A REAL MOMENT-HSOS HIERARCHY FOR COMPLEX POLYNOMIAL OPTIMIZATION WITH REAL COEFFICIENTS[*]

JIE WANG[†] AND VICTOR MAGRON[‡]

Abstract. This paper proposes a real moment-HSOS hierarchy for complex polynomial optimization problems with real coefficients. We show that this hierarchy provides the same sequence of lower bounds as the complex analogue, yet is much cheaper to solve. In addition, we prove that global optimality is achieved when the ranks of the moment matrix and certain submatrix equal two in case that a sphere constraint is present, and as a consequence, the complex polynomial optimization problem has either two real optimal solutions or a pair of conjugate optimal solutions. A simple procedure for extracting a pair of conjugate optimal solutions is given in the latter case. Various numerical examples are presented to demonstrate the efficiency of this new hierarchy, and an application to polyphase code design is also provided.

Key words. complex polynomial optimization, semidefinite relaxation, moment-HSOS hierarchy, polyphase code design, conjugate invariance

AMS subject classifications. Primary, 90C23; Secondary, 90C22,90C26

1. Introduction. A complex polynomial optimization problem (CPOP) takes the following general form:

$$\text{(CPOP)} \qquad f_{\min} := \begin{cases} \inf_{\mathbf{z} \in \mathbb{C}^n} & f(\mathbf{z}, \overline{\mathbf{z}}) \\ \text{s.t.} & g_i(\mathbf{z}, \overline{\mathbf{z}}) \geq 0, \quad i = 1, \ldots, m, \\ & h_i(\mathbf{z}, \overline{\mathbf{z}}) = 0, \quad i = 1, \ldots, l, \end{cases}$$

where $\overline{\mathbf{z}} := (\overline{z}_1, \ldots, \overline{z}_n)$ stands for the conjugate of complex variables $\mathbf{z} := (z_1, \ldots, z_n)$, and $f, g_1, \ldots, g_m, h_1, \ldots, h_l$ are polynomials in $\mathbf{z}, \overline{\mathbf{z}}$ that are invariant under conjugation (hence take real values). (CPOP) arises naturally from diverse areas, such as imaging science [10], signal processing [1, 3, 8, 16], automatic control [24], quantum mechanics [11], power systems [4]. Note that by introducing real variables for the real and imaginary parts of each complex variable, respectively, (CPOP) can be converted into a polynomial optimization problem (POP) involving only real variables.

Nowadays a popular scheme to globally solve POPs is the so-called moment-sum-of-squares (moment-SOS) hierarchy (also known as Lasserre's hierarchy) [13] which consists of a sequence of increasingly tight semidefinite relaxations. In a similar spirit, the complex moment-Hermitian-sum-of-squares (moment-HSOS) hierarchy involving positive semidefinite Hermitian matrices has also been established to handle CPOPs [12]. We remark that to solve a CPOP, one can either apply the real moment-SOS hierarchy after converting it into an equivalent real POP or directly apply the complex

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

[‡] LAAS CNRS & IMT, Toulouse, France (vmagron@laas.fr)

moment-HSOS hierarchy. It is known that the complex moment-HSOS hierarchy may produce weaker bounds than the real moment-SOS hierarchy at the same relaxation order, which is, however, still of interest because of its lower computational complexity [12, 28].

A common drawback of the real moment-SOS hierarchy and the complex moment-HSOS hierarchy is that the size of the related semidefinite relaxations grows rapidly with both the number of polynomial variables and the relaxation order, which makes them hardly applicable to large-scale problems. In order to improve scalability, it is crucial to exploit structures encoded in problem data to develop structured hierarchies. Such structures include correlative sparsity [12, 25], term sparsity [29, 30, 32, 28], symmetry [21]. Structured hierarchies have been successfully employed to tackle lots of practical applications, for instance, estimating roundoff errors of computer arithmetic [15], bounding joint spectral radius [27], computing optimal power flow [31], quantum optimal control [5], neural network verification [18], just to name a few.

In this paper, we are concerned with a particular subclass of (CPOP), where the polynomials $f, g_1, \ldots, g_m, h_1, \ldots, h_l$ all have real coefficients, or equivalently, the polynomials $f, g_1, \ldots, g_m, h_1, \ldots, h_l$ are invariant under the conjugate transformation: $\mathbf{z} \mapsto \overline{\mathbf{z}}$. For this type of (CPOP), we propose a new moment-HSOS hierarchy that involves positive semidefinite **real** matrices. We show that, in comparison to the complex moment-HSOS hierarchy, this new hierarchy (1) has **lower complexity** (see Table 1), and (2) provides the **same bound** at each relaxation order (Theorem 3.3). Furthermore, as another contribution of this paper, we prove that global optimality of the new hierarchy is attained when the ranks of the moment matrix and certain submatrix equal two in the presence of a sphere constraint. Note that here we do not require the "joint hyponormality" condition to declare global optimality (cf. [12, Proposition 4.1]). The rank-two condition also implies that the related complex polynomial optimization problem has either two real optimal solutions or a pair of conjugate optimal solutions. In the latter case, we provide a simple procedure for extracting a pair of conjugate global minimizers from the moment matrix of order one.

The efficiency of the new hierarchy is demonstrated on randomly generated CPOP instances as well as two CPOPs (respectively related to Smale's Mean Value conjecture and the Mordell inequality conjecture) from the literature. Finally, as an application, we apply the new hierarchy to the problem of polyphase code design arising from signal processing.

2. Notation and preliminaries. Let $\mathbb{N}$ be the set of nonnegative integers. For $n \in \mathbb{N} \setminus \{0\}$, let $[n] := \{1, 2, \ldots, n\}$. For $\boldsymbol{\alpha} = (\alpha_i) \in \mathbb{N}^n$, let $|\boldsymbol{\alpha}| := \sum_{i=1}^n \alpha_i$. For $r \in \mathbb{N}$, let $\mathbb{N}_r^n := \{\boldsymbol{\alpha} \in \mathbb{N}^n \mid |\boldsymbol{\alpha}| \le r\}$. We use $A \succeq 0$ to indicate that the matrix A is positive semidefinite (PSD). Let $\mathbf{i}$ be the imaginary unit, satisfying $\mathbf{i}^2 = -1$. Let $\mathbf{z} = (z_1, \ldots, z_n)$ be a tuple of complex variables and $\overline{\mathbf{z}} = (\overline{z}_1, \ldots, \overline{z}_n)$ its conjugate. 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 denote by $\mathbb{C}[\mathbf{z}, \overline{\mathbf{z}}] := \mathbb{C}[z_1, \ldots, z_n, \overline{z}_1, \ldots, \overline{z}_n]$, $\mathbb{R}[\mathbf{z}, \overline{\mathbf{z}}] := \mathbb{R}[z_1, \ldots, z_n, \overline{z}_1, \ldots, \overline{z}_n]$, the complex and real polynomial rings in $\mathbf{z}, \overline{\mathbf{z}}$, respectively. A polynomial $f \in \mathbb{C}[\mathbf{z}, \overline{\mathbf{z}}]$ (resp. $\mathbb{R}[\mathbf{z}, \overline{\mathbf{z}}]$) can be written as $f = \sum_{(\boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathbb{N}^n \times \mathbb{N}^n} f_{\boldsymbol{\beta}, \boldsymbol{\gamma}} \mathbf{z}^{\boldsymbol{\beta}} \overline{\mathbf{z}}^{\boldsymbol{\gamma}}$ with $f_{\boldsymbol{\beta}, \boldsymbol{\gamma}} \in \mathbb{C}$ (resp. $f_{\boldsymbol{\beta}, \boldsymbol{\gamma}} \in \mathbb{R}$), $\mathbf{z}^{\boldsymbol{\beta}} = z_1^{\beta_1} \cdots z_n^{\beta_n}, \overline{\mathbf{z}}^{\boldsymbol{\gamma}} = \overline{z}_1^{\gamma_1} \cdots \overline{z}_n^{\gamma_n}$. The *conjugate* of f is $\overline{f} = \sum_{(\boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathbb{N}^n \times \mathbb{N}^n} \overline{f}_{\boldsymbol{\beta}, \boldsymbol{\gamma}} \mathbf{z}^{\boldsymbol{\gamma}} \overline{\mathbf{z}}^{\boldsymbol{\beta}}$. The polynomial f is *self-conjugate* if $\overline{f} = f$. It is clear that self-conjugate polynomials take only real values. The set of self-conjugate polynomials in $\mathbb{C}[\mathbf{z}, \overline{\mathbf{z}}]$ (resp. $\mathbb{R}[\mathbf{z}, \overline{\mathbf{z}}]$)

is denoted by $\mathbb{C}[\mathbf{z}, \overline{\mathbf{z}}]^c$ (resp. $\mathbb{R}[\mathbf{z}, \overline{\mathbf{z}}]^c$). For $r \in \mathbb{N}$, let us denote

$$\mathbb{C}_r[\mathbf{z}, \overline{\mathbf{z}}] := \left\{ \sum_{|\boldsymbol{\beta}|, |\boldsymbol{\gamma}| \le r} f_{\boldsymbol{\beta}, \boldsymbol{\gamma}} \mathbf{z}^{\boldsymbol{\beta}} \overline{\mathbf{z}}^{\boldsymbol{\gamma}} \,\Big|\, f_{\boldsymbol{\beta}, \boldsymbol{\gamma}} \in \mathbb{C} \right\},$$

and

$$\mathbb{R}_r[\mathbf{z}, \overline{\mathbf{z}}] := \left\{ \sum_{|\boldsymbol{\beta}|, |\boldsymbol{\gamma}| \le r} f_{\boldsymbol{\beta}, \boldsymbol{\gamma}} \mathbf{z}^{\boldsymbol{\beta}} \overline{\mathbf{z}}^{\boldsymbol{\gamma}} \,\Big|\, f_{\boldsymbol{\beta}, \boldsymbol{\gamma}} \in \mathbb{R} \right\}.$$

Furthermore, let $\mathbb{C}_r[\mathbf{z}, \overline{\mathbf{z}}]^c := \mathbb{C}[\mathbf{z}, \overline{\mathbf{z}}]^c \cap \mathbb{C}_r[\mathbf{z}, \overline{\mathbf{z}}]$ and $\mathbb{R}_r[\mathbf{z}, \overline{\mathbf{z}}]^c := \mathbb{R}[\mathbf{z}, \overline{\mathbf{z}}]^c \cap \mathbb{R}_r[\mathbf{z}, \overline{\mathbf{z}}]$.

A polynomial $\sigma = \sum_{(\boldsymbol{\beta}, \boldsymbol{\gamma})} \sigma_{\boldsymbol{\beta}, \boldsymbol{\gamma}} \mathbf{z}^{\boldsymbol{\beta}} \overline{\mathbf{z}}^{\boldsymbol{\gamma}} \in \mathbb{C}[\mathbf{z}, \overline{\mathbf{z}}]^c$ is called an *Hermitian sum of squares* (HSOS) if there exist polynomials $f_i \in \mathbb{C}[\mathbf{z}], i \in [t]$ such that $\sigma = |f_1|^2 + \cdots + |f_t|^2$. The set of HSOS polynomials is denoted by $\Sigma^{\mathbb{C}}[\mathbf{z}, \overline{\mathbf{z}}]$. For $\sigma \in \mathbb{R}[\mathbf{z}, \overline{\mathbf{z}}]^c$, one can further require that there exist polynomials $f_i \in \mathbb{R}[\mathbf{z}], i \in [t]$ such that $\sigma = |f_1|^2 + \cdots + |f_t|^2$, and the set of such σ is denoted by $\Sigma^{\mathbb{R}}[\mathbf{z}, \overline{\mathbf{z}}]$. We also denote $\Sigma_r^{\mathbb{C}}[\mathbf{z}, \overline{\mathbf{z}}] := \Sigma^{\mathbb{C}}[\mathbf{z}, \overline{\mathbf{z}}] \cap \mathbb{C}_r[\mathbf{z}, \overline{\mathbf{z}}]$ and $\Sigma_r^{\mathbb{R}}[\mathbf{z}, \overline{\mathbf{z}}] := \Sigma^{\mathbb{R}}[\mathbf{z}, \overline{\mathbf{z}}] \cap \mathbb{R}_r[\mathbf{z}, \overline{\mathbf{z}}]$.

2.1. The complex moment-HSOS hierarchy. By invoking Borel measures, (CPOP) admits the following reformulation:

$$(2.1) \qquad \begin{cases} \inf_{\mu \in \mathcal{M}_+(K)} & \int_K f \, d\mu \\ \text{s.t.} & \int_K d\mu = 1, \end{cases}$$

where

$$(2.2) \qquad K := \{ \mathbf{z} \in \mathbb{C}^n \mid g_i(\mathbf{z}, \overline{\mathbf{z}}) \ge 0, i \in [m], h_i(\mathbf{z}, \overline{\mathbf{z}}) = 0, i \in [l] \},$$

and $\mathcal{M}_+(K)$ denotes the set of finite positive Borel measures on K. This reformulation leads to a complex moment hierarchy for (CPOP). To this end, let $\mathbf{y} = (y_{\boldsymbol{\beta}, \boldsymbol{\gamma}})_{(\boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathbb{N}^n \times \mathbb{N}^n} \subseteq \mathbb{C}$ be a pseudo-moment sequence indexed by $(\boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathbb{N}^n \times \mathbb{N}^n$ satisfying $y_{\boldsymbol{\beta}, \boldsymbol{\gamma}} = \overline{y}_{\boldsymbol{\gamma}, \boldsymbol{\beta}}$. Let $L_{\mathbf{y}} : \mathbb{C}[\mathbf{z}, \overline{\mathbf{z}}] \to \mathbb{R}$ be the linear Riesz functional associated with $\mathbf{y}$ given by

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

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

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

Let $g = \sum_{(\boldsymbol{\beta}', \boldsymbol{\gamma}')} g_{\boldsymbol{\beta}', \boldsymbol{\gamma}'} \mathbf{z}^{\boldsymbol{\beta}'} \overline{\mathbf{z}}^{\boldsymbol{\gamma}'} \in \mathbb{C}[\mathbf{z}, \overline{\mathbf{z}}]^c$. The *(complex) localizing* matrix $\mathbf{M}_r(g\mathbf{y})$ (of order r) associated with g and $\mathbf{y}$ is the matrix indexed by $\mathbb{N}_r^n$ such that

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

Both the moment and localizing matrices are Hermitian matrices by definition.

Let $d^f := \max\{|\boldsymbol{\beta}|, |\boldsymbol{\gamma}| : f_{\boldsymbol{\beta},\boldsymbol{\gamma}} \neq 0\}$, $d_i^g := \max\{|\boldsymbol{\beta}|, |\boldsymbol{\gamma}| : g_{\boldsymbol{\beta},\boldsymbol{\gamma}}^i \neq 0\}$ for $i \in [m]$, $d_i^h := \max\{|\boldsymbol{\beta}|, |\boldsymbol{\gamma}| : h_{\boldsymbol{\beta},\boldsymbol{\gamma}}^i \neq 0\}$ for $i \in [l]$, where $f = \sum_{(\boldsymbol{\beta},\boldsymbol{\gamma})} f_{\boldsymbol{\beta},\boldsymbol{\gamma}} \mathbf{z}^{\boldsymbol{\beta}} \overline{\mathbf{z}}^{\boldsymbol{\gamma}}$, $g_i = \sum_{(\boldsymbol{\beta},\boldsymbol{\gamma})} g_{\boldsymbol{\beta},\boldsymbol{\gamma}}^i \mathbf{z}^{\boldsymbol{\beta}} \overline{\mathbf{z}}^{\boldsymbol{\gamma}}$, $h_i = \sum_{(\boldsymbol{\beta},\boldsymbol{\gamma})} h_{\boldsymbol{\beta},\boldsymbol{\gamma}}^i \mathbf{z}^{\boldsymbol{\beta}} \overline{\mathbf{z}}^{\boldsymbol{\gamma}}$. Moreover, let

$$(2.3) \qquad r_K := \max\left\{2, d_1^g, \ldots, d_m^g, d_1^h, \ldots, d_l^h\right\},$$

and $r_{\min} := \max\{d^f, d_1^g, \ldots, d_m^g, d_1^h, \ldots, d_l^h\}$. The complex moment hierarchy indexed by $r \geq r_{\min}$ (called the relaxation order) for (CPOP) is then given by

$$(\text{CMom-}r) \qquad \theta_r := \begin{cases} \inf_{\mathbf{y} \subseteq \mathbb{C}} & L_{\mathbf{y}}(f) \\ \text{s.t.} & \mathbf{M}_r(\mathbf{y}) \succeq 0, \\ & \mathbf{M}_{r-d_i^g}(g_i \mathbf{y}) \succeq 0, \quad i \in [m], \\ & \mathbf{M}_{r-d_i^h}(h_i \mathbf{y}) = 0, \quad i \in [l], \\ & y_{\mathbf{0},\mathbf{0}} = 1, \end{cases}$$

which is a complex semidefinite program (SDP). The dual of (CMom-r) can be formulized as follows:
(CSOS-r)

$$\hat{\theta}_r := \begin{cases} \sup_{\gamma, \sigma_i, \tau_i} & \gamma \\ \text{s.t.} & f - \gamma = \sigma_0 + \sigma_1 g_1 + \cdots + \sigma_m g_m + \tau_1 h_1 + \cdots + \tau_l h_l, \\ & \sigma_0 \in \Sigma_r^{\mathbb{C}}[\mathbf{z}, \overline{\mathbf{z}}], \sigma_i \in \Sigma_{r-d_i^g}^{\mathbb{C}}[\mathbf{z}, \overline{\mathbf{z}}], i \in [m], \tau_i \in \mathbb{C}_{r-d_i^h}[\mathbf{z}, \overline{\mathbf{z}}]^c, i \in [l]. \end{cases}$$

To reformulate (CMom-r) or (CSOS-r) as an SDP over real numbers, we refer the reader to [26].

Remark 2.1. By the complex version of Putinar's Positivstellensatz theorem due to D'Angelo and Putinar [9], global asymptotic convergence of the complex moment-HSOS hierarchy is guaranteed when a sphere constraint is present.

3. The real moment-HSOS hierarchy. Our starting point is the following theorem stating that a self-conjugate polynomial with real coefficients is an HSOS if and only if it is also an HSOS with real coefficients.

THEOREM 3.1. *Let $f = \sum_{(\boldsymbol{\beta},\boldsymbol{\gamma})} f_{\boldsymbol{\beta},\boldsymbol{\gamma}} \mathbf{z}^{\boldsymbol{\beta}} \overline{\mathbf{z}}^{\boldsymbol{\gamma}} \in \mathbb{R}[\mathbf{z}, \overline{\mathbf{z}}]^c$. Then it holds*

$$(3.1) \qquad f \in \Sigma_d^{\mathbb{C}}[\mathbf{z}, \overline{\mathbf{z}}] \iff f \in \Sigma_d^{\mathbb{R}}[\mathbf{z}, \overline{\mathbf{z}}].$$

Proof. We shall prove only the direction "$\Rightarrow$" as the opposite direction is trivial. Let $[\mathbf{z}]_d = (\mathbf{z}^{\boldsymbol{\alpha}})_{|\boldsymbol{\alpha}| \leq d}$ be the standard complex monomial basis. Let G be the symmetric real matrix indexed by $|\boldsymbol{\alpha}| \leq d$ satisfying $G_{\boldsymbol{\beta},\boldsymbol{\gamma}} = f_{\boldsymbol{\beta},\boldsymbol{\gamma}}$ such that $f = [\mathbf{z}]_d^* G [\mathbf{z}]_d$. Since $f \in \Sigma_d^{\mathbb{C}}[\mathbf{z}, \overline{\mathbf{z}}]$, there exists $F \in \mathbb{C}^{s \times t}$ ($t := \binom{n+d}{d}$) such that $f = (F[\mathbf{z}]_d)^*(F[\mathbf{z}]_d)$. It follows $G = F^* F$. Let $\mathbf{a} \in \mathbb{R}^t$ be arbitrary. We have $\mathbf{a}^{\mathsf{T}} G \mathbf{a} = \mathbf{a}^* F^* F \mathbf{a} = (F\mathbf{a})^*(F\mathbf{a}) \geq 0$. Hence G is PSD and so $f \in \Sigma_d^{\mathbb{R}}[\mathbf{z}, \overline{\mathbf{z}}]$. $\qquad\square$

Theorem 3.1 can be further extended to Putinar type representations of positive polynomials on a semialgebraic set.

THEOREM 3.2. *Let $f, g_1, \ldots, g_m, h_1, \ldots, h_l \in \mathbb{R}[\mathbf{z}, \overline{\mathbf{z}}]^c$. Assume that f admits a representation*

$$(3.2) \qquad f = \sigma_0 + \sum_{i=1}^m \sigma_i g_i + \sum_{i=1}^l \tau_i h_i,$$

where $\sigma_i \in \Sigma_{s_i}^{\mathbb{C}}[\mathbf{z}, \overline{\mathbf{z}}], i \in \{0\} \cup [m]$ and $\tau_i \in \mathbb{C}_{t_i}[\mathbf{z}, \overline{\mathbf{z}}]^{\mathrm{c}}, i \in [l]$. Then there exist $\sigma_i' \in \Sigma_{s_i}^{\mathbb{R}}[\mathbf{z}, \overline{\mathbf{z}}], i \in \{0\} \cup [m]$ and $\tau_i' \in \mathbb{R}_{t_i}[\mathbf{z}, \overline{\mathbf{z}}]^{\mathrm{c}}, i \in [l]$ such that

$$(3.3) \qquad f = \sigma_0' + \sum_{i=1}^{m} \sigma_i' g_i + \sum_{i=1}^{l} \tau_i' h_i.$$

Proof. Since $f, g_1, \ldots, g_m, h_1, \ldots, h_l \in \mathbb{R}[\mathbf{z}, \overline{\mathbf{z}}]^{\mathrm{c}}$, we have

$$f(\mathbf{z}, \overline{\mathbf{z}}) = \frac{1}{2}\left(f(\mathbf{z}, \overline{\mathbf{z}}) + f(\overline{\mathbf{z}}, \mathbf{z})\right)$$

$$= \frac{1}{2}\left(\sigma_0(\mathbf{z}, \overline{\mathbf{z}}) + \sigma_0(\overline{\mathbf{z}}, \mathbf{z})\right) + \sum_{i=1}^{m} \frac{1}{2}\left(\sigma_i(\mathbf{z}, \overline{\mathbf{z}}) + \sigma_i(\overline{\mathbf{z}}, \mathbf{z})\right)g_i(\mathbf{z}, \overline{\mathbf{z}})$$

$$+ \sum_{i=1}^{l} \frac{1}{2}\left(\tau_i(\mathbf{z}, \overline{\mathbf{z}}) + \tau_i(\overline{\mathbf{z}}, \mathbf{z})\right)h_i(\mathbf{z}, \overline{\mathbf{z}}).$$

Let $\sigma_i' = \frac{1}{2}(\sigma_i(\mathbf{z}, \overline{\mathbf{z}}) + \sigma_i(\overline{\mathbf{z}}, \mathbf{z})) \in \Sigma_{s_i}^{\mathbb{C}}[\mathbf{z}, \overline{\mathbf{z}}]$ for $i \in \{0\} \cup [m]$ and $\tau_i' = \frac{1}{2}(\tau_i(\mathbf{z}, \overline{\mathbf{z}}) + \tau_i(\overline{\mathbf{z}}, \mathbf{z})) \in \mathbb{C}_{t_i}[\mathbf{z}, \overline{\mathbf{z}}]^{\mathrm{c}}$ for $i \in [l]$. One can easily check that $\sigma_i', \tau_i' \in \mathbb{R}[\mathbf{z}, \overline{\mathbf{z}}]^{\mathrm{c}}$ for all i. Thus by Theorem 3.1, we have $\sigma_i' \in \Sigma_{s_i}^{\mathbb{R}}[\mathbf{z}, \overline{\mathbf{z}}], i \in \{0\} \cup [m]$ and $\tau_i' \in \mathbb{R}_{t_i}[\mathbf{z}, \overline{\mathbf{z}}]^{\mathrm{c}}, i \in [l]$, which completes the proof. $\qquad\square$

Theorem 3.2 allows us to consider the following alternative to the HSOS relaxation (CSOS-r) of (CPOP):
(RSOS-r)

$$\hat{\eta}_r := \begin{cases} \sup\limits_{\gamma, \sigma_i, \tau_i} & \gamma \\ \text{s.t.} & f - \gamma = \sigma_0 + \sigma_1 g_1 + \cdots + \sigma_m g_m + \tau_1 h_1 + \cdots + \tau_l h_l, \\ & \sigma_0 \in \Sigma_r^{\mathbb{R}}[\mathbf{z}, \overline{\mathbf{z}}], \sigma_i \in \Sigma_{r-d_i^g}^{\mathbb{R}}[\mathbf{z}, \overline{\mathbf{z}}], i \in [m], \tau_i \in \mathbb{R}_{r-d_i^h}[\mathbf{z}, \overline{\mathbf{z}}]^{\mathrm{c}}, i \in [l]. \end{cases}$$

On the other hand, the assumption that the coefficients of $f, g_1, \ldots, g_m, h_1, \ldots, h_l$ are all real implies that (CPOP) is invariant under the conjugate transformation: $\mathbf{z} \mapsto \overline{\mathbf{z}}$. Therefore, in (2.1) it is more natural to seek a Borel measure that is invariant under the conjugate transformation. A conjugate-invariant Borel measure has real moments. Hence we can consider the following moment relaxation for (CPOP) defined over real numbers:

$$(\text{RMom-}r) \qquad \eta_r := \begin{cases} \inf\limits_{\mathbf{y} \subseteq \mathbb{R}} & L_{\mathbf{y}}(f) \\ \text{s.t.} & \mathbf{M}_r(\mathbf{y}) \succeq 0, \\ & \mathbf{M}_{r-d_i^g}(g_i\mathbf{y}) \succeq 0, \quad i \in [m], \\ & \mathbf{M}_{r-d_i^h}(h_i\mathbf{y}) = 0, \quad i \in [l], \\ & y_{\mathbf{0},\mathbf{0}} = 1. \end{cases}$$

(RMom-r) and (RSOS-r) form a pair of primal-dual SDPs.

THEOREM 3.3. *For any $r \geq r_{\min}$, one has $\theta_r = \eta_r$ and $\hat{\theta}_r = \hat{\eta}_r$.*

Proof. Suppose that $\mathbf{y}$ is a feasible solution of (RMom-r). Clearly $\mathbf{y}$ is also a feasible solution of (CMom-r). So $\theta_r \leq \eta_r$. Conversely, suppose that $\mathbf{y}$ is a feasible solution of (CMom-r). Then $\mathbf{y}' = \frac{\mathbf{y} + \overline{\mathbf{y}}}{2}$ is a feasible solution of (RMom-r) and $L_{\mathbf{y}'}(f) = \frac{1}{2}(L_{\mathbf{y}}(f) + L_{\overline{\mathbf{y}}}(f)) = L_{\mathbf{y}}(f)$. It follows $\theta_r \geq \eta_r$ and we then conclude $\theta_r = \eta_r$. The equality $\hat{\theta}_r = \hat{\eta}_r$ follows from Theorem 3.2. $\qquad\square$

In Table 1, we list the sizes of SDPs involved in the real HSOS hierarchy (R-HSOS), the complex HSOS hierarchy (C-HSOS)[1], and the real SOS hierarchy (R-SOS) (i.e., the usual SOS hierarchy applied to the equivalent real POP) for (CPOP), respectively.

TABLE 1
Complexity comparison of different hierarchies for (CPOP). *n: number of complex variables, r: relaxation order, n_{sdp}: maximal size of PSD blocks, m_{sdp}: number of affine constraints.*

	R-HSOS	C-HSOS	R-SOS
n_{sdp}	$\binom{n+r}{r}$	$2\binom{n+r}{r}$	$\binom{2n+r}{r}$
m_{sdp}	$\frac{1}{2}\left(\binom{n+r}{r}+1\right)\binom{n+r}{r}$	$\binom{n+r}{r}^2$	$\binom{2n+2r}{2r}$

Remark 3.4. If (CPOP) additionally exhibits correlative and/or term sparsity, then one can construct a sparsity-adapted real moment-HSOS hierarchy in a similar spirit with [28].

4. Detecting global optimality and extracting solutions. In general, detecting global optimality and extracting solutions for (RMom-r) can be carried out in the same way as for (CMom-r). For details, we refer the reader to Proposition 4.1 of [12] and the discussions therein. However, we observe that the 'joint hyponormality' condition in [12, Proposition 4.1] could be actually replaced with the following simpler 'hyponormality' condition.

COROLLARY 4.1. *Consider* (CPOP) *with complex (not necessarily real) coefficients. Suppose that* (CPOP) *involves a sphere constraint $|z_1|^2 + \cdots + |z_n|^2 = R$ for some $R > 0$. Let $\mathbf{y}$ be an optimal solution of* (CMom-r).
- *If there is an integer t such that $r_{\min} \le t \le r$ and $\operatorname{rank} \mathbf{M}_t(\mathbf{y}) = 1$, then $\theta_r = f_{\min}$ and there is at least one global minimizer.*
- *If there exists an integer t such that $\max\{r_K, r_{\min}\} \le t \le r$ and the following two conditions hold*
  1. *(Flatness)* $\operatorname{rank} \mathbf{M}_t(\mathbf{y}) = \operatorname{rank} \mathbf{M}_{t-d_K}(\mathbf{y})$ $(=: s)$;
  2. *(Hyponormality) For all $i \in [n]$,*
  $$\begin{bmatrix} \mathbf{M}_{t-d_K}(\mathbf{y}) & \mathbf{M}_{t-d_K}(\overline{z_i}\mathbf{y}) \\ \mathbf{M}_{t-d_K}(z_i\mathbf{y}) & \mathbf{M}_{t-d_K}(|z_i|^2\mathbf{y}) \end{bmatrix} \succeq 0.$$

Then $\theta_r = f_{\min}$ and there are at least s global minimizers.

Proof. The proof is almost the same as the one of [12, Proposition 4.1], which relies on [12, Theorem 5.1]. One just needs to note that the condition *(Hyponormality)* implies that the so-called *shift* operators are all normal, which together with pairwise commutativity suffices to guarantee that these shift operators (and their adjoint operators) are simultaneously diagonalizable. □

Next, we concentrate on a particular case when $\operatorname{rank} \mathbf{M}_r(\mathbf{y}) = 2$, which frequently occurs in practice. We will see that the rank-two condition allows us to remove the hyponormality assumption in Corollary 4.1. To begin with, let us prove the following proposition which is crucial for the proof of our main theorem and may also have independent interests [19].

---

[1]Here, the size of a complex SDP is measured after converting it to an equivalent real SDP; see [26].

PROPOSITION 4.2. *Suppose that the matrices $A_1, \ldots, A_n \in \mathbb{R}^{2 \times 2}$ commute pairwise and the equality*

$$(4.1) \qquad\qquad A_1^{\mathsf{T}} A_1 + \cdots + A_n^{\mathsf{T}} A_n = R I_2$$

holds for some $R > 0$. Here, I_2 denotes the 2×2 identity matrix. Then one of the following two statements holds true:
(1) $A_i = \begin{bmatrix} a_i & b_i \\ b_i & d_i \end{bmatrix}$ for $i = 1, \ldots, n$;
(2) $A_i = \begin{bmatrix} a_i & -b_i \\ b_i & a_i \end{bmatrix}$ for $i = 1, \ldots, n$.
As a result, $A_1, \ldots, A_n$ are simultaneously diagonalizable.

Proof. Let $A_i = \begin{bmatrix} a_i & c_i \\ b_i & d_i \end{bmatrix}$ for $i = 1, \ldots, n$. Since $A_1, \ldots, A_n$ commute pairwise, we have

$$(4.2a) \qquad\qquad a_i b_j + b_i d_j = a_j b_i + b_j d_i,$$
$$(4.2b) \qquad\qquad a_i c_j + c_i d_j = a_j c_i + c_j d_i,$$
$$(4.2c) \qquad\qquad b_i c_j = b_j c_i,$$

for all $1 \leq i < j \leq n$. Moreover, the equality (4.1) yields

$$(4.3a) \qquad\qquad \sum_{i=1}^{n} \left(a_i^2 + b_i^2 \right) = R,$$

$$(4.3b) \qquad\qquad \sum_{i=1}^{n} \left(c_i^2 + d_i^2 \right) = R,$$

$$(4.3c) \qquad\qquad \sum_{i=1}^{n} \left(a_i c_i + b_i d_i \right) = 0.$$

First we consider the case that not all c_i are zeros (i.e., $\sum_{i=1}^{n} c_i^2 \neq 0$). By (4.2c), we may assume $b_i = t c_i$ for some t and for $i = 1, \ldots, n$. By (4.2b), we may assume $a_i = c_i s + d_i$ for some s and for $i = 1, \ldots, n$. Then (4.3a) and (4.3c) become

$$(4.4a) \qquad\qquad \sum_{i=1}^{n} \left(c_i^2 s^2 + 2 c_i d_i s + d_i^2 + t^2 c_i^2 \right) = R,$$

$$(4.4b) \qquad\qquad \sum_{i=1}^{n} \left(c_i^2 s + c_i d_i + t c_i d_i \right) = 0.$$

(4.4a)−(4.4b)×s−(4.3b) yields

$$(4.5) \qquad\qquad (1 - t) \left(s \sum_{i=1}^{n} c_i d_i - (1 + t) \sum_{i=1}^{n} c_i^2 \right) = 0.$$

It follows from (4.5) that one of the following cases holds true: (i) $t = 1$, which implies that A_i are all of form in (1); (ii) $t = -1$ and then by (4.4b) $s = 0$, which implies that A_i are all of form in (2); (iii) $t \neq \pm 1$ and

$$(4.6) \qquad\qquad s = \frac{(1 + t) \sum_{i=1}^{n} c_i^2}{\sum_{i=1}^{n} c_i d_i}.$$

For case (iii), substituting into (4.4a) with (4.6), we obtain

$$\sum_{i=1}^{n} \left(c_i^2 s^2 + 2c_i d_i s + d_i^2 + t^2 c_i^2 \right) - R$$

$$= \frac{(1+t)^2 \left(\sum_{i=1}^{n} c_i^2 \right)^3}{\left(\sum_{i=1}^{n} c_i d_i \right)^2} + 2(1+t) \sum_{i=1}^{n} c_i^2 + \sum_{i=1}^{n} d_i^2 + t^2 \sum_{i=1}^{n} c_i^2 - R$$

$$= \frac{(1+t)^2 \left(\sum_{i=1}^{n} c_i^2 \right)^3}{\left(\sum_{i=1}^{n} c_i d_i \right)^2} + (1+t)^2 \left(\sum_{i=1}^{n} c_i^2 \right)$$

$$= (1+t)^2 \left(\sum_{i=1}^{n} c_i^2 \right) \left(\frac{\left(\sum_{i=1}^{n} c_i^2 \right)^2}{\left(\sum_{i=1}^{n} c_i d_i \right)^2} + 1 \right) = 0,$$

a contradictory.

Now suppose $c_i = 0$ for all i. Then we must have $b_i = 0$ for all i and so A_i fits into (1). Indeed, if not all b_i are zeros, then by (4.2a), we may assume $a_i = b_i s + d_i$ for some s and for $i = 1, \ldots, n$. (4.3b) and (4.3c) become $\sum_{i=1}^{n} d_i^2 = R$ and $\sum_{i=1}^{n} b_i d_i = 0$. Thus (4.3a) yields

$$\sum_{i=1}^{n} \left(b_i^2 s^2 + 2b_i d_i s + d_i^2 + b_i^2 \right) - R = (1+s^2) \sum_{i=1}^{n} b_i^2 = 0,$$

which implies $\sum_{i=1}^{n} b_i^2 = 0$, a contradictory.   □

Now we are ready to prove the promised theorem.

THEOREM 4.3. *Suppose that* (CPOP) *involves a sphere constraint* $|z_1|^2 + \cdots + |z_n|^2 = R$ *for some* $R > 0$, *and let* $\mathbf{y}$ *be an optimal solution of* (RMom-r). *If* $\operatorname{rank} \mathbf{M}_r(\mathbf{y}) = \operatorname{rank} \mathbf{M}_{r-d_K}(\mathbf{y}) = 2$, *then global optimality is attained, i.e.,* $\theta_r = f_{\min}$. *Moreover, one can extract either two real solutions or a pair of complex conjugate solutions for* (CPOP).

Proof. The proof is an adaptation of the one of [12, Theorem 5.1].

The real PSD matrix $\mathbf{M}_r(\mathbf{y})$ of rank 2 can be factorized in the Grammian form such that $y_{\boldsymbol{\beta}, \boldsymbol{\gamma}} = a_{\boldsymbol{\beta}}^{\mathsf{T}} a_{\boldsymbol{\gamma}}$ for all $|\boldsymbol{\beta}|, |\boldsymbol{\gamma}| \leq r$, where $a_{\boldsymbol{\beta}} \in \mathbb{R}^2$. Since $\operatorname{rank} \mathbf{M}_{r-1}(\mathbf{y}) = 2$, we have $\mathbb{R}^2 = \operatorname{span}(a_{\boldsymbol{\alpha}})_{|\boldsymbol{\alpha}| \leq r-1}$. Let us consider the shift operators[2] $T_1, \ldots, T_n : \mathbb{R}^2 \to \mathbb{R}^2$, given by

$$(4.7) \qquad T_i : \sum_{|\boldsymbol{\alpha}| \leq r-1} w_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}} \mapsto \sum_{|\boldsymbol{\alpha}| \leq r-1} w_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}+\mathbf{e}_i}, \quad i \in [n],$$

where $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is the standard vector basis of $\mathbb{R}^n$. To ensure that the shifts are well-defined, we have to check that each element of $\mathbb{R}^2$ has a unique image under T_i, that is, if $\sum_{|\boldsymbol{\alpha}| \leq r-1} u_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}} = \sum_{|\boldsymbol{\alpha}| \leq r-1} v_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}}$, then $\sum_{|\boldsymbol{\alpha}| \leq r-1} u_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}+\mathbf{e}_i} = \sum_{|\boldsymbol{\alpha}| \leq r-1} v_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}+\mathbf{e}_i}$ for all $i \in [n]$. Actually, the sphere constraint yields

$$\sum_{i=1}^{n} y_{\boldsymbol{\beta}+\mathbf{e}_i, \boldsymbol{\gamma}+\mathbf{e}_i} = R y_{\boldsymbol{\beta}, \boldsymbol{\gamma}}$$

---

[2] For convenience, we identify the operators with their matrices in the standard basis of $\mathbb{R}^2$.

for all $|\boldsymbol{\beta}|, |\boldsymbol{\gamma}| \leq r - 1$. Thus,

$$\sum_{i=1}^{n} \left\| \sum_{|\boldsymbol{\alpha}| \leq r-1} w_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}+\mathbf{e}_i} \right\|^2 = \sum_{|\boldsymbol{\beta}|,|\boldsymbol{\gamma}| \leq r-1} w_{\boldsymbol{\beta}} w_{\boldsymbol{\gamma}} a_{\boldsymbol{\beta}+\mathbf{e}_i} a_{\boldsymbol{\gamma}+\mathbf{e}_i}^{\mathsf{T}}$$

$$= \sum_{i=1}^{n} \sum_{|\boldsymbol{\beta}|,|\boldsymbol{\gamma}| \leq r-1} w_{\boldsymbol{\beta}} w_{\boldsymbol{\gamma}} \sum_{i=1}^{n} y_{\boldsymbol{\beta}+\mathbf{e}_i, \boldsymbol{\gamma}+\mathbf{e}_i}$$

$$= R \sum_{|\boldsymbol{\beta}|,|\boldsymbol{\gamma}| \leq r-1} w_{\boldsymbol{\beta}} w_{\boldsymbol{\gamma}} y_{\boldsymbol{\beta},\boldsymbol{\gamma}}$$

$$= R \left\| \sum_{|\boldsymbol{\alpha}| \leq r-1} w_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}} \right\|^2 .$$

It follows that if $\sum_{|\boldsymbol{\alpha}| \leq r-1} w_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}} = 0$, then $\sum_{|\boldsymbol{\alpha}| \leq r-1} w_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}+\mathbf{e}_i} = 0$ for all $i \in [n]$, from which we immediately obtain the desired result.

We now prove that $T_1, \ldots, T_n$ commute pairwise. First note that for any $\boldsymbol{\alpha} \in \mathbb{N}^n_{r-2}$, we have $T_i T_j (a_{\boldsymbol{\alpha}}) = T_i (a_{\boldsymbol{\alpha}+\mathbf{e}_j}) = a_{\boldsymbol{\alpha}+\mathbf{e}_i+\mathbf{e}_j} = a_{\boldsymbol{\alpha}+\mathbf{e}_j+\mathbf{e}_i} = T_j T_i (a_{\boldsymbol{\alpha}})$. Therefore, for any $u \in \mathbb{R}^2$ with $u = \sum_{|\boldsymbol{\alpha}| \leq r-2} u_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}}$ (this is always possible as rank $\mathbf{M}_{r-2}(\mathbf{y}) = 2$), we have

$$T_i T_j (u) = T_i T_j \left(\sum_{|\boldsymbol{\alpha}| \leq r-1} u_{\boldsymbol{\alpha}} a_{\boldsymbol{\alpha}} \right) = \sum_{|\boldsymbol{\alpha}| \leq r-1} u_{\boldsymbol{\alpha}} T_i T_j (a_{\boldsymbol{\alpha}})$$

$$= \sum_{|\boldsymbol{\alpha}| \leq r-1} u_{\boldsymbol{\alpha}} T_j T_i (a_{\boldsymbol{\alpha}}) = T_j T_i (u),$$

which proves the pairwise commutativity of $T_1, \ldots, T_n$. Moreover, we have

$$\sum_{i=1}^{n} u^{\mathsf{T}} T_i^{\mathsf{T}} T_i u = \sum_{i=1}^{n} \sum_{|\boldsymbol{\beta}|,|\boldsymbol{\gamma}| \leq r-2} u_{\boldsymbol{\beta}} u_{\boldsymbol{\gamma}} a_{\boldsymbol{\beta}+\mathbf{e}_i}^{\mathsf{T}} a_{\boldsymbol{\gamma}+\mathbf{e}_i}$$

$$= \sum_{|\boldsymbol{\beta}|,|\boldsymbol{\gamma}| \leq r-2} u_{\boldsymbol{\beta}} u_{\boldsymbol{\gamma}} \sum_{i=1}^{n} y_{\boldsymbol{\beta}+\mathbf{e}_i, \boldsymbol{\gamma}+\mathbf{e}_i}$$

$$= R \sum_{|\boldsymbol{\beta}|,|\boldsymbol{\gamma}| \leq r-2} u_{\boldsymbol{\beta}} u_{\boldsymbol{\gamma}} y_{\boldsymbol{\beta},\boldsymbol{\gamma}}$$

$$= R \sum_{|\boldsymbol{\beta}|,|\boldsymbol{\gamma}| \leq r-2} u_{\boldsymbol{\beta}} u_{\boldsymbol{\gamma}} a_{\boldsymbol{\beta}}^{\mathsf{T}} a_{\boldsymbol{\gamma}} = R u^{\mathsf{T}} u,$$

which implies

$$T_1^{\mathsf{T}} T_1 + \cdots + T_n^{\mathsf{T}} T_n = R I_2.$$

Then by Proposition 4.2, $T_1, \ldots, T_n$ are simultaneously diagonalizable. That is, there exists a unitary matrix P such that $T_i = P D_i P^*$, $i = 1, \ldots, n$, where $D_i = \mathrm{diag}\,(d_{i1}, d_{i2})$ is a diagonal matrix. For convenience, we let $T^{\boldsymbol{\alpha}} := T_1^{\alpha_1} \cdots T_n^{\alpha_n}$. Then

for $|\boldsymbol{\beta}|, |\boldsymbol{\gamma}| \le r$, we have

$$y_{\boldsymbol{\beta},\boldsymbol{\gamma}} = a_{\boldsymbol{\beta}}^{\mathsf{T}} a_{\boldsymbol{\gamma}} = \left(T^{\boldsymbol{\beta}} a_{\mathbf{0}}\right)^{\mathsf{T}} \left(T^{\boldsymbol{\gamma}} a_{\mathbf{0}}\right) = a_{\mathbf{0}}^{\mathsf{T}} \left(T^{\boldsymbol{\beta}}\right)^{\mathsf{T}} T^{\boldsymbol{\gamma}} a_{\mathbf{0}}$$

$$= a_{\mathbf{0}}^{\mathsf{T}} \left(PD^{\boldsymbol{\beta}} P^*\right)^* \left(PD^{\boldsymbol{\gamma}} P^*\right) a_{\mathbf{0}} = a_{\mathbf{0}}^{\mathsf{T}} \left(\sum_{j=1}^{2} p_j \overline{d}_j^{\boldsymbol{\beta}} d_j^{\boldsymbol{\gamma}} p_j^*\right) a_{\mathbf{0}}$$

$$= \sum_{j=1}^{2} a_{\mathbf{0}}^{\mathsf{T}} p_j p_j^* a_{\mathbf{0}} \overline{d}_j^{\boldsymbol{\beta}} d_j^{\boldsymbol{\gamma}} = \sum_{j=1}^{2} |a_{\mathbf{0}}^{\mathsf{T}} p_j|^2 \overline{d}_j^{\boldsymbol{\beta}} d_j^{\boldsymbol{\gamma}},$$

where p_1, p_2 denote the columns of P and $d_j := (d_{1j}, \ldots, d_{nj})$. As a result, we obtain that the measure $\mu = |a_{\mathbf{0}}^{\mathsf{T}} p_1|^2 \delta_{d_1} + |a_{\mathbf{0}}^{\mathsf{T}} p_2|^2 \delta_{d_2}$ satisfies $y_{\boldsymbol{\beta},\boldsymbol{\gamma}} = \int_{\mathbb{C}^n} \mathbf{z}^{\boldsymbol{\beta}} \overline{\mathbf{z}}^{\boldsymbol{\gamma}} \, \mathrm{d}\mu$ for all $|\boldsymbol{\beta}|, |\boldsymbol{\gamma}| \le r$.

Finally, we check whether the measure μ is supported on K. Indeed, we have

$$g_i(d_j, \overline{d}_j) = p_j^* p_j \sum_{\boldsymbol{\beta},\boldsymbol{\gamma}} g_{\boldsymbol{\beta},\boldsymbol{\gamma}}^i d_j^{\boldsymbol{\beta}} \overline{d}_j^{\boldsymbol{\gamma}} \qquad (\text{since } p_j^* p_j = 1)$$

$$= \sum_{\boldsymbol{\beta},\boldsymbol{\gamma}} g_{\boldsymbol{\beta},\boldsymbol{\gamma}}^i \left(d_j^{\boldsymbol{\gamma}} p_j\right)^* \left(d_j^{\boldsymbol{\beta}} p_j\right)$$

$$= \sum_{\boldsymbol{\beta},\boldsymbol{\gamma}} g_{\boldsymbol{\beta},\boldsymbol{\gamma}}^i \left(T^{\boldsymbol{\gamma}} p_j\right)^* \left(T^{\boldsymbol{\beta}} p_j\right) \qquad (\text{let } p_j = \sum_{|\boldsymbol{\alpha}| \le r - d_K} p_{\boldsymbol{\alpha}}^j a_{\boldsymbol{\alpha}})$$

$$= \sum_{|\boldsymbol{\beta}'|, |\boldsymbol{\gamma}'| \le r - d_K} p_{\boldsymbol{\beta}'}^j \overline{p}_{\boldsymbol{\gamma}'}^j \left(\sum_{\boldsymbol{\beta},\boldsymbol{\gamma}} g_{\boldsymbol{\beta},\boldsymbol{\gamma}}^i \left(T^{\boldsymbol{\gamma}} a_{\boldsymbol{\gamma}'}\right)^{\mathsf{T}} \left(T^{\boldsymbol{\beta}} a_{\boldsymbol{\beta}'}\right)\right)$$

$$= \sum_{|\boldsymbol{\beta}'|, |\boldsymbol{\gamma}'| \le r - d_K} p_{\boldsymbol{\beta}'}^j \overline{p}_{\boldsymbol{\gamma}'}^j \left(\sum_{\boldsymbol{\beta},\boldsymbol{\gamma}} g_{\boldsymbol{\beta},\boldsymbol{\gamma}}^i a_{\boldsymbol{\gamma}'+\boldsymbol{\gamma}}^{\mathsf{T}} a_{\boldsymbol{\beta}'+\boldsymbol{\beta}}\right)$$

$$= \sum_{|\boldsymbol{\beta}'|, |\boldsymbol{\gamma}'| \le r - d_K} p_{\boldsymbol{\beta}'}^j \overline{p}_{\boldsymbol{\gamma}'}^j \left(\sum_{\boldsymbol{\beta},\boldsymbol{\gamma}} g_{\boldsymbol{\beta},\boldsymbol{\gamma}}^i y_{\boldsymbol{\beta}'+\boldsymbol{\beta}, \boldsymbol{\gamma}'+\boldsymbol{\gamma}}\right) \ge 0.$$

The last inequality is because $\mathbf{M}_{r-d_i^g}(g_i \mathbf{y}) \succeq 0$ and $r - d_K \le r - d_i^g$. Similarly, we can show $h_i(d_j, \overline{d}_j) = 0$ by using $\mathbf{M}_{r-d_i^h}(h_i \mathbf{y}) = 0$.

By Proposition 4.2, either all T_i are symmetric matrices or all T_i are matrices of similarity transformations. In the former case, two real solutions for (CPOP) can be extracted; in the latter case, a pair of complex conjugate solutions for (CPOP) can be extracted. $\qquad \square$

The sphere constraint in Theorem 4.3 can be replaced by the unit-norm constraint on each variable as the latter together with $\operatorname{rank} \mathbf{M}_r(\mathbf{y}) = \operatorname{rank} \mathbf{M}_{r-2}(\mathbf{y})$ can also ensure that $T_1, \ldots, T_n$ are simultaneously diagonalizable; see [12, Appendix B].

THEOREM 4.4. *Suppose that* (CPOP) *involves the unit-norm constraint* $|z_i|^2 = 1$ *for all* $i \in [n]$, *and let* $\mathbf{y}$ *be an optimal solution of* (RMom-r). *If* $\operatorname{rank} \mathbf{M}_r(\mathbf{y}) = \operatorname{rank} \mathbf{M}_{r-d_K}(\mathbf{y}) = 2$, *then global optimality is attained, i.e.,* $\theta_r = f_{\min}$. *Moreover, one can extract either two real solutions or a pair of complex conjugate solutions for* (CPOP).

Remark 4.5. The condition $\operatorname{rank} \mathbf{M}_r(\mathbf{y}) = \operatorname{rank} \mathbf{M}_{r-d_K}(\mathbf{y}) = 2$ in Theorems 4.3 and 4.4 can be weakened as: there exists an integer t such that $\max\{r_K, r_{\min}\} \le t \le r$ such that $\operatorname{rank} \mathbf{M}_t(\mathbf{y}) = \operatorname{rank} \mathbf{M}_{t-d_K}(\mathbf{y}) = 2$.

Theorems 4.3 and 4.4 guarantee that (CPOP) has either two real solutions or a pair of complex conjugate solutions. For the latter case, we provide a simple procedure for extracting a pair of conjugate optimal solutions (denoted by $\mathbf{z}, \overline{\mathbf{z}}$) from the first-order moment matrix $\mathbf{M}_1(\mathbf{y})$. First of all, note that as all T_i are matrices of similarity transformations, we can take the unitary matrix $P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{\mathbf{i}}{\sqrt{2}} & -\frac{\mathbf{i}}{\sqrt{2}} \end{bmatrix}$ such that $T_i = PD_iP^*$, $i = 1, \ldots, n$. Let $a_{\mathbf{0}} = [a_{\mathbf{0}}^1, a_{\mathbf{0}}^2]^\mathsf{T}$ and we have $a_{\mathbf{0}}^\mathsf{T} a_{\mathbf{0}} = y_{0,0} = 1$. Then $|a_{\mathbf{0}}^\mathsf{T} p_1|^2 = \left| \frac{1}{\sqrt{2}} a_{\mathbf{0}}^1 + \frac{\mathbf{i}}{\sqrt{2}} a_{\mathbf{0}}^2 \right|^2 = \frac{1}{2} a_{\mathbf{0}}^\mathsf{T} a_{\mathbf{0}} = \frac{1}{2}$. Similarly, $|a_{\mathbf{0}}^\mathsf{T} p_2|^2 = \frac{1}{2}$. Therefore, $\mu = \frac{1}{2} \delta_{\mathbf{z}} + \frac{1}{2} \delta_{\overline{\mathbf{z}}}$. Let $[\mathbf{z}]_1 = [1, z_1, \ldots, z_n]^\mathsf{T}$. We then have $\mathbf{M}_1(\mathbf{y}) = \frac{1}{2}[\mathbf{z}]_1[\mathbf{z}]_1^* + \frac{1}{2}[\overline{\mathbf{z}}]_1[\overline{\mathbf{z}}]_1^* = \mathscr{R}([\mathbf{z}]_1[\mathbf{z}]_1^*)$ ($\mathscr{R}$ means taking the real part). Let $\mathbf{z} = (\rho_i(\cos(\phi_i) + \sin(\phi_i)))_{i \in [n]}$. Because

$$[\mathbf{z}]_1[\mathbf{z}]_1^* = \begin{bmatrix} 1 & \overline{z}_1 & \cdots & \overline{z}_n \\ z_1 & |z_1|^2 & \cdots & z_1 \overline{z}_n \\ \vdots & \vdots & \ddots & \vdots \\ z_n & z_n \overline{z}_1 & \cdots & |z_n|^2 \end{bmatrix},$$

we obtain

$$\mathbf{M}_1(\mathbf{y}) = \begin{bmatrix} 1 & \rho_1 \cos(\phi_1) & \cdots & \rho_n \cos(\phi_n) \\ \rho_1 \cos(\phi_1) & \rho_1^2 & \cdots & \rho_1 \rho_n \cos(\phi_1 - \phi_n) \\ \vdots & \vdots & \ddots & \vdots \\ \rho_n \cos(\phi_n) & \rho_1 \rho_n \cos(\phi_1 - \phi_n) & \cdots & \rho_n^2 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 \\ \rho_1 \cos(\phi_1) & \rho_1 \sin(\phi_1) \\ \vdots & \vdots \\ \rho_n \cos(\phi_n) & \rho_n \sin(\phi_n) \end{bmatrix} \cdot \begin{bmatrix} 1 & \rho_1 \cos(\phi_1) & \cdots & \rho_n \cos(\phi_n) \\ 0 & \rho_1 \sin(\phi_1) & \cdots & \rho_n \sin(\phi_n) \end{bmatrix}.$$

Thus, we can retrieve the pair of conjugate optimal solutions from the above 'Cholesky decomposition' of $\mathbf{M}_1(\mathbf{y})$.

EXAMPLE 4.6. *Let*

$$f(\mathbf{z}, \overline{\mathbf{z}}) = 0.5z_1\overline{z}_2 + 0.5z_1\overline{z}_3 + 0.5z_2\overline{z}_1 + 0.25|z_2|^2 + 0.25z_2\overline{z}_3 + 0.5z_3\overline{z}_1$$
$$+ 0.25z_3\overline{z}_2 + z_1 + z_2 + z_3 + \overline{z}_1 + \overline{z}_2 + \overline{z}_3,$$

and consider the CPOP:

(4.8) $$\inf_{\mathbf{z} \in \mathbb{C}^3} \left\{ f(\mathbf{z}, \overline{\mathbf{z}}) \text{ s.t. } |z_i|^2 = 1, i = 1, 2, 3 \right\}.$$

The first-order moment-HSOS relaxation yields the value -3.75 *and*

$$\mathbf{M}_1(\mathbf{y}) = \begin{bmatrix} 1 & -0.250013 & -0.875003 & -0.875003 \\ -0.250013 & 1 & -0.249981 & -0.249981 \\ -0.875003 & -0.249981 & 1 & 1 \\ -0.875003 & -0.249981 & 1 & 1 \end{bmatrix}.$$

We have

$$\mathbf{M}_1(\mathbf{y}) = \begin{bmatrix} 1 & 0 \\ -0.250013 & 0.968242 \\ -0.875003 & -0.484117 \\ -0.875003 & -0.484117 \end{bmatrix} \cdot \begin{bmatrix} 1 & -0.250013 & -0.875003 & -0.875003 \\ 0 & 0.968242 & -0.484117 & -0.484117 \end{bmatrix}.$$

Let

$$\mathbf{z} = (-0.250013 + 0.968242\mathbf{i}, -0.875003 - 0.484117\mathbf{i}, -0.875003 - 0.484117\mathbf{i}).$$

One can check that $\mathbf{z}$ is feasible to (4.8) *and $f(\mathbf{z}, \overline{\mathbf{z}}) = -3.75$. Therefore, $\mathbf{z}$ and $\overline{\mathbf{z}}$ are a pair of conjugate optimal solutions to* (4.8)*. Note that here we successfully extract two solutions even without the validity of* $\operatorname{rank} \mathbf{M}_r(\mathbf{y}) = \operatorname{rank} \mathbf{M}_{r-d_K}(\mathbf{y})$.

5. Numerical experiments. The real moment-HSOS hierarchy has been implemented in the Julia package `TSSOS`[3]. In this section, we compare the performance of the real moment-HSOS hierarchy (R-HSOS), the complex moment-HSOS hierarchy (C-HSOS), and the real moment-SOS hierarchy (R-SOS) on CPOPs with real coefficients using `TSSOS` where `Mosek` 10.0 [2] is employed as an SDP solver with default settings. In presenting the results, 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.

• **Minimizing a random quadratic polynomial over the unit sphere.** Our first example is to minimize a random quadratic polynomial over the unit sphere:

$$(5.1) \qquad \begin{cases} \inf\limits_{\mathbf{z} \in \mathbb{C}^n} & [\mathbf{z}]_1^* Q [\mathbf{z}]_1 \\ \text{s.t.} & |z_1|^2 + \cdots + |z_n|^2 = 1, \end{cases}$$

where $[\mathbf{z}]_1 = (\mathbf{z}^{\boldsymbol{\alpha}})_{|\boldsymbol{\alpha}| \leq 1}$ and Q is a random symmetric matrix whose entries are selected with respect to the uniform probability distribution on $[-1, 1]$.

We approach (5.1) for $n = 50, 100, \ldots, 300$ with the first-order relaxation of the three hierarchies. To measure the quality of obtained lower bounds, we also compute an upper bound (indicated by 'ub' in the table) for (5.1) using the nonlinear programming solver `Ipopt`. Related results are shown in Table 2, from which we see that R-HSOS is the fastest and is more scalable with the size of the problem. Interestingly, we can also see that the lower bounds produced by the SDP relaxations coincide with the corresponding upper bounds, and thus are actually globally optimal.

• **Minimizing a random quartic polynomial over the unit sphere.** Our second example is to minimize a quartic polynomial over the unit sphere:

$$(5.2) \qquad \begin{cases} \inf\limits_{\mathbf{z} \in \mathbb{C}^n} & [\mathbf{z}]_2^* Q [\mathbf{z}]_2 \\ \text{s.t.} & |z_1|^2 + \cdots + |z_n|^2 = 1, \end{cases}$$

where $[\mathbf{z}]_2 = (\mathbf{z}^{\boldsymbol{\alpha}})_{|\boldsymbol{\alpha}| \leq 2}$ and Q is a random symmetric matrix whose entries are selected with respect to the uniform probability distribution on $[-1, 1]$.

We approach (5.2) for $n = 8, 10, \ldots, 22$ with the second-order relaxations of the three hierarchies. To measure the quality of obtained lower bounds, we also compute

---

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

TABLE 2
Minimizing a random quadratic polynomial over the unit sphere ($r = 1$).

n	ub	R-HSOS		C-HSOS		R-SOS	
		opt	time	opt	time	opt	time
50	-4.2477	-4.2477	0.12	-4.2477	0.79	-4.2477	2.57
100	-6.3120	-6.3120	2.90	-6.3120	17.3	-6.3120	108
150	-7.0808	-7.0808	21.2	-7.0808	134	-7.0808	1216
200	-8.5434	-8.5434	104	-8.5434	731	-	-
250	-8.7356	-8.7356	421	-	-	-	-
300	-9.7941	-9.7941	1163	-	-	-	-

an upper bound (indicated by 'ub' in the table) for (5.2) using `Ipopt`. Related results are shown in Table 3, from which we see that R-HSOS is the fastest and is more scalable with the size of the problem. Besides, we can see that the lower bounds produced by R-SOS coincide with the corresponding upper bounds, and thus are globally optimal, while the lower bounds produced by R-HSOS and C-HSOS are not globally optimal.

TABLE 3
Minimizing a random quartic polynomial over the unit sphere ($r = 2$).

n	ub	R-HSOS		C-HSOS		R-SOS	
		opt	time	opt	time	opt	time
8	-2.4039	-3.5594	0.10	-3.5594	0.51	-2.4039	3.98
10	-2.8058	-4.3101	0.43	-4.3101	3.10	-2.8058	22.3
12	-2.8804	-5.0127	2.00	-5.0127	11.1	-2.8804	119
14	-3.2747	-5.7066	6.59	-5.7066	50.9	-3.2747	630
16	-3.6565	-6.3434	27.0	-6.3434	180	-	-
18	-4.2988	-7.3378	91.3	-7.3378	596	-	-
20	-4.1186	-7.4586	269	-7.4586	1844	-	-
22	-4.5440	-8.3169	717	-	-	-	-

• **Smale's Mean Value conjecture.** The next example is adapted from Smale's Mean Value conjecture. Let us begin with a result of Smale [22].

THEOREM 5.1 ([22]). *Let $P \in \mathbb{C}[\mathbf{z}]$ be a polynomial of degree $n + 1 \geq 2$ and $a \in \mathbb{C}$ such that $P'(a) \neq 0$. Then there exists a critical point b of P such that*

(5.3)
$$\left| \frac{P(a) - P(b)}{a - b} \right| \leq 4|P'(a)|.$$

Theorem 5.1 was proved by Smale in 1981, and he further asked whether one can replace the factor 4 in (5.3) by 1, or even possibly by $\frac{n}{n+1}$, which is known as Smale's Mean Value conjecture and also appears in [23]. This conjecture is verified for $n = 1, 2, 3$ [20, Chapter 7.2] and has the following equivalent (normalized) form.

CONJECTURE 5.2. *Let $P \in \mathbb{C}[\mathbf{z}]$ be a polynomial of degree $n + 1 \geq 2$ such that $P(0) = 0$ and $P'(0) = 1$, and let $b_1, \ldots, b_n$ be its critical points. Then,*

$$(5.4) \qquad \min \left\{ \left| \frac{P(b_i)}{b_i} \right| : i = 1, \ldots, n \right\} \leq \frac{n}{n+1}.$$

The upper bound $\frac{n}{n+1}$ in Conjecture 5.2 is attained by the polynomial $P(z) = z^{n+1} + z$.

We can reformulate Conjecture 5.2 as the following CPOP by introducing a new variable u:

$$(5.5) \qquad \begin{cases} \sup_{(\mathbf{z}, u) \in \mathbb{C}^{n+1}} & |u| \\ \quad \text{s.t.} & |H(z_i)| \geq |u|, \quad i = 1, \ldots, n, \\ & z_1 \cdots z_n = \frac{(-1)^n}{n+1}, \end{cases}$$

where $H(y) := \frac{1}{y} \int_0^y p(z) \, dz$ and $p(z) := (n+1)(z - z_1) \cdots (z - z_n)$ with $p(0) = 1$. The optimum of (5.5) is conjectured to be $\frac{n}{n+1}$. To ensure global convergence of the moment-HSOS hierarchies, we add a sphere constraint to (5.5):

$$(5.6) \qquad \begin{cases} \sup_{(\mathbf{z}, u) \in \mathbb{C}^{n+1}} & |u| \\ \quad \text{s.t.} & |H(z_i)| \geq |u|, \quad i = 1, \ldots, n, \\ & z_1 \cdots z_n = \frac{(-1)^n}{n+1}, \\ & |z_1|^2 + |z_2|^2 + \cdots + |z_n|^2 = n \left(\frac{1}{n+1} \right)^{\frac{2}{n}}. \end{cases}$$

We approach (5.6) for $n = 2, 3, 4, 5$ with R-HSOS and C-HSOS. Related results are shown in Table 4. For $n = 2$, the 2nd order relaxation (numerically) retrieves the desired value $\frac{2}{3}$; for $n = 3$, the 5th order relaxation (numerically) retrieves the desired value $\frac{3}{4}$; for $n = 4$, the 9th order relaxation (numerically) retrieves the desired value $\frac{4}{5}$; for $n = 5$, the hierarchies are not able to retrieve the desired value $\frac{5}{6}$ due to memory limitation. It can be seen from the table that R-HSOS is more scalable with the size of the problem.

TABLE 4
Results for Smale's Mean Value conjecture.

n	r	R-HSOS		C-HSOS	
		opt	time	opt	time
2	2	0.66666	0.005	0.66666	0.006
3	5	0.75000	0.15	0.75000	0.45
4	9	0.80000	6257	-	-
5	7	1.20615	16497	-	-

• **The Mordell inequality conjecture.** Our final example is the Mordell inequality conjecture. In 1958, Birch proposed the following conjecture: Suppose that the complex numbers $z_1, \ldots, z_n \in \mathbb{C}$ satisfies $|z_1|^2 + \cdots + |z_n|^2 = n$. Then the maximum of $\Delta := \prod_{1 \leq i < j \leq n} |z_i - z_j|^2$ is n^n. In 1960, Mordell proved affirmatively the conjecture for $n = 3$ [17]. In the same year, J. H. H. Chalk gave a counterexample

in [6] which disproves the conjecture for $n \geq 6$. The case of $n = 4$ was affirmatively solved later in 2011 [14]. So the only open case is when $n = 5$. We reformulate the Mordell inequality conjecture as the following CPOP:

$$(5.7) \qquad \begin{cases} \sup\limits_{\mathbf{z} \in \mathbb{C}^n} & \prod_{1 \leq i < j \leq n} |z_i - z_j|^2 \\ \text{s.t.} & |z_1|^2 + \cdots + |z_n|^2 = n, \end{cases}$$

whose optimum is conjectured to be n^n $(n \leq 5)$.

The maximum of Δ is achieved only if $z := \frac{1}{n}(z_1 + \cdots + z_n) = 0$. To see this, note

$$(5.8) \qquad \sum_{i=1}^n |z_i - z|^2 = \sum_{i=1}^n |z_i|^2 - n|z|^2.$$

If $z \neq 0$, then by substituting $y_i = \lambda(z_i - z)$ for z_i with $\lambda = \sqrt{\frac{\sum_{i=1}^n |z_i|^2}{\sum_{i=1}^n |z_i - z|^2}}$, it holds $|y_1|^2 + \cdots + |y_n|^2 = n$ and the value of Δ gets larger. Therefore we can eliminate z_n in (5.7) using $z_1 + \cdots + z_n = 0$ without changing its optimum:

$$(5.9) \qquad \begin{cases} \sup\limits_{\mathbf{z} \in \mathbb{C}^{n-1}} & \prod_{1 \leq i < j \leq n-1} |z_i - z_j|^2 \prod_{i=1}^{n-1} |z_i + z_1 + \ldots + z_{n-1}|^2 \\ \text{s.t.} & |z_1|^2 + \cdots + |z_{n-1}|^2 + |z_1 + \ldots + z_{n-1}|^2 = n. \end{cases}$$

We approach (5.9) with R-HSOS and C-HSOS. Related results are shown in Table 5 ($n = 3$) and Table 6 ($n = 4$), from which we see that the upper bounds provided by the two hierarchies get close to the desired value 27 or 256 as the relaxation order r grows while R-HSOS is faster and more scalable.

TABLE 5
Results for the Mordell inequality conjecture with $n = 3$.

r	R-HSOS		C-HSOS	
	opt	time	opt	time
8	27.658	0.02	27.658	0.04
10	27.348	0.03	27.348	0.06
12	27.228	0.05	27.228	0.11
14	27.144	0.11	27.144	0.26
16	27.104	0.18	27.104	0.49
18	27.080	0.19	27.079	0.81
20	27.074	0.28	27.064	1.07
22	27.059	0.44	27.055	1.35

6. Application in polyphase code design. In this section, we provide an application of the proposed real moment-HSOS hierarchy to polyphase code design, which is a fundamental problem in signal processing [7]. A polyphase code set consists of n signals represented by n complex numbers $z_1, \ldots, z_n$ of unit norm. Given a polyphase code set $z_1, \ldots, z_n$, the related aperiodic autocorrelation functions are defined by

$$(6.1) \qquad A(j) = \sum_{i=1}^{n-j} z_i \overline{z}_{i+j}, \quad j = 1, \ldots, n - 2.$$

TABLE 6
Results for the Mordell inequality conjecture with $n = 4$.

r	R-HSOS		C-HSOS	
	opt	time	opt	time
8	497.37	1.38	497.37	5.41
10	343.67	8.93	343.67	52.5
12	326.85	44.9	326.85	279
14	292.89	215	292.87	643
16	277.65	792	277.55	1979
18	274.10	1685	274.16	7048
20	269.07	4245	-	-

In polyphase code design, the goal is to minimize the norms of aperiodic autocorrelation functions $A(j), j = 1, \ldots, n-2$ (called sidelobes). If we take the total autocorrelation sidelobe energy as the cost function, then the polyphase code design problem can be formulized as the following CPOP:

$$(6.2) \qquad \begin{cases} \inf_{\mathbf{z} \in \mathbb{C}^n} & \sum_{j=1}^{n-2} \left| \sum_{i=1}^{n-j} z_i \overline{z}_{i+j} \right|^2 \\ \text{s.t.} & |z_i|^2 = 1, \quad i = 1, \ldots, n. \end{cases}$$

We approach (6.2) for $n = 4, 5, \ldots, 11$ with the 5th order relaxations of the three hierarchies. Related results are shown in Table 7. We make the following observations: (1) for $n = 4, 5, 6$, R-HSOS, C-HSOS, and R-SOS provide the same lower bound; (2) for $n \geq 7$, R-HSOS and C-HSOS provide the same lower bound while R-SOS runs out of memory; (3) R-HSOS is the fastest.

TABLE 7
Results for the polyphase code design problem (6.2) ($r = 5$).

n	opt	R-HSOS	C-HSOS	R-SOS
4	0.5000	0.01	0.02	2.04
5	1.0000	0.02	0.04	22.6
6	4.0000	0.06	0.18	497
7	1.1418	0.21	0.38	-
8	1.7428	0.57	1.94	-
9	0.0594	2.74	8.95	-
10	3.4932	13.4	68.7	-
11	2.2096	80.4	401	-

Alternatively, if we take the autocorrelation sidelobe peak as the cost function, then the polyphase code design problem can be formulized as the following optimization problem:

$$(6.3) \qquad \begin{cases} \inf_{\mathbf{z} \in \mathbb{C}^n} & \max \left\{ \left| \sum_{i=1}^{n-j} z_i \overline{z}_{i+j} \right| : j = 1, \ldots, n-2 \right\} \\ \text{s.t.} & |z_i|^2 = 1, \quad i = 1, \ldots, n. \end{cases}$$

By introducing an auxiliary variable u, we can reformulate (6.3) as the following CPOP:

$$(6.4) \quad \begin{cases} \inf\limits_{(\mathbf{z},u)\in\mathbb{C}^{n+1}} & |u| \\ \text{s.t.} & \left|\sum_{i=1}^{n-j} z_i \bar{z}_{i+j}\right| \le |u|, \quad j=1,\ldots,n-2, \\ & |z_i|^2 = 1, \quad i=1,\ldots,n. \end{cases}$$

We approach (6.4) for $n = 4, 5, \ldots, 15$ with the 4th order relaxations of the three hierarchies. Related results are shown in Table 8. We make the following observations: (1) for $n = 4, 5, 6, 7$, R-HSOS, C-HSOS, and R-SOS provide the same lower bound; (2) for $n \ge 8$, R-HSOS and C-HSOS provide the same lower bound while R-SOS runs out of memory; (3) R-HSOS is the fastest.

TABLE 8
Results for the polyphase code design problem (6.4) *(r = 4).*

n	opt	R-HSOS	C-HSOS	R-SOS
4	0.5000	0.01	0.02	1.43
5	0.7703	0.02	0.04	10.5
6	1.0000	0.05	0.12	73.2
7	0.5219	0.14	0.33	732
8	0.6483	0.37	0.77	-
9	0.1119	1.03	2.67	-
10	0.5805	2.30	7.21	-
11	0.4943	6.37	24.2	-
12	0.4928	19.3	87.8	-
13	0.4189	105	440	-
14	0.3309	324	1548	-
15	0.3098	1096	5546	-

7. Conclusions. In this paper, we have presented a real moment-HSOS hierarchy for CPOPs with real coefficients, which offers a new item in the arsenal of structure-exploiting techniques for moment-HSOS hierarchies of CPOPs. One possible topic for future research is developing a structured moment-HSOS hierarchy of CPOPs for which the polynomial data are invariant under conjugation of subsets of variables. Besides, inspired by the complexity comparison presented in Table 1, we are wondering whether it is advantageous to solve certain real POPs via the complex moment-HSOS hierarchy after converting them to equivalent complex POPs? We give an illustrative example below, and leave a thorough investigation to future research.

EXAMPLE 7.1. *Consider the following real POP:*

$$(7.1) \quad \begin{cases} \inf\limits_{x_1,x_2,x_3,x_4\in\mathbb{R}} & 3 - x_1^2 - x_3^2 + x_1 x_2^2 + 2x_2 x_3 x_4 - x_1 x_4^2 \\ \text{s.t.} & x_2 \ge 0, \\ & x_1^2 + 3x_3^2 - 2 = 0, \\ & x_4 = 0, \\ & x_1^2 + x_2^2 + x_3^2 + x_4^2 - 3 = 0. \end{cases}$$

The second-order (moment-SOS) relaxation gives a lower bound -0.414213 which is globally optimal as the moment matrix is of rank one.

Note that (7.1) can be equivalently expressed as the following CPOP by letting $z_1 = x_1 + x_3\mathbf{i}, z_2 = x_2 + x_4\mathbf{i}$:

$$
(7.2) \quad
\begin{cases}
\inf\limits_{z_1, z_2 \in \mathbb{C}} & 3 - |z_1|^2 + \frac{1}{2}z_1\overline{z}_2^2 + \frac{1}{2}z_2^2\overline{z}_1 \\
\text{s.t.} & z_2 + \overline{z}_2 \geq 0, \\
& |z_1|^2 - \frac{1}{4}z_1^2 - \frac{1}{4}\overline{z}_1^2 - 1 = 0, \\
& z_2^2 + \overline{z}_2^2 - 2|z_2|^2 = 0, \\
& |z_1|^2 + |z_2|^2 - 3 = 0.
\end{cases}
$$

The second-order (real moment-HSOS) relaxation gives a lower bound -0.909535, and the third-order relaxation achieves the global optimum -0.414213. Despite the real moment-HSOS hierarchy needs a higher order relaxation to attain global optimality, we note that the third-order complex moment matrix is of size 10 while the second-order real moment matrix is of size 15, which means for this specific problem, the complex reformulation is more preferable from the point of view of computational complexity.

REFERENCES

[1] T. Aittomaki and V. Koivunen, *Beampattern optimization by minimization of quartic polynomial*, in 2009 IEEE/SP 15th Workshop on Statistical Signal Processing, IEEE, 2009, pp. 437–440.

[2] 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, H. Frenk, K. Roos, T. Terlaky, and S. Zhang, eds., vol. 33 of Applied Optimization, Springer US, 2000, pp. 197–232, https://doi.org/10.1007/978-1-4757-3216-0_8, http://dx.doi.org/10.1007/978-1-4757-3216-0_8.

[3] A. Aubry, A. De Maio, B. Jiang, and S. Zhang, *Ambiguity function shaping for cognitive radar via complex quartic optimization*, IEEE Transactions on Signal Processing, 61 (2013), pp. 5603–5619.

[4] D. Bienstock, M. Escobar, C. Gentile, and L. Liberti, *Mathematical programming formulations for the alternating current optimal power flow problem*, 4OR, 18 (2020), pp. 249–292.

[5] D. I. Bondar, K. Jacobs, G. Korpas, J. Marecek, et al., *Quantum optimal control via polynomial optimization: A globally convergent approach*, arXiv preprint arXiv:2209.05790, (2022).

[6] J. Chalk, *A note on a discriminant inequality*, Canadian Mathematical Bulletin, 3 (1960), pp. 7–9.

[7] H. Deng, *Polyphase code design for orthogonal netted radar systems*, IEEE Transactions on Signal Processing, 52 (2004), pp. 3126–3135.

[8] B. Dumitrescu, *Positive trigonometric polynomials and signal processing applications*, vol. 103, Springer, 2007.

[9] J. P. D'Angelo and M. Putinar, *Polynomial optimization on odd-dimensional spheres*, in Emerging Applications of Algebraic Geometry, Springer, 2009, pp. 1–15.

[10] F. Fogel, I. Waldspurger, and A. d'Aspremont, *Phase retrieval for imaging problems*, Mathematical Programming Computation, 8 (2016), pp. 311–335.

[11] J. J. Hilling and A. Sudbery, *The geometric measure of multipartite entanglement and the singular values of a hypermatrix*, Journal of Mathematical Physics, 51 (2010), p. 072102.

[12] 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.

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

[14] K. Lin, *On mordell's problem*, Math Media (in Chinese), 35 (2011), pp. 51–58.

[15] V. Magron, G. Constantinides, and A. Donaldson, *Certified Roundoff Error Bounds Using Semidefinite Programming*, ACM Trans. Math. Softw., 43 (2017), pp. 1–34.

[16] B. MARIERE, Z.-Q. LUO, AND T. N. DAVIDSON, *Blind constant modulus equalization via convex optimization*, IEEE Transactions on Signal Processing, 51 (2003), pp. 805–818.

[17] L. J. MORDELL, *On a discriminant inequality*, Canadian Journal of Mathematics, 12 (1960), pp. 699–704.

[18] M. NEWTON AND A. PAPACHRISTODOULOU, *Sparse polynomial optimisation for neural network verification*, arXiv preprint arXiv:2202.02241, (2022).

[19] D. PIEKARZ, *The von neumann inequality for 3x3 matricesin the unit euclidean ball*, (2023).

[20] Q. I. RAHMAN, G. SCHMEISSER, ET AL., *Analytic theory of polynomials*, no. 26 in Oxford: New York: Clarendon Press, Oxford University Press, 2002.

[21] 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.

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

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

[24] O. TOKER AND H. OZBAY, *On the complexity of purely complex μ computation and related problems in multidimensional systems*, IEEE Transactions on Automatic Control, 43 (1998), pp. 409–414.

[25] H. WAKI, S. KIM, M. KOJIMA, AND M. MURAMATSU, *Sums of Squares and Semidefinite Programming Relaxations for Polynomial Optimization Problems with Structured Sparsity*, SIAM Journal on Optimization, 17 (2006), pp. 218–242.

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

[27] J. WANG, M. MAGGIO, AND V. MAGRON, *SparseJSR: A fast algorithm to compute joint spectral radius via sparse SOS decompositions*, in 2021 American Control Conference (ACC), IEEE, 2021, pp. 2254–2259.

[28] J. WANG AND V. MAGRON, *Exploiting sparsity in complex polynomial optimization*, Journal of Optimization Theory and Applications, 192 (2022), pp. 335–359.

[29] J. WANG, V. MAGRON, AND J.-B. LASSERRE, *Chordal-TSSOS: a moment-SOS hierarchy that exploits term sparsity with chordal extension*, SIAM Journal on Optimization, 31 (2021), pp. 114–141.

[30] 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.

[31] J. WANG, V. MAGRON, AND J. B. LASSERRE, *Certifying global optimality of AC-OPF solutions via sparse polynomial optimization*, Electric Power Systems Research, 213 (2022), p. 108683.

[32] J. WANG, V. MAGRON, J. B. LASSERRE, AND N. H. A. MAI, *CS-TSSOS: Correlative and term sparsity for large-scale polynomial optimization*, ACM Transactions on Mathematical Software, 48 (2022), pp. 1–26.