STRENGTHENING LASSERRE'S HIERARCHY IN REAL AND COMPLEX POLYNOMIAL OPTIMIZATION*

JIE WANG†

Abstract. We establish a connection between multiplication operators and shift operators. Moreover, we derive positive semidefinite conditions of finite rank moment sequences and use these conditions to strengthen Lasserre's hierarchy for real and complex polynomial optimization. Integration of the strengthening technique with sparsity is considered. Extensive numerical experiments show that our strengthening technique can significantly improve the bound (especially for complex polynomial optimization) and allows to achieve global optimality at lower relaxation orders, thus providing substantial computational savings.

Key words. polynomial optimization, complex polynomial optimization, semidefinite relaxation, Lasserre's hierarchy, multiplication operator, shift operator

MSC codes. Primary, 90C23; Secondary, 90C22,90C26

1. Introduction.

Lasserre's hierarchy [6] is a well-established scheme for globally solving (real) polynomial optimization problems and attracts a lot of attentions of researchers from diverse fields due to its nice theoretical properties in recent years [3, 11]. There is also a complex variant of Lasserre's hierarchy for globally solving complex polynomial optimization problems [5].

A bottleneck of Lasserre's hierarchy is its limited scalability as the size of associated semidefinite relaxations grows rapidly with relaxation orders. One way for overcoming this is exploiting structures (sparsity, symmetry) of polynomial optimization problems to obtain structured semidefinite relaxations of reduced sizes. We refer the reader to the recent works [12, 18, 19] on this topic. Another practical idea is strengthening Lasserre's hierarchy to accelerate its convergence, for instance, using Lagrange multiplier expressions as done in [10].

In this paper we propose to strengthen Lasserre's hierarchy using positive semidefinite (PSD) optimality conditions for any real and complex polynomial optimization problem. These PSD optimality conditions arise from the characterization of normality of *shift operators* which is closely related to *multiplication operators*. Both operators have applications in extractions of optimal solutions when solving polynomial optimization problems with Lasserre's hierarchy [4, 5]. We establish a connection between shift operators and multiplication operators. Further, we derive PSD conditions of finite rank moment sequences via shift operators. These PSD conditions are then employed to strengthen Lasserre's hierarchy. In particular, for real polynomial optimization, we present an intermediate relaxation between two successive moment relaxations; for complex polynomial optimization, we present a two-level hierarchy of moment relaxations which thus offers one more level of flexibility. To improve scalability, the strengthening technique is further integrated into different sparse versions of Lasserre's hierarchy. Diverse numerical experiments are performed. It is shown that the strengthening technique can indeed improve the bound provided by the usual Lasserre's hierarchy and very likely allows to achieve global optimality at

*Submitted to the editors DATE.

Funding: This work was funded by National Key R&D Program of China under grant No. 2023YFA1009401, Natural Science Foundation of China under grant No. 12201618 and 12171324.

†Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China (wangjie212@amss.ac.cn, https://wangjie212.github.io/jiewang/)

lower relaxation orders, especially in complex polynomial optimization.

2. Notation and preliminaries. Let $\mathbb{N}$ be the set of nonnegative integers. For $n \in \mathbb{N}\setminus\{0\}$, let $[n] \coloneqq \{1,2,\ldots,n\}$. For $\boldsymbol{\alpha} = (\alpha_i) \in \mathbb{N}^n$, let $|\boldsymbol{\alpha}| \coloneqq \sum_{i=1}^n \alpha_i$. For $r \in \mathbb{N}$, let $\mathbb{N}^n_r \coloneqq \{\boldsymbol{\alpha} \in \mathbb{N}^n \mid |\boldsymbol{\alpha}| \le r\}$ and $|\mathbb{N}^n_r|$ stands for its cardinality. We use $A \succeq 0$ to indicate that the matrix A is positive semidefinite (PSD). Let $\mathbf{i}$ be the imaginary unit, satisfying $\mathbf{i}^2 = -1$. Throughout the paper, let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Let $\mathbb{F}[\mathbf{x}] \coloneqq \mathbb{F}[x_1,\ldots,x_n]$ be the ring of multivariate polynomials in n variables over the field $\mathbb{F}$, and $\mathbb{F}[\mathbf{x}]_d$ denote the subset of polynomials of degree no greater than d. A polynomial $f \in \mathbb{F}[\mathbf{x}]$ can be written as $f = \sum_{\boldsymbol{\alpha}\in\mathbb{N}^n} f_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}}$ with $f_{\boldsymbol{\alpha}} \in \mathbb{F}$ and $\mathbf{x}^{\boldsymbol{\alpha}} \coloneqq x_1^{\alpha_1}\cdots x_n^{\alpha_n}$. For $d \in \mathbb{N}$, $[\mathbf{x}]_d$ stands for the standard monomial basis of degree up to d, and $[\mathbf{x}]$ stands for the standard monomial basis.

Let $\overline{a}$ denote the conjugate of a complex number a and $\boldsymbol{v}^*$ (resp. A^*) denote the conjugate transpose of a complex vector $\boldsymbol{v}$ (resp. a complex matrix A). We use $\overline{\mathbf{x}} = (\overline{x}_1,\ldots,\overline{x}_n)$ to denote the conjugate of the tuple of complex variables $\mathbf{x}$. We denote by $\mathbb{C}[\mathbf{x},\overline{\mathbf{x}}] \coloneqq \mathbb{C}[x_1,\ldots,x_n,\overline{x}_1,\ldots,\overline{x}_n]$ the complex polynomial rings in $\mathbf{x},\overline{\mathbf{x}}$. A polynomial $f \in \mathbb{C}[\mathbf{x},\overline{\mathbf{x}}]$ can be written as $f = \sum_{(\boldsymbol{\beta},\boldsymbol{\gamma})\in\mathbb{N}^n\times\mathbb{N}^n} f_{\boldsymbol{\beta},\boldsymbol{\gamma}}\mathbf{x}^{\boldsymbol{\beta}}\overline{\mathbf{x}}^{\boldsymbol{\gamma}}$ with $f_{\boldsymbol{\beta},\boldsymbol{\gamma}} \in \mathbb{C}$. The *conjugate* of f is defined as $\overline{f} = \sum_{(\boldsymbol{\beta},\boldsymbol{\gamma})\in\mathbb{N}^n\times\mathbb{N}^n} \overline{f}_{\boldsymbol{\beta},\boldsymbol{\gamma}}\mathbf{x}^{\boldsymbol{\gamma}}\overline{\mathbf{x}}^{\boldsymbol{\beta}}$. The polynomial f is *self-conjugate* if $\overline{f} = f$. It is clear that self-conjugate polynomials take only real values.

2.1. The real Lasserre's hierarchy for real polynomial optimization. Consider the real polynomial optimization problem:

$$\text{(RPOP)} \qquad f_{\min} \coloneqq \inf\{f(\mathbf{x}) : \mathbf{x} \in \mathbf{K}\},$$

where $f \in \mathbb{R}[\mathbf{x}]$ and the feasible set $\mathbf{K}$ is given by

$$\mathbf{K} \coloneqq \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \ge 0, \ldots, g_m(\mathbf{x}) \ge 0\}, \tag{2.1}$$

for some polynomials $g_1,\ldots,g_m \in \mathbb{R}[\mathbf{x}]$. By invoking Borel measures, (RPOP) admits the following reformulation:

$$\begin{cases} \inf\limits_{\mu\in\mathcal{M}_+(\mathbf{K})} & \int_{\mathbf{K}} f\,\mathrm{d}\mu \\ \quad\text{s.t.} & \int_{\mathbf{K}} \mathrm{d}\mu = 1, \end{cases} \tag{2.2}$$

where $\mathcal{M}_+(\mathbf{K})$ denotes the set of finite positive Borel measures on $\mathbf{K}$.

Suppose that $\mathbf{y} = (y_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}\in\mathbb{N}^n}$ is a (pseudo-moment) sequence in $\mathbb{R}$. We associate it with a linear functional $L_{\mathbf{y}} : \mathbb{R}[\mathbf{x}] \to \mathbb{R}$ by

$$f = \sum_{\boldsymbol{\alpha}} f_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} \longmapsto L_{\mathbf{y}}(f) = \sum_{\boldsymbol{\alpha}} f_{\boldsymbol{\alpha}} y_{\boldsymbol{\alpha}}.$$

For $r \in \mathbb{N}$, the r-th order *real moment matrix* $\mathbf{M}^{\mathbb{R}}_r(\mathbf{y})$ is the matrix indexed by $\mathbb{N}^n_r$ such that

$$[\mathbf{M}^{\mathbb{R}}_r(\mathbf{y})]_{\boldsymbol{\beta}\boldsymbol{\gamma}} \coloneqq L_{\mathbf{y}}(\mathbf{x}^{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\gamma}}) = y_{\boldsymbol{\beta}+\boldsymbol{\gamma}}, \quad \forall \boldsymbol{\beta},\boldsymbol{\gamma} \in \mathbb{N}^n_r.$$

The real moment matrix $\mathbf{M}^{\mathbb{R}}(\mathbf{y})$ indexed by $\mathbb{N}^n$ is defined similarly. For a polynomial $g = \sum_{\boldsymbol{\alpha}} g_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} \in \mathbb{R}[\mathbf{x}]$, the r-th order *real localizing matrix* $\mathbf{M}^{\mathbb{R}}_r(g\mathbf{y})$ associated with g is the matrix indexed by $\mathbb{N}^n_r$ such that

$$[\mathbf{M}^{\mathbb{R}}_r(g\mathbf{y})]_{\boldsymbol{\beta}\boldsymbol{\gamma}} \coloneqq L_{\mathbf{y}}(g\mathbf{x}^{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\gamma}}) = \sum_{\boldsymbol{\alpha}} g_{\boldsymbol{\alpha}} y_{\boldsymbol{\alpha}+\boldsymbol{\beta}+\boldsymbol{\gamma}}, \quad \forall \boldsymbol{\beta},\boldsymbol{\gamma} \in \mathbb{N}^n_r.$$

The sequence $\mathbf{y}$ is called a *real moment sequence* if it can be realized by a Borel measure μ, i.e., $y_{\boldsymbol{\alpha}} = \int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} \, \mathrm{d}\mu$ for any $\boldsymbol{\alpha} \in \mathbb{N}^n$, and $\mathbf{y}$ is said to be of finite rank if μ is a finitely atomic measure (that is, a linear positive combination of finitely many Dirac measures), where the rank of $\mathbf{y}$ is defined as the number of atoms.

LEMMA 2.1 ([8], Lemma 4.2). *If $\mathbf{y}$ is a real moment sequence of finite rank, then $\mathbf{M}^{\mathbb{R}}(\mathbf{y}) \succeq 0$ and the rank of $\mathbf{y}$ is equal to* $\operatorname{rank} M^{\mathbb{R}}(\mathbf{y})$.

Let $d_i := \lceil \deg(g_i)/2 \rceil, i = 1, \ldots, m$, $d_{\min} := \max\{\lceil \deg(f)/2 \rceil, d_1, \ldots, d_m\}$. With $r \geq d_{\min}$, the real Lasserre's hierarchy of moment relaxations for (RPOP) [6] is given by

$$\rho_r := \begin{cases} \inf\limits_{\mathbf{y}} & L_{\mathbf{y}}(f) \\ \text{s.t.} & \mathbf{M}_r^{\mathbb{R}}(\mathbf{y}) \succeq 0, \quad y_{\mathbf{0}} = 1, \\ & \mathbf{M}_{r-d_i}^{\mathbb{R}}(g_i \mathbf{y}) \succeq 0, \quad i \in [m]. \end{cases} \tag{2.3}$$

2.2. The complex Lasserre's hierarchy for complex polynomial optimization. Consider the complex polynomial optimization problem:

$$f_{\min} := \inf \{ f(\mathbf{x}, \overline{\mathbf{x}}) : \mathbf{x} \in \mathbf{K} \}, \tag{CPOP}$$

where

$$\mathbf{K} := \{ \mathbf{x} \in \mathbb{C}^n \mid g_i(\mathbf{x}, \overline{\mathbf{x}}) \geq 0, i \in [m] \}, \tag{2.4}$$

and $f, g_1, \ldots, g_m \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$ are self-conjugate polynomials. By invoking Borel measures, (CPOP) also admits the following reformulation:

$$\begin{cases} \inf\limits_{\mu \in \mathcal{M}_+(\mathbf{K})} & \int_{\mathbf{K}} f \, \mathrm{d}\mu \\ \text{s.t.} & \int_{\mathbf{K}} \mathrm{d}\mu = 1, \end{cases} \tag{2.5}$$

where $\mathcal{M}_+(\mathbf{K})$ denotes the set of finite positive Borel measures on $\mathbf{K}$.

Suppose that $\mathbf{y} = (y_{\boldsymbol{\beta},\boldsymbol{\gamma}})_{(\boldsymbol{\beta},\boldsymbol{\gamma}) \in \mathbb{N}^n \times \mathbb{N}^n}$ in $\mathbb{C}$ is a (pseudo-moment) sequence satisfying $y_{\boldsymbol{\beta},\boldsymbol{\gamma}} = \overline{y}_{\boldsymbol{\gamma},\boldsymbol{\beta}}$. We associate it with a linear functional $L_{\mathbf{y}} : \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}] \to \mathbb{C}$ by

$$f = \sum_{(\boldsymbol{\beta},\boldsymbol{\gamma})} f_{\boldsymbol{\beta},\boldsymbol{\gamma}} \mathbf{x}^{\boldsymbol{\beta}} \overline{\mathbf{x}}^{\boldsymbol{\gamma}} \longmapsto L_{\mathbf{y}}(f) = \sum_{(\boldsymbol{\beta},\boldsymbol{\gamma})} f_{\boldsymbol{\beta},\boldsymbol{\gamma}} y_{\boldsymbol{\beta},\boldsymbol{\gamma}}.$$

For $r \in \mathbb{N}$, the r-th order *complex moment matrix* $\mathbf{M}_r^{\mathbb{C}}(\mathbf{y})$ is the matrix indexed by $\mathbb{N}_r^n$ such that

$$[\mathbf{M}_r^{\mathbb{C}}(\mathbf{y})]_{\boldsymbol{\beta}\boldsymbol{\gamma}} := L_{\mathbf{y}}(\mathbf{x}^{\boldsymbol{\beta}} \overline{\mathbf{x}}^{\boldsymbol{\gamma}}) = y_{\boldsymbol{\beta},\boldsymbol{\gamma}}, \quad \forall \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}_r^n.$$

The complex moment matrix $\mathbf{M}^{\mathbb{C}}(\mathbf{y})$ indexed by $\mathbb{N}^n$ is defined similarly. For a self-conjugate polynomial $g = \sum_{(\boldsymbol{\beta}',\boldsymbol{\gamma}')} g_{\boldsymbol{\beta}',\boldsymbol{\gamma}'} \mathbf{x}^{\boldsymbol{\beta}'} \overline{\mathbf{x}}^{\boldsymbol{\gamma}'} \in \mathbb{C}[\mathbf{x}, \overline{\mathbf{x}}]$, the r-th order *complex localizing matrix* $\mathbf{M}_r^{\mathbb{C}}(g\mathbf{y})$ associated with g is the matrix indexed by $\mathbb{N}_r^n$ such that

$$[\mathbf{M}_r^{\mathbb{C}}(g\mathbf{y})]_{\boldsymbol{\beta}\boldsymbol{\gamma}} := L_{\mathbf{y}}(g \mathbf{x}^{\boldsymbol{\beta}} \overline{\mathbf{x}}^{\boldsymbol{\gamma}}) = \sum_{(\boldsymbol{\beta}',\boldsymbol{\gamma}')} g_{\boldsymbol{\beta}',\boldsymbol{\gamma}'} y_{\boldsymbol{\beta}+\boldsymbol{\beta}',\boldsymbol{\gamma}+\boldsymbol{\gamma}'}, \quad \forall \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}_r^n.$$

The sequence $\mathbf{y}$ is called a *complex moment sequence* if it can be realized by a Borel measure μ, i.e., $y_{\boldsymbol{\beta},\boldsymbol{\gamma}} = \int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\beta}} \overline{\mathbf{x}}^{\boldsymbol{\gamma}} \, \mathrm{d}\mu$ for any $\boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}^n$, and $\mathbf{y}$ is said to be of finite rank if μ is a finitely atomic measure.

LEMMA 2.2 ([5], Theorem 5.1). *If* $\mathbf{y}$ *is a complex moment sequence of finite rank, then* $\mathbf{M}^{\mathbb{C}}(\mathbf{y}) \succeq 0$ *and the rank of* $\mathbf{y}$ *is equal to* $\operatorname{rank} M^{\mathbb{C}}(\mathbf{y})$.

Let $d_0 := \max\{|\boldsymbol{\beta}|, |\boldsymbol{\gamma}| : f_{\boldsymbol{\beta},\boldsymbol{\gamma}} \neq 0\}$, $d_i := \max\{|\boldsymbol{\beta}|, |\boldsymbol{\gamma}| : g^i_{\boldsymbol{\beta},\boldsymbol{\gamma}} \neq 0\}$ for $i \in [m]$, where $f = \sum_{(\boldsymbol{\beta},\boldsymbol{\gamma})} f_{\boldsymbol{\beta},\boldsymbol{\gamma}} \mathbf{x}^{\boldsymbol{\beta}} \overline{\mathbf{x}}^{\boldsymbol{\gamma}}, g_i = \sum_{(\boldsymbol{\beta},\boldsymbol{\gamma})} g^i_{\boldsymbol{\beta},\boldsymbol{\gamma}} \mathbf{x}^{\boldsymbol{\beta}} \overline{\mathbf{x}}^{\boldsymbol{\gamma}}$. Set $d_{\min} := \max\{d_0, d_1, \ldots, d_m\}$. With $r \geq d_{\min}$, the complex Lasserre's hierarchy of moment relaxations for (CPOP) [5] is given by

$$
\tau_r := \begin{cases} \inf\limits_{\mathbf{y}} & L_{\mathbf{y}}(f) \\ \text{s.t.} & \mathbf{M}^{\mathbb{C}}_r(\mathbf{y}) \succeq 0, \quad y_{\mathbf{0},\mathbf{0}} = 1 \\ & \mathbf{M}^{\mathbb{C}}_{r-d_i}(g_i \mathbf{y}) \succeq 0, \quad i \in [m]. \end{cases} \tag{2.6}
$$

Note that (2.6) is a complex semidefinite program (SDP). To reformulate it as a real SDP, we refer the reader to [15].

3. Multiplication operators and shift operators. In this section, we establish an interesting connection between multiplication operators and shift operators.

For $p \in \mathbb{F}[\mathbf{x}]_r$ (resp. $\mathbb{F}[\mathbf{x}]$), we write $\mathbf{p}$ for the coefficient vector of p such that $p = \mathbf{p}^{\intercal}[\mathbf{x}]_r$ (resp. $p = \mathbf{p}^{\intercal}[\mathbf{x}]$).

LEMMA 3.1 ([8], Lemma 5.2). *The kernel* $I := \{p \in \mathbb{R}[\mathbf{x}] \mid \mathbf{M}^{\mathbb{R}}(\mathbf{y})\mathbf{p} = \mathbf{0}\}$ *of a moment matrix* $\mathbf{M}^{\mathbb{R}}(\mathbf{y})$ *is an ideal in* $\mathbb{R}[\mathbf{x}]$. *Moreover, if* $\mathbf{M}^{\mathbb{R}}(\mathbf{y}) \succeq 0$, *then* I *is a real radical ideal.*

LEMMA 3.2 ([7]). *Let* $\mathbf{y}$ *be a complex moment sequence of finite rank. The kernel* $I := \{p \in \mathbb{C}[\mathbf{x}] \mid \mathbf{M}^{\mathbb{C}}(\mathbf{y})\mathbf{p} = \mathbf{0}\}$ *of the moment matrix* $\mathbf{M}^{\mathbb{C}}(\mathbf{y})$ *is a radical ideal in* $\mathbb{C}[\mathbf{x}]$.

Suppose that $\mathbf{y}$ is a (real or complex) moment sequence of rank t. Let $I := \{p \in \mathbb{F}[\mathbf{x}] \mid \mathbf{M}^{\mathbb{F}}(\mathbf{y})\mathbf{p} = \mathbf{0}\}$. Then $\mathbb{F}[\mathbf{x}]/I$ is a linear space over $\mathbb{F}$ of dimension t. The multiplication operators $M_i, i \in [n]$ acting on $\mathbb{F}[\mathbf{x}]/I$ are defined by

$$
\begin{aligned} M_i : \mathbb{F}[\mathbf{x}]/I &\longrightarrow \mathbb{F}[\mathbf{x}]/I, \\ p &\longmapsto x_i p. \end{aligned} \tag{3.1}
$$

Since the moment matrix $\mathbf{M}^{\mathbb{F}}(\mathbf{y})$ is PSD with $\operatorname{rank} \mathbf{M}^{\mathbb{F}}(\mathbf{y}) = t$, it can be factorized in the Grammian form such that

$$
[\mathbf{M}^{\mathbb{F}}(\mathbf{y})]_{\boldsymbol{\beta}\boldsymbol{\gamma}} = \mathbf{a}^*_{\boldsymbol{\beta}} \mathbf{a}_{\boldsymbol{\gamma}}, \quad \forall \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}^n, \tag{3.2}
$$

where $\{\mathbf{a}_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in \mathbb{N}^n} \subseteq \mathbb{F}^t$. The shift operators $T_1, \ldots, T_n : \mathbb{F}^t \to \mathbb{F}^t$ are defined by

$$
T_i : \sum_{\boldsymbol{\alpha}} p_{\boldsymbol{\alpha}} \mathbf{a}_{\boldsymbol{\alpha}} \longmapsto \sum_{\boldsymbol{\alpha}} p_{\boldsymbol{\alpha}} \mathbf{a}_{\boldsymbol{\alpha}+\mathbf{e}_i}, \tag{3.3}
$$

where $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is the standard vector basis of $\mathbb{N}^n$.

Let us define the following linear map

$$
\theta : \mathbb{F}[\mathbf{x}] \longrightarrow \mathbb{F}^t, \quad p = \sum_{\boldsymbol{\alpha}} p_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} \longmapsto \sum_{\boldsymbol{\alpha}} p_{\boldsymbol{\alpha}} \mathbf{a}_{\boldsymbol{\alpha}}. \tag{3.4}
$$

LEMMA 3.3. *The linear map* θ *induces an isomorphism:* $\mathbb{F}[\mathbf{x}]/I \cong \mathbb{F}^t$.

Proof. It is clear that θ is surjective. We remain to show that the kernel of θ is I. First, let $p = \sum_{\boldsymbol\beta} p_{\boldsymbol\beta} \mathbf{x}^{\boldsymbol\beta} \in I$. It follows from

$$\mathbf{M}^{\mathbb{F}}(\mathbf{y})\mathbf{p} = \left(\sum_{\boldsymbol\beta} \mathbf{a}_{\boldsymbol\alpha}^* \mathbf{a}_{\boldsymbol\beta} p_{\boldsymbol\beta} \right)_{\boldsymbol\alpha \in \mathbb{N}^n} = \mathbf{0} \tag{3.5}$$

that $\mathbf{a}_{\boldsymbol\alpha}^*(\sum_{\boldsymbol\beta} p_{\boldsymbol\beta} \mathbf{a}_{\boldsymbol\beta}) = 0$ for all $\mathbf{a}_{\boldsymbol\alpha}$. Since $\{\mathbf{a}_{\boldsymbol\alpha}\}_{\boldsymbol\alpha \in \mathbb{N}^n}$ spans $\mathbb{F}^t$, we obtain $\sum_{\boldsymbol\beta} p_{\boldsymbol\beta} \mathbf{a}_{\boldsymbol\beta} = \mathbf{0}$. This proves $p \in \ker(\theta)$ and hence $I \subseteq \ker(\theta)$. Conversely, let $p \in \mathbb{F}[\mathbf{x}]$ such that $\sum_{\boldsymbol\beta} p_{\boldsymbol\beta} \mathbf{a}_{\boldsymbol\beta} = \mathbf{0}$. Then we see $p \in I$. This proves $\ker(\theta) \subseteq I$. ∎

THEOREM 3.4. *Let* $\mathbf{y}$ *be a moment sequence of finite rank. Then the multiplication operator* M_i *is similar to the shift operator* T_i *for* $i \in [n]$*. More concretely, we have* $T_i = \theta \circ M_i \circ \theta^{-1}$ *for* $i \in [n]$.

Proof. Let $p = \sum_{\boldsymbol\alpha} p_{\boldsymbol\alpha} \mathbf{x}^{\boldsymbol\alpha} \in \mathbb{F}[\mathbf{x}]$. We have

$$T_i(\theta(p)) = T_i \left(\sum_{\boldsymbol\alpha} p_{\boldsymbol\alpha} \mathbf{a}_{\boldsymbol\alpha} \right) = \sum_{\boldsymbol\alpha} p_{\boldsymbol\alpha} \mathbf{a}_{\boldsymbol\alpha + \mathbf{e}_i}. \tag{3.6}$$

On the other hand, we have

$$\theta(M_i(p)) = \theta(x_i p) = \theta \left(\sum_{\boldsymbol\alpha} p_{\boldsymbol\alpha} \mathbf{x}^{\boldsymbol\alpha + \mathbf{e}_i} \right) = \sum_{\boldsymbol\alpha} p_{\boldsymbol\alpha} \mathbf{a}_{\boldsymbol\alpha + \mathbf{e}_i}. \tag{3.7}$$

Thus, $T_i(\theta(p)) = \theta(M_i(p))$. It follows $T_i \circ \theta = \theta \circ M_i$. As θ is invertible by Lemma 3.3, we obtain $T_i = \theta \circ M_i \circ \theta^{-1}$. ∎

COROLLARY 3.5. *Let* $\mathbf{y}$ *be a moment sequence of finite rank. Then the shift operators* $T_1, \ldots, T_n$ *are well-defined.*

Proof. We need to show that $T_i(\sum_{\boldsymbol\alpha} p_{\boldsymbol\alpha} \mathbf{a}_{\boldsymbol\alpha}) = \mathbf{0}$ if $\sum_{\boldsymbol\alpha} p_{\boldsymbol\alpha} \mathbf{a}_{\boldsymbol\alpha} = \mathbf{0}$. The assumption $\sum_{\boldsymbol\alpha} p_{\boldsymbol\alpha} \mathbf{a}_{\boldsymbol\alpha} = \mathbf{0}$ implies $\theta(\sum_{\boldsymbol\alpha} p_{\boldsymbol\alpha} \mathbf{x}^{\boldsymbol\alpha}) = \mathbf{0}$ and so $\sum_{\boldsymbol\alpha} p_{\boldsymbol\alpha} \mathbf{x}^{\boldsymbol\alpha} \in I$. By Theorem 3.4, we have

$$\begin{aligned}
T_i \left(\sum_{\boldsymbol\alpha} p_{\boldsymbol\alpha} \mathbf{a}_{\boldsymbol\alpha} \right) &= \theta \circ M_i \circ \theta^{-1} \left(\sum_{\boldsymbol\alpha} p_{\boldsymbol\alpha} \mathbf{a}_{\boldsymbol\alpha} \right) \\
&= \theta \circ M_i \left(\sum_{\boldsymbol\alpha} p_{\boldsymbol\alpha} \mathbf{x}^{\boldsymbol\alpha} \right) \\
&= \theta \left(\sum_{\boldsymbol\alpha} p_{\boldsymbol\alpha} \mathbf{x}^{\boldsymbol\alpha + \mathbf{e}_i} \right) \\
&= \theta \left(x_i \sum_{\boldsymbol\alpha} p_{\boldsymbol\alpha} \mathbf{x}^{\boldsymbol\alpha} \right) = \mathbf{0},
\end{aligned}$$

where the last equality follows from the fact that I is an ideal and so $x_i \sum_{\boldsymbol\alpha} p_{\boldsymbol\alpha} \mathbf{x}^{\boldsymbol\alpha} \in I$. ∎

For the remainder of the paper, we assume a basis of $\mathbb{R}^t$ is given and identify the shift operators with their representing matrices for convenience.

The real shift operators have the distinguished property of being symmetric.

LEMMA 3.6. *Let* $\mathbf{y}$ *be a real moment sequence of finite rank. The shift operators* $T_i, i \in [n]$ *are symmetric.*

Proof. Suppose that $\operatorname{rank}\mathbf{M}^{\mathbb{R}}(\mathbf{y}) = t$ and $[\mathbf{M}^{\mathbb{R}}(\mathbf{y})]_{\boldsymbol{\beta}\boldsymbol{\gamma}} = \mathbf{a}_{\boldsymbol{\beta}}^{\intercal}\mathbf{a}_{\boldsymbol{\gamma}}$ for $\boldsymbol{\beta},\boldsymbol{\gamma} \in \mathbb{N}^n$, where $\{\mathbf{a}_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathbb{N}^n} \subseteq \mathbb{R}^t$. Let $\mathbf{u} \in \mathbb{R}^t$ be arbitrary and we may write

$$\mathbf{u} = \sum_{\boldsymbol{\alpha}} u_{\boldsymbol{\alpha}}\mathbf{a}_{\boldsymbol{\alpha}}, \quad \vec{\mathbf{u}} := (u_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}}.$$

From

$$\begin{aligned}
\mathbf{u}^{\intercal}T_i\mathbf{u} &= \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}} u_{\boldsymbol{\alpha}}u_{\boldsymbol{\beta}}\mathbf{a}_{\boldsymbol{\alpha}}^{\intercal}(T_i\mathbf{a}_{\boldsymbol{\beta}}) = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}} u_{\boldsymbol{\alpha}}u_{\boldsymbol{\beta}}\mathbf{a}_{\boldsymbol{\alpha}}^{\intercal}\mathbf{a}_{\boldsymbol{\beta}+\mathbf{e}_i} = \vec{\mathbf{u}}^{\intercal}\mathbf{M}^{\mathbb{R}}(x_i\mathbf{y})\,\vec{\mathbf{u}},\\
\mathbf{u}^{\intercal}T_i^{\intercal}\mathbf{u} &= \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}} u_{\boldsymbol{\alpha}}u_{\boldsymbol{\beta}}(T_i\mathbf{a}_{\boldsymbol{\alpha}})^{\intercal}\mathbf{a}_{\boldsymbol{\beta}} = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}} u_{\boldsymbol{\alpha}}u_{\boldsymbol{\beta}}\mathbf{a}_{\boldsymbol{\alpha}+\mathbf{e}_i}^{\intercal}\mathbf{a}_{\boldsymbol{\beta}} = \vec{\mathbf{u}}^{\intercal}\mathbf{M}^{\mathbb{R}}(x_i\mathbf{y})\,\vec{\mathbf{u}},
\end{aligned}$$

we obtain $\mathbf{u}^{\intercal}(T_i - T_i^{\intercal})\mathbf{u} = 0$. Thus, $T_i = T_i^{\intercal}$. ∎

4. Strengthening Lasserre's hierarchy. The study of shift operators enables us to give the following PSD optimality conditions for the pseudo-moment sequence $\mathbf{y}$.

THEOREM 4.1.

(i) Suppose that $\mathbf{M}_r^{\mathbb{R}}(\mathbf{y}) \succeq 0$ for some $r \in \mathbb{N}$. Then for any $s \in \mathbb{N}$ with $s < r$,

$$\begin{bmatrix}\mathbf{M}_s^{\mathbb{R}}(\mathbf{y}) & \mathbf{M}_s^{\mathbb{R}}(x_i\mathbf{y})\\ \mathbf{M}_s^{\mathbb{R}}(x_i\mathbf{y}) & \mathbf{M}_s^{\mathbb{R}}(x_i^2\mathbf{y})\end{bmatrix} \succeq 0, \quad i \in [n]. \tag{4.1}$$

(ii) Suppose that $\mathbf{y}$ is a complex moment sequence admitting a Dirac representing measure. Then for any $s \in \mathbb{N}$,

$$\begin{bmatrix}\mathbf{M}_s^{\mathbb{C}}(\mathbf{y}) & \mathbf{M}_s^{\mathbb{C}}(x_i\mathbf{y})\\ \mathbf{M}_s^{\mathbb{C}}(\overline{x}_i\mathbf{y}) & \mathbf{M}_s^{\mathbb{C}}(|x_i|^2\mathbf{y})\end{bmatrix} \succeq 0, \quad i \in [n]. \tag{4.2}$$

Proof. (i). Assume that $\operatorname{rank}\mathbf{M}_r^{\mathbb{R}}(\mathbf{y}) = t$ and $[\mathbf{M}_r^{\mathbb{R}}(\mathbf{y})]_{\boldsymbol{\beta}\boldsymbol{\gamma}} = \mathbf{a}_{\boldsymbol{\beta}}^{\intercal}\mathbf{a}_{\boldsymbol{\gamma}}$ for $|\boldsymbol{\beta}|,|\boldsymbol{\gamma}| \le r$, where $\{\mathbf{a}_{\boldsymbol{\alpha}}\}_{|\boldsymbol{\alpha}|\le r} \subseteq \mathbb{R}^t$. Let

$$A := \left[\{\mathbf{a}_{\boldsymbol{\alpha}}\}_{|\boldsymbol{\alpha}|\le s}, \{\mathbf{a}_{\boldsymbol{\alpha}+\mathbf{e}_i}\}_{|\boldsymbol{\alpha}|\le s}\right] \in \mathbb{R}^{t\times 2|\mathbb{N}_s^n|}. \tag{4.3}$$

Then one can easily see that

$$\begin{bmatrix}\mathbf{M}_s^{\mathbb{R}}(\mathbf{y}) & \mathbf{M}_s^{\mathbb{R}}(x_i\mathbf{y})\\ \mathbf{M}_s^{\mathbb{R}}(x_i\mathbf{y}) & \mathbf{M}_s^{\mathbb{R}}(x_i^2\mathbf{y})\end{bmatrix} = A^{\intercal}A \succeq 0, \quad \forall i \in [n].$$

(ii). Since $\mathbf{y}$ has a Dirac representing measure, the moment matrix $\mathbf{M}_s^{\mathbb{C}}(\mathbf{y})$ has rank one and the shift operators $T_i, i \in [n]$ are complex numbers. It follows that

$$\begin{bmatrix}1 & \overline{T}_i\\ T_i & \overline{T}_iT_i\end{bmatrix} \succeq 0, \quad i \in [n]. \tag{4.4}$$

Assume that $\mathbf{M}^{\mathbb{C}}(\mathbf{y}) = \mathbf{a}^*\mathbf{a}$, where $\mathbf{a} = (a_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}\in\mathbb{N}^n} \in \mathbb{C}^{\mathbb{N}^n}$. For any $\mathbf{u} = (u_{\boldsymbol{\alpha}})_{|\boldsymbol{\alpha}|\le s}, \boldsymbol{v} = (v_{\boldsymbol{\beta}})_{|\boldsymbol{\beta}|\le s} \in \mathbb{C}^{|\mathbb{N}_s^n|}$, let $u = \sum_{|\boldsymbol{\alpha}|\le s} u_{\boldsymbol{\alpha}}a_{\boldsymbol{\alpha}}, v = \sum_{|\boldsymbol{\beta}|\le s} v_{\boldsymbol{\beta}}a_{\boldsymbol{\beta}}$. We have

$$\begin{aligned}
\overline{u}u &= \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}} \overline{u}_{\boldsymbol{\alpha}}u_{\boldsymbol{\beta}}\overline{a}_{\boldsymbol{\alpha}}a_{\boldsymbol{\beta}} = \mathbf{u}^*\mathbf{M}_s^{\mathbb{C}}(\mathbf{y})\mathbf{u},\\
\overline{u}\overline{T}_iv &= \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}} \overline{u}_{\boldsymbol{\alpha}}v_{\boldsymbol{\beta}}\overline{T_ia_{\boldsymbol{\alpha}}}a_{\boldsymbol{\beta}} = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}} \overline{u}_{\boldsymbol{\alpha}}v_{\boldsymbol{\beta}}\overline{a}_{\boldsymbol{\alpha}+\mathbf{e}_i}a_{\boldsymbol{\beta}} = \mathbf{u}^*\mathbf{M}_s^{\mathbb{C}}(x_i\mathbf{y})\boldsymbol{v},\\
\overline{v}T_iu &= \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}} \overline{v}_{\boldsymbol{\alpha}}u_{\boldsymbol{\beta}}\overline{a}_{\boldsymbol{\alpha}}(T_ia_{\boldsymbol{\beta}}) = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}} \overline{v}_{\boldsymbol{\alpha}}u_{\boldsymbol{\beta}}\overline{a}_{\boldsymbol{\alpha}}a_{\boldsymbol{\beta}+\mathbf{e}_i} = \boldsymbol{v}^*\mathbf{M}_s^{\mathbb{C}}(\overline{x}_i\mathbf{y})\mathbf{u},
\end{aligned}$$

and

$$\overline{v}\overline{T}_iT_iv = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}} \overline{v}_{\boldsymbol{\alpha}} v_{\boldsymbol{\beta}} \overline{T_i a_{\boldsymbol{\alpha}}}(T_i a_{\boldsymbol{\beta}}) = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}} \overline{v}_{\boldsymbol{\alpha}} v_{\boldsymbol{\beta}} \overline{a}_{\boldsymbol{\alpha}+\mathbf{e}_i} a_{\boldsymbol{\beta}+\mathbf{e}_i} = \boldsymbol{v}^*\mathbf{M}_s^{\mathbb{C}}(|x_i|^2\mathbf{y})\boldsymbol{v},$$

which gives

$$\begin{bmatrix}\overline{u} & \overline{v}\end{bmatrix}\begin{bmatrix} I & \overline{T}_i \\ T_i & \overline{T}_iT_i\end{bmatrix}\begin{bmatrix}\mathbf{u}\\ \boldsymbol{v}\end{bmatrix} = \begin{bmatrix}\mathbf{u}^* & \boldsymbol{v}^*\end{bmatrix}\begin{bmatrix}\mathbf{M}_s^{\mathbb{C}}(\mathbf{y}) & \mathbf{M}_s^{\mathbb{C}}(x_i\mathbf{y}) \\ \mathbf{M}_s^{\mathbb{C}}(\overline{x}_i\mathbf{y}) & \mathbf{M}_s^{\mathbb{C}}(|x_i|^2\mathbf{y})\end{bmatrix}\begin{bmatrix}\mathbf{u}\\ \boldsymbol{v}\end{bmatrix}. \tag{4.5}$$

From this and (4.4), we obtain (4.2) as desired. ∎

We say that an operator T is *normal* if $T^*T = TT^*$. In case that T is of finite dimension, it is not hard to see that the normality of T is equivalent to the PSD condition $T^*T - TT^* \succeq 0$, which is further equivalent to

$$\begin{bmatrix} I & T^* \\ T & T^*T\end{bmatrix} \succeq 0. \tag{4.6}$$

Suppose that $\mathbf{y}$ is a complex moment sequence such that $\operatorname{rank}\mathbf{M}^{\mathbb{C}}(\mathbf{y}) = \operatorname{rank}\mathbf{M}_s^{\mathbb{C}}(\mathbf{y})$. In a similar manner as the proof of Theorem 4.1 (ii), we can show that the shift operators $T_i, i \in [n]$ are normal if and only if the PSD conditions (4.2) hold. It would be interesting to ask: if $\mathbf{y}$ is a complex moment sequence of finite rank, do we have that the shift operators $T_i, i \in [n]$ are normal? We will explore this question in the future work.

Using the PSD optimality conditions in Theorem 4.1, we can strengthen Lasserre's hierarchy of moment relaxations. In particular, for real polynomial optimization, we consider

$$\rho_r' := \begin{cases} \inf\limits_{\mathbf{y}} & L_{\mathbf{y}}(f) \\ \text{s.t.} & y_{\mathbf{0}} = 1, \\ & \mathbf{M}_{r-d_i}^{\mathbb{R}}(g_i\mathbf{y}) \succeq 0, \quad i \in [m], \\ & \begin{bmatrix}\mathbf{M}_r^{\mathbb{R}}(\mathbf{y}) & \mathbf{M}_r^{\mathbb{R}}(x_i\mathbf{y}) \\ \mathbf{M}_r^{\mathbb{R}}(x_i\mathbf{y}) & \mathbf{M}_r^{\mathbb{R}}(x_i^2\mathbf{y})\end{bmatrix} \succeq 0, \quad i \in [n]. \end{cases} \tag{4.7}$$

THEOREM 4.2. *It holds $\rho_r \le \rho_r' \le \rho_{r+1} \le f_{\min}$ for any $r \ge d_{\min}$.*

Proof. Since (4.7) is a strengthening of (2.3), it follows $\rho_r \le \rho_r'$. The inequality $\rho_r' \le \rho_{r+1}$ follows from the fact that the second PSD constraints of (4.7) are implied by $\mathbf{M}_{r+1}^{\mathbb{R}}(\mathbf{y}) \succeq 0$ due to Theorem 4.1 (i). ∎

By Theorem 4.2, (4.7) provides an intermediate relaxation between the r-th and $(r+1)$-th moment relaxations for (RPOP).

For complex polynomial optimization, we consider

$$\tau_{r,s}' := \begin{cases} \inf\limits_{\mathbf{y}} & L_{\mathbf{y}}(f) \\ \text{s.t.} & \mathbf{M}_r^{\mathbb{C}}(\mathbf{y}) \succeq 0, \quad y_{\mathbf{0},\mathbf{0}} = 1, \\ & \mathbf{M}_{r-d_i}^{\mathbb{C}}(g_i\mathbf{y}) \succeq 0, \quad i \in [m], \\ & \begin{bmatrix}\mathbf{M}_s^{\mathbb{C}}(\mathbf{y}) & \mathbf{M}_s^{\mathbb{C}}(x_i\mathbf{y}) \\ \mathbf{M}_s^{\mathbb{C}}(\overline{x}_i\mathbf{y}) & \mathbf{M}_s^{\mathbb{C}}(|x_i|^2\mathbf{y})\end{bmatrix} \succeq 0, \quad i \in [n]. \end{cases} \tag{4.8}$$

Here $s \in \mathbb{N}$ is a tunable parameter which we call the *normal order*.

THEOREM 4.3. *It hold $\tau_r \leq \tau'_{r,s} \leq \tau'_{r,s+1} \leq f_{\min}$ and $\tau'_{r,s} \leq \tau'_{r+1,s}$ for any $r \geq d_{\min}$ and any $s \in \mathbb{N}$.*

Proof. Since (4.8) is a strengthening of (2.6), it follows $\tau_r \leq \tau'_r$. If the infimum of (CPOP) is attained, let $\mathbf{w}$ be a minimizer of (CPOP) and $\mathbf{y}$ be the moment sequence of the Dirac measure $\delta_{\mathbf{w}}$. By Theorem 4.1 (ii), $\mathbf{y}$ is a feasible solution of (4.8) and $L_{\mathbf{y}}(f) = f_{\min}$. Thus, $\tau'_r \leq f_{\min}$. If the infimum of (CPOP) is not attained, let $\{\mathbf{w}^{(k)}\}_{k\geq 1}$ be a minimizing sequence of (CPOP) and $\mathbf{y}^{(k)}$ be the moment sequence of the Dirac measure $\delta_{\mathbf{w}^{(k)}}$, respectively. We have that every $\mathbf{y}^{(k)}$ is a feasible solution of (4.8) and $\lim_{k\to\infty} L_{\mathbf{y}^{(k)}}(f) = f_{\min}$. Thus, $\tau'_r \leq f_{\min}$. The inequalities $\tau'_{r,s} \leq \tau'_{r,s+1}$ and $\tau'_{r,s} \leq \tau'_{r+1,s}$ are easily obtained from the constructions. ∎

By Theorem 4.3, (4.8) is a two-level hierarchy indexed by the relaxation order r and the normal order s, and hence allows one more level of flexibility by playing with the two parameters.

5. Integration with sparsity. The strengthening technique discussed in Section 4 can be integrated into different sparse versions of Lasserre's hierarchy to improve scalability. We refer the reader to [9] for relevant details on different sparse versions of Lasserre's hierarchy.

5.1. Correlative sparsity. Consider (RPOP) (resp. (CPOP)). Suppose that the two index sets $[n]$ and $[m]$ can be decomposed into $\{I_1, \ldots, I_p\}$ and $\{J_1, \ldots, J_p\}$, respectively, such that 1) $f = f_1 + \cdots + f_p$ with $f_k \in \mathbb{R}[\mathbf{x}_{I_k}]$ (resp. $\mathbb{C}[\mathbf{x}_{I_k}, \overline{\mathbf{x}}_{I_k}]$) for $k \in [p]$; 2) for all $k \in [p]$ and $i \in J_k$, $g_i \in \mathbb{R}[\mathbf{x}_{I_k}]$ (resp. $\mathbb{C}[\mathbf{x}_{I_k}, \overline{\mathbf{x}}_{I_k}]$), where $\mathbb{R}[\mathbf{x}_{I_k}]$ (resp. $\mathbb{C}[\mathbf{x}_{I_k}, \overline{\mathbf{x}}_{I_k}]$) denotes the polynomial ring in those variables indexed by I_k. Let $\mathbf{M}^{\mathbb{R}}_r(\mathbf{y}, I_k)$ (resp. $\mathbf{M}^{\mathbb{R}}_r(g\mathbf{y}, I_k)$) be the submatrix obtained from $\mathbf{M}^{\mathbb{R}}_r(\mathbf{y})$ (resp. $\mathbf{M}^{\mathbb{R}}_r(g\mathbf{y})$) by retaining only those rows and columns indexed by $\boldsymbol{\beta} \in \mathbb{N}^n_r$ of $\mathbf{M}^{\mathbb{R}}_r(\mathbf{y})$ (resp. $\mathbf{M}^{\mathbb{R}}_r(g\mathbf{y})$) with $\beta_i = 0$ if $i \notin I_k$. Then, we can strengthen the correlative sparse Lasserre's hierarchy of moment relaxations for real polynomial optimization by considering

$$
\text{(5.1)}\qquad \begin{cases}
\inf\limits_{\mathbf{y}} & L_{\mathbf{y}}(f)\\
\text{s.t.} & \mathbf{M}^{\mathbb{R}}_r(\mathbf{y}, I_k) \succeq 0, \quad k \in [p],\\
& \mathbf{M}^{\mathbb{R}}_{r-d_i}(g_i\mathbf{y}, I_k) \succeq 0, \quad i \in J_k, k \in [p],\\
& \begin{bmatrix} \mathbf{M}^{\mathbb{R}}_1(\mathbf{y}) & \mathbf{M}^{\mathbb{R}}_1(x_i\mathbf{y})\\ \mathbf{M}^{\mathbb{R}}_1(x_i\mathbf{y}) & \mathbf{M}^{\mathbb{R}}_1(x_i^2\mathbf{y}) \end{bmatrix} \succeq 0, \quad i \in [n],\\
& y_{\mathbf{0}} = 1.
\end{cases}
$$

Also, we can strengthen the correlative sparse Lasserre's hierarchy of moment relaxations for complex polynomial optimization by considering

$$
\text{(5.2)}\qquad \begin{cases}
\inf\limits_{\mathbf{y}} & L_{\mathbf{y}}(f)\\
\text{s.t.} & \mathbf{M}^{\mathbb{C}}_r(\mathbf{y}, I_k) \succeq 0, \quad k \in [p],\\
& \mathbf{M}^{\mathbb{C}}_{r-d_i}(g_i\mathbf{y}, I_k) \succeq 0, \quad i \in J_k, k \in [p],\\
& \begin{bmatrix} \mathbf{M}^{\mathbb{C}}_s(\mathbf{y}, I_k) & \mathbf{M}^{\mathbb{C}}_s(x_i\mathbf{y}, I_k)\\ \mathbf{M}^{\mathbb{C}}_s(\overline{x}_i\mathbf{y}, I_k) & \mathbf{M}^{\mathbb{C}}_s(|x_i|^2\mathbf{y}, I_k) \end{bmatrix} \succeq 0, \quad i \in I_k, k \in [p],\\
& y_{\mathbf{0},\mathbf{0}} = 1.
\end{cases}
$$

5.2. Sign symmetry. For $p \in \mathbb{R}[\mathbf{x}]$ and a binary vector $\mathbf{s} \in \{0,1\}^n$, let $[p]_{\mathbf{s}} \in \mathbb{R}[\mathbf{x}]$ be defined by $[p]_{\mathbf{s}}(x_1, \ldots, x_n) := p((-1)^{s_1}x_1, \ldots, (-1)^{s_n}x_n)$. Then p is said to have the sign symmetry represented by $\mathbf{s} \in \{0,1\}^n$ if $[p]_{\mathbf{s}} = p$. We use $S(p) \subseteq \{0,1\}^n$ to denote all sign symmetries of p. Consider (RPOP) and let $U := S(f) \cap \bigcap_{i=1}^m S(g_i)$. We define an equivalence relation $\sim$ on $[x]$ by

$$\mathbf{x}^{\boldsymbol{\alpha}} \sim \mathbf{x}^{\boldsymbol{\beta}} \iff U \subseteq S(\mathbf{x}^{\boldsymbol{\alpha}+\boldsymbol{\beta}}). \tag{5.3}$$

For each $i \in [m]$, the equivalence relation $\sim$ gives rise to a partition of $[\mathbf{x}]_{r-d_i}$:

$$[\mathbf{x}]_{r-d_i} = \bigsqcup_{k=1}^{p_i} [\mathbf{x}]_{r-d_i,k}. \tag{5.4}$$

We then build the submatrix $\mathbf{M}^{\mathbb{R}}_{r-d_i,k}(g_i\mathbf{y})$ of $\mathbf{M}^{\mathbb{R}}_{r-d_i}(g_i\mathbf{y})$ with respect to the sign symmetry by retaining only those rows and columns indexed by $[\mathbf{x}]_{r-d_i,k}$ for each $k \in [p_i]$. Moreover, for each $i \in [n]$, the equivalence relation $\sim$ gives rise to a partition of $[\mathbf{x}]_r \cup x_i[\mathbf{x}]_r$: $[\mathbf{x}]_r \cup x_i[\mathbf{x}]_r = \bigsqcup_{k=1}^{q_i} [\mathbf{x}]_{r,i,k}$. We build the submatrix $\mathbf{N}^{\mathbb{R}}_{r,i,k}(\mathbf{y})$ of the second PSD matrix in (4.7) by retaining only those rows and columns indexed by $[\mathbf{x}]_{r,i,k}$ for each $k \in [q_i]$. Then, we can strengthen the sign-symmetry Lasserre's hierarchy of moment relaxations for real polynomial optimization by considering

$$\begin{cases} \inf\limits_{\mathbf{y}} & L_{\mathbf{y}}(f) \\ \text{s.t.} & \mathbf{M}^{\mathbb{R}}_{r-d_i,k}(g_i\mathbf{y}) \succeq 0, \quad k \in [p_i], \\ & \mathbf{N}^{\mathbb{R}}_{r,i,k}(\mathbf{y}) \succeq 0, \quad k \in [q_i], i \in [n], \\ & y_{\mathbf{0}} = 1. \end{cases} \tag{5.5}$$

The complex case proceeds in a similar way, which we omit for conciseness.

6. Numerical experiments.

The strengthened real and complex Lasserre's hierarchies have been implemented in the Julia package `TSSOS`[1]. In this section, we evaluate their performance on diverse polynomial optimization problems using `TSSOS` and `Mosek` 10.0 [1] is employed as an SDP solver with default settings. When presenting the results, 'LAS' means the usual Lasserre's hierarchy and 'S-LAS' means the strengthened Lasserre's hierarchy; the column labelled by 'opt' records optima of SDPs and the column labelled by 'time' records running time in seconds. Moreover, the symbol '-' means that `Mosek` runs out of memory. All numerical experiments were performed on a desktop computer with Intel(R) Core(TM) i9-10900 CPU@2.80GHz and 64G RAM.

6.1. Minimizing a random real quadratic polynomial with binary variables.

Let us minimize a random real quadratic polynomial with binary variables:

$$\begin{cases} \inf\limits_{\mathbf{x}\in\mathbb{R}^n} & [\mathbf{x}]_1^{\intercal} Q [\mathbf{x}]_1 \\ \text{s.t.} & x_i^2 = 1, \quad i = 1, \ldots, n, \end{cases} \tag{6.1}$$

where $Q \in \mathbb{R}^{(n+1)\times(n+1)}$ is a random symmetric matrix whose entries are selected with respect to the uniform probability distribution on $[0,1]$. For each $n \in \{10, 20, 30, 40\}$, we solve three instances using LAS ($r = 1, 2$) and S-LAS ($r = 1, s = 1$), respectively. The results are presented in Table 1. For this problem, we empirically observe that LAS at $r = 2$ achieves global optimality. It can be seen from the table that the strengthening technique significantly improves the bound provided by LAS at $r = 1$ while it is much cheaper than going to LAS at $r = 2$.

[1] `TSSOS` is freely available at https://github.com/wangjie212/TSSOS.

TABLE 1
Minimizing a random real quadratic polynomial with binary variables.

n	trial	LAS ($r=1$)		LAS ($r=2$)		S-LAS ($r=1$)	
		opt	time	opt	time	opt	time
10	1	-6.9868	0.006	-6.6118	0.04	-6.6118	0.03
	2	-9.9016	0.006	-9.6732	0.04	-9.6732	0.03
	3	-8.6265	0.007	-6.6963	0.04	-6.8216	0.03
20	1	-26.613	0.01	-23.407	5.95	-23.521	0.43
	2	-28.474	0.01	-24.330	6.08	-26.575	0.44
	3	-30.996	0.01	-27.657	5.61	-27.657	0.47
30	1	-51.429	0.08	-44.597	382	-47.817	6.29
	2	-57.277	0.03	-49.871	435	-53.539	5.74
	3	-49.950	0.03	-42.548	479	-46.970	5.30
40	1	-79.672	0.09	-	-	-74.532	43.1
	2	-83.814	0.13	-	-	-81.274	36.9
	3	-85.887	0.09	-	-	-79.748	41.0

6.2. The point cloud registration problem. Given two sets of 3D points $\{\mathbf{a}_i\}_{i=1}^N, \{\mathbf{b}_i\}_{i=1}^N$ with putative correspondences $\mathbf{a}_i \leftrightarrow \mathbf{b}_i$, the point cloud registration problem in computer vision is to find the best 3D rotation R and translation $\mathbf{t}$ to align them while explicitly tolerating outliers. It can be formulated as the nonlinear optimization problem:

$$\min_{R\in SO(3),\mathbf{t}\in\mathbb{R}^3} \sum_{i=1}^N \min\left\{\frac{\|\mathbf{b}_i - R\mathbf{a}_i - \mathbf{t}\|^2}{\beta_i^2}, 1\right\}, \tag{6.2}$$

where $\beta_i > 0$ is a given threshold that determines the maximum inlier residual. By introducing N binary variables $\{\theta_i\}_{i=1}^N$, (6.2) can be equivalently reformulated as a polynomial optimization problem:

$$\min_{\substack{R\in SO(3),\mathbf{t}\in\mathbb{R}^3,\\ \theta_i\in\{-1,1\}}} \sum_{i=1}^N \frac{1+\theta_i}{2}\frac{\|\mathbf{b}_i - R\mathbf{a}_i - \mathbf{t}\|^2}{\beta_i^2} + \frac{1-\theta_i}{2}. \tag{6.3}$$

Note that in (6.3), the rotation matrix R can be parametrized by its entries which we denote by $\mathbf{r}$ and the constraint $R \in SO(3)$ can be expressed by polynomial constraints in $\mathbf{r}$. Yang and Carlone [20] proposed a customized monomial basis for the dense Lasserre's hierarchy for (6.3) which is $[1, \mathbf{x}, \boldsymbol{\theta}, \mathbf{r}\otimes\mathbf{t}, \mathbf{x}\otimes\boldsymbol{\theta}]$ with $\mathbf{x} \coloneqq [\mathbf{r}, \mathbf{t}]$ and $\boldsymbol{\theta} \coloneqq \{\theta_i\}_{i=1}^N$. Moreover, they also proposed a sparse Lasserre's hierarchy for (6.3) in which the variables are decomposed into N cliques: $[\mathbf{x}, \theta_i], i \in [N]$ and for the i-th clique, the monomial basis $[1, \mathbf{x}, \theta_i, \mathbf{r}\otimes\mathbf{t}, \mathbf{x}\otimes\theta_i]$ is used. It was empirically shown in [20] that the dense Lasserre's hierarchy achieves global optimality at relaxation order $r = 2$ while the sparse Lasserre's hierarchy is usually not tight at the same relaxation order.

For each $N \in \{10, 20, 30, 40\}$, we randomly generate three instances of (6.3) with 60% outliers. We solve each instance using the dense LAS (with the above monomial basis) at $r = 2$, the sparse LAS (with the above monomial basis) at $r = 2, s = 1$, and the sparse S-LAS (with the above monomial basis) at $r = 2$, respectively. The results are presented in Table 2 from which we can see that the strengthening technique improves the bound provided by the sparse LAS while it is much cheaper than the dense LAS.

TABLE 2
The point cloud registration problem.

N	trial	dense LAS		sparse LAS		sparse S-LAS	
		opt	time	opt	time	opt	time
10	1	6.5437	18.3	6.1294	1.32	6.2392	4.05
	2	6.4687	17.8	6.2538	1.33	6.4461	4.56
	3	6.3971	21.0	6.1144	1.32	6.2634	4.50
20	1	14.062	424	12.345	2.28	13.007	28.3
	2	14.256	350	12.423	2.79	13.053	29.1
	3	13.780	321	12.279	2.57	12.851	26.8
30	1	20.870	2461	18.670	3.47	19.696	138
	2	20.263	2808	18.522	4.98	19.381	139
	3	20.452	2435	18.459	3.64	19.792	136
40	1	-	-	24.942	4.84	26.495	662
	2	-	-	24.783	4.62	26.751	630
	3	-	-	24.888	4.25	27.295	632

6.3. Minimizing a random complex quadratic polynomial with unit-norm variables. Let us now minimize a random complex quadratic polynomial with unit-norm variables:

$$\begin{cases} \inf\limits_{\mathbf{x}\in\mathbb{C}^n} & [\mathbf{x}]_1^\star Q[\mathbf{x}]_1 \\ \text{s.t.} & |x_i|^2 = 1, \quad i = 1,\ldots,n, \end{cases} \tag{6.4}$$

where $Q \in \mathbb{C}^{(n+1)\times(n+1)}$ is a random Hermitian matrix whose entries (both real and imaginary parts) are selected with respect to the uniform probability distribution on $[0,1]$. For each $n \in \{10, 20, 30\}$, we solve three instances using LAS ($r = 1, 2$) and S-LAS ($r = 1, s = 1$), respectively. The results are presented in Table 3. For this problem, we empirically observe that LAS at $r = 2$ achieves global optimality. It is evident from the table that the strengthening technique significantly improves the bound (indeed, achieving global optimality for $n \le 20$) provided by LAS at $r = 1$ while it is much cheaper than going to LAS at $r = 2$.

6.4. Minimizing a random complex quartic polynomial on a sphere. Let us minimize a random complex quartic polynomial on a unit sphere:

$$\begin{cases} \inf\limits_{\mathbf{x}\in\mathbb{C}^n} & [\mathbf{x}]_2^\star Q[\mathbf{x}]_2 \\ \text{s.t.} & |x_1|^2 + \cdots + |x_n|^2 = 1, \end{cases} \tag{6.5}$$

where $Q \in \mathbb{C}^{|[\mathbf{x}]_2|\times|[\mathbf{x}]_2|}$ ($|[\mathbf{x}]_2|$ is the cardinality of $[\mathbf{x}]_2$) is a random Hermitian matrix whose entries (both real and imaginary parts) are selected with respect to the uniform

Table 3
Minimizing a random complex quadratic polynomial with unit-norm variables.

n	trial	LAS ($r=1$)		LAS ($r=2$)		S-LAS ($r=1$)	
		opt	time	opt	time	opt	time
10	1	-10.830	0.01	-10.474	1.57	-10.474	0.15
	2	-14.005	0.01	-13.905	1.76	-13.905	0.15
	3	-14.308	0.01	-13.751	1.71	-13.751	0.16
20	1	-39.274	0.03	-38.323	1227	-38.323	6.39
	2	-44.009	0.03	-43.911	1076	-43.911	5.51
	3	-43.043	0.03	-42.017	1061	-42.017	5.76
30	1	-75.249	0.14	-	-	-72.948	234
	2	-79.995	0.13	-	-	-79.382	161
	3	-74.888	0.12	-	-	-73.680	148

probability distribution on $[0,1]$. For each $n \in \{5, 10, 15\}$, we solve three instances using LAS ($r = 2, 3$) and S-LAS ($r = 2, s = 1$), respectively. The results are presented in Table 4. We can see from the table that the strengthening technique significantly improves the bound provided by LAS at both $r = 2$ and $r = 3$ while it is much cheaper than going to LAS at $r = 3$.

Table 4
Minimizing a random complex quartic polynomial on a unit sphere.

n	trial	LAS ($r=2$)		LAS ($r=3$)		S-LAS ($r=2$)	
		opt	time	opt	time	opt	time
5	1	-4.4125	0.04	-4.1976	2.09	-4.0517	0.06
	2	-2.9632	0.04	-2.5182	1.94	-2.3767	0.05
	3	-3.9058	0.04	-3.3651	1.97	-3.1354	0.05
10	1	-5.9950	3.08	-	-	-4.6231	4.50
	2	-5.9757	2.93	-	-	-4.5794	4.08
	3	-5.6221	3.05	-	-	-4.1087	4.18
15	1	-8.5265	82.3	-	-	-6.5370	130
	2	-8.0241	87.4	-	-	-6.3118	121
	3	-8.0791	85.7	-	-	-6.1881	123

6.5. Minimizing a random complex quartic polynomial with correlative sparsity on multi-spheres. Let us minimize a random complex quartic polynomial with correlative sparsity on multi-spheres:

$$\begin{cases} \inf\limits_{\mathbf{x}\in\mathbb{C}^n} & \sum_{i=1}^{l}[\mathbf{x}_i]_2^{\star}Q_i[\mathbf{x}_i]_2 \\ \text{s.t.} & \|\mathbf{x}_i\|^2 = 1, \quad i \in [l], \end{cases} \tag{6.6}$$

where $n = 4l + 2$, $\mathbf{x}_i \coloneqq \{x_{4i-3}, \ldots, x_{4i+2}\}$, and $Q_i \in \mathbb{C}^{|[\mathbf{x}_i]_2|\times|[\mathbf{x}_i]_2|}$ is a random Hermitian matrix whose entries (both real and imaginary parts) are selected with respect to the uniform probability distribution on $[0, 1]$. For each $l \in \{5, 10, 50, 100\}$, we solve three instances using the sparse LAS ($r = 2, 3$) and the sparse S-LAS ($r = 2, s = 1$), respectively. The results are presented in Table 5. Again, we can conclude from the table that the strengthening technique significantly improves the bound provided by the sparse LAS at both $r = 2$ and $r = 3$ while it is much cheaper than going to the sparse LAS at $r = 3$.

TABLE 5
Minimizing a random complex quartic polynomial on multi-spheres.

n	trial	LAS ($r = 2$)		LAS ($r = 3$)		S-LAS ($r = 2$)	
		opt	time	opt	time	opt	time
22	1	-16.561	0.48	-13.190	87.7	-12.911	0.74
	2	-17.891	0.47	-14.468	89.5	-13.918	0.72
	3	-18.119	0.51	-14.408	90.7	-14.094	0.71
42	1	-34.424	1.28	-27.404	122	-26.607	1.68
	2	-35.052	1.30	-28.862	124	-27.896	1.81
	3	-34.392	1.24	-27.796	122	-27.071	1.66
202	1	-168.10	5.15	-133.07	645	-132.14	9.81
	2	-168.90	5.14	-135.14	597	-133.14	8.34
	3	-166.92	4.35	-135.01	612	-132.40	8.89
402	1	-339.50	10.7	-	-	-268.95	23.5
	2	-328.91	12.1	-	-	-259.32	23.5
	3	-333.95	11.0	-	-	-264.59	21.8

6.6. Smale's Mean Value conjecture. The following complex polynomial optimization problem is borrowed from [17]:

$$
\begin{cases}
\sup\limits_{(\mathbf{z},u)\in\mathbb{C}^{n+1}} & |u| \\
\text{s.t.} & |H(z_i)| \geq |u|, \quad i = 1, \ldots, n, \\
& z_1 \cdots z_n = \frac{(-1)^n}{n+1}, \\
& |z_1|^2 + |z_2|^2 + \cdots + |z_n|^2 = n\left(\frac{1}{n+1}\right)^{\frac{2}{n}},
\end{cases}
\tag{6.7}
$$

where $H(y) \coloneqq \frac{1}{y}\int_0^y p(z)\,\mathrm{d}z$ and $p(z) \coloneqq (n+1)(z-z_1)\cdots(z-z_n)$ with $p(0) = 1$. This problem is used in [17] to verify Smale's Mean Value conjecture [13, 14] which is open for $n \geq 4$ since 1981. The optimum of (6.7) is conjectured to be $\frac{n}{n+1}$. We refer the reader to [17] for more details. Here we solve (6.7) with $n = 4$ using LAS ($r = 4, 6, 8$) and S-LAS ($r = 4, s = 1, 2, 3$). The results are presented in Table 6, from which we see that the strengthening technique enables us to achieve global optimality at lower relaxation orders so that the computational cost is significantly reduced.

6.7. The Mordell inequality conjecture. Our next example concerns the Mordell inequality conjecture due to Birch in 1958: if the numbers $z_1, \ldots, z_n \in \mathbb{C}$ satisfies $|z_1|^2 + \cdots + |z_n|^2 = n$, then the maximum of $\prod_{1\leq i<j\leq n} |z_i - z_j|^2$ is n^n. This conjecture was proved for $n \leq 4$ and disproved for $n \geq 6$, and so the only remaining open case is when $n = 5$. The reader is referred to [17] for more details. Without

TABLE 6
The results for (6.7) *with* $n = 4$.

	$r=4$		$r=6$		$r=8$	
LAS	opt	time	opt	time	opt	time
	1.4218	0.16	0.8404	22.8	-	-
	$r=4, s=1$		$r=4, s=2$		$r=4, s=3$	
S-LAS	opt	time	opt	time	opt	time
	1.4218	0.17	1.2727	0.45	0.8000	18.2

loss of generality, we may eliminate one variable and reformulate the conjecture as the following complex polynomial optimization problem:

$$\begin{cases} \sup\limits_{\mathbf{z}\in\mathbb{C}^{n-1}} & \prod_{1\le i<j\le n-1}|z_i-z_j|^2\prod_{i=1}^{n-1}|z_i+z_1+\ldots+z_{n-1}|^2 \\ \text{s.t.} & |z_1|^2+\cdots+|z_{n-1}|^2+|z_1+\ldots+z_{n-1}|^2=n. \end{cases} \tag{6.8}$$

Here we solve (6.8) with $n = 3, 4$ using LAS and S-LAS. The results are presented in Tables 7 and 8, respectively. From the tables, we see that the strengthening technique enables us to achieve global optimality at much lower relaxation orders so that the computational cost is significantly reduced.

TABLE 7
The results for (6.8) *with* $n = 3$.

	$r=4$		$r=6$		$r=8$	
LAS	opt	time	opt	time	opt	time
	27.347	0.04	27.122	0.19	27.074	0.35
	$r=3, s=0$		$r=3, s=1$		$r=3, s=2$	
S-LAS	opt	time	opt	time	opt	time
	54.000	0.005	54.000	0.008	27.000	0.01

TABLE 8
The results for (6.8) *with* $n = 4$.

	$r=10$		$r=12$		$r=14$		$r=16$		$r=18$	
LAS	opt	time	opt	time	opt	time	opt	time	opt	time
	343.66	8.58	326.85	50.1	292.89	212	277.64	790	-	-
	$r=6, s=1$		$r=6, s=2$		$r=6, s=3$		$r=6, s=4$		$r=6, s=5$	
S-LAS	opt	time	opt	time	opt	time	opt	time	opt	time
	1638.4	0.13	1337.5	0.20	932.20	0.25	582.86	0.76	256.00	3.10

6.8. The AC-OPF problem. The AC optimal power flow (AC-OPF) is a central problem in power systems, which aims to minimize the generation cost of an alternating current transmission network under the physical constraints. Mathematically, it can be formulated as the following complex polynomial optimization problem:

$$
\left\{
\begin{array}{ll}
\inf\limits_{\{V_i\}_{i\in N}} & \sum_{k\in G}\Big(\mathbf{c}_{2k}\Re\big(\mathbf{S}^d_{i_k}+\mathbf{Y}^{sh}_{i_k}|V_{i_k}|^2+\sum_{(i_k,j)\in E_{i_k}\cup E^R_{i_k}} S_{i_kj}\big)^2 \\
& \qquad +\mathbf{c}_{1k}\Re\big(\mathbf{S}^d_{i_k}+\mathbf{Y}^{sh}_{i_k}|V_{i_k}|^2+\sum_{(i_k,j)\in E_{i_k}\cup E^R_{i_k}} S_{i_kj}\big)+\mathbf{c}_{0k}\Big) \\
\text{s.t.} & \angle V_{ref}=0, \\
& \mathbf{S}^{gl}_k \le \mathbf{S}^d_{i_k}+\mathbf{Y}^{sh}_{i_k}|V_{i_k}|^2+\sum_{(i_k,j)\in E_{i_k}\cup E^R_{i_k}} S_{i_kj} \le \mathbf{S}^{gu}_k, \quad \forall k\in G, \\
& \boldsymbol{v}^l_i \le |V_i| \le \boldsymbol{v}^u_i, \quad \forall i\in N, \\
& S_{ij}=\left(\mathbf{Y}^*_{ij}-\mathbf{i}\frac{\mathbf{b}^c_{ij}}{2}\right)\frac{|V_i|^2}{|\mathbf{T}_{ij}|^2}-\mathbf{Y}^*_{ij}\frac{V_iV_j^*}{\mathbf{T}_{ij}}, \quad \forall (i,j)\in E, \\
& S_{ji}=\left(\mathbf{Y}^*_{ij}-\mathbf{i}\frac{\mathbf{b}^c_{ij}}{2}\right)|V_j|^2-\mathbf{Y}^*_{ij}\frac{V_i^*V_j}{\mathbf{T}^*_{ij}}, \quad \forall (i,j)\in E, \\
& |S_{ij}|\le \mathbf{s}^u_{ij}, \quad \forall (i,j)\in E\cup E^R, \\
& \boldsymbol{\theta}^{\Delta l}_{ij} \le \angle(V_iV_j^*) \le \boldsymbol{\theta}^{\Delta u}_{ij}, \quad \forall (i,j)\in E.
\end{array}
\right. \tag{6.9}
$$

For a full description on the AC-OPF problem, the reader may refer to [2] as well as [16]. For an AC-OPF instance, we can obtain an upper bound ('ub') on the optimum from a local solver. Then the *optimality gap* between the upper bound and the lower bound ('lb') provided by SDP relaxations is defined by

$$\text{gap} \coloneqq \frac{\text{ub}-\text{lb}}{\text{ub}} \times 100\%.$$

For our purpose, we select instances from the AC-OPF library *PGLiB* [2] that exhibit significant optimality gaps. The number appearing in each case name stands for the number of buses, which is equal to the number of complex variables involved in (6.9). We solve each instance using the sparse LAS and the sparse S-LAS with minimum relaxation order [16]. The results are presented in Table 9, from which we see that the strengthening technique substantially reduce the optimality gap in most cases.

REFERENCES

[1] E. D. Andersen and K. D. Andersen, *The Mosek Interior Point Optimizer for Linear Programming: An Implementation of the Homogeneous Algorithm*, in High Performance Optimization, vol. 33 of Applied Optimization, Springer US, 2000, pp. 197–232, https://doi.org/10.1007/978-1-4757-3216-0_8.

[2] S. Babaeinejadsarookolaee, A. Birchfield, R. D. Christie, C. Coffrin, C. DeMarco, R. Diao, M. Ferris, S. Fliscounakis, S. Greene, R. Huang, et al., *The power grid library for benchmarking AC optimal power flow algorithms*, arXiv preprint arXiv:1908.02788, (2019).

[3] D. Henrion, M. Korda, and J. B. Lasserre, *Moment-sos Hierarchy, The: Lectures In Probability, Statistics, Computational Geometry, Control And Nonlinear Pdes*, vol. 4, World Scientific, 2020.

[4] D. Henrion and J.-B. Lasserre, *Detecting global optimality and extracting solutions in gloptipoly*, in Positive Polynomials in Control, Springer, 2005, pp. 293–310.

[5] C. Josz and D. K. Molzahn, *Lasserre hierarchy for large scale polynomial optimization in real and complex variables*, SIAM Journal on Optimization, 28 (2018), pp. 1017–1048.

[6] J.-B. Lasserre, *Global Optimization with Polynomials and the Problem of Moments*, SIAM Journal on Optimization, 11 (2001), pp. 796–817.

[7] J. B. Lasserre, M. Laurent, and P. Rostalski, *Semidefinite characterization and computation of zero-dimensional real radical ideals*, Foundations of Computational Mathematics, 8 (2008), pp. 607–647.

[8] M. Laurent, *Sums of squares, moment matrices and optimization over polynomials*, in Emerging applications of algebraic geometry, Springer, 2009, pp. 157–270.

[9] V. Magron and J. Wang, *Sparse polynomial optimization: theory and practice*, World Scientific, 2023.

Table 9
The results for the AC-OPF problem.

case name	LAS			S-LAS		
	opt	time	gap	opt	time	gap
30_as	7.3351e2	0.35	8.66%	7.4426e2	1.13	7.32%
39_epri	1.3749e5	0.27	0.66%	1.3841e5	0.71	0.00%
162_ieee_dtc	1.0176e5	13.5	5.84%	1.0647e5	114	1.49%
179_goc	7.5150e5	2.26	0.36%	7.5386e5	7.69	0.05%
30_as_sad	8.6569e2	0.33	3.52%	8.7496e2	1.33	2.49%
118_ieee_sad	9.6760e4	2.10	7.98%	1.0294e5	5.79	2.10%
162_ieee_dtc_sad	1.0176e5	10.8	6.36%	1.0738e5	118	1.20%
179_goc_sad	7.5279e5	2.43	1.27%	7.5581e5	7.41	0.88%
30_as_api	2.6237e3	0.41	47.4%	4.9935e3	1.19	0.05%
39_epri_api	2.4511e5	0.23	1.82%	2.4963e5	0.73	0.01%
89_pegase_api	1.0139e5	15.2	22.1%	1.0507e5	44.1	19.2%
118_ieee_api	1.7571e5	2.01	27.4%	2.2293e5	5.99	7.96%
162_ieee_dtc_api	1.1526e5	9.81	4.73%	1.1956e5	119	1.17%
179_goc_api	1.8603e6	2.74	3.70%	1.9226e6	8.11	0.48%

[10] J. Nie, *Tight relaxations for polynomial optimization and lagrange multiplier expressions*, Mathematical Programming, 178 (2019), pp. 1–37.

[11] J. Nie, *Moment and Polynomial Optimization*, 2023.

[12] C. Riener, T. Theobald, L. J. Andrén, and J. B. Lasserre, *Exploiting symmetries in sdp-relaxations for polynomial optimization*, Mathematics of Operations Research, 38 (2013), pp. 122–141.

[13] S. Smale, *The fundamental theorem of algebra and complexity theory*, Bulletin (New Series) of the American Mathematical Society, 4 (1981), pp. 1–36.

[14] S. Smale, *Mathematical problems for the next century*, The mathematical intelligencer, 20 (1998), pp. 7–15.

[15] J. Wang, *A more efficient reformulation of complex SDP as real SDP*, 2023. Preprint https: arXiv.org/abs/2307.11599.

[16] J. Wang and V. Magron, *Exploiting sparsity in complex polynomial optimization*, Journal of Optimization Theory and Applications, 192 (2022), pp. 335–359.

[17] J. Wang and V. Magron, *A real moment-hsos hierarchy for complex polynomial optimization with real coefficients*, arXiv preprint arXiv:2308.14631, (2023).

[18] J. Wang, V. Magron, and J.-B. Lasserre, *TSSOS: A moment-SOS hierarchy that exploits term sparsity*, SIAM Journal on Optimization, 31 (2021), pp. 30–58.

[19] J. Wang, V. Magron, J.-B. Lasserre, and N. H. A. Mai, *CS-TSSOS: Correlative and term sparsity for large-scale polynomial optimization*, arXiv:2005.02828, (2020).

[20] H. Yang and L. Carlone, *One ring to rule them all: Certifiably robust geometric perception with outliers*, in Advances in Neural Information Processing Systems, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, eds., vol. 33, Curran Associates, Inc., 2020, pp. 18846–18859.