$$(4.3) \qquad \min_{y \in \mathscr{M}_{2d}} \quad \langle f, y \rangle \quad \text{s.t.} \quad M_d(y) \succeq 0, \, \langle 1, y \rangle = 1.$$

Lasserre [6] showed that (4.2) and (4.3) have the same optimal value (there is no duality gap), which we denote by f_{sos}, and (4.2) achieves its optimum if $f_{sos} > -\infty$, because (4.3) has an interior point. Clearly, $f_{sos} \leq f_{min}$. As demonstrated by the numerical experiments of [17], it occurs quite a lot that $f_{min} = f_{sos}$. Thus, one is

wondering: how do we check $f_{sos} = f_{min}$, and if so how do we get minimizers of (4.1)? This issue could be solved by using flat truncation.

The whole space $\mathbb{R}^n$ would be defined by the trivial inequality $1 \geq 0$. Thus, the associated k-th truncated quadratic module and preordering coincide and

$$Q_k(1) = Pr_k(1) = \Sigma_{n,2k},$$

where $\Sigma_{n,2k}$ denotes the cone of SOS polynomials having n variables and degree $2k$. Like (1.2) and (1.4), the k-th order Lasserre's relaxation for (4.1) is

(4.4)
$$\max \quad \gamma \quad \text{s.t. } f - \gamma \in \Sigma_{n,2k},$$

and its dual optimization problem is

(4.5)
$$\min_{y \in \mathscr{M}_{2k}} \quad \langle f, y \rangle \quad \text{s.t. } M_k(y) \succeq 0, \langle 1, y \rangle = 1.$$

Clearly, for every $k \geq d$, (4.4) is equivalent to (4.2). However, (4.5) and (4.3) are different in satisfying flat truncation, though they have the same optimal value. Theorem 2.6 implies the following.

Corollary 4.1. *Let f_{sos} be the optimal value of (4.2). Suppose $f_{min} = f_{sos}$ and (4.1) has a nonempty set of finitely many global minimizers. Then, for k big enough, every minimizer y^* of (4.5) has a flat truncation.*

If $f_{min} > -\infty$, then f generically has finitely many minimizers (cf. [15, Prop. 2.1]). Note that for the case $f_{sos} = f_{min}$, if y^* is an optimizer of (4.3) instead of (4.5), then y^* might not have a flat truncation. In this sense, (4.5) is stronger than (4.3), though their optimal values are same. We show the difference between (4.5) and (4.3) in satisfying flat truncation as follows.

Example 4.2. Consider the polynomial optimization problem

$$\min_{x \in \mathbb{R}^3} \quad (x_1 x_2 - 1)^2 + (x_1 x_3 - 1)^2 + (x_2 x_3 - 1)^2.$$

Let f be the objective. Clearly, $f_{sos} = f_{min} = 0$. The global minimizers are $\pm(1, 1, 1)$. However, not every minimizer of the corresponding (4.3) has a flat truncation. For instance, for all $a > 0$, the tms [2]

$$y^*(a) := (1, \ 0, \ 0, \ 0, \ 1, \ 1, \ 1, \ 1, \ 1, \ 1, \ 0, \ 0, \ 0, \ 0, \ 0, \ 0, \ 0, \ 0, \ 0, \ 0,$$
$$1 + a, \ 1, \ 1, \ 1, \ 1, \ 1, \ 1, \ 1, \ 1, \ 1, \ 1 + a, \ 1, \ 1, \ 1, \ 1 + a)$$

is a minimizer of (4.3). For all $a > 0$, one can verify that

$$\operatorname{rank} M_0(y^*(a)) = 1, \quad \operatorname{rank} M_1(y^*(a)) = 2, \quad \operatorname{rank} M_2(y^*(a)) = 5.$$

Thus, $y^*(a)$ does not have a flat truncation if $a > 0$. If we solve (4.3) by `SeDuMi`, a numerical value 2.7232 of a is returned, and the computed minimizer does not have a flat truncation. For $k = 3$, the numerical optimizer of (4.5) returned by `SeDuMi` does not have a flat truncation. However, for $k = 4$, the numerical optimizer of (4.5) returned by `SeDuMi` has a flat truncation, and we are able to get the two global minimizers. □

---

[2]Its entries are listed in graded lexicographical ordering.

4.2. **Jacobian SDP relaxation.** Now we consider Jacobian SDP relaxations for polynomial optimization introduced by the author in the earlier work [16]. It is a refining of relaxation (2.6). Its basic idea is to introduce some redundant polynomial equalities, say,

$$\varphi_1(x) = \cdots = \varphi_L(x) = 0,$$

which are constructed from the minors of Jacobians of $f, g_1, \ldots, g_m$. For instance, when (1.1) has a single inequality constraint, say $g(x) \geq 0$, the newly introduced equalities are

$$g(x) \cdot f_{x_i}(x) = 0, \ i = 1, \ldots, n,$$

$$\sum_{1 \leq i < j \leq n, i+j=\ell} \Big(f_{x_i}(x)g_{x_j}(x) - f_{x_j}(x)g_{x_i}(x) \Big) = 0, \ \ell = 3, \ldots, 2n - 1.$$

When (1.1) achieves its minimum and some generic nonsingularity conditions (see Assumption 2.2 of [16]) holds, (1.1) is equivalent to

$$(4.6) \qquad \begin{cases} \min_{x \in \mathbb{R}^n} & f(x) \\ \quad \text{s.t.} & \varphi_1(x) = \cdots = \varphi_L(x) = 0, \\ & g_1(x) \geq 0, \ldots, g_m(x) \geq 0. \end{cases}$$

Let PJ_k denote the k-th truncated preordering for (4.6) (cf. [16, Sec. 2.2]). Then, the resulting version of relaxation (2.6) for (4.6) is

$$(4.7) \qquad\qquad\qquad \max \quad \gamma \quad \text{s.t.} \quad f - \gamma \in PJ_k.$$

Similarly, its dual optimization problem is (cf. [16, Sec. 2.2])

$$(4.8) \qquad \begin{cases} \min_{y \in \mathscr{M}_{2k}} & \langle f, y \rangle \\ \quad \text{s.t.} & L_{\varphi_j}^{(k)}(y) = 0, j = 1, \ldots, L, \\ & L_{g_\nu}^{(k)}(y) \succeq 0 \, (\forall \nu \in \{0,1\}^m), \, \langle 1, y \rangle = 1. \end{cases}$$

An attractive property of the hierarchy of (4.7) is that it always has finite convergence. Thus, a practical concern in applications is: how does one identify its finite convergence? Interestingly, flat truncation could serve as a certificate for this purpose when (1.1) has finitely many minimizers.

Corollary 4.3. *Suppose (1.1) has a nonempty set of finitely many global minimizers and its feasible set is nonsingular (Assumption 2.2 of [16] is satisfied). Then, for all k big enough, the optimal value of (4.7) equals the global minimum of (1.1) and every minimizer of (4.8) has a flat truncation.*

Proof. As before, denote by $\widetilde{f}_k$ the optimal value of (4.7). By Theorem 2.3 of [16], there exists k_1 such that $\widetilde{f}_{k_1}$ is equal to the global minimum of (4.6), which is also equal to f_{min}, when Assumption 2.2 of [16] is satisfied. So, there is no duality gap between (4.7) and (4.8) when k is big. Clearly, (4.6) also has finitely many global minimizers, since it is equivalent to (1.1). Thus, the conclusion of this corollary just follows from Theorem 2.6. $\square$

In particular, if a polynomial has finitely many global minimizers in the whole space $\mathbb{R}^n$, then its minimizers could be obtained by using gradient SOS relaxation

[13], which is a special case of Jacobian SDP relaxation (see Corollary 2.6 of [16]). Its dual version of (4.8) becomes

(4.9)
$$
\begin{cases}
\min_{y \in \mathscr{M}_{2k}} & \langle f, y \rangle \\
\text{s.t.} & L_{f_{x_j}}^{(k)}(y) = 0, j = 1, \ldots, n, \\
& M_k(y) \succeq 0, \ \langle 1, y \rangle = 1.
\end{cases}
$$

Hence, we could also get the following.

Corollary 4.4. *Suppose a polynomial f has finitely many global minimizers in $\mathbb{R}^n$. Then, for all k big enough, the optimal value of (4.9) equals the global minimum of f in $\mathbb{R}^n$, and its every minimizer has a flat truncation.*

5. Some discussions

Our main result is that, in the generic case (e.g., Assumption 2.1 holds), Putinar type Lasserre's hierarchy has finite convergence if and only if the flat truncation condition (1.7) is satisfied at every minimizer of the dual optimization problem (1.4) for k big enough. Under the archimedean condition, we also showed that (1.7) is always asymptotically satisfied if there are finitely many global minimizers. (We have similar conclusions for Schmüdgen type Lasserre's hierarchy.) This result has applications in solving polynomial optimization problems of the form (1.1). One could solve (1.4), starting with an order $k \geq \max\{d_f, d_g\}$. If (1.7) is satisfied, then we stop and can get one or several global minimizers; if (1.7) is not satisfied, then increase k by one and solve (1.4) again. By repeating this procedure, we solve (1.4) and then check (1.7) iteratively. In the generic case, when Lasserre's hierarchy has finite convergence, our result implies that this procedure must terminate within finitely many steps; when Lasserre's hierarchy has only asymptotic convergence, our result shows that this procedure will terminate asymptotically. Therefore, flat truncation would be used in the solution extraction mechanism in polynomial optimization packages like `GloptiPoly 3` [8], `YALMIP` [12] and `SOSTOOLS` [18].

We would like to make some further remarks about the difference between the flat truncation condition (1.7) and the flat extension condition (1.5). Consider the case that Lasserre's hierarchy has finite convergence, say, $f_k^* = f_k = f_{min}$ for all $k \geq k_0$. Assume $k_0 \geq d := \max\{d_f, d_g\}$. Suppose y^* is a minimizer of (1.4) with order $k > k_0$ and (1.7) is satisfied for $t \in [k_0 + d_g, k]$. Clearly, the truncation $\hat{y} := y^*|_{2k_0}$ is also a minimizer of (1.4) with order k_0, and $\hat{y}$ has a flat extension $y^*|_{2t}$. One is thus interested in the question whether every minimizer of (1.4) with order k_0 is a truncation of a minimizer of (1.4) with order $k > k_0$ that satisfies (1.7)? Or equivalently, does every minimizer of (1.4) with order k_0 has a flat extension? Unfortunately, this may not be true. We show this by using Example 1.1. By `SeDuMi`, the computed optimizer z^* of (1.4) with order 2 is $z^* := (1.0000, 0.0000, 0.0000, 0.0000, 0.7908)$, and

$$
M_2(z^*) = \begin{bmatrix} 1.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.7908 \end{bmatrix}.
$$

The tms z^* is not *recursively generated* (cf. [3, Section 2]), because the coefficient vector of the polynomial $x - 1$ belongs to the kernel of $M_2(z^*)$ but the coefficient vector of the multiple $x(x - 1)$ does not. So, z^* does not admit any representing measure (cf. [3, Corollary 2.11]), and hence z^* has no flat extension.

It is interesting to know for how big k every minimizer of the dual optimization problem (1.4) satisfies the flat truncation condition (1.7). The existence of such a k is guaranteed by Theorem 2.2 for the generic case, but there is no estimate on it. Clearly, any bound on such k depends on f. Thus, one is interested to know whether there is a uniform bound N, which only depends on the defining polynomials g_i of the set K and the degree of f, such that every minimizer of (1.4) with order N satisfies (1.7) for all generic f with the given degree? If such a uniform bound N would exist, it would be very useful in applications. This is because one would only need to solve (1.4) for a fixed order N; if (1.7) is satisfied, we know Lasserre's hierarchy has finite convergence; if not, then Lasserre's hierarchy has no finite convergence in the generic case. However, such a uniform bound N typically does not exist. For instance, there is no such an N when K is the three dimensional unit ball and $\deg(f) = 6$. Consider the Motzkin polynomial $M := x_1^4 x_2^2 + x_1^2 x_2^4 + x_3^6 - 3x_1^2 x_2^2 x_3^2$. There exists a sequence of generic polynomials p_k of degree six such that $p_k \to M$ as $k \to \infty$. Let γ_k be the minimum value of p_k on the unit ball. Clearly, $\gamma_k \to 0$. The unit ball has nonempty interior, so (1.2) achieves its optimal value for every k (cf. [6]). If, otherwise, such N would exist, then $p_k - \gamma_k \in Q_N(1 - \|x\|_2^2)$ for all k. Since the cone $Q_N(1 - \|x\|_2^2)$ is closed (cf. [11, Theorem 3.33]), the Motzkin polynomial M, which is the limit of $p_k - \gamma_k$ as $k \to \infty$, must also belong to $Q_N(1 - \|x\|_2^2)$. However, this is not true (cf. [16, Example 5.3]). Therefore, such a uniform bound N does not exist. Any bound on k, for which every minimizer of (1.4) satisfies (1.7), must depend on f, even if f has generic coefficients. The dependence relation between such a bound on k and a generic polynomial f is an interesting future research question.

References

[1] J. Bochnak, M. Coste and M-F. Roy. *Real Algebraic Geometry*, Springer, 1998.

[2] Conway, John B. A course in Functional Analysis, Springer-Verlag, 1990, Second Edition.

[3] R. Curto and L. Fialkow. Truncated K-moment problems in several variables. *Journal of Operator Theory*, 54(2005), pp. 189-226.

[4] J. W. Helton and J. Nie. A Semidefinite Approach for Truncated K-Moment Problem. *Preprint*, 2011.

[5] D. Henrion and J. Lasserre. Detecting global optimality and extracting solutions in GloptiPoly. *Positive polynomials in control*, 293310, Lecture Notes in Control and Inform. Sci., 312, Springer, Berlin, 2005.

[6] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, 11(3): 796-817, 2001.

[7] J.B. Lasserre, M. Laurent, and P. Rostalski. Semidefinite characterization and computation of zero-dimensional real radical ideals. *Found. Comput. Math.* 8, pp. 607–647, 2008.

[8] D. Henrion, J. Lasserre and J. Loefberg. `GloptiPoly 3`: moments, optimization and semidefinite programming. *Optimization Methods and Software*, Vol. 24, Nos. 4-5, pp. 761-779, 2009.

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

[10] M. Laurent. Semidefinite representations for finite varieties. *Mathematical Programming*, Vol. 109, pp. 1–26, 2007.

[11] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. *Emerging Applications of Algebraic Geometry, Vol. 149 of IMA Volumes in Mathematics and its Applications* (Eds. M. Putinar and S. Sullivant), Springer, pages 157-270, 2009.

[12] J. Löfberg. YALMIP: A Toolbox for Modeling and Optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[13] J. Nie, J. Demmel and B. Sturmfels. Minimizing polynomials via sum of squares over the gradient ideal. *Mathematical Programming*, Series A, Vol. 106, No. 3, pp. 587–606, 2006.

[14] J.Nie and M. Schweighofer. On the complexity of putinar's positivstellensatz. *Journal of Complexity* 23(2007), pp.135-150.

[15] J. Nie and K. Ranestad. Algebraic Degree of Polynomial Optimization. *SIAM Journal on Optimization*, Vol. 20, No. 1, pp. 485–502, 2009.

[16] J. Nie. An exact Jacobian SDP relaxation for polynomial optimization. *Mathematical Programming*, Series A, to appear.

[17] P. A. Parrilo and B. Sturmfels. Minimizing polynomial functions. In S. Basu and L. Gonzalez-Vega, editors, *Algorithmic and Quantitative Aspects of Real Algebraic Geometry in Mathematics and Computer Science, volume 60 of DIMACS Series in Discrete Mathematics and Computer Science*, pages 83-99. AMS, 2003.

[18] S. Prajna, A. Papachristodoulou,P. Seiler and P. A. Parrilo. SOSTOOLS: Sum of squares optimization toolbox for MATLAB. http://www.mit.edu/~parrilo/sostools, 2004.

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

[20] C. Scheiderer. Sums of squares of regular functions on real algebraic varieties. *Trans. Am. Math. Soc.*, 352, 1039-1069 (1999).

[21] K. Schmüdgen. The K-moment problem for compact semialgebraic sets. *Math. Ann.* 289 (1991), 203–206.

[22] M. Schweighofer. On the complexity of Schmüdgen's Positivstellensatz. *Journal of Complexity* 20, 529-543 (2004).

[23] M. Schweighofer. Optimization of polynomials on compact semialgebraic sets. *SIAM Journal on Optimization*, Vol. 15, No. 3, pp. 805–825, 2005.

[24] J.F. Sturm. SeDuMi 1.02:a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11 & 12 (1999), pp. 625–653.

Department of Mathematics, University of California, San Diego

E-mail address: njw@math.ucsd.edu