MOMENT/SUM-OF-SQUARES HIERARCHY FOR COMPLEX POLYNOMIAL OPTIMIZATION

CÉDRIC JOSZ[†§] AND DANIEL K. MOLZAHN[‡]

Abstract. We consider the problem of finding the global optimum of a real-valued complex polynomial $f : z \in \mathbb{C}^n \longmapsto \sum_{\alpha,\beta} f_{\alpha,\beta}\bar{z}^\alpha z^\beta \in \mathbb{R}$ ($z^\alpha := z_1^{\alpha_1} \dots z_n^{\alpha_n}$, $\overline{f_{\alpha,\beta}} = f_{\beta,\alpha}$) on a compact set defined by real-valued complex polynomial inequalities. It reduces to solving a sequence of complex semidefinite programming relaxations that grow tighter and tighter thanks to D'Angelo's and Putinar's Positivstellenstatz discovered in 2008. In other words, the Lasserre hierarchy may be transposed to complex numbers. We propose a method for exploiting sparsity and apply the complex hierarchy to problems with several thousand complex variables. These problems consist of computing optimal power flows in the European high-voltage transmission network.

Key words. Quillen property, Lasserre hierarchy, Shor relaxation, complex moment problem, sparse semidefinite programming, optimal power flow.

AMS subject classifications.

1. Introduction. Multivariate polynomial optimization where variables and data are complex numbers is a non-deterministic polynomial-time hard problem that arises in various applications such as electric power systems (Section 4), imaging science [10, 15, 32, 80], signal processing [1, 8, 20, 49, 52, 54], automatic control [87], and quantum mechanics [35]. Complex numbers are typically used to model oscillatory phenomena which are omnipresent in physical systems. Although complex polynomial optimization problems can readily be converted into real polynomial optimization problems where variables and data are real numbers, efforts have been made to find *ad hoc* solutions to complex problems [38, 39, 81]. The observation that relaxing non-convex constraints and converting from complex to real numbers are two non-commutative operations motivates our work. This leads us to transpose to complex numbers Lasserre's moment/sum-of-squares hierarchy [44] for real polynomial optimization.

The moment/sum-of-squares hierarchy succeeds to the vast development of real algebraic geometry during the twentieth century [67]. In 1900, Hilbert's seventeenth problem [77] raised the question of whether a non-negative polynomial in multiple real variables can be decomposed as a sum of squares of fractions of polynomials, to which Artin [6] answered in the affirmative in 1927. Later, positive polynomials on sets defined by a finite number of polynomial inequality constraints were investigated by Krivine [42], Stengle [82], Schmüdgen [76], and Putinar [68]. A theorem concerning such polynomials is referred to as *Positivstellensatz* [75]. For instance, Putinar proved under an assumption slightly stronger than compactness that they can be decomposed as a weighted sum of the constraints where the weights are sums of squares of polynomials. Lasserre [43–45] used this result in 2001 to develop a hierarchy of semidefinite programs to solve real polynomial optimization problems with compact feasible sets, with Parrilo [65, 66] making a similar contribution independently. In order to satisfy the assumption made by Putinar, Lasserre proposed to add a redundant ball con-

[†]French National Research Institute in Scientific Computing INRIA, Paris-Rocquencourt, BP 105, F-78153 Le Chesnay, France.

[‡]Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109, USA (molzahn@umich.edu). Support from Dow Sustainability Fellowship, ARPA-E grant DE-AR0000232 and Los Alamos National Laboratory subcontract 270958.

[§]French Transmission System Operator RTE, 9, rue de la Porte de Buc, BP 561, F-78000 Versailles, France (cedric.josz@rte-france.com). Support from CIFRE ANRT contract 2013/0179.

straint $x_1^2 + \ldots + x_n^2 \leqslant R^2$ to the description of the feasible set when it is included in a ball of radius R. Subsequent work on the hierarchy includes its comparison with lift-and-project methods [46], a new proof of Putinar's Positivstellensatz via a 1928 theorem of Pólya [78], and a proof of generically finite convergence [63].

In 1968, Quillen [73] showed that a real-valued bihomogenous complex polynomial that is positive away from the origin can be decomposed as a sum of squared moduli of holomorphic polynomials when it is multiplied by $(|z_1|^2 + \ldots + |z_n|^2)^r$ for some $r \in \mathbb{N}$. The result was rediscovered years later by Catlin and D'Angelo [19] and ignited a search for complex analogues of Hilbert's seventeenth problem [26, 27] and the ensuing Positivstellensätze [29, 70–72]. Notably, D'Angelo and Putinar [28] proved in 2008 that a positive complex polynomial on a sphere intersected by a finite number of polynomial inequality constraints can be decomposed as a weighted sum of the constraints where the weights are sums of squared moduli of holomorphic polynomials. Similar to Lasserre, we use D'Angelo's and Putinar's Positivstellensatz to construct a complex moment/sum-of-squares hierarchy of semidefinite programs to solve complex polynomial optimization problems with compact feasible sets. To satisfy the assumption in the Positivstellensatz, we propose to add a slack variable $z_{n+1} \in \mathbb{C}$ and a redundant constraint $|z_1|^2 + \ldots + |z_{n+1}|^2 = R^2$ to the description of the feasible set when it is in a ball of radius R. The complex hierarchy is more tractable than the real hierarchy yet produces potentially weaker bounds. Computational advantages are shown using the optimal power flow problem in electrical engineering.

The paper is organized as follows. Section 2 uses Shor and second-order conic relaxations to motivate the construction of a complex moment/sum-of-squares hierarchy in Section 3. Using a sparsity-exploiting method, numerical experiments on the optimal power flow problem are presented in Section 4. Section 5 concludes our work.

2. Motivation. Let $\mathbb{N}$, $\mathbb{N}^*$, $\mathbb{R}$, $\mathbb{R}_+$ and $\mathbb{C}$ denote the set of natural, positive natural, real, non-negative real, and complex numbers respectively. Also, let "**i**" denote the imaginary unit and $\mathbb{H}_n$ denote the set of Hermitian matrices of order $n \in \mathbb{N}^*$. Let's begin with the subclass of complex polynomial optimization composed of quadratically-constrained quadratic programs

$$\text{(2.1a)} \qquad \text{QCQP-}\mathbb{C}: \quad \inf_{z \in \mathbb{C}^n} \; z^H H_0 z,$$

$$\text{(2.1b)} \qquad \text{s.t.} \quad z^H H_i z \leqslant h_i, \quad i = 1, \ldots, m,$$

where $m \in \mathbb{N}^*$, $H_0, \ldots, H_m \in \mathbb{H}_n$, $h_0, \ldots, h_m \in \mathbb{R}$, and $(\cdot)^H$ denotes the conjugate transpose. The feasible set is not assumed to contain a point (i.e. it may be empty). The Shor [79] and second-order conic relaxations of QCQP-$\mathbb{C}$ share the following property: it is better to relax non-convex constraints before converting from complex to real numbers rather than to do the two operations in the opposite order.

2.1. Shor Relaxation. For $H \in \mathbb{H}_n$ and $z \in \mathbb{C}^n$, the relationship $z^H H z = \mathrm{Tr}(H z z^H)$ holds where $\mathrm{Tr}\,(\cdot)$ denotes the trace[1] of a complex square matrix. Relaxing the rank of $Z = zz^H$ in (2.1) yields

$$\text{(2.2a)} \qquad \text{SDP-}\mathbb{C}: \quad \inf_{Z \in \mathbb{H}_n} \; \mathrm{Tr}(H_0 Z),$$

$$\text{(2.2b)} \qquad \text{s.t.} \quad \mathrm{Tr}(H_i Z) \leqslant h_i, \quad i = 1, \ldots, m,$$

$$\text{(2.2c)} \qquad Z \succcurlyeq 0,$$

[1] For all matrices $A, B \in \mathbb{C}^{n \times n}$, $\mathrm{Tr}(AB) = \sum_{1 \leqslant i,j \leqslant n} A_{ij} B_{ji}$.

where $\succcurlyeq 0$ indicates positive semidefiniteness.

Let $\mathrm{Re}Z$ and $\mathrm{Im}Z$ denote the real and imaginary parts of the matrix $Z \in \mathbb{C}^{n\times n}$ respectively. Consider the ring homomorphism $\Lambda : (\mathbb{C}^{n\times n}, +, \times) \longrightarrow (\mathbb{R}^{2n\times 2n}, +, \times)$ defined by

$$\Lambda(Z) := \begin{pmatrix} \mathrm{Re}Z & -\mathrm{Im}Z \\ \mathrm{Im}Z & \mathrm{Re}Z \end{pmatrix}, \tag{2.3}$$

whose relevant properties are proven in Appendix A. To convert SDP-$\mathbb{C}$ into real numbers, real and imaginary parts of the complex matrix variable are identified using two properties: (1) a complex matrix Z is positive semidefinite if and only if the real matrix $\Lambda(Z)$ is positive semidefinite, and (2) if $Z_1, Z_2 \in \mathbb{H}_n$, then $\mathrm{Tr}\left[\Lambda(Z_1)\Lambda(Z_2)\right] = \mathrm{Tr}\left[\Lambda(Z_1 Z_2)\right] = 2\mathrm{Tr}(Z_1 Z_2)$. This yields the converted problem

$$\text{CSDP-}\mathbb{R}: \quad \inf_{X\in\mathbb{S}_{2n}} \quad \mathrm{Tr}(\Lambda(H_0)X), \tag{2.4a}$$

$$\text{s.t.} \quad \mathrm{Tr}(\Lambda(H_i)X) \leqslant h_i, \qquad i = 1, \ldots, m, \tag{2.4b}$$

$$X \succcurlyeq 0, \tag{2.4c}$$

$$X = \begin{pmatrix} A & B^T \\ B & C \end{pmatrix} \quad \& \quad \begin{array}{rcr} A & = & C, \\ B^T & = & -B, \end{array} \tag{2.4d}$$

where $\mathbb{S}_{2n}$ denotes the set of real symmetric matrices of order $2n$ and $(\cdot)^T$ indicates the transpose. A global solution to QCQP-$\mathbb{C}$ can be retrieved from CSDP-$\mathbb{R}$ if and only if $\mathrm{rank}(X) \in \{0, 2\}$ at optimality (proof in Appendix B).

In order to convert QCQP-$\mathbb{C}$ into real numbers, real and imaginary parts of the complex vector variable are identified. This is done by considering a new variable $x = \left((\mathrm{Re}z)^T \; (\mathrm{Im}z)^T \right)^T$ and observing that if $H \in \mathbb{H}_n$, then $z^H H z = x^T \Lambda(H) x = \mathrm{Tr}(\Lambda(H)xx^T)$. This gives rise to a problem which we will call QCQP-$\mathbb{R}$. Relaxing the rank of $X = xx^T$ yields

$$\text{SDP-}\mathbb{R}: \quad \inf_{X\in\mathbb{S}_{2n}} \quad \mathrm{Tr}(\Lambda(H_0)X), \tag{2.5a}$$

$$\text{s.t.} \quad \mathrm{Tr}(\Lambda(H_i)X) \leqslant h_i, \qquad i = 1, \ldots, m, \tag{2.5b}$$

$$X \succcurlyeq 0. \tag{2.5c}$$

A global solution to QCQP-$\mathbb{C}$ can be retrieved from SDP-$\mathbb{R}$ if and only if $\mathrm{rank}(X) \in \{0, 1\}$ or $\mathrm{rank}(X) = 2$ and (2.4d) holds at optimality.

We have val(SDP-$\mathbb{C}$) = val(CSDP-$\mathbb{R}$) = val(SDP-$\mathbb{R}$) where "val" is the optimal value of a problem (proof in Appendix C). The number of scalar variables of CSDP-$\mathbb{R}$ is half that of SDP-$\mathbb{R}$ due to constraint (2.4d). This constraint also halves the possible ranks of the matrix variable, which must be an even integer in CSDP-$\mathbb{R}$ whereas it can be any integer between 0 and $2n$ in SDP-$\mathbb{R}$. The number of variables in SDP-$\mathbb{R}$ can be reduced by a small fraction ($\frac{2}{2n+1}$ to be exact) by setting a diagonal element of X to 0. This does not affect the optimal value (proof in Appendix D). Figure 1 summarizes this section.

2.2. Second-Order Conic Relaxation. In SDP-$\mathbb{C}$ of Section 2.1, assume that the semidefinite constraint (2.2c) is relaxed to the second-order cones

$$\begin{pmatrix} Z_{ii} & Z_{ij} \\ Z_{ij}^H & Z_{jj} \end{pmatrix} \succcurlyeq 0 \quad , \quad 1 \leqslant i \neq j \leqslant n. \tag{2.6}$$

$$
\begin{tikzcd}[column sep=huge, row sep=huge]
\text{QCQP-}\mathbb{C} \arrow[r, "Z = zz^H", "\text{Relax rk } Z = 1"'] \arrow[d, "\begin{array}{l}\text{Identify real and}\\ \text{imaginary parts:}\\ x = \begin{pmatrix} \mathrm{Re} z \\ \mathrm{Im} z \end{pmatrix}\end{array}"] & \text{SDP-}\mathbb{C} \arrow[d, "\begin{array}{l}\text{Identify real and}\\ \text{imaginary parts:}\\ X = \begin{pmatrix} \frac{\mathrm{Re} Z}{2} & -\frac{\mathrm{Im} Z}{2} \\ \frac{\mathrm{Im} Z}{2} & \frac{\mathrm{Re} Z}{2} \end{pmatrix}\end{array}"'] \\
\text{QCQP-}\mathbb{R} \arrow[r, "X = xx^T", "\text{Relax rk } X = 1"'] & \begin{array}{cc}\text{SDP-}\mathbb{R} \neq & \text{CSDP-}\mathbb{R} \\ \text{Global} & \text{Global} \\ \text{solution if} & \text{solution if} \\ \text{rk } X = 1 & \text{rk } X = 2 \end{array}
\end{tikzcd}
$$

FIG. 1. *Non-Commutativity of Complex-to-Real Conversion and Relaxation*

Equation (2.6) is equivalent to constraining the determinant $Z_{ii}Z_{jj} - Z_{ij}Z_{ij}^H$ and diagonal elements Z_{ii} to be non-negative. This yields

$$
\begin{aligned}
&\text{(2.7a)} && \text{SOCP-}\mathbb{C}: \quad \inf_{Z \in \mathbb{H}_n} \ \mathrm{Tr}(H_0 Z), \\
&\text{(2.7b)} && \text{s.t.} \quad \mathrm{Tr}(H_i Z) \leqslant h_i, \quad i = 1, \ldots, m, \\
&\text{(2.7c)} && |Z_{ij}|^2 \leqslant Z_{ii} Z_{jj}, \quad 1 \leqslant i \neq j \leqslant n, \\
&\text{(2.7d)} && Z_{ii} \geqslant 0, \quad i = 1, \ldots, n,
\end{aligned}
$$

where $|\cdot|$ denotes the complex modulus. Identifying real and imaginary parts of the matrix variable Z leads to

$$
\begin{aligned}
&\text{(2.8a)} && \text{CSOCP-}\mathbb{R}: \quad \inf_{X \in \mathbb{S}_{2n}} \ \mathrm{Tr}(\Lambda(H_0) X), \\
&\text{(2.8b)} && \text{s.t.} \quad \mathrm{Tr}(\Lambda(H_i) X) \leqslant h_i, \quad i = 1, \ldots, m, \\
&\text{(2.8c)} && X_{ij}^2 + X_{n+i,j}^2 \leqslant X_{ii} X_{jj}, \quad 1 \leqslant i \neq j \leqslant n, \\
&\text{(2.8d)} && X_{ii} + X_{n+i,n+i} \geqslant 0, \quad i = 1, \ldots, n, \\
&\text{(2.8e)} && X = \begin{pmatrix} A & B^T \\ B & C \end{pmatrix} \quad \& \quad \begin{array}{rcr} A & = & C, \\ B^T & = & -B. \end{array}
\end{aligned}
$$

In SDP-$\mathbb{R}$ of Section 2.1, assume that the semidefinite constraint (2.5c) is relaxed to the second-order cones

$$
\begin{pmatrix} X_{ii} & X_{ij} \\ X_{ij} & X_{jj} \end{pmatrix} \succcurlyeq 0 \ , \quad 1 \leqslant i \neq j \leqslant 2n. \tag{2.9}
$$

This leads to

$$
\begin{aligned}
&\text{(2.10a)} && \text{SOCP-}\mathbb{R}: \quad \inf_{X \in \mathbb{S}_{2n}} \ \mathrm{Tr}(\Lambda(H_0) X), \\
&\text{(2.10b)} && \text{s.t.} \quad \mathrm{Tr}(\Lambda(H_i) X) \leqslant h_i, \quad i = 1, \ldots, m, \\
&\text{(2.10c)} && X_{ij}^2 \leqslant X_{ii} X_{jj}, \quad 1 \leqslant i \neq j \leqslant 2n, \\
&\text{(2.10d)} && X_{ii} \geqslant 0, \quad i = 1, \ldots, 2n.
\end{aligned}
$$

Unlike in Section 2.1, we have val(SOCP-$\mathbb{C}$) = val(CSOCP-$\mathbb{R}$) $\geqslant$ val(SOCP-$\mathbb{R}$) (proof in Appendix E). The number of scalar variables of CSOCP-$\mathbb{R}$ is half that of SOCP-$\mathbb{R}$ due to constraint (2.8e). The number of second-order conic constraints in CSOCP-$\mathbb{R}$, equal to $\frac{n(n-1)}{2}$, is roughly a fourth of that in SOCP-$\mathbb{R}$, equal to $\frac{2n(2n-1)}{2}$.

2.3. Exploiting Sparsity. Given an undirected graph $(\mathcal{V},\mathcal{E})$ where $\mathcal{V} \subset \{1,\ldots,n\}$ and $\mathcal{E} \subset \mathcal{V}\times\mathcal{V}$, define for all $Z \in \mathbb{H}_n$

$$\Psi_{(\mathcal{V},\mathcal{E})}(Z)_{ij} := \begin{cases} Z_{ij} & \text{if } \ (i,j)\in\mathcal{E} \ \text{ or } \ i=j\in\mathcal{V}, \\ 0 & \text{else.} \end{cases} \tag{2.11}$$

We associate an undirected graph $\mathcal{G}$ to QCQP-$\mathbb{C}$ whose nodes are $\{1,\ldots,n\}$ and that satisfies $H_i = \Psi_{\mathcal{G}}(H_i)$ for $i = 0,\ldots,m$. Let $\mathbb{H}_n^+$ denote the set of positive semidefinite Hermitian matrices of size n and let "Ker" denote the kernel of a linear application. Given the definition of $\mathcal{G}$, constraint (2.2c) of SDP-$\mathbb{C}$ can be relaxed to $Z \in \mathbb{H}_n^+ + \text{Ker } \Psi_{\tilde{\mathcal{G}}}$ without changing its optimal value for any graph $\tilde{\mathcal{G}}$ whose nodes are $\{1,\ldots,n\}$ and where $\mathcal{G} \subset \tilde{\mathcal{G}}$. Consider a chordal extension $\mathcal{G} \subset \mathcal{G}^{\text{ch}}$, that is to say that all cycles of length four or more have a chord (edge between two non-consecutive nodes of the cycle). Let $\mathcal{C}_1,\ldots,\mathcal{C}_p \subset \mathcal{G}^{\text{ch}}$ denote the maximal cliques of $\mathcal{G}^{\text{ch}}$. (A clique is a subgraph where all nodes are linked to one another. The set of maximally sized cliques of a given graph can be computed in linear time [84]). A chordal extension has a useful property for exploiting sparsity [34]: for all $Z \in \mathbb{H}_n$, we have that $Z \in \mathbb{H}_n^+ + \text{Ker } \Psi_{\mathcal{G}^{\text{ch}}}$ if and only if $\Psi_{\mathcal{C}_i}(Z) \succcurlyeq 0$ for $i = 1,\ldots,p$. Note that $\Psi_{\mathcal{C}_i}(Z) \succcurlyeq 0$ if and only if $\Lambda \circ \Psi_{\mathcal{C}_i}(Z) \succcurlyeq 0$, where "$\circ$" is the composition of functions. Given a graph $(\mathcal{V},\mathcal{E})$, define for $X \in \mathbb{S}_{2n}$

$$\tilde{\Psi}_{(\mathcal{V},\mathcal{E})}(X) := \begin{pmatrix} \Psi_{(\mathcal{V},\mathcal{E})}(A) & \Psi_{(\mathcal{V},\mathcal{E})}(B^T) \\ \Psi_{(\mathcal{V},\mathcal{E})}(B) & \Psi_{(\mathcal{V},\mathcal{E})}(C) \end{pmatrix}, \tag{2.12}$$

using the block decomposition in the left hand part of (2.4d). Notice that $\Lambda \circ \Psi_{(\mathcal{V},\mathcal{E})} = \tilde{\Psi}_{(\mathcal{V},\mathcal{E})} \circ \Lambda$. As a result, (2.4c) can be replaced by $\tilde{\Psi}_{\mathcal{C}_i}(X) \succcurlyeq 0$ for $i = 1,\ldots,p$ without changing the optimal value of CSDP-$\mathbb{R}$, with an analogous replacement for constraint (2.5c) in SDP-$\mathbb{R}$. If in SDP-$\mathbb{R}$ we exploit the sparsity of matrices $\Lambda(H_i)$ instead of that of H_i, the resulting graph has twice as many nodes. Computing a chordal extension and maximal cliques is hence more costly.

Sparsity in the second-order conic relaxations is exploited using the fact that applying (2.8c) and (2.10c) only for (i,j) that are edges of $\mathcal{G}$ does not change the optimal values of CSOCP-$\mathbb{R}$ and SOCP-$\mathbb{R}$.

3. Complex Moment/Sum-of-Squares Hierarchy. We now transpose the work of Lasserre [44] from real to complex numbers. Let z^α denote the monomial $z_1^{\alpha_1}\cdots z_n^{\alpha_n}$ where $z \in \mathbb{C}^n$ and $\alpha \in \mathbb{N}^n$ for some integer $n \in \mathbb{N}^*$. Define $|\alpha| := \alpha_1 + \ldots + \alpha_n$ and $\overline{w}$ as the conjugate of $w \in \mathbb{C}$. Define $\bar{z} := (\bar{z}_1,\ldots,\bar{z}_n)^T$ where $z \in \mathbb{C}^n$. Consider the sets

$$\begin{array}{rl} \mathbb{C}[z] := & \{\ p:\mathbb{C}^n \to \mathbb{C} \mid p(z) = \sum_{|\alpha|\leqslant l} p_\alpha z^\alpha,\ l\in\mathbb{N},\ p_\alpha \in \mathbb{C}\ \}, \\ \mathbb{C}[\bar{z},z] := & \{\ f:\mathbb{C}^n \to \mathbb{C} \mid f(z) = \sum_{|\alpha|,|\beta|\leqslant l} f_{\alpha,\beta}\bar{z}^\alpha z^\beta,\ l\in\mathbb{N},\ f_{\alpha,\beta}\in\mathbb{C}\ \}, \\ \mathbb{R}[\bar{z},z] := & \{\ f \in \mathbb{C}[\bar{z},z] \mid \overline{f(z)} = f(z),\ \forall z \in \mathbb{C}^n\ \}, \\ \Sigma[z] := & \{\ \sigma:\mathbb{C}^n\to\mathbb{C} \mid \sigma = \sum_{j=1}^r |p_j|^2,\ r\in\mathbb{N}^*,\ p_j\in\mathbb{C}[z]\ \}, \end{array} \tag{3.1}$$

and for all $d \in \mathbb{N}$

$$
\begin{array}{rl}
\mathbb{C}_d[z] := & \{ \ p : \mathbb{C}^n \to \mathbb{C} \mid p(z) = \sum_{|\alpha| \leqslant d} p_\alpha z^\alpha, \ p_\alpha \in \mathbb{C} \ \}, \\
\mathbb{C}_d[\bar{z}, z] := & \{ \ f : \mathbb{C}^n \to \mathbb{C} \mid f(z) = \sum_{|\alpha|,|\beta| \leqslant d} f_{\alpha,\beta} \bar{z}^\alpha z^\beta, \ f_{\alpha,\beta} \in \mathbb{C} \ \}, \\
\mathbb{R}_d[\bar{z}, z] := & \{ \ f \in \mathbb{C}_d[\bar{z}, z] \mid \overline{f(z)} = f(z), \ \forall z \in \mathbb{C}^n \ \}, \\
\Sigma_d[z] := & \{ \ \sigma : \mathbb{C}^n \to \mathbb{C} \mid \sigma = \sum_{j=1}^r |p_j|^2, \ r \in \mathbb{N}^*, \ p_j \in \mathbb{C}_d[z] \ \}.
\end{array} \tag{3.2}
$$

Note that the coefficients of a function $f \in \mathbb{R}[\bar{z}, z]$ satisfy $\overline{f_{\alpha,\beta}} = f_{\beta,\alpha}$ for all $|\alpha|, |\beta| \leqslant l$ for some $l \in \mathbb{N}$. The set of complex polynomials $\mathbb{C}[\bar{z}, z]$ is a $\mathbb{C}$-algebra (i.e. commutative ring and vector space over $\mathbb{C}$) and the set of holomorphic polynomials $\mathbb{C}[z]$ is a subalgebra of it (i.e. subspace closed under sum and product). The set of real-valued complex polynomials $\mathbb{R}[\bar{z}, z]$ is an $\mathbb{R}$-algebra. The set of sums of squared moduli of holomorphic polynomials $\Sigma[z]$ and the set $\Sigma_d[z] \subset \mathbb{R}_d[z]$ are pointed cones (i.e. closed under multiplication by elements of $\mathbb{R}_+$) that are convex (i.e. $tu + (1-t)v$ with $0 \leqslant t \leqslant 1$ belongs to them if u and v do). Let $C(K, \mathbb{C})$ denote the Banach (i.e. complete) $\mathbb{C}$-algebra of continuous functions from a compact set $K \subset \mathbb{C}^n$ to $\mathbb{C}$ equipped with the norm $\|\varphi\|_\infty := \sup_{z \in K} |\varphi(z)|$. Consider $R_K : \mathbb{C}[\bar{z}, z] \longrightarrow C(K, \mathbb{C})$ defined by $f \longmapsto f_{|K}$ where $f_{|K}$ denotes the restriction of f to K. $R_K(\mathbb{C}[\bar{z}, z])$ is a unital subalgebra of $C(K, \mathbb{C})$ (i.e. contains multiplicative unit) that separates points of K (i.e. $u \neq v \in K \Longrightarrow \exists \varphi \in R_K(\mathbb{C}[\bar{z}, z]) : \varphi(u) \neq \varphi(v)$) and that is closed under complex conjugation. It is hence a dense subalgebra due to the Complex Stone-Weiestrass Theorem. Likewise, $C(K, \mathbb{R}) := \{\varphi \in C(K, \mathbb{C}) \mid \overline{\varphi(z)} = \varphi(z), \ \forall z \in \mathbb{C}^n\}$ is a Banach $\mathbb{R}$-algebra of which $R_K(\mathbb{R}[\bar{z}, z])$ is a dense subalgebra. In other words, a continuous real-valued function of multiple complex variables can be approximated as close as desired by real-valued complex polynomials when restricted to a compact set. They are hence a powerful modeling tool in optimization. Speaking of which, let $m \in \mathbb{N}^*$ and $k, k_1, \ldots, k_m \in \mathbb{N}$. Consider $(f, g_1, \ldots, g_m) \in \mathbb{R}_k[\bar{z}, z] \times \mathbb{R}_{k_1}[\bar{z}, z] \times \ldots \times \mathbb{R}_{k_m}[\bar{z}, z]$ where there exists $|\alpha| = k$ and $|\beta| \leqslant k$ such that $f_{\alpha,\beta} \neq 0$. In addition, for $i = 1, \ldots, m$, there exists $|\alpha| = k_i$ and $|\beta| \leqslant k_i$ such that $g_{i,\alpha,\beta} \neq 0$. Consider the complex multivariate polynomial optimization problem

$$
\boxed{f^{\text{opt}} \quad := \quad \inf_{z \in \mathbb{C}^n} \quad f(z) \quad \text{s.t.} \quad g_i(z) \geqslant 0, \quad i = 1, ..., m,} \tag{3.3}
$$

where by convention $f^{\text{opt}} := +\infty$ if the feasible set is empty. The feasible set is a closed semi-algebraic set on which we make the following assumption from now on:

$$
\boxed{K := \{ \ z \in \mathbb{C}^n \mid g_i(z) \geqslant 0, \ i = 1, ..., m \ \} \ \text{ is compact.}} \tag{3.4}
$$

Let K^{opt} denote the set of optimal solutions to (3.3). It may be empty because we do not assume K to be non-empty. (Note that in practice, it is often hard to know whether there exists a feasible solution, as for the application of Section 4.)

Let $\mathcal{M}(K)$ denote the Banach space over $\mathbb{R}$ of Radon measures on K. Bear in mind that since K is compact, $\mathcal{M}(K)$ may be identified with the topological dual of $C(K, \mathbb{R})$ i.e. the Banach space over $\mathbb{R}$ of linear continuous applications from $C(K, \mathbb{R})$ to $\mathbb{R}$ equipped with the operator norm. (This is due to the Riesz-Markov-Kakutani Representation Theorem.) For $\varphi \in C(K, \mathbb{C})$, define $\int_K \varphi d\mu := \int_K \text{Re}(\varphi) d\mu + \mathbf{i} \int_K \text{Im}(\varphi) d\mu$ [74, 1.31 Definition][2]. Next, consider the convex pointed cone $\mathcal{P}(K) := \{ \ \varphi \in C(K, \mathbb{R}) \mid \varphi(z) \geqslant 0, \ \forall z \in K \ \}$. A Radon measure μ is positive

[2]We wish to thank Bruno Nazaret for bringing this reference to our attention.

(denoted $\mu \geqslant 0$) if $\varphi \in \mathcal{P}(K)$ implies that $\int_K \varphi d\mu \geqslant 0$. Let $\mathcal{M}_+(K)$ denote the set of positive Radon measures. With these definitions, we have

$$f^{\mathrm{opt}} \quad = \quad \inf_{\mu \in \mathcal{M}(K)} \quad \int_K f d\mu \quad \text{s.t.} \quad \int_K d\mu = 1 \;\;\&\;\; \mu \geqslant 0. \tag{3.5}$$

Indeed, if $z \in K$, then the Dirac[3] measure δ_z is a feasible point of (3.5) for which the objective value is equal to $f(z)$. Hence the optimal value of (3.5) is less than or equal to f^{opt}. Conversely, if μ is a feasible point of (3.5), then $\int_K (f - f^{\mathrm{opt}}) d\mu \geqslant 0$ and hence $\int_K f d\mu \geqslant \int_K f^{\mathrm{opt}} d\mu = f^{\mathrm{opt}} \int_K d\mu = f^{\mathrm{opt}}$.

Proposition 3.1. *The set of optimal solutions to* (3.5) *is*

$$\{\ \mu \in \mathcal{M}_+(K) \mid \mu(K^{\mathrm{opt}}) = 1 \;\;\&\;\; \mu(K \setminus K^{\mathrm{opt}}) = 0\ \}. \tag{3.6}$$

As a consequence, if K^{opt} is a finite set of $S \in \mathbb{N}^*$ points $z(1), \ldots, z(S) \in \mathbb{C}^n$, then the set optimal solutions to (3.5) *is* $\{\ \sum_{j=1}^S \lambda_j \delta_{z(j)} \mid \sum_{j=1}^S \lambda_j = 1 \;\;\&\;\; \lambda_1, \ldots, \lambda_S \in \mathbb{R}_+\}$.

Proof. Consider μ an optimal solution to (3.5). It must be that $\int_K (f - f^{\mathrm{opt}}) d\mu = 0$. Thus $\int_{K \setminus K^{\mathrm{opt}}} (f - f^{\mathrm{opt}}) d\mu = 0$ and $\mu(K \setminus K^{\mathrm{opt}}) = \int_{K \setminus K^{\mathrm{opt}}} d\mu = 0$. Therefore $\mu(K^{\mathrm{opt}}) = \int_{K^{\mathrm{opt}}} d\mu = \mu(K) - \mu(K \setminus K^{\mathrm{opt}}) = 1$. Conversely, if μ belongs to the set in (3.6), then it is feasible for (3.5) and $\int_K (f - f^{\mathrm{opt}}) d\mu = \int_{K \setminus K^{\mathrm{opt}}} (f - f^{\mathrm{opt}}) d\mu = 0$. Hence $\int_K f d\mu = \int_K f^{\mathrm{opt}} d\mu = f^{\mathrm{opt}} \int_K d\mu = f^{\mathrm{opt}}$. □

In order to dualize the equality constraint in (3.5), consider the Lagrange function $\mathcal{L} : \mathcal{M}_+(K) \times \mathbb{R} \longrightarrow \mathbb{R}$ defined by $(\mu, \lambda) \longmapsto \int_K f d\mu + \lambda \left(1 - \int_K d\mu\right)$. We have $\mathcal{L}(\mu, \lambda) = \lambda + \int_K (f - \lambda) d\mu$ and

$$\inf_{\mu \in \mathcal{M}_+(K)} \int_K (f - \lambda) d\mu = \left\{ \begin{array}{cl} 0 & \text{if } f(z) - \lambda \geqslant 0, \quad \forall z \in K, \\ -\infty & \text{else,} \end{array} \right. \tag{3.7}$$

since, in the second case, we may consider $t\delta_z$ for a $z \in K$ such that $f(z) - \lambda < 0$ and $t \to +\infty$. This leads to the dual problem

$$f^{\mathrm{opt}} \quad = \quad \sup_{\lambda \in \mathbb{R}} \quad \lambda \quad \text{s.t.} \quad f(z) - \lambda \geqslant 0, \;\; \forall z \in K. \tag{3.8}$$

Primal problem (3.5) gives rise to the complex moment hierarchy in Section 3.1. Dual problem (3.8) gives rise to the complex sum-of-squares hierarchy in Section 3.2.

3.1. Complex Moment Hierarchy. Let $\mathcal{H}$ (respectively $\mathcal{H}_d$) denote the set of sequences of complex numbers $(y_{\alpha,\beta})_{\alpha,\beta \in \mathbb{N}^n}$ (respectively $(y_{\alpha,\beta})_{|\alpha|,|\beta| \leqslant d}$) such that $\overline{y_{\alpha,\beta}} = y_{\beta,\alpha}$ for all $\alpha, \beta \in \mathbb{N}^n$ (respectively $|\alpha|, |\beta| \leqslant d$).

Definition 3.2. *An element $y \in \mathcal{H}$ is said to have a representing measure μ on K if $\mu \in \mathcal{M}_+(K)$ and $y_{\alpha,\beta} = \int_K \bar{z}^\alpha z^\beta d\mu$ for all $\alpha, \beta \in \mathbb{N}^n$. In that case, $y_{\alpha,\beta}$ is called the (α, β)-moment of μ.*

When $y \in \mathcal{H}$ has a representing measure on K, the measure is unique because $R_K(\mathbb{C}[\bar{z}, z])$ is dense in $C(K, \mathbb{C})$. The complex moment problem consists in characterizing the sequences that are representable by a measure on K and is connected to other branches of mathematics such as functional analysis and spectral theory of operators [2]. It has been studied by Atzmon [7], Schmüdgen [76], Putinar [69], Curto and Fialkow [22–24], Stochel [83], and Vasilescu [90]. For example, Atzmon [7, Theorem 2.1] proved that the solutions to the complex moment problem where $K = \{z \in$

[3]The Dirac measure δ_z with $z \in K$ may be identified with the continuous linear application from $C(K, \mathbb{R})$ to $\mathbb{R}$ defined by $\varphi \longmapsto \varphi(z)$. This is one way to interpret the fact that $\int_K f d\delta_z = f(z)$.

$\mathbb{C} \mid |z| = 1\}$ are the sequences $y \in \mathcal{H}$ such that $\sum_{m,n,j,k\in\mathbb{N}} c_{n,j}\ \overline{c}_{m,k}\ y_{m+j,n+k} \geqslant 0$ and $\sum_{m,n\in\mathbb{N}} w_m \overline{w}_n\ (y_{m,n} - y_{m+1,n+1}) \geqslant 0$ for all complex numbers $(c_{j,k})_{j,k\in\mathbb{N}}$ and $(w_n)_{n\in\mathbb{N}}$ with only finitely many non-zero terms. A generalization to the multidimensional case is considered in Section 3.3. We conclude our presentation of the complex moment problem by noting that the case where K is not compact is an open problem.

Consider a feasible point μ of (3.5) and the sequence $y \in \mathcal{H}$ that has representation measure μ on K. Notice that $\int_K f d\mu = \int_K \sum_{|\alpha|,|\beta|\leqslant k} f_{\alpha,\beta} \bar{z}^\alpha z^\beta d\mu = \sum_{|\alpha|,|\beta|\leqslant k} f_{\alpha,\beta} \int_K \bar{z}^\alpha z^\beta d\mu = \sum_{|\alpha|,|\beta|\leqslant k}\ f_{\alpha,\beta} y_{\alpha,\beta} =: L_y(f)$ and $\int_K d\mu = \int_K \bar{z}^0 z^0 d\mu = y_{0,0} = 1$. For all $p \in \mathbb{C}[z]$, we have $|p|^2 g_i \geqslant 0$ on K. Since $\mu \geqslant 0$, this implies that $\int_K |p|^2 g_i d\mu \geqslant 0$. Naturally, we also have $\int_K |p|^2 g_0 d\mu \geqslant 0$ if we define $g_0 := 1$. Define $k_0 := 0$ and $d^{\min} := \max\{k, k_1 \ldots, k_m\}$. Consider $d \geqslant d^{\min}$, $0 \leqslant i \leqslant m$, and $p \in \mathbb{C}_{d-k_i}[z]$. We have $\int_K |p|^2 g_i d\mu = \ldots$

$$
\begin{aligned}
&= \textstyle\int_K |\sum_{|\alpha|\leqslant d-k_i}\ p_\alpha z^\alpha|^2\ (\sum_{|\gamma|,|\delta|\leqslant k_i}\ g_{i,\gamma,\delta}\ \bar{z}^\gamma z^\delta)\ d\mu \\
&= \textstyle\int_K (\sum_{|\alpha|,|\beta|\leqslant d-k_i}\ \overline{p}_\alpha p_\beta \bar{z}^\alpha z^\beta)\ (\sum_{|\gamma|,|\delta|\leqslant k_i}\ g_{i,\gamma,\delta}\ \bar{z}^\gamma z^\delta)\ d\mu \\
&= \textstyle\int_K \sum_{|\alpha|,|\beta|\leqslant d-k_i} \overline{p}_\alpha p_\beta\ \sum_{|\gamma|,|\delta|\leqslant k_i}\ g_{i,\gamma,\delta}\ \bar{z}^{\alpha+\gamma} z^{\beta+\delta}\ d\mu \\
&= \textstyle\sum_{|\alpha|,|\beta|\leqslant d-k_i} \overline{p}_\alpha p_\beta\ \sum_{|\gamma|,|\delta|\leqslant k_i}\ g_{i,\gamma,\delta}\ \int_K \bar{z}^{\alpha+\gamma} z^{\beta+\delta}\ d\mu \\
&= \textstyle\sum_{|\alpha|,|\beta|\leqslant d-k_i} \overline{p}_\alpha p_\beta\ (\sum_{|\gamma|,|\delta|\leqslant k_i}\ g_{i,\gamma,\delta}\ y_{\alpha+\gamma,\beta+\delta}) =: M_{d-k_i}(g_i y)(\alpha,\beta) \\
&= \textstyle\sum_{|\alpha|,|\beta|\leqslant d-k_i} \overline{p}_\alpha p_\beta\ M_{d-k_i}(g_i y)(\alpha,\beta) \\
&= \vec{p}^H M_{d-k_i}(g_i y) \vec{p},
\end{aligned}
\tag{3.9}
$$

where $\vec{p} := (p_\alpha)_{|\alpha|\leqslant d-k_i}$ and $M_{d-k_i}(g_i y)$ is a Hermitian matrix indexed by $|\alpha|, |\beta| \leqslant d - k_i$. As a result

$$M_{d-k_i}(g_i y) \succcurlyeq 0, \qquad i = 0, \ldots, m, \qquad \forall d \geqslant d^{\min}. \tag{3.10}$$

To sum up, y is a feasible point of

$$
\begin{array}{rcll}
\rho & := & \inf_{y\in\mathcal{H}} & L_y(f), \\
 & & \text{s.t.} & y_{0,0} = 1, \\
 & & & M_{d-k_i}(g_i y) \succcurlyeq 0, \quad i = 0, \ldots, m, \quad \forall d \geqslant d^{\min},
\end{array}
\tag{3.11}
$$

with same objective value as μ in (3.5). Automatically, $\rho \leqslant f^{\text{opt}}$. Consider the relaxation of (3.11) defined by

$$
\boxed{
\begin{array}{rcll}
\rho_d & := & \inf_{y\in\mathcal{H}_d} & L_y(f), \\
 & & \text{s.t.} & y_{0,0} = 1, \\
 & & & M_{d-k_i}(g_i y) \succcurlyeq 0, \quad i = 0, \ldots, m,
\end{array}
}
\tag{3.12}
$$

which we name the *complex moment relaxation of order d* for reasons that will become clear with Theorem 3.8. In Section 3.2, we will introduce its dual counterpart.

Remark 3.1. Given $y \in \mathcal{H}$, the function L_y in this section can be formally be defined by the $\mathbb{C}$-linear operator $L_y : \mathbb{C}[\bar{z}, z] \longrightarrow \mathbb{C}$ such that $L_y(\bar{z}^\alpha z^\beta) = y_{\alpha,\beta}$ for all $\alpha, \beta \in \mathbb{N}$. If $\varphi \in \mathbb{C}[\bar{z}, z]$ and $\overline{\varphi} = \varphi$, then $\overline{L_y(\varphi)} = L_y(\varphi)$. Given $l, d \in \mathbb{N}$ and $\varphi \in \mathbb{R}_l[\bar{z}, z]$, the matrix M_d in (3.9) can be formally be defined as the Hermitian matrix indexed by $|\alpha|, |\beta| \leqslant d$ such that $M_d(\varphi y)(\alpha, \beta) := L_y(\varphi(z)\bar{z}^\alpha z^\beta) = \sum_{|\gamma|,|\delta|\leqslant l}\ \varphi_{\gamma,\delta}\ y_{\alpha+\gamma,\beta+\delta}$. Notice that $M_d(\varphi y)(0, 0) = L_y(\varphi)$. Lastly, define $M_d(y) := M_d(g_0 y)$ which we refer to as *complex moment matrix of order d*.

3.2. Complex Sum-of-Squares Hierarchy. We introduced notation $\vec{p}$ for $p \in \mathbb{C}_d[z]$ where $d \in \mathbb{N}$ in Section 3.1 and will now extend it to $\sigma \in \Sigma_d[z]$. For such an element, there exists $r \in \mathbb{N}^*$ and $p_j \in \mathbb{C}_d[z]$ such that $\sigma = \sum_{j=1}^r |p_j|^2$. Let $\vec{\sigma} := \sum_{j=1}^r \vec{p}_j \vec{p}_j^H$. Also, define $\langle A, B \rangle_{\mathcal{H}_d} := \mathrm{Tr}(AB)$ where $A, B \in \mathcal{H}_d$. Given $d \geqslant d^{\min}$, consider the Lagrange function $\mathcal{L}_d : \mathcal{H}_d \times \mathbb{R} \times \Sigma_{d-k_0}[z] \times \ldots \times \Sigma_{d-k_m}[z] \longrightarrow \mathbb{R}$ defined by $(y, \lambda, \sigma_0, \ldots, \sigma_m) \longmapsto L_y(f) + \lambda(1 - y_{0,0}) - \sum_{i=0}^m \langle M_{d-k_i}(g_i y), \vec{\sigma}_i \rangle_{\mathcal{H}_{d-k_i}}$. Compute $\mathcal{L}_d(y, \lambda, \sigma_0, \ldots, \sigma_m) = \lambda + L_y(f - \lambda) - \sum_{i=0}^m \sum_{j=0}^{r_i} (\vec{p}_j^{\,i})^H M_{d-k_i}(g_i y) \vec{p}_j^{\,i} = \lambda + L_y(f - \lambda) - \sum_{i=0}^m \sum_{j=0}^{r_i} L_y(|p_j^i|^2 g_i) = \lambda + L_y(f - \lambda - \sum_{i=0}^m \sigma_i g_i)$. Observe that

$$\inf_{y \in \mathcal{H}} \; L_y\left(f - \lambda - \sum_{i=0}^m \sigma_i g_i \right) = \left\{ \begin{array}{cl} 0 & \text{if } f(z) - \lambda - \sum_{i=0}^m \sigma_i(z) g_i(z) = 0, \\ & \text{for all } z \in \mathbb{C}^n, \\ -\infty & \text{else.} \end{array} \right. \tag{3.13}$$

Indeed, in the second case, there exists $z \in \mathbb{C}^n$ such that $f(z) - \lambda - \sum_{i=0}^m \sigma_i(z) g_i(z) \neq 0$. With $(y_{\alpha,\beta})_{\alpha,\beta\in\mathbb{N}} := (\bar{z}^\alpha z^\beta)_{\alpha,\beta\in\mathbb{N}}$, $L_{ty}(f - \lambda - \sum_{i=0}^m \sigma_i g_i) \longrightarrow -\infty$ for either $t \longrightarrow -\infty$ or $t \longrightarrow +\infty$. The associated dual problem of (3.12) is thus

$$\boxed{\begin{array}{llll} \rho_d^* & := & \sup_{\lambda,\sigma} & \lambda, \\ & & \text{s.t.} & f - \lambda = \sum_{i=0}^m \sigma_i g_i, \\ & & & \lambda \in \mathbb{R},\; \sigma_i \in \Sigma_{d-k_i}[z],\; i = 0, \ldots, m, \end{array}} \tag{3.14}$$

which we name the *complex sum-of-squares relaxation of order d*. Consider

$$\begin{array}{llll} \rho^* & := & \sup_{\lambda,\sigma} & \lambda, \\ & & \text{s.t.} & f - \lambda = \sum_{i=0}^m \sigma_i g_i, \\ & & & \lambda \in \mathbb{R},\; \sigma_i \in \Sigma[z],\; i = 0, \ldots, m, \end{array} \tag{3.15}$$

whose relationship with (3.11) is touched upon in Proposition 3.3 below.

Proposition 3.3. *We have* $\rho_d^* \leqslant \rho_d$ *for all* $d \geqslant d^{\min}$ *and* $\rho_d^* \longrightarrow \rho^* \leqslant \rho \leqslant f^{\mathrm{opt}}$.

Proof. The sequence $(\rho_d^*)_{d \geqslant d^{\min}}$ is non-decreasing and upper bounded by $\rho^* \in \mathbb{R} \cup \{\pm\infty\}$. Thus it converges towards some limit $\rho_{\lim}^* \in \mathbb{R} \cup \{\pm\infty\}$ such that $\rho_{\lim}^* \leqslant \rho^*$. If $\rho^* = -\infty$, then $\rho_d^* = -\infty$ for all $d \geqslant d^{\min}$ and $\rho_d^* \longrightarrow \rho^*$. If not, by definiton of the optimum ρ^*, there exists a sequence $(\lambda^l, \sigma_0^l, \ldots, \sigma_m^l)$ of feasible points such that $\lambda^l \leqslant \rho^*$ and $\lambda^l \longrightarrow \rho^*$. To each $l \in \mathbb{N}$, we may associate an integer $d(l) \in \mathbb{N}$ such that $(\lambda^l, \sigma_0^l, \ldots, \sigma_m^l)$ is a feasible point of the complex sum-of-squares relaxation of order $d(l)$. Thus $\lambda^l \leqslant \rho_{d(l)}^* \leqslant \rho^*$. As a result, $\rho_{\mathrm{limit}}^* = \rho^*$. Moreover, $(\rho_d)_{d \geqslant d^{\min}}$ is non-decreasing and upper bounded by $\rho \in \mathbb{R} \cup \{\pm\infty\}$. Thus it converges towards some limit $\rho_{\lim} \in \mathbb{R} \cup \{\pm\infty\}$ such that $\rho_{\lim} \leqslant \rho$. Moreover, weak duality implies that $\rho_d^* \leqslant \rho_d$ $(\leqslant \rho)$. Thus $\rho^* \leqslant \rho_{\lim} \leqslant \rho$. It was shown in Section 3.1 that $\rho \leqslant f^{\mathrm{opt}}$. ∎

Remark 3.2. Problems (3.15) and (3.11) may be interpreted as a pair of primal-dual linear programs in infinite-dimensional spaces [5]. Indeed, consider the duality bracket $\langle ., . \rangle$ defined from $\mathbb{R}[\bar{z}, z] \times \mathcal{H}$ to $\mathbb{R}$ by $\langle \varphi, y \rangle := L_y(\varphi)$. A sequence $(\varphi^n)_{n\in\mathbb{N}}$ in $\mathbb{R}[\bar{z}, z]$ is said to converge weakly towards $\varphi \in \mathbb{R}[\bar{z}, z]$ if for all $y \in \mathcal{H}$, we have $\langle \varphi^n, y \rangle \longrightarrow \langle \varphi, y \rangle$. Consider the weakly continuous $\mathbb{R}$-linear operator $A : \mathbb{R}[\bar{z}, z] \longrightarrow \mathbb{R}[\bar{z}, z]$ defined by $\varphi \longmapsto \varphi - \varphi_{0,0}$. Its dual $A^* : \mathcal{H} \longrightarrow \mathcal{H}$ is defined by $y \longmapsto y - y_{0,0}\delta_{0,0}$ where $(\delta_{0,0})_{0,0} = 1$ and $(\delta_{0,0})_{\alpha,\beta} = 0$ if $(\alpha, \beta) \neq (0, 0)$. Indeed, $\langle A\varphi, y \rangle = \langle \varphi, A^* y \rangle$ for all $(\varphi, y) \in \mathbb{R}[\bar{z}, z] \times \mathcal{H}$. Consider the convex pointed cone defined by $C := \Sigma[z] g_0 + \ldots + \Sigma[z] g_m$ and its dual cone $C^* := \{y \in \mathcal{H} \mid \forall \varphi \in C,\; \langle \varphi, y \rangle \geqslant 0\}$. With $b := Af$, notice that

$$\begin{array}{lllll} f_{0,0} - \rho^* & = & \inf_{\varphi \in \mathbb{R}[\bar{z}, z]} & \langle \varphi, \delta_{0,0} \rangle & \text{s.t. } A\varphi = b \;\;\&\;\; \varphi \in C, \\ f_{0,0} - \rho & = & \sup_{y \in \mathcal{H}} & \langle b, y \rangle & \text{s.t. } \delta_{0,0} - A^* y \in C^*. \end{array} \tag{3.16}$$

Let $\mathrm{cl}(C)$ denote the weak closure of C in $\mathbb{R}[\bar{z}, z]$. According to [3, 5.91 Bipolar Theorem][4], we have $\mathrm{cl}(C) = C^{**}$. In the next section, Theorem 3.4 and Theorem 3.8 provide a sufficient condition ensuring no duality gap in (3.16) and $\mathrm{cl}(C) = \{\varphi \in \mathbb{R}[\bar{z}, z] \mid \varphi_{|K} \geqslant 0\}$ respectively.

3.3. Convergence of the Complex Hierarchy. We now turn our attention to a result from algebraic geometry discovered in 2008.

THEOREM 3.4 (D'Angelo's and Putinar's Positivstellenstatz [28]). *If one of the constraints that define K in* (3.4) *is a sphere constraint $|z_1|^2 + \ldots + |z_n|^2 = 1$, then*

$$f_{|K} > 0 \quad \Longrightarrow \quad \exists \sigma_0, \ldots, \sigma_m \in \Sigma[z] : \quad f = \sum_{i=0}^{m} \sigma_i g_i. \tag{3.17}$$

Proof. D'Angelo and Putinar wrote the theorem slightly differently so we provide an explanation. Say that constraints g_{m-1} and g_m are such that $g_{m-1} = s$ and $g_m = -s$ where $s(z) := 1 - |z_1|^2 - \ldots - |z_n|^2$. With the assumptions of Theorem 3.4, the authors of [28, Theorem 3.1] show that there exists $\sigma_0, \ldots, \sigma_{m-2} \in \Sigma[z]$ and $r \in \mathbb{R}[\bar{z}, z]$ such that $f(z) = \sum_{i=0}^{m-2} \sigma_i(z) g_i(z) + r(z) s(z)$ for all $z \in \mathbb{C}^n$. Thanks to [27, Proposition 1.2], there exists $\sigma_{m-1}, \sigma_m \in \Sigma[z]$ such that $r = \sigma_{m-1} - \sigma_m$ hence the desired result. ∎

Theorem 3.4 can easily be generalized to any sphere $|z_1|^2 + \ldots + |z_n|^2 = R^2$ of radius $R > 0$. With scaled variable $w = \frac{z}{R} \in \mathbb{C}^n$, the sphere constraint has radius 1 and a monomial of (3.3) with coefficient $c_{\alpha,\beta} \in \mathbb{C}$ reads $c_{\alpha,\beta} \bar{z}^\alpha z^\beta = c_{\alpha,\beta} (R\bar{w})^\alpha (Rw)^\beta = R^{|\alpha|+|\beta|} c_{\alpha,\beta} \bar{w}^\alpha w^\beta$. With the scaled coefficients $R^{|\alpha|+|\beta|} c_{\alpha,\beta}$, Theorem 3.4 can then be applied. Reverting back to the old scale $z = Rw$ in (3.17) leads to the desired result. Accordingly, we define the following statement which we will consider true only when explicitly stated:

(3.18) **Sphere Assumption:** One of the constraints of polynomial optimization problem (3.3) is a sphere constraint $|z_1|^2 + \ldots + |z_n|^2 = R^2$ for some radius $R > 0$.

Under the sphere assumption, K is compact so assumption (3.4) holds.

COROLLARY 3.5. *Under the sphere assumption* (3.18)*, $\rho_d^* \to f^{opt}$ and $\rho_d \to f^{opt}$.*

Proof. Theorem 3.4 implies that $\rho^* = f^{\mathrm{opt}}$ because for all $\epsilon > 0$, function $f - (f^{\mathrm{opt}} - \epsilon)$ is positive on K. The sequences $(\rho_d^*)_{d \geqslant d^{\min}}$ and $(\rho_d)_{d \geqslant d^{\min}}$ converge towards f^{opt} due to Proposition 3.3. ∎

To require a sphere constraint in a complex polynomial optimization problem seems very restrictive and irrelevant for many problems. But in fact, a sphere constraint can be applied to any complex polynomial optimization problem (3.3) with a feasible set contained in a ball $|z_1|^2 + \ldots + |z_n|^2 \leqslant R^2$ of known radius $R > 0$. Indeed, simply add a slack variable $z_{n+1} \in \mathbb{C}$ and the constraint

$$\boxed{|z_1|^2 + \ldots + |z_{n+1}|^2 = R^2.} \tag{3.19}$$

Let $\hat{K}$ denote the feasible set of the problem in $n+1$ variables. If $(z_1, \ldots, z_{n+1}) \in \hat{K}$, then $(z_1, \ldots, z_n) \in K$ and has the same objective value. Conversely, if $(z_1, \ldots, z_n) \in K$,

[4] We wish to thank Jean-Bernard Baillon for bringing this reference to our attention.

then $(z_1,\dots,z_{n+1}) \in \hat{K}$ for all $z_{n+1} \in \mathbb{C}$ such that $|z_{n+1}|^2 = R^2 - |z_1|^2 \dots - |z_n|^2$. Again, the objective value is unchanged. To ensure a bijection between K and $\hat{K}$, add yet two more constraints $\mathbf{i}z_{n+1} - \mathbf{i}\overline{z}_{n+1} = 0$ and $z_{n+1} + \overline{z}_{n+1} \geqslant 0$, thereby preserving the number of global solutions. In that case, the application from K to $\hat{K}$ defined by $(z_1,\dots,z_n) \longmapsto (z_1,\dots,z_n,\sqrt{R^2 - |z_1|^2 - \dots - |z_n|^2})$ is a bijection. Adding the two extra constraints is optional and not required for convergence of optimal values.

As seen in Theorem 3.4, an equality constraint may be enforced via two opposite inequality constraints. Let $h_1,\dots,h_e$ denote $e \in \mathbb{N}^*$ equality constraints in polynomial optimization problem (3.3). Putinar and Scheiderer [71, Propositions 6.6 and 3.2 (iii)] show that the sphere assumption in D'Angelo's and Putinar's Positivstellensatz may be weakened to the existence of $r_1,\dots,r_e \in \mathbb{R}[\bar{z},z]$, $\sigma \in \Sigma[z]$, and $a \in \mathbb{R}$ such that

$$\sum_{j=1}^{e} r_j(z)h_j(z) = \sum_{i=1}^{n} |z_i|^2 + \sigma(z) + a, \qquad \forall z \in \mathbb{C}^n. \tag{3.20}$$

If a problem contains the constraints $|z_1|^2 - 1 = \dots = |z_n|^2 - 1 = 0$, then the assumption is satisfied by $r_1 = \dots = r_n = 1$, $\sigma = 0$ and $a = -n$. In particular, there is no need to add a slack variable in the non-bipartite Grothendieck problem over the complex numbers [10].

Example 3.1. D'Angelo and Putinar [28] consider $\frac{1}{3} < a < \frac{4}{9}$ and problem

$$\begin{array}{lrcl} \inf_{z\in\mathbb{C}} & f(z) & := & 1 - \frac{4}{3}|z|^2 + a|z|^4, \\ \text{s.t.} & g(z) & := & 1 - |z|^2 \geqslant 0, \end{array} \tag{3.21}$$

whose set of global solutions is $K^{\text{opt}} = \{z \in \mathbb{C} \mid |z| = 1\}$ and $f^{\text{opt}} = a - \frac{1}{3} > 0$. They prove that the decomposition $f = \sigma_0 + \sigma_1 g$ $(\sigma_0,\sigma_1 \in \Sigma[z])$ of Theorem 3.4 does not hold. As a result, the optimal values of the complex sum-of-squares relaxations cannot exceed 0 even though $f^{\text{opt}} > 0$. Indeed, if $\rho_d^* > 0$ for some order $d \geqslant d^{\min}$, then there exists $\lambda \geqslant \frac{\rho_d^*}{2}$ and $\sigma_0,\sigma_1 \in \Sigma_d[z]$ such that $f - \lambda = \sigma_0 + \sigma_1 g$. Thus $f = \lambda + \sigma_0 + \sigma_1 g$ where $\lambda + \sigma_0 \in \Sigma_d[z]$, which is a contradiction. We suggest solving

$$\begin{array}{lrcl} \inf_{z_1,z_2\in\mathbb{C}} & \hat{f}(z_1,z_2) & := & 1 - \frac{4}{3}|z_1|^2 + a|z_1|^4, \\ \text{s.t.} & \hat{g}(z_1,z_2) & := & 1 - |z_1|^2 - |z_2|^2 = 0, \end{array} \tag{3.22}$$

for which the decomposition of Theorem 3.4 holds thereby ensuring convergence of the complex moment/sum-of-squares hierarchy (Corollary 3.5). In other words, for all $\lambda < f^{\text{opt}}$ there exists $\hat{\sigma}_0 \in \Sigma[z_1,z_2]$ and $\hat{r} \in \mathbb{R}[\overline{z}_1,\overline{z}_2,z_1,z_2]$ such that

$$\hat{f}(z_1,z_2) - \lambda = \hat{\sigma}_0(z_1,z_2) + \hat{r}(z_1,z_2)\hat{g}(z_1,z_2), \qquad \forall z_1,z_2 \in \mathbb{C}. \tag{3.23}$$

Plug in $z_1 = z$ and $z_2 = 0$ and obtain $f(z) - \lambda = \hat{\sigma}_0(z,0) + \hat{r}(z,0)g(z)$ for all $z \in \mathbb{C}$. While function $z \longmapsto \hat{\sigma}_0(z,0)$ belongs to $\Sigma[z]$, function $z \longmapsto \hat{r}(z,0)$ does not! Hence we do not contradict the fact that $f = \sigma_0 + \sigma_1 g$ $(\sigma_0,\sigma_1 \in \Sigma[z])$ is impossible. Consider $a = \frac{1}{2}(\frac{1}{3} + \frac{4}{9}) = \frac{7}{18}$ so that $f^{\text{opt}} = \frac{1}{18}$. Notice that $d^{\min} = 2$ for (3.21) and (3.22). The complex relaxations of orders $2 \leqslant d \leqslant 3$ of (3.21) both yield[5] the value -0.3333. The complex relaxation of order 2 of (3.22) yields the value 0.0556 ($\approx f^{\text{opt}}$) and optimal polynomials $\hat{\sigma}_0(z_1,z_2) = 0.2780|z_2|^2 + 0.2776|z_1z_2|^2 + 0.6667|z_2|^4$ and $\hat{r}(z_1,z_2) = 0.9444 - 0.3889|z_1|^2 + 0.6665|z_2|^2$, all of which satisfy (3.23).

Proposition 3.6. *Assume that $y \in \mathcal{H}_d$ is an optimal solution to the complex moment relaxation of order $d \geqslant d^{\min}$. If* rank $M_d(y) =$ rank $M_{d-d^{\min}}(y)$ $(=: S)$, *then*

[5] MATLAB 2013a, YALMIP 2015.06.26 [50], and MOSEK are used for the numerical experiments.

$\rho_d = f^{opt}$ *and the complex polynomial optimization problem* (3.3) *has at least* S *global solutions.*

Proof. Thanks to [25, Theorem 5.1], the rank condition implies that the truncated sequence $y \in \mathcal{H}_d$ can be represented by a measure μ on K (i.e. $y_{\alpha,\beta} = \int_K \bar{z}^\alpha z^\beta d\mu,\ \forall |\alpha|, |\beta| \leqslant d$) and can thus be extended to $y \in \mathcal{H}$. The same theorem implies that $\mu = \sum_{j=1}^S \lambda_j \delta_{z(j)}$ for some S different point $z(1), \ldots, z(S)$ in K and some $\lambda_1, \ldots, \lambda_S \in \mathbb{R}_+$. In addition, $y_{0,0} = \int_K \bar{z}^0 z^0 d\mu = \sum_{j=1}^S \lambda_j = 1$ and thus $f^{\text{opt}} \geqslant \rho_d = L_y(f) = \int_K f d\mu = \sum_{j=1}^S \lambda_j f(z(j)) \geqslant \sum_{j=1}^S \lambda_j f^{\text{opt}} = f^{\text{opt}}$. We simultaneously deduce that $\rho_d = f^{\text{opt}} = f(z(1)) = \ldots = f(z(S))$. ☐

In particular, if y is an optimal solution to the complex moment relaxation of order d and rank $M_d(y) = 1$, Proposition 3.6 implies that $y_{\alpha,\beta} = \int_K \bar{z}^\alpha z^\beta d\delta_z = \bar{z}^\alpha z^\beta,\ \forall |\alpha|, |\beta| \leqslant d$ where $z \in K^{\text{opt}}$. A global solution can simply be read from y because $z = (y_{0,\beta})_{|\beta|=1}$. Future work consists in transposing [45, Algorithm 4.2] from real to complex numbers in order to extract global solutions when $S > 1$ in Proposition 3.6.

Example 3.2. Putinar and Scheiderer [72] consider parameters $0 < a < \frac{1}{2}$ and $C > \frac{1}{1-2a}$, and problem

$$\begin{array}{llll} & \inf_{z\in\mathbb{C}} & f(z) & := & C - |z|^2, \\ \text{(3.24)} & \text{s.t.} & g(z) & := & |z|^2 - az^2 - a\bar{z}^2 - 1 = 0, \end{array}$$

whose set of global solutions is $K^{\text{opt}} = \left\{ \pm \frac{1}{\sqrt{1-2a}} \right\}$ and $f^{\text{opt}} = C - \frac{1}{1-2a} > 0$. They prove that the decomposition of Theorem 3.4 does not hold. Since the feasible set is included in the Euclidean ball of radius $\sqrt{C}$, we suggest solving

$$\begin{array}{lllll} & \inf_{z_1,z_2\in\mathbb{C}} & \hat{f}(z_1,z_2) & := & C - |z_1|^2, \\ & \text{s.t.} & \hat{g}_1(z_1,z_2) & := & |z_1|^2 - az_1^2 - a\bar{z}_1^2 - 1 = 0, \\ \text{(3.25)} & & \hat{g}_2(z_1,z_2) & := & C - |z_1|^2 - |z_2|^2 = 0, \\ & & \hat{g}_3(z_1,z_2) & := & \mathbf{i}z_2 - \mathbf{i}\bar{z}_2 = 0, \\ & & \hat{g}_4(z_1,z_2) & := & z_2 + \bar{z}_2 \geqslant 0. \end{array}$$

Consider $a = \frac{1}{4}$ and $C = \frac{1}{1-2a} + 1 = 3$ so that $f^{\text{opt}} = 1$. Notice that $d^{\min} = 2$ for (3.24) and (3.25). The complex relaxations of orders $2 \leqslant d \leqslant 3$ of (3.24) are unbounded. The complex relaxation of order 2 of (3.25) yields the value 0.6813. That of order 3 yields the value 1.0000 and the complex moment matrix[6] with 10^{-4} precision

$$\text{(3.26)} \qquad M_3(y) \quad = \quad \begin{array}{c|cccccccccc} & (0,0) & (1,0) & (0,1) & (2,0) & (1,1) & (0,2) & (3,0) & (2,1) & (1,2) & (0,3) \\ \hline (0,0) & 1 & 0 & 1 & 2 & 0 & 1 & 0 & 2 & 0 & 1 \\ (1,0) & 0 & 2 & 0 & 0 & 2 & 0 & 4 & 0 & 2 & 0 \\ (0,1) & 1 & 0 & 1 & 2 & 0 & 1 & 0 & 2 & 0 & 1 \\ (2,0) & 2 & 0 & 2 & 4 & 0 & 2 & 0 & 4 & 0 & 2 \\ (1,1) & 0 & 2 & 0 & 0 & 2 & 0 & 4 & 0 & 2 & 0 \\ (0,2) & 1 & 0 & 1 & 2 & 0 & 1 & 0 & 2 & 0 & 1 \\ (3,0) & 0 & 4 & 0 & 0 & 4 & 0 & 8 & 0 & 4 & 0 \\ (2,1) & 2 & 0 & 2 & 4 & 0 & 2 & 0 & 4 & 0 & 2 \\ (1,2) & 0 & 2 & 0 & 0 & 2 & 0 & 4 & 0 & 2 & 0 \\ (0,3) & 1 & 0 & 1 & 2 & 0 & 1 & 0 & 2 & 0 & 1 \end{array}$$

[6] It so happens that the Hermitian matrix $M_3(y)$, indexed by $(\alpha, \beta) \in \mathbb{N}^2 \times \mathbb{N}^2$ with $|\alpha|, |\beta| \leqslant 3$, is real-valued in this example.

which satisfies rank $M_3(y) = $ rank $M_1(y) = 2$. Proposition 3.6 implies that $f^{\text{opt}} \approx 1.000$ and that there exists at least two global solutions $u, v \in \mathbb{C}^2$ to (3.25). It also implies that there exists $\lambda_1, \lambda_2 \in \mathbb{R}_+$ such that $y_{\alpha,\beta} = \lambda_1 \overline{u}^\alpha u^\beta + \lambda_2 \overline{v}^\alpha v^\beta$ for all $|\alpha|, |\beta| \leqslant 3$. One may check that $\lambda_1 = \lambda_2 = 0.5000$, $u = (1.4142 + \mathbf{i}0.0000, 1.0000 + \mathbf{i}0.0000)$, and $v = (-1.4142 + \mathbf{i}0.0000, 1.0000 + \mathbf{i}0.0000)$ approximately satisfy those equations. We suspect that the set of optimal solutions to the complex moment relaxation of order 3 is $\{\lambda_1 (\overline{u}^\alpha u^\beta)_{|\alpha|,|\beta|\leqslant 3} + \lambda_2 (\overline{v}^\alpha v^\beta)_{|\alpha|,|\beta|\leqslant 3} \mid \lambda_1, \lambda_2 \geqslant 0 \;\&\; \lambda_1 + \lambda_2 = 1\}$ and that the path-following interior-point solver returns the analytic center of it. This suggests that $\pm 1.4142 + \mathbf{i}0.0000$ are two global solutions to (3.24).

As a by-product of Corollary 3.5, we propose a solution to the complex moment problem in Theorem 3.8 below. To that end, consider Lemma 3.7 below where we transpose [40, Lemma 3] from real to complex numbers.

Lemma 3.7. Let $s : \mathbb{C}^n \longrightarrow \mathbb{R}$ be defined by $s(z) := R^2 - |z_1|^2 - \ldots - |z_n|^2$. Given $d \in \mathbb{N}^*$ and $y \in \mathcal{H}_d$, we have

$$(3.27) \qquad (\; M_d(g_0 y) \succcurlyeq 0 \quad \& \quad M_{d-1}(sy) = 0 \;) \qquad \Longrightarrow \qquad \text{Tr}(M_d(g_0 y)) \leqslant y_{0,0} \sum_{l=0}^{d} R^{2l}.$$

Proof. Given $1 \leqslant l \leqslant d$, we have $\text{Tr}(M_{l-1}(sy)) = \ldots$

$$(3.28) \qquad \begin{array}{rl} = & \sum_{|\alpha|\leqslant l-1} \; M_{l-1}(sy)(\alpha,\alpha) \\ = & \sum_{|\alpha|\leqslant l-1} \; L_y(s(z)\bar{z}^\alpha z^\alpha) \\ = & \sum_{|\alpha|\leqslant l-1} \sum_{|\gamma|\leqslant 1} \; s_{\gamma,\gamma} \; y_{\gamma+\alpha,\gamma+\alpha} \\ = & \sum_{|\alpha|\leqslant l-1, |\gamma|=0} \; s_{\gamma,\gamma} \; y_{\gamma+\alpha,\gamma+\alpha} \; + \; \sum_{|\alpha|\leqslant l-1,|\gamma|=1} \; s_{\gamma,\gamma} \; y_{\gamma+\alpha,\gamma+\alpha} \\ = & \sum_{|\alpha|\leqslant l-1} R^2 \; y_{\alpha,\alpha} - \sum_{|\alpha|\leqslant l-1,|\gamma|=1} y_{\gamma+\alpha,\gamma+\alpha}. \end{array}$$

$M_{d-1}(sy) = 0$ implies that $M_{l-1}(sy) = 0$ for all $1 \leqslant l \leqslant d$ and hence $\text{Tr}(M_{l-1}(sy)) = 0$. In addition, $\sum_{0<|\alpha|\leqslant l} y_{\alpha,\alpha} \leqslant \sum_{|\alpha|\leqslant l-1,|\gamma|=1} y_{\gamma+\alpha,\gamma+\alpha}$. Thus

$$(3.29) \qquad \sum_{|\alpha|\leqslant l} y_{\alpha,\alpha} \leqslant y_{0,0} + R^2 \sum_{|\alpha|\leqslant l-1} y_{\alpha,\alpha}, \qquad l = 1, \ldots, d,$$

which proves the lemma. $\square$

Theorem 3.8. *Under the sphere assumption* (3.18)*, a sequence $y \in \mathcal{H}$ has a representing measure on K if and only if*

$$(3.30) \qquad M_d(g_i y) \succcurlyeq 0, \qquad i = 0, \ldots, m, \qquad \forall d \in \mathbb{N}.$$

Proof. In Section 3.1, it was shown that if $y \in \mathcal{H}$ has a representing measure on K, then (3.10) holds. Notice that (3.10) and (3.30) are equivalent, hence the “only if” part. Concerning the “if” part, assume that $y \in \mathcal{H}$ satisfies (3.30). If $y_{0,0} = 0$, then Lemma 3.7 implies that $y = 0$ which can be represented by $\mu = 0$ on K. Otherwise $y_{0,0} > 0$ and $y/y_{0,0}$ is a feasible point of problem (3.11) whose optimal value is f^{opt} for all $f \in \mathbb{R}[\bar{z}, z]$ according to Corollary 3.5. If moreover $f_{|K} \geqslant 0$, then $L_{y/y_{0,0}}(f) \geqslant f^{\text{opt}} \geqslant 0$. In particular, if $f_{|K} = 0$, then $L_{y/y_{0,0}}(f) = 0$. We may therefore define $\tilde{L}_{y/y_{0,0}} : R_K(\mathbb{C}[\bar{z}, z]) \longrightarrow \mathbb{C}$ such that $\tilde{L}_{y/y_{0,0}}(\varphi_{|K}) := L_{y/y_{0,0}}(\varphi)$ (similarily to Schweighöfer [78, Proof of Theorem 2]). If $\varphi \in R_K(\mathbb{R}[\bar{z}, z])$, then $\tilde{L}_{y/y_{0,0}}(\|\varphi\|_\infty - \varphi) \geqslant 0$ and $\tilde{L}_{y/y_{0,0}}(\varphi) \leqslant \|\varphi\|_\infty$. Linearity implies that $|\tilde{L}_{y/y_{0,0}}(\varphi)| \leqslant$

$\|\varphi\|_\infty$. As a result, for all $\varphi \in R_K(\mathbb{C}[\bar{z}, z])$, we have $|\tilde{L}_{y/y_{0,0}}(\varphi)| = |\tilde{L}_{y/y_{0,0}}(\mathrm{Re}(\varphi) + \mathbf{i}\mathrm{Im}(\varphi))| = |\tilde{L}_{y/y_{0,0}}(\mathrm{Re}(\varphi)) + \mathbf{i}\tilde{L}_{y/y_{0,0}}(\mathrm{Im}(\varphi))| \leqslant |\tilde{L}_{y/y_{0,0}}(\mathrm{Re}(\varphi))| + |\tilde{L}_{y/y_{0,0}}(\mathrm{Im}(\varphi))| \leqslant \|\mathrm{Re}(\varphi)\|_\infty + \|\mathrm{Im}(\varphi)\|_\infty \leqslant 2\|\varphi\|_\infty$. Moreover, $R_K(\mathbb{C}[\bar{z}, z])$ is dense in $C(K, \mathbb{C})$. Therefore $\tilde{L}_{y/y_{0,0}}$ may be extended to a continous linear functional on $C(K, \mathbb{C})$ (we preserve the same name for the extension). K is compact thus the Riesz-Markov-Kakutani Representation Theorem implies that there exists a unique Radon measure μ such that $\tilde{L}_{y/y_{0,0}}(\varphi) = \int_K \varphi d\mu$ for all $\varphi \in C(K, \mathbb{C})$. It is positive because $\varphi \in \mathcal{P}(K)$ implies that $\tilde{L}_{y/y_{0,0}}(\varphi) \geqslant 0$ (density argument). Finally, if $\alpha, \beta \in \mathbb{N}^n$, $y_{\alpha,\beta}/y_{0,0} = L_{y/y_{0,0}}(\bar{z}^\alpha z^\beta)$ (c.f. Remark 3.1) hence y has representing measure $y_{0,0}\mu$ on K. □

Vasilescu [90, Theorem I.2.17] has already proposed a solution to the complex moment problem on K but it is different from ours. We now transpose the proof of [40, Theorem 1] from real to complex numbers.

Proposition 3.9. *Under the sphere assumption* (3.18), $\rho_d^* = \rho_d \in \mathbb{R} \cup \{+\infty\}$ *for all* $d \geqslant d^{\min}$.

Proof. Given $A \in \mathcal{H}_d$, consider the Frobenius norm $\|A\|_F := \lambda_{\max}(A)$, the greatest eigenvalue of A, and the norm $\|A\|_{\mathcal{H}_d} := \sqrt{\langle A, A\rangle_{\mathcal{H}_d}}$. Consider $d \geqslant d^{\min}$. Two cases can occur. The first is that the feasible set of the complex moment relaxation of order d is non-empty, in which case we consider a feasible point $(y_{\alpha,\beta})_{|\alpha|,|\beta|\leqslant d}$. All norms are equivalent in finite dimension so there exists a constant $C_d \in \mathbb{R}$ such that $\sqrt{\sum_{|\alpha|,|\beta|\leqslant d} |y_{\alpha,\beta}|^2} = \|M_d(g_0 y)\|_{\mathcal{H}_d} \leqslant C_d \; \|M_d(g_0 y)\|_F \leqslant C_d \; \sum_{l=0}^d R^{2l}$, according to Lemma 3.7. As a result, the feasible set of the complex moment relaxation of order d is a non-empty compact set and so is its image by Λ (defined in (2.3)). We can thus apply Trnovská's result [88] which states that in a semidefinite program in real numbers, if the primal feasible set is non-empty and compact, then there exists a dual interior point and there is no duality gap.

The second case is that the feasible set of the complex moment relaxation of order d is empty, i.e. $\rho_d = +\infty$. It must be strongly infeasible because it cannot be weakly infeasible (see [30, Section 5.2] for definitions). Indeed, if it is weakly infeasible, then there exists a sequence $(y^j)_{j\in\mathbb{N}}$ of elements of $\mathcal{H}$ such that for all $j \in \mathbb{N}$, we have $|y^j_{0,0} - 1| \leqslant \frac{1}{j+1}$ and $\lambda_{\min}(M_{d-k_i}(g_i y^j)) \geqslant -\frac{1}{j+1}$ where $i = 0, \ldots, m$. Define $c := (n + d)!/(n!d!)$. We now mimick the computations in Lemma 3.7 using $y^j_{0,0} \leqslant 1 + \frac{1}{j+1} \leqslant 2$ and $|\mathrm{Tr}(M_{l-1}(sy^j))| \leqslant \frac{c}{j+1} \leqslant c$ if $1 \leqslant l \leqslant d$. Consider $j_0 \in \mathbb{N}$ such that for all $j \geqslant j_0$ and $1 \leqslant l \leqslant d$, we have $\sum_{|\alpha|\leqslant l-1, |\gamma|=1} y^j_{\gamma+\alpha,\gamma+\alpha} - \sum_{0<|\alpha|\leqslant l} y^j_{\alpha,\alpha} \geqslant -1$. Equation (3.29) then becomes $\sum_{|\alpha|\leqslant l} y^j_{\alpha,\alpha} \leqslant 2 + R^2 \left(\sum_{|\alpha|\leqslant l-1} y^j_{\alpha,\alpha}\right) + c + 1$. As a result, $\mathrm{Tr}(M_d(g_0 y^j)) = \sum_{|\alpha|\leqslant d} y^j_{\alpha,\alpha} \leqslant (3 + c) \sum_{l=0}^d R^{2l}$, which, together with $\lambda_{\min}(M_d(g_0 y^j)) \geqslant -\frac{1}{j+1} \geqslant -1$, yields $\lambda_{\max}(M_d(g_0 y^j)) \leqslant (3 + c) \sum_{l=0}^d R^{2l} + c - 1$. Hence for all $j \geqslant j_0$, the spectrum of $M_d(g_0 y^j)$ is lower bounded by -1 and upper bounded by $B_d := (3+c) \sum_{l=0}^d R^{2l} + c - 1 \geqslant 1$. We therefore have $\sqrt{\sum_{|\alpha|,|\beta|\leqslant d} |y^j_{\alpha,\beta}|^2} \leqslant C_d \; \|M_d(g_0 y)\|_F \leqslant C_d \times B_d$. The sequence $(y^j)_{j\geqslant j_0}$ is thus included in a compact set. Hence there exists a subsequence that converges towards a limit $y^{\lim}$ which satisfies $y^{\lim}_{0,0} = 1$ and the constraints $\lambda_{\min}(M_{d-k_i}(g_i y^{\lim})) \geqslant 0$, $i = 0, \ldots, m$. Therefore $y^{\lim}$ is a feasible point of the complex moment relaxation of order d, which is a contradiction. Strong infeasibility means that the dual feasible set contains an improving ray [30, Definition 5.2.2]. Moreover, $\inf_{y\in\mathcal{H}_d} L_y(f)$ subject to $y_{0,0} = 1$, $M_d(g_0 y) \succcurlyeq 0$, and $M_{d-1}(sy) = 0$ is a semidefinite program with a non-empty compact feasible set hence the dual feasible set contains a point $(\lambda, \sigma_0, \sigma_1)$. As result $(\lambda, \sigma_0, \sigma_1, 0, \ldots, 0)$ is a feasible point of the complex sum-of-squares relaxation of order d. Together with the improving ray, this means that $\rho_d^* = +\infty$. To conclude, $\rho_d^* = \rho_d$ in both cases. $\square$

Proposition 3.10. *Assume that complex polynomial optimization problem* (3.3) *satisfies* (3.20) *and has a global solution* $z^{\text{opt}} \in K^{\text{opt}}$. *In addition, assume that* $(\sigma_0^{\text{opt}}, \ldots, \sigma_m^{\text{opt}}) \in \Sigma[z]^{m+1}$ *is an optimal solution to the sum-of-squares problem* (3.15). *Then* $(z^{\text{opt}}, \sigma_1^{\text{opt}}, \ldots, \sigma_m^{\text{opt}})$ *is a saddle point of* $\phi : \mathbb{C}^n \times \Sigma[z]^m \longrightarrow \mathbb{R}$ *defined by* $(z, \sigma) \longmapsto f(z) - \sum_{i=1}^m \sigma_i(z) g_i(z)$.

Proof. The optimality of $(\sigma_0^{\text{opt}}, \ldots, \sigma_m^{\text{opt}})$ means that $f - f^{\text{opt}} = \sum_{i=0}^m \sigma_i^{\text{opt}} g_i$. With $f(z^{\text{opt}}) - f^{\text{opt}} = \sum_{i=0}^m \sigma_i^{\text{opt}}(z^{\text{opt}}) g_i(z^{\text{opt}}) = 0$, $\sigma_i^{\text{opt}}(z^{\text{opt}}) \geqslant 0$, and $g_i(z^{\text{opt}}) \geqslant 0$, we have $\sigma_i^{\text{opt}}(z^{\text{opt}}) g_i(z^{\text{opt}}) = 0$ for $i = 0, \ldots, m$. It follows that $\phi(z^{\text{opt}}, \sigma) \leqslant \phi(z^{\text{opt}}, \sigma^{\text{opt}})$ for all $\sigma \in \Sigma[z]$. For all $z \in \mathbb{C}^n$, $\phi(z^{\text{opt}}, \sigma^{\text{opt}}) \leqslant \phi(z, \sigma^{\text{opt}})$ because $f(z) - f^{\text{opt}} - \sum_{i=1}^m \sigma_i^{\text{opt}}(z) g_i(z) = \sigma_0^{\text{opt}}(z) \geqslant 0$. $\square$

Given an application $\varphi : \mathbb{C}^n \longrightarrow \mathbb{R}$, define $\tilde{\varphi} : \mathbb{R}^{2n} \longrightarrow \mathbb{R}$ by $(x, y) \longmapsto \varphi(x + \mathbf{i}y)$. If $\tilde{\varphi}$ is $\mathbb{R}$-differentiable at point $(x, y) \in \mathbb{R}^{2n}$, consider the Wirtinger derivative [92] defined by $\nabla\varphi(x + \mathbf{i}y) := \frac{1}{2}(\nabla_x \tilde{\varphi}(x, y) - \mathbf{i}\nabla_y \tilde{\varphi}(x, y)) \in \mathbb{C}^n$.

Corollary 3.11. *With the same assumptions as in* Proposition 3.10, *we have*

$$\boxed{\begin{array}{l} \nabla f(z^{\text{opt}}) = \sum_{i=1}^m \sigma_i^{\text{opt}}(z^{\text{opt}}) \nabla g_i(z^{\text{opt}}), \\ \sigma_i^{\text{opt}}(z^{\text{opt}}), g_i(z^{\text{opt}}) \geqslant 0, \quad i = 1, \ldots, m, \\ \sigma_i^{\text{opt}}(z^{\text{opt}}) g_i(z^{\text{opt}}) = 0, \quad i = 1, \ldots, m. \end{array}} \tag{3.31}$$

Proof. z^{opt} is a minimizer of $z \in \mathbb{C}^n \longmapsto \phi(z, \sigma^{\text{opt}})$ thus $\nabla_z \phi(z^{\text{opt}}, \sigma^{\text{opt}}) = \nabla f(z^{\text{opt}}) - \sum_{i=1}^m \nabla\sigma_i^{\text{opt}}(z^{\text{opt}}) g_i(z^{\text{opt}}) - \sum_{i=1}^m \sigma_i^{\text{opt}}(z^{\text{opt}}) \nabla g_i(z^{\text{opt}}) = 0$. Consider $1 \leqslant i \leqslant m$. Since $\sigma_i^{\text{opt}}(z^{\text{opt}}) = 0$ and $\sigma_i^{\text{opt}} \in \Sigma[z]$, it must be that $|z_k - z_k^{\text{opt}}|^2$ divides $\sigma_{i,k}^{\text{opt}} : z_k \in \mathbb{C} \longmapsto \sigma_i^{\text{opt}}(z_1^{\text{opt}}, \ldots, z_{k-1}^{\text{opt}}, z_k, z_{k+1}^{\text{opt}}, \ldots, z_n^{\text{opt}})$. With $z_k^{\text{opt}} =: x_k^{\text{opt}} + \mathbf{i}y_k^{\text{opt}}$, the real number x_k^{opt} is a root of multiplicity 2 of $x_k \in \mathbb{R} \longmapsto \sigma_{i,k}^{\text{opt}}(x_k + \mathbf{i}y_k^{\text{opt}})$, with an analogous remark for y_k^{opt}. Thus $\nabla\sigma_i^{\text{opt}}(z^{\text{opt}}) = 0$ which leads to the desired result. $\square$

3.4. Comparison of Real and Complex Hierarchies. Similar to Sections 2.1 and 2.2, the following notations will be used: POP-$\mathbb{C}$ denotes the complex polynomial optimization problem (3.3); POP-$\mathbb{R}$ denotes the real polynomial optimization problem after conversion of POP-$\mathbb{C}$ into real numbers; MSOS$_d$-$\mathbb{C}$ denotes the complex moment/sum-of-squares relaxation of order d applied to POP-$\mathbb{C}$; CMSOS$_d$-$\mathbb{R}$ denotes the conversion of MSOS$_d$-$\mathbb{C}$ into real numbers; and MSOS$_d$-$\mathbb{R}$ denotes the real moment/sum-of-squares relaxation of order d applied to POP-$\mathbb{R}$. Let $d^{\min}$-$\mathbb{R}$ and $d^{\min}$-$\mathbb{C}$ respectively denote the minimum orders of the real and complex hierarchies. Consider the sets

$$\begin{array}{rl} \mathbb{R}[x,y] := & \{\, q : \mathbb{R}^{2n} \to \mathbb{R} \mid q(x,y) = \sum_{|\kappa| \leqslant j} q_\kappa (x,y)^\kappa, j \in \mathbb{N}, q_\kappa \in \mathbb{R}\}, \\ \mathbb{R}_d[x,y] := & \{\, q : \mathbb{R}^{2n} \to \mathbb{R} \mid q(x,y) = \sum_{|\kappa| \leqslant d} q_\kappa (x,y)^\kappa,\ q_\kappa \in \mathbb{R} \,\}, \\ \Sigma_d[x,y] := & \{\, \sigma : \mathbb{R}^{2n} \to \mathbb{R} \mid \sigma = \sum_{i=1}^r q_i^2, \text{ with } r \in \mathbb{N}^*, q_i \in \mathbb{R}_d[x,y] \,\}, \end{array} \tag{3.32}$$

where $\kappa \in \mathbb{N}^{2n}$ and $(x,y)^\kappa := x_1^{\kappa_1} \ldots x_n^{\kappa_n} y_1^{\kappa_{n+1}} \ldots y_n^{\kappa_{2n}}$.

Proposition 3.12. *Under the sphere assumption* (3.18), *for all integer d greater than or equal to* $\max\{d^{\min}\text{-}\mathbb{R}, d^{\min}\text{-}\mathbb{C}\}$, *we have*

$$\text{val}(\text{MSOS}_d\text{-}\mathbb{C}) = \text{val}(\text{CMSOS}_d\text{-}\mathbb{R}) \leqslant \text{val}(\text{MSOS}_d\text{-}\mathbb{R}). \tag{3.33}$$

Proof. It suffices to compare the optimal values of the real and complex sum-of-squares relaxations. This is due to Proposition 3.9 and [40] where the ball constraint can be replaced by the sphere constraint to ensure no duality gap. We have

$$
\begin{array}{rcll}
\mathrm{val}(\text{POP-}\mathbb{C}) & = & \sup_{\lambda\in\mathbb{R}} & \lambda, \\
 & & \text{s.t.} & f(z)-\lambda \geqslant 0,\ \forall z\in K, \\[1em]
\mathrm{val}(\text{MSOS}_d\text{-}\mathbb{C}) & = & \sup_{\lambda,\sigma} & \lambda, \\
 & & \text{s.t.} & f-\lambda = \sum_{i=0}^m \sigma_i g_i, \\
 & & & \lambda\in\mathbb{R},\ \sigma_i\in\Sigma_{d-k_i}[z],\ i=0,\dots,m, \\[1em]
\mathrm{val}(\text{CMSOS}_d\text{-}\mathbb{R}) & = & \sup_{\lambda,\sigma} & \lambda, \\
 & & \text{s.t.} & \tilde{f}-\lambda = \sum_{i=0}^m \sigma_i \tilde{g}_i, \\
 & & & \lambda\in\mathbb{R},\ \sigma_i\in\Sigma_{d-k_i}[x+\mathbf{i}y],\ i=0,\dots,m, \\[1em]
\mathrm{val}(\text{POP-}\mathbb{R}) & = & \sup_{\lambda\in\mathbb{R}} & \lambda, \\
 & & \text{s.t.} & \tilde{f}(x,y)-\lambda \geqslant 0,\ \forall (x+\mathbf{i}y)\in K, \\[1em]
\mathrm{val}(\text{MSOS}_d\text{-}\mathbb{R}) & = & \sup_{\lambda,\sigma} & \lambda, \\
 & & \text{s.t.} & \tilde{f}-\lambda = \sum_{i=0}^m \sigma_i \tilde{g}_i, \\
 & & & \lambda\in\mathbb{R},\ \sigma_i\in\Sigma_{d-k_i}[x,y],\ i=0,\dots,m.
\end{array}
\tag{3.34}
$$

We now conclude because for all $d\in\mathbb{N}$, $\Sigma_d[x+\mathbf{i}y]\subset\Sigma_d[x,y]$. Indeed, if $\sigma = \sum_{j=1}^r |p_j|^2$ with $r\in\mathbb{N}^*$ and $p_1,\dots,p_r\in\mathbb{C}_d[z]$, then $\tilde{\sigma}(x,y) = \sum_{j=1}^r |\tilde{p}_j(x,y)|^2 = \frac{1}{4}\sum_{j=1}^r \left(\tilde{p}_j(x,y)+\overline{\tilde{p}_j(x,y)}\right)^2 + \left(\mathbf{i}\tilde{p}_j(x,y)+\overline{\mathbf{i}\tilde{p}_j(x,y)}\right)^2 \in \Sigma_d[x,y]$. ∎

We may suspect the inequality in (3.33) to be strict in some cases because $\Sigma_d[x+\mathbf{i}y]$ is a strict subset of $\Sigma_d[x,y]$ for all $d\in\mathbb{N}^*$. Indeed, for $i=1,\dots,n$, we have $x_i^2 = \left(\frac{z_i+\bar{z}_i}{2}\right)^2 = \frac{1}{4}(z_i^2+2|z_i|^2+\bar{z}_i^2) \in \Sigma_d[x,y]\setminus\Sigma_d[x+\mathbf{i}y]$. According to numerical experiments[7], the inequality is strict for (3.25) in Example 3.2 (val(CMSOS$_2$-$\mathbb{R}$) $\approx$ 0.6813 and val(MSOS$_2$-$\mathbb{R}$) $\approx$ 1.0000).

Proposition 3.33 seems to imply that the real moment/sum-of-squares hierarchy is better than the complex one. However, the size of the largest semidefinite constraint of CMSOS$_d$-$\mathbb{R}$, equal to $2(n+d)!/(n!d!)$, is far inferior to that of MSOS$_d$-$\mathbb{R}$, equal to $(2n+d)!/((2n)!d!)$. For instance, if $n=10$ and $d=3$, the former is 572 and the latter is 1,771.

Proposition 3.13. *Given $l\in\mathbb{N}$ and $\varphi\in\mathbb{C}_l[\bar{z},z]$, we have*

$$\forall z\in\mathbb{C}^n,\ \forall\theta\in\mathbb{R},\ \varphi(e^{\mathbf{i}\theta}z)=\varphi(z) \quad\Longleftrightarrow\quad \forall|\alpha|,|\beta|\leqslant l,\ |\alpha-\beta|\varphi_{\alpha,\beta}=0. \tag{3.35}$$

Proof. ($\Longrightarrow$) Notice that $\varphi(e^{\mathbf{i}\theta}z) = \sum_{|\alpha|,|\beta|\leqslant l}\varphi_{\alpha,\beta}\overline{(e^{\mathbf{i}\theta}z)}^{\alpha}(e^{\mathbf{i}\theta}z)^{\beta} = \dots$ $\sum_{|\alpha|,|\beta|\leqslant l}\varphi_{\alpha,\beta}e^{\mathbf{i}(|\alpha|-|\beta|)\theta}\bar{z}^{\alpha}z^{\beta}$. Polarization implies that for all $z,w\in\mathbb{C}^n$, we have $\sum_{|\alpha|,|\beta|\leqslant l}\varphi_{\alpha,\beta}e^{\mathbf{i}(|\alpha|-|\beta|)\theta}\bar{z}^{\alpha}w^{\beta} = \sum_{|\alpha|,|\beta|\leqslant l}\varphi_{\alpha,\beta}\bar{z}^{\alpha}w^{\beta}$ and hence for all $|\alpha|,|\beta|\leqslant l$, $\varphi_{\alpha,\beta}e^{\mathbf{i}(|\alpha|-|\beta|)\theta} = \varphi_{\alpha,\beta}$. If $\varphi_{\alpha,\beta}\neq 0$, then for all $\theta\in\mathbb{R}$, $|\alpha-\beta|\theta\equiv 0[2\pi]$ and thus $|\alpha-\beta|=0$. ($\Longleftarrow$) Simply compute $\varphi(e^{\mathbf{i}\theta}z)$. ∎

Definition 3.14. *Complex polynomial optimization problem* (3.3) *is said to be*

[7] We attempted a formal proof but it is difficult even on such a small example.

oscillatory if $f, g_1, \ldots,$ and g_m satisfy either of the two equivalent properties in (3.35)*.*

Proposition 3.15. *If complex polynomial optimization problem* (3.3) *is oscillatory, then $d^{\min}$-$\mathbb{R} = d^{\min}$-$\mathbb{C}$.*

Proof. Observe that $d^{\min}$-$\mathbb{C} = \max\{|\alpha|, |\beta| \mid f_{\alpha,\beta}\ g_{1,\alpha,\beta}\ \ldots\ g_{m,\alpha,\beta} \neq 0\}$ and $d^{\min}$-$\mathbb{R} = \max\{\lceil(|\alpha| + |\beta|)/2\rceil \mid f_{\alpha,\beta}\ g_{1,\alpha,\beta}\ \ldots\ g_{m,\alpha,\beta} \neq 0\}$ where $\lceil . \rceil$ denotes the ceiling of a real number. Both are equal if the problem is oscillatory. ∎

Conjecture 3.1. *Under the sphere assumption* (3.18)*, if complex polynomial optimization problem* (3.3) *is oscillatory, then for all $d \geqslant d^{\min}$-$\mathbb{R} = d^{\min}$-$\mathbb{C}$, we have*

$$\text{val}(\text{MSOS}_d\text{-}\mathbb{C}) = \text{val}(\text{CMSOS}_d\text{-}\mathbb{R}) = \text{val}(\text{MSOS}_d\text{-}\mathbb{R}). \tag{3.36}$$

In Section 4, we consider problems for which Conjecture 3.1 seems to hold numerically. This suggests that for oscillatory problems, the complex hierarchy is more tractable than the real hierarchy at no loss of bound quality.

3.5. Exploiting Sparsity in Real and Complex Hierarchies. The chordal sparsity technique described in Section 2.3 has been extended to the real hierarchy by Waki [91] and may readily be transposed to the complex hierarchy. Each positive semidefinite constraint in (3.12) is replaced by a set of positive semidefinite constraints on certain submatrices of $M_{d-k_i}(g_i y)$. These submatrices are defined by the maximal cliques of a chordal extension of the graph associated with the objective and constraint equations. Equivalently, the sum-of-squares variables σ_i in the dual formulation (3.14) are restricted to be functions of a subset (defined by the same maximal cliques) of the decision variables $z_1, \ldots, z_n$. These sparse relaxation hierarchies provide potentially lower bounds than their dense counterparts yet retain convergence guarantees [45]. However, further size reduction is often necessary. We propose to selectively apply computationally intensive higher-order constraints in the sparse relaxations. In other words, rather than a single relaxation order applied to all constraints, each constraint has an associated relaxation order. This allows for solving many large-scale problems.

We now formalize our approach applied to the complex hierarchy.[8] Objective function f and constraints $(g_i)_{1\leqslant i\leqslant m}$ in (3.3) have an associated undirected sparsity graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ with nodes $\mathcal{N} = \{1, \ldots, n\}$ corresponding to each variable and edges $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$ for each pair of variables that appear together in any monomial that has a non-zero coefficient in the objective function or constraints.

Each constraint function g_i has an associated relaxation order d_i so that $d \in \mathbb{N}^m$. When $d_i > 1$, there must exist at least one clique of a chordal extension of $\mathcal{G}$ (cf Section 2.3) that contains all variables with non-zero coefficients in g_i. To ensure this, define a supergraph $\hat{\mathcal{G}} = (\mathcal{N}, \hat{\mathcal{E}})$ where $\hat{\mathcal{E}}$ is composed of $\mathcal{E}$ augmented with edges connecting all variables with non-zero coefficients in g_i, not necessarily in the same monomial. For example, $g_1(z) = \bar{z}_1 z_2 + \bar{z}_2 z_1 + \bar{z}_1 z_3 + \bar{z}_3 z_1 + \bar{z}_1 z_4 + \bar{z}_4 z_1$ with $d_1 > 1$ implies $\mathcal{E} \supset \{(1,2), (1,3), (1,4)\}$ and $\hat{\mathcal{E}} \supset \{(1,2), (1,3), (1,4), (2,3), (2,4), (3,4)\}$.

To exploit sparsity, construct a chordal extension $\mathcal{G}^{\text{ch}}$ of $\hat{\mathcal{G}}$.[9] Denote the set of maximally sized cliques of the chordal extension by $\mathcal{C}_1, \ldots, \mathcal{C}_p \subset \mathcal{G}^{\text{ch}}$. By construction of $\hat{\mathcal{G}}$, each constraint function g_i for which $d_i > 1$ has all associated variables contained in at least one clique. For each g_i for which $d_i > 1$, denote as $\mathcal{C}^{(i)}$ the *minimal covering*

[8] See [58] for the details of this approach as applied to MSOS_d-$\mathbb{R}$ in the context of the optimal power flow problem.

[9] One approach to creating a chordal extension is to use the sparsity pattern of a Cholesky factorization (employing a minimum degree ordering to maintain sparsity) of the Laplacian matrix associated with $\hat{\mathcal{G}}$ plus an identity matrix.

clique (i.e., the smallest clique in $\{\mathcal{C}_1, \ldots, \mathcal{C}_p\}$ that contains all variables in g_i). (If not unique, a single clique $\mathcal{C}^{(i)}$ is chosen arbitrarily among the smallest cliques.) Associate an order $\tilde{d} \in \mathbb{N}^p$ with each clique $\mathcal{C}_\gamma$, $\gamma = 1, \ldots, p$ defined such that $\tilde{d}_\gamma$ is the maximum relaxation order d_i among all constraints for which the clique $\mathcal{C}_\gamma$ is the minimal covering clique. If a clique $\mathcal{C}_\gamma$ is not a minimal covering clique for any constraints, then $\tilde{d}_\gamma = 1$. See Appendix F for a small illustrative example.

For all $1 \leqslant i \leqslant m$ such that $d_i > 1$, the positive semidefinite constraints $M_{d-k_i}(g_i y) \succcurlyeq 0$ in the moment hierarchy (3.12) are replaced by $N_i(g_i y) \succcurlyeq 0$, $i = 1, \ldots, m$, where $N_i(g_i y)(\alpha, \beta) := M_{d_i - k_i}(g_i y)(\alpha, \beta)$ such that all non-zero entries of α and β correspond to variables in $\mathcal{C}^{(i)}$. For $i = 0$, the positive semidefinite constraint $M_d(y) \succcurlyeq 0$ (recall that $g_0 = 1$ and $k_0 = 0$) is replaced by constraints defined by each maximal clique: $\tilde{N}_\gamma(y) \succcurlyeq 0$, $\gamma = 1, \ldots, p$, where $\tilde{N}_\gamma(y)(\alpha, \beta) =: M_{\tilde{d}_\gamma}(y)(\alpha, \beta)$ such that all non-zero entries of α and β correspond to variables in $\mathcal{C}_\gamma$.

For the sum-of-squares representation of the hierarchy, the polynomials $\sigma_i \in \Sigma_{d-k_i}[z]$, $i = 1, \ldots, m$ in (3.14) are replaced by sums-of-squares polynomials $\omega_i \in \Sigma_{d_i - k_i}[z_{\mathcal{C}^{(i)}}]$, $i = 1, \ldots, m$, where $z_{\mathcal{C}^{(i)}}$ denotes the subset of variables z that are in the clique $\mathcal{C}^{(i)}$. The polynomial $\sigma_0 \in \Sigma_{d-k_0}[z]$ is replaced by the polynomial $\sum_{\gamma=1}^{p} \tau_\gamma$ where $\tau_\gamma \in \Sigma_{\tilde{d}_\gamma}[z_{\mathcal{C}_\gamma}]$, $\gamma = 1, \ldots, p$.

The sparse version of the real hierarchy MSOS$_d$-$\mathbb{R}$ converges to the global optimum of a polynomial optimization problem when the constraints include ball constraints on all decision variables x included in each clique: $\sum_{k \in \mathcal{C}_i} x_k^2 \leqslant (R_{\mathcal{C}_i})^2$, $i = 1, \ldots, p$, where $R_{\mathcal{C}_i}$ is the radius of a ball enclosing all decision variables in clique $\mathcal{C}_i$ [45]. A similar result holds for the complex hierarchy MSOS$_d$-$\mathbb{C}$ with sphere constraints enforced for the variables included in each clique. Due to (3.20), the sparse version of the complex hierarchy is guaranteed to converge to the global optimum of (3.3) with increasing relaxation order when the constraints include $(\sum_{k \in \mathcal{C}_i} |z_k|^2) + |z_{n+i}|^2 = (R_{\mathcal{C}_i})^2$, $i = 1, \ldots, p$, where z_{n+i} is a slack variable associated with clique $\mathcal{C}_i$.

Selectively applying the higher-order constraints requires a method for determining the relaxation order d_i for each constraint. We use a heuristic based on "mismatches" to the closest rank-one matrix [58]. Consider a complex polynomial optimization problem with a single global optimum. Both MSOS$_d$-$\mathbb{R}$ and MSOS$_d$-$\mathbb{C}$ yield the global optimum upon satisfaction of a rank condition (Proposition 3.6). Upon satisfaction of this condition with $S = 1$, a global optimum z^{opt} may be obtained using an eigendecomposition $z^{\text{opt}} = \sqrt{\lambda}\, \eta$ where η is the unit-length eigenvector associated with the non-zero eigenvalue λ of $(y_{\alpha,\beta})_{|\alpha|=|\beta|=1}$. This suggests that an "approximate" solution z^{approx} to the polynomial optimization problem can be obtained from the largest eigenvalue λ_1 and associated unit-length eigenvector η_1 when the rank condition is not satisfied: $z^{\text{approx}} = \sqrt{\lambda_1}\, \eta_1$. Define "mismatches" $\zeta \in \mathbb{R}$ and $\Delta \in \mathbb{R}^m$ between the solution y to the relaxation and z^{approx}:

$$\zeta := \left| f\left(z^{\text{approx}}\right) - L_y(f) \right|, \tag{3.37a}$$

$$\Delta_i := \left| g_i\left(z^{\text{approx}}\right) - L_y(g_i) \right|, \qquad i = 1, \ldots, m. \tag{3.37b}$$

We use the iteration in Algorithm 1 to determine relaxation orders d_i, $i = 1, \ldots, m$. Each iteration solves the moment/sum-of-squares relaxation after increasing the relaxation orders d_i in a manner that is dependent on the largest associated Δ_i values. Denote $d^{\max} := \max_i \{d_i\}$.[10] At each iteration of the algorithm, increment d_i at up to

[10] Note that $d^{\max}$ is not a specified maximum relaxation order but can change at each iteration.

h constraints, where h is a specified parameter, that have the largest mismatches Δ_i among constraints satisfying two conditions: (1) $d_i < d^{\max}$ and (2) $\Delta_i > \epsilon_g$, where ϵ_g is a specified mismatch tolerance. If no constraints satisfy both of these conditions, increment d_i at up to h constraints with the largest Δ_i greater than the specified tolerance and increment $d^{\max}$. That is, in order to prevent unnecessarily increasing the size of the matrices, the heuristic avoids incrementing the maximum relaxation order $d^{\max}$ until $d_i = d^{\max}$ at all constraints g_i with mismatch $\Delta_i > \epsilon_g$.

There is a computational trade-off in choosing the value of h. Larger values of h likely result in fewer iterations of the algorithm but each iteration is slower if more buses than necessary have high-order relaxations. Smaller values of h result in faster solution at each iteration, but may require more iterations.

The algorithm terminates upon satisfaction of two conditions: First, $|\Delta|_\infty \leqslant \epsilon_g$, where $|\cdot|_\infty$ denotes the infinity norm (maximum absolute value), which indicates that the iterate is a numerically feasible point of polynomial optimization problem (3.3). Second, $\zeta \leqslant \epsilon_f$, which indicates global optimality to within a relative tolerance ϵ_f. If the relaxation satisfies the former but not the latter termination condition (which was never observed in practice for the problem in Section 4), the algorithm increases d_i at the h constraints with largest mismatch Δ_i and continues iterating.

The moment/sum-of-squares hierarchy is successively tightened in a manner that preserves computational tractability. For sufficiently small tolerances ϵ_f and ϵ_g, Algorithm 1 eventually proceeds to build the complete moment/sum-of-squares hierarchies. Thus, Algorithm 1 inherits the theoretical convergence guarantees of MSOS$_d$-$\mathbb{C}$. The same can be said of the real version of Algorithm 1 applied to MSOS$_d$-$\mathbb{R}$.

Algorithm 1 Iterative Solution for Sparse Moment/Sum-of-Squares Relaxations

```
Set d_i = 1, i = 1, ..., m.
repeat
    Solve relaxation with order d_i for constraints g_i(z) ⩾ 0, i = 1, ..., m.
    Calculate mismatches Δ_i, i = 1, ..., m using (3.37b).
    Increase entries of d according to the mismatch heuristic.
until |Δ|_∞ < ε_g and ζ < ε_f
Extract solution z^opt.
```

4. Application to Electric Power Systems. The optimal power flow is a central problem in electric power systems introduced half a century ago by Carpentier [16]. While many non-linear methods [18,36,64,94] have been developed to solve this notoriously difficult problem, there is a strong motivation for producing more robust and reliable tools. First, power systems are growing in complexity due to the increase in the share of renewables, the increase in the peak load, and the expected wider use of demand response and storage. Second, new tools are needed to profit from high-performance computing and advanced telecommunications (phasor measurement units, dynamic line ratings, etc.). Finally, the ultimate goal is to solve large problems with combinatorial complexity due to phase-shifting transformers, high-voltage direct current, and special protection schemes. Solving the continuous case (i.e., optimal power flow) to global optimality would be of great benefit to that end.

The optimal power flow problem is an instance of complex polynomial optimization. Since 2006, the power systems literature has been studying the ability of the Shor and second-order conic relaxations to find global solutions [4,9,11,12,14,21,37,47,51,53,57,59–62,85,86,93]. Some relaxations are presented in real numbers [47,59] and some in complex numbers [11, 12, 93]. Nevertheless, in all numerical applications, standard solvers such as SeDuMi, SDPT3, and MOSEK are used which currently handle only real numbers. Modeling languages such as YALMIP and CVX do handle inputs in complex numbers, but the data is transformed into real numbers before calling the solver [13, Example 4.42].

We use the European network to illustrate the fact that it is beneficial to relax non-convex constraints before converting from complex to real numbers. The Shor relaxation, the second-order conic relaxation, and the moment/sum-of-squares hierarchy are considered. Network sparsity is exploited in each case.

4.1. Optimal Power Flow. A transmission network can be modeled using an undirected graph $\mathcal{G} = (\mathcal{B}, \mathcal{L})$ where buses $\mathcal{B} = \{1, \dots n\}$ are linked to one another via lines $\mathcal{L} \subset \mathcal{B} \times \mathcal{B}$. Power flows are governed by the admittance matrix $Y \in \mathbb{C}^{n \times n}$ whose extra diagonal terms $(l, m) \in \mathcal{L}$ are equal to $y_{lm}/(\rho_{ml}\rho_{lm}^H)$ and whose diagonal terms (l, l) are equal to $\sum_{(l,m)\in\mathcal{L}} (y_{lm} + y_{lm}^{\mathrm{gr}})/|\rho_{lm}|^2$. All others terms are equal to zero. Here, $y_{lm} \in \mathbb{C}$ denotes the mutual admittance between buses $(l, m) \in \mathcal{L}$, $y_{lm}^{\mathrm{gr}} \in \mathbb{C}$ denotes the admittance-to-ground at end l of line $(l, m) \in \mathcal{L}$, and $\rho_{lm} \in \mathbb{C}$ denotes the ratio of the ideal phase-shifting transformer at end l of line $(l, m) \in \mathcal{L}$.

Each bus injects power $p_k^{\mathrm{gen}} + \mathbf{i} q_k^{\mathrm{gen}}$ into the network with capacity limits $p_k^{\min}, p_k^{\max}$, $q_k^{\min}, q_k^{\max}$ (potentially all equal to 0) and extracts power demand $p_k^{\mathrm{dem}} + \mathbf{i} q_k^{\mathrm{dem}}$ from the network. Each bus operates at a voltage $v_k \in \mathbb{C}$. Finding power flows that minimize active power loss is a problem that can be cast as an instance of QCQP-$\mathbb{C}$:

$$\inf_{v\in\mathbb{C}^n} \; v^H \frac{Y^H + Y}{2} v, \tag{4.1}$$

subject to

$$\forall k \in \mathcal{B}, \quad p_k^{\min} - p_k^{\mathrm{dem}} \leqslant v^H H_k v \leqslant p_k^{\max} - p_k^{\mathrm{dem}}, \tag{4.2}$$

$$\forall k \in \mathcal{B}, \quad q_k^{\min} - q_k^{\mathrm{dem}} \leqslant v^H \tilde{H}_k v \leqslant q_k^{\max} - q_k^{\mathrm{dem}}, \tag{4.3}$$

$$\forall k \in \mathcal{B}, \quad (v_k^{\min})^2 \leqslant v^H e_k e_k^T v \leqslant (v_k^{\max})^2, \tag{4.4}$$

where $H_k = \frac{Y^H e_k e_k^T + e_k e_k^T Y}{2}$ and $\tilde{H}_k = \frac{Y^H e_k e_k^T - e_k e_k^T Y}{2\mathbf{i}}$ are Hermitian and e_k is the k^{th} column of the identity matrix. In Section 4.2, power flows are computed that seek to minimize either power loss or generation costs $\sum_{k\in\mathcal{B}} a_k (p_k^{\mathrm{gen}})^2 + b_k p_k^{\mathrm{gen}} + c_k$ where $a_k, b_k, c_k \in \mathbb{R}$, $a_k \geqslant 0$, and $p_k^{\mathrm{gen}} = v^H H_k v + p_k^{\mathrm{dem}}$. In the case of generation costs, new real variables $(t_k)_{k\in\mathcal{B}}$ are introduced, objective (4.1) is replaced by $\sum_{k\in\mathcal{B}} t_k$, and new constraints are added for all $k \in \mathcal{B}$:

$$a_k (v^H H_k v + p_k^{\mathrm{dem}})^2 + b_k (v^H H_k v + p_k^{\mathrm{dem}}) + c_k \leqslant t_k. \tag{4.5}$$

In Section 4.2 apparent power flow limits $|v_l i_{lm}^H| \leqslant s_{lm}^{\max}$ are enforced where $v_l i_{lm}^H = v^H F_{lm} v$ and $F_{lm} = a_{lm}^H e_l e_l^T + b_{lm}^H e_m e_l^T$, with $a_{lm} = (y_{lm} + y_{lm}^{\mathrm{gr}})/|\rho_{lm}|^2$ and $b_{lm} = -y_{lm}/(\rho_{ml}\rho_{lm}^H)$. These can be written for all $(l, m) \in \mathcal{L}$:

$$\left(v^H \frac{F_{lm} + F_{lm}^H}{2} v \right)^2 + \left(v^H \frac{F_{lm} - F_{lm}^H}{2\mathbf{i}} v \right)^2 \leqslant (s_{lm}^{\max})^2 \tag{4.6}$$

Note that (4.5)-(4.6) yield second-order conic constraints for all the relaxations considered in this paper as well as semidefinite constraints for higher orders of the moment/sum-of-squares hierarchies.

4.2. Numerical Results. We consider large test cases representing portions of European electric power systems. They represent Great Britain (GB) [89] and Poland (PL) [94] power systems as well as other European systems from the PEGASE project [31]. The test cases were preprocessed to remove low-impedance lines as described in [61] in order to improve the solver's numerical convergence, which is a typical procedure in power system analyses.[11] A 1×10^{-3} per unit low-impedance line threshold was used for all test cases except for PEGASE-1354 and PEGASE-2869 which use a 3×10^{-3} per unit threshold. The processed data is described in Table 1. This table also includes the at-least-locally-optimal objective values obtained from the interior point solver in MATPOWER [94] for the problems after preprocessing. Note that the PEGASE systems specify generation costs that minimize active power losses, so the objective values in both columns are the same.

TABLE 1
Size of Data (After Low-Impedance Line Preprocessing)

Test Case Name	**Number of Complex Variables**	**Number of Edges in Graph**	MATPOWER **Solution [94]** Gen. Cost ($/hr)	MATPOWER **Solution [94]** Loss Min. (MW)
GB-2224	2,053	2,581	1,942,260	60,614
PL-2383wp	2,177	2,651	1,868,350	24,911
PL-2736sp	2,182	2,675	1,307,859	18,336
PL-2737sop	2,183	2,675	777,617	11,397
PL-2746wop	2,189	2,708	1,208,257	19,212
PL-2746wp	2,192	2,686	1,631,737	25,269
PL-3012wp	2,292	2,805	2,592,462	27,646
PL-3120sp	2,314	2,835	2,142,720	21,513
PEGASE-89	70	185	5,819	5,819
PEGASE-1354	983	1,526	74,043	74,043
PEGASE-2869	2,120	3,487	133,945	133,945
PEGASE-9241	7,154	12,292	315,749	315,749
PEGASE-9241R[12]	7,154	12,292	315,785	315,785

Implementations use YALMIP 2015.06.26 [50], Mosek 7.1.0.28, and MATLAB 2013a on a computer with a quad-core 2.70 GHz processor and 16 GB of RAM. The results do not include the typically small formulation times.

4.2.1. Shor Relaxation. Table 2 shows the results of applying SDP-$\mathbb{R}$ and SDP-$\mathbb{C}$ to the test cases. For some problems, the Shor relaxation is *exact* and yields the globally optimal decision variables and objective values. To practically identify such problems, solutions for which all power injection mismatches $S_k^{\text{inj mis}}$ (see Section 4.2.3) are less than a tolerance of 1 MVA are considered exact. These problems are identified with an asterisk (*) in Table 2.

The lower bounds in Table 2 suggest that the corresponding MATPOWER solutions in Table 1 are at least very close to being globally optimal. The gap between the MATPOWER solutions and the lower bounds from SDP-$\mathbb{C}$ for the generation cost minimizing problems are less than 0.72% for GB-2224, 0.29% for the Polish systems,

[11] Low-impedance lines often model connections between buses in the same physical location.
[12] PEGASE-9241 contains negative resistances to account for generators at lower voltage levels. In PEGASE-9241R these are set to 0.

and 0.02% for the PEGASE systems with the exception of PEGASE-9241. The non-physical negative resistances in PEGASE-9241 result in weaker lower bounds from the relaxations, yielding a gap of 1.64% for this test case.

TABLE 2
Real and Complex SDP (Generation Cost Minimization)

Case Name	SDP-$\mathbb{R}$		SDP-$\mathbb{C}$	
	Val. ($/hr)	Time (sec)	Val. ($/hr)	Time (sec)
GB-2224	1,928,194	10.9	1,928,444	6.2
PL-2383wp	1,862,979	48.1	1,862,985	23.0
PL-2736sp*	1,307,749	35.7	1,307,764	22.0
PL-2737sop*	777,505	41.7	777,539	19.5
PL-2746wop*	1,208,168	51.1	1,208,182	22.8
PL-2746wp	1,631,589	43.8	1,631,655	20.0
PL-3012wp	2,588,249	52.8	2,588,259	24.3
PL-3120sp	2,140,568	64.4	2,140,605	25.5
PEGASE-89*	5,819	1.5	5,819	0.9
PEGASE-1354	74,035	11.2	74,035	5.6
PEGASE-2869	133,936	38.2	133,936	20.6
PEGASE-9241	310,658	369.7	310,662	136.1
PEGASE-9241R	315,848	317.2	315,731	95.9

As shown in Appendices C and D, the optimal objective values for SDP-$\mathbb{R}$ and SDP-$\mathbb{C}$ should be identical. With all objective values in Table 2 matching to within 0.037%, this is numerically validated.

For these test cases, SDP-$\mathbb{C}$ is significantly faster (between a factor of 1.60 and 3.31) than SDP-$\mathbb{R}$. This suggests that exploiting the matrix variable's structure in SDP-$\mathbb{C}$ is generally more computationally beneficial than eliminating the row and column corresponding to the angle reference in SDP-$\mathbb{R}$.

4.2.2. Second-Order Conic Relaxation. Table 3 shows the results of applying SOCP-$\mathbb{R}$ and SOCP-$\mathbb{C}$ to the test cases. Unlike the Shor relaxation, the second-order conic relaxation is not exact for any of the test cases. (SOCP-$\mathbb{C}$ is generally not exact with the exception of radial systems for which the relaxation is provably exact when certain non-trivial technical conditions are satisfied [51].)

SOCP-$\mathbb{C}$ provides better lower bounds and is computationally faster than SOCP-$\mathbb{R}$. Specifically, lower bounds from SOCP-$\mathbb{C}$ are between 0.87% and 3.96% larger and solver times are faster by between a factor of 1.24 and 6.76 than those from SOCP-$\mathbb{R}$.

4.2.3. Moment/Sum-of-Squares Hierarchy. Relaxations from the real moment/sum-of-squares hierarchy globally solve a broad class of optimal power flow problems [33,41,56,58]. Previous work uses MSOS$_d$-$\mathbb{R}$ by first converting the complex formulation of the optimal power flow problem to real numbers.

We next summarize computational aspects of both the real and complex hierarchies. The dense formulations of the hierarchies solve small problems (up to approximately ten buses). Without also selectively applying the higher-order relaxations' constraints (i.e., $d_1 = d_2 = \ldots = d_m = d$), exploiting network sparsity enables solution of the second-order relaxations for problems with up to approximately 40 buses.

Scaling to larger problems is accomplished by both exploiting network sparsity and selectively applying the computationally intensive higher-order relaxation con-

Table 3
Real and Complex SOCP (Generation Cost Minimization)

Case Name	**SOCP-$\mathbb{R}$**		**SOCP-$\mathbb{C}$**	
	Val. ($/hr)	Time (sec)	Val. ($/hr)	Time (sec)
GB-2224	1,855,393	3.5	1,925,723	1.4
PL-2383wp	1,776,726	8.5	1,849,906	2.4
PL-2736sp	1,278,926	4.8	1,303,958	1.7
PL-2737sop	765,184	5.5	775,672	1.6
PL-2746wop	1,180,352	5.1	1,203,821	1.7
PL-2746wp	1,586,226	5.5	1,626,418	1.7
PL-3012wp	2,499,097	5.9	2,571,422	2.0
PL-3120sp	2,080,418	6.2	2,131,258	2.2
PEGASE-89	5,744	0.5	5,810	0.4
PEGASE-1354	73,102	3.4	73,999	1.5
PEGASE-2869	132,520	9.0	133,869	2.7
PEGASE-9241	306,050	35.3	309,309	10.0
PEGASE-9241R	312,682	36.7	315,411	5.4

straints to specific "problematic" buses. To better match the structure of the optimal power flow constraint equations, we use the algorithm in [58], which is slightly different than that described in Section 3.5. Rather than consider each constraint individually, we use the mismatch in apparent power injections at each bus rather than the active and reactive power injection equations separately. The relaxation orders d_i associated with all constraints at a bus are changed together.

Specifically, mismatches for the active and reactive power injection constraints (4.2) and (4.3) at bus i, denoted as $P_i^{\text{inj mis}}$ and $Q_i^{\text{inj mis}}$, are calculated using (3.37b). Problematic buses are identified as those with large apparent power injection mismatch $S_i^{\text{inj mis}} = |P_i^{\text{inj mis}} + \mathbf{i}Q_i^{\text{inj mis}}|$. Application of the higher-order relaxation's constraints to these problematic buses using the iterative algorithm described in [58] (cf Section 3.5) results in global solutions to many optimal power flow problems and enables computational scaling to systems with thousands of buses [58, 61]. This section extends this approach to the complex hierarchy.

Tables 4 and 5 show the results of applying the algorithm from [58] for both the real and complex hierarchies to several test cases with tolerances $\epsilon_g = 1$ MVA and $\epsilon_f = 0.05\%$.[13] The optimal objective values in these tables match to at least 0.007%, which is within the expected solver tolerance. Further, the solutions for both the real and complex hierarchies match the optimal objective values for the loss minimizing problems obtained from Matpower shown in Table 1 to within 0.013%, providing an additional numerical proof that these solutions are globally optimal. Note, however, that local solvers do not always globally solve optimal power flow problems [14,17,58].

The test cases considered in Tables 4 and 5 minimize active power losses rather than generation costs. Although the moment/sum-of-squares hierarchy solves many small- and medium-size test cases which minimize generation cost, application of the algorithm in [58] to larger generation-cost-minimizing test cases often requires too

[13] The algorithm in [58] has a parameter h specifying the maximum number of buses to increase the relaxation order d_i at each iteration. This parameter is set to two for these examples. Additionally, bounds on the lifted variables y derived from the voltage magnitude limits (4.4) are enforced to improve numeric convergence.

many higher-order constraints for tractability. See [53, 60, 61] for related algorithms which often find feasible points that are nearly globally optimal for such problems.

The feasible set of the optimal power flow problem is included is the ball of radius $\sum_{k\in\mathcal{B}} (v_k^{\max})^2$ so a slack variable and a sphere constraint may be added as suggested in Section 3.3. In order to preserve sparsity, a slack variable and a sphere constraint may be added *for each maximal clique* of the chordal extension of the network graph. Global convergence is then guaranteed due to (3.20). However, the sphere constraint tends to introduce numerical convergence challenges in problems with several thousand buses, resulting in the need for higher-order constraints at more buses and correspondingly longer solver times.

Interestingly, the examples in Table 5 converged without the slack variables and sphere constraints, and the results therein correspond to relaxations without sphere constraints. A potential way to account for the success of the complex hierarchy without sphere constraints would be to compute the Hermitian complexity [29] of the ideal generated by the polynomials associated with equality constraints. A step in that direction would be to assess the greatest number of distinct points (possibly infinite) $v^i \in \mathbb{C}^n, 1 \leqslant i \leqslant p$, such that $(v^i)^H(H_k + \mathbf{i}\tilde{H}_k)v^j = -p_k^{\text{dem}} - \mathbf{i}q_k^{\text{dem}}$ for all buses k not connected to a generator and for all $1 \leqslant i, j \leqslant p$. Note that the Hermitian complexity of the ideal generated by $\sum_{i=1}^n |z_i|^2 + \sigma(z) + a$ as defined in (3.20) with $a < 0$ is equal to 1.

Despite being unnecessary for convergence of the hierarchies in Table 5, the sphere constraint can tighten the relaxations of some optimal power flow problems. Consider, for instance, the 9-bus example in [14]. The dense second-order relaxations from the real and complex hierarchies (both with and without the sphere constraint) yield the global optimum of \$3088/hour. Likewise, with second-order constraints enforced at all buses, the sparse versions of the real hierarchy and the complex hierarchy with the sphere constraint yield the global optimum. However, the sparse version of the second-order complex hierarchy without the sphere constraint only provided a lower bound of \$2939/hour. Thus, the sphere constraint tightens the sparse version of the second-order complex hierarchy for this test case. Since the sparse version of the third-order complex hierarchy without the sphere constraint yields the global optimum, the sphere constraint is unnecessary for convergence in this example.

TABLE 4
Real Moment/Sum-of-Squares Hierarchy $MSOS_d$-$\mathbb{R}$ (Active Power Loss Minimization)

Case Name	**Num. Iter.**	**Global Obj. Val. (MW)**	**Max S^{mis} (MVA)**	**Solver Time (sec)**
PL-2383wp	3	24,990	0.25	583.4
PL-2736sp	1	18,334	0.39	44.0
PL-2737sop	1	11,397	0.45	52.4
PL-2746wop	2	19,210	0.28	2,662.4
PL-2746wp	1	25,267	0.40	45.9
PL-3012wp	5	27,642	1.00	318.7
PL-3120sp	7	21,512	0.77	386.6
PEGASE-1354	5	74,043	0.85	406.9
PEGASE-2869	6	133,944	0.63	921.3

Similar to the second-order conic relaxation, the results in Tables 4 and 5 show that the complex hierarchy generally has computational advantages over the real

Table 5
Complex Moment/Sum-of-Squares Hierarchy $MSOS_d$-$\mathbb{C}$ (Active Power Loss Minimization)

Case Name	**Num. Iter.**	**Global Obj. Val. (MW)**	**Max S^{mis} (MVA)**	**Solver Time (sec)**
PL-2383wp	3	24,991	0.10	53.9
PL-2736sp	1	18,335	0.11	17.8
PL-2737sop	1	11,397	0.07	25.7
PL-2746wop	2	19,212	0.12	124.3
PL-2746wp	1	25,269	0.05	18.5
PL-3012wp	7	27,644	0.91	141.0
PL-3120sp	9	21,512	0.27	193.9
PEGASE-1354	11	74,042	1.00	1,132.6
PEGASE-2869	9	133,939	0.97	700.8

hierarchy. For all the test cases except PEGASE-1354, MSOS_d-$\mathbb{C}$ solves between a factor of 1.31 and 21.42 faster than MSOS_d-$\mathbb{R}$. The most significant computational speed improvements for the complex hierarchy over the real hierarchy are seen for cases (e.g., PL-2383wp and PL-2746wop) where the higher-order constraints account for a large portion of the solver times. The complex hierarchy for these cases has significantly fewer terms in the higher-order constraints than the real hierarchy.

Observe that several of the test cases (PL-3012wp, PL-3120sp, PEGASE-1354, and PEGASE-2869) require more iterations of the algorithm from [58] for MSOS_d-$\mathbb{C}$ than for MSOS_d-$\mathbb{R}$. Nevertheless, the improved speed per iteration results in faster overall solution times for all of these test cases except for PEGASE-1354, for which six additional iterations result in a factor of 2.78 slower solver time.

Both hierarchies were also applied to a variety of small test cases (less than ten buses) from [14, 48, 55, 62] for which the first-order relaxations failed to yield the global optima. For all these test cases, the dense versions of both MSOS_d-$\mathbb{C}$ and MSOS_d-$\mathbb{R}$ converged at the same relaxation order. Section 3.4 demonstrates that the MSOS_d-$\mathbb{R}$ is at least as tight as MSOS_d-$\mathbb{C}$. The results for small problems suggest that the hierarchies have the same tightness for some class of polynomial optimization problems which includes the optimal power flow problem with the sphere constraint (cf Conjecture 3.1). The numerical results for some large test cases have different numbers of iterations between the real and complex hierarchies. Rather than differences in the theoretical tightness of the relaxation hierarchies, we attribute this discrepancy in the number of iterations to numerical convergence inaccuracies; not enforcing the sphere constraint for the sparse complex hierarchy; and, in some cases, the algorithm from [58] selecting different buses at which to enforce the higher-order constraints.

5. Conclusion. We construct a complex moment/sum-of-squares hierarchy for complex polynomial optimization and prove convergence toward the global optimum. Theoretical and experimental evidence suggest that relaxing non-convex constraints before converting from complex to real numbers is better than doing the operations in the opposite order. We conclude with the question: is it possible to gain efficiency by transposing convex optimization algorithms from real to complex numbers?

Appendix A. Ring Homomorphism. Let I_p denote the identity matrix of order $p \in \mathbb{N}$. $\Lambda(I_n) = I_{2n}$ and if $Z_1, Z_2 \in \mathbb{C}^{n\times n}$, $\Lambda(Z_1 + Z_2) = \Lambda(Z_1) + \Lambda(Z_2)$ and

$$\begin{array}{rcl} \Lambda(Z_1)\Lambda(Z_2) & = & \begin{pmatrix} \mathrm{Re}Z_1 & -\mathrm{Im}Z_1 \\ \mathrm{Im}Z_1 & \mathrm{Re}Z_1 \end{pmatrix} \begin{pmatrix} \mathrm{Re}Z_2 & -\mathrm{Im}Z_2 \\ \mathrm{Im}Z_2 & \mathrm{Re}Z_2 \end{pmatrix} \\ & = & \begin{pmatrix} \mathrm{Re}Z_1\mathrm{Re}Z_2 - \mathrm{Im}Z_1\mathrm{Im}Z_2 & -\mathrm{Re}Z_1\mathrm{Im}Z_2 - \mathrm{Im}Z_1\mathrm{Re}Z_2 \\ \mathrm{Im}Z_1\mathrm{Re}Z_2 + \mathrm{Re}Z_1\mathrm{Im}Z_2 & \mathrm{Re}Z_1\mathrm{Re}Z_2 - \mathrm{Im}Z_1\mathrm{Im}Z_2 \end{pmatrix} \\ & = & \Lambda[\mathrm{Re}Z_1\mathrm{Re}Z_2 - \mathrm{Im}Z_1\mathrm{Im}Z_2 + \mathbf{i}(\mathrm{Im}Z_1\mathrm{Re}Z_2 + \mathrm{Re}Z_1\mathrm{Im}Z_2)] \\ & = & \Lambda[(\mathrm{Re}Z_1 + \mathbf{i}\mathrm{Im}Z_1)(\mathrm{Re}Z_2 + \mathbf{i}\mathrm{Im}Z_2)] \\ & = & \Lambda(Z_1 Z_2). \end{array}$$

Appendix B. Proof of Rank-2 Condition. It is proven here that a Hermitian matrix Z is positive semidefinite and has rank 1 if and only if $\Lambda(Z)$ is positive semidefinite and has rank 2.

($\Longrightarrow$) Say $Z = zz^H$ where real and imaginary parts are defined by $z = x_1 + \mathbf{i}x_2$ and $(x_1, x_2) \neq (0, 0)$. Then

$$\Lambda(Z) = \begin{pmatrix} x_1x_1^T + x_2x_2^T & x_1x_2^T - x_2x_1^T \\ x_2x_1^T - x_1x_2^T & x_1x_1^T + x_2x_2^T \end{pmatrix} \tag{B.1a}$$

$$= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T + \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}^T . \tag{B.1b}$$

The rank of $\Lambda(Z)$ is equal to 2 since $(\; x_1^T \; x_2^T \;)^T$ and $(\; (-x_2)^T \; x_1^T \;)^T$ are non-zero orthogonal vectors.

($\Longleftarrow$) Say $\Lambda(Z) = xx^T + yy^T$ where x and y are non-zero real vectors. Consider the block structure $x = (\; x_1^T \; x_2^T \;)^T$ and $y = (\; y_1^T \; y_2^T \;)^T$. For $i = 1, \ldots, n$, it must be that

$$x_{1i}^2 + y_{1i}^2 = x_{2i}^2 + y_{2i}^2, \tag{B.2a}$$

$$x_{1i}x_{2i} + y_{1i}y_{2i} = 0. \tag{B.2b}$$

Two cases can occur. The first is that $x_{1i}x_{2i} \neq 0$ in which case there exists a real number $\lambda_i \neq 0$ such that

$$\left\{ \begin{array}{rcl} y_{1i} & = & -\lambda_i \; x_{2i}, \\ y_{2i} & = & \frac{1}{\lambda_i} \; x_{1i}. \end{array} \right. \tag{B.3}$$

Equation (B.2a) implies that $(1 - \lambda_i^2)x_{1i}^2 = (1 - \frac{1}{\lambda_i^2})x_{2i}^2$ thus $(1 - \lambda_i^2)(1 - \frac{1}{\lambda_i^2}) \geqslant 0$ and $\lambda_i = \pm 1$. The second case is that $x_{1i}x_{2i} = 0$. Then, according to (B.2b), $y_{1i}y_{2i} = 0$. If either $x_{1i} = y_{1i} = 0$ or $x_{2i} = y_{2i} = 0$, then (B.2a) implies that $x_{1i} = x_{2i} = y_{1i} = y_{2i} = 0$. If $x_{1i} = y_{2i} = 0$, then (B.2a) implies that $y_{1i} = \pm x_{2i}$. If $x_{2i} = y_{1i} = 0$, then (B.2a) implies that $y_{2i} = \pm x_{1i}$.

In any case, there exists $\epsilon_i = \pm 1$ such that

$$\left\{ \begin{array}{rcl} y_{1i} & = & -\epsilon_i \; x_{2i}, \\ y_{2i} & = & \epsilon_i \; x_{1i}. \end{array} \right. \tag{B.4}$$

For $i, j = 1, \ldots, n$ it must be that

$$(1 - \epsilon_i\epsilon_j)(x_{1i}x_{1j} - x_{2i}x_{2j}) = 0, \tag{B.5a}$$

$$(1-\epsilon_i\epsilon_j)(x_{1j}x_{2i}+x_{1i}x_{2j})=0. \tag{B.5b}$$

Moreover

$$\begin{cases} x_{1i}x_{1j}+y_{1i}y_{1j} &= x_{1i}x_{1j}+\epsilon_i\epsilon_j x_{2i}x_{2j}, \\ x_{1i}x_{2j}+y_{1i}y_{2j} &= x_{1i}x_{2j}-\epsilon_i\epsilon_j x_{2i}x_{1j}. \end{cases} \tag{B.6}$$

It will now be shown that

$$\begin{cases} x_{1i}x_{1j}+y_{1i}y_{1j} &= x_{1i}x_{1j}+x_{2i}x_{2j}, \\ x_{1i}x_{2j}+y_{1i}y_{2j} &= x_{1i}x_{2j}-x_{2i}x_{1j}. \end{cases} \tag{B.7}$$

It is obvious if $\epsilon_i\epsilon_j=1$. If $\epsilon_i\epsilon_j=-1$, then (B.5a)–(B.5b) imply

$$x_{1i}x_{1j}-x_{2i}x_{2j}=0, \tag{B.8a}$$

$$x_{1j}x_{2i}+x_{1i}x_{2j}=0. \tag{B.8b}$$

If $x_{1i}x_{1j}x_{2i}x_{2j}=0$, it can be seen that (B.7) holds. If not, (B.8a) implies that there exists a real number $\mu_{ij}\neq 0$ such that

$$\begin{cases} x_{2i} &= \mu_{ij}\ x_{1i}, \\ x_{2j} &= \frac{1}{\mu_{ij}}\ x_{1j}. \end{cases} \tag{B.9}$$

Further, (B.8b) implies that $(\mu_{ij}+\frac{1}{\mu_{ij}})x_{1j}x_{2i}=0$. This is impossible ($\mu_{ij}+\frac{1}{\mu_{ij}}\neq 0$ and $x_{1j}x_{2i}\neq 0$). Thus, (B.7) holds.

With the left hand side corresponding to $\Lambda(Z)=xx^T+yy^T$ and the right hand side corresponding to (B.1b), equation (B.7) implies that $\Lambda(Z)$ is equal to (B.1b). Since the function Λ is injective, it must be that $Z=(x_1+\mathbf{i}x_2)(x_1+\mathbf{i}x_2)^H$.

Appendix C. Invariance of Shor Relaxation Bound. We have val(CSDP-$\mathbb{R}$) $\geqslant$ val(SDP-$\mathbb{R}$) since the feasible set is more tightly constrained due to (2.4d). To prove the opposite inequality, define $\tilde{\Lambda}(X):=(A+C)/2+\mathbf{i}(B-B^T)/2$ for all $X\in\mathbb{S}_{2n}$ using the block decomposition in the left hand part of (2.4d). It is proven here that if X is a feasible point of SDP-$\mathbb{R}$, then $\Lambda\circ\tilde{\Lambda}(X)$ is a feasible point of CSDP-$\mathbb{R}$ with same objective value as X. Firstly, $\Lambda\circ\tilde{\Lambda}(X)$ satisfies (2.4d) because $\tilde{\Lambda}(X)$ is a Hermitian matrix. Secondly, in order to show that $\Lambda\circ\tilde{\Lambda}(X)$ satisfies (2.4c), notice that if $x=(\ x_1^T\ x_2^T\ )^T$ then

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} C & -B \\ -B^T & A \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}^T \begin{pmatrix} A & B^T \\ B & C \end{pmatrix} \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}. \tag{C.1}$$

Hence $\Lambda\circ\tilde{\Lambda}(X)$ is equal to the sum of two positive semidefinite matrices. Finally, to prove that $\Lambda\circ\tilde{\Lambda}(X)$ satisfies (2.4b) and has same objective value as X, notice that if $H\in\mathbb{H}_n$ and $Y\in\mathbb{S}_{2n}$, then $\mathrm{Tr}\,[\Lambda(H)Y]=\sum_{1\leqslant i,j\leqslant 2n}\Lambda(H)_{ij}Y_{ji}=\sum_{1\leqslant i,j\leqslant 2n}\Lambda(H)_{ij}Y_{ij}=\sum_{1\leqslant i,j\leqslant n}\mathrm{Re}(H)_{ij}A_{ij}+\mathrm{Im}(H)_{ij}B_{ij}+(-\mathrm{Im}(H)_{ij})(B^T)_{ij}+\mathrm{Re}(H)_{ij}C_{ij}=\sum_{1\leqslant i,j\leqslant n}\mathrm{Re}(H_{ij})(A+C)_{ij}+\mathrm{Im}(H_{ij})(B-B^T)_{ij}=\dots$
$2\sum_{1\leqslant i,j\leqslant n}\mathrm{Re}[H_{ij}(\tilde{\Lambda}(Y)_{ij})^H]=2\sum_{1\leqslant i,j\leqslant n}H_{ij}(\tilde{\Lambda}(Y)_{ij})^H=2\mathrm{Tr}[H\tilde{\Lambda}(Y)]$. Completing the proof, for all $H\in\mathbb{H}_n$, $\mathrm{Tr}[\Lambda(H)\ \Lambda\circ\tilde{\Lambda}(X)]=2\mathrm{Tr}[H\tilde{\Lambda}(X)]=\mathrm{Tr}\,[\Lambda(H)X]$.

Appendix D. Invariance of SDP-$\mathbb{R}$ Relaxation Bound. We assume that X is a feasible point of SDP-$\mathbb{R}$ and construct a feasible point of SDP-$\mathbb{R}$ with same objective value and first diagonal entry equal to 0. Consider the eigenvalue decomposition

$X = \sum_{k=1}^{p} x_k x_k^T$ for some $x_k \in \mathbb{R}^{2n}$ and $p \in \mathbb{N}$. For all $\theta \in \mathbb{R}$, define

$$R_\theta := \Lambda[\cos(\theta) I_n + \mathbf{i}\sin(\theta) I_n] = \begin{pmatrix} \cos(\theta) I_n & -\sin(\theta) I_n \\ \sin(\theta) I_n & \cos(\theta) I_n \end{pmatrix}. \tag{D.1}$$

For $k = 1, \ldots, p$, define $\theta_k \in \mathbb{R}$ such that $x_{k,n+1} + \mathbf{i} x_{k,1} =: \sqrt{x_{k,n+1}^2 + x_{k1}^2}\, e^{\mathbf{i}\theta_k}$. Construct $\tilde{X} := \sum_{k=1}^{p} (R_{\theta_k} x_k)(R_{\theta_k} x_k)^T \succcurlyeq 0$ whose first diagonal entry is equal to 0. If $H \in \mathbb{H}_n$, $\mathrm{Tr}(\Lambda(H)\tilde{X}) = \sum_{k=1}^{p} \mathrm{Tr}[\Lambda(H) R_{\theta_k} x_k x_k^T R_{\theta_k}^T] = \sum_{k=1}^{p} \mathrm{Tr}[R_{\theta_k}^T \Lambda(H) R_{\theta_k} x_k x_k^T] = \sum_{k=1}^{p} \mathrm{Tr}[\Lambda\{(\cos(\theta_k) I_n - \mathbf{i}\sin(\theta_k) I_n) H (\cos(\theta_k) I_n + \mathbf{i}\sin(\theta_k) I_n)\} x_k x_k^T] = \ldots$
$\sum_{k=1}^{p} \mathrm{Tr}[\Lambda(H) x_k x_k^T] = \mathrm{Tr}(\Lambda(H) X)$.

Appendix E. Discrepancy Between Second-Order Conic Relaxation Bounds. We have val(CSOCP-$\mathbb{R}$) $\geqslant$ val(SOCP-$\mathbb{R}$) since the feasible set is more tightly constrained. Indeed, (2.8c)–(2.8e) imply (2.10c)–(2.10d). The opposite inequality between optimal values does not hold, and this can be proven by considering the example QCQP-$\mathbb{C}$ defined by $\inf_{z_1,z_2\in\mathbb{C}} (1+\mathbf{i})\bar{z}_1 z_2 + (1-\mathbf{i})\bar{z}_2 z_1$ s.t. $\bar{z}_1 z_1 \leqslant 1$, $\bar{z}_2 z_2 \leqslant 1$. CSOCP-$\mathbb{R}$ yields the globally optimal value of $-2\sqrt{2}$, while SOCP-$\mathbb{R}$ yields -4.

Appendix F. Five-Bus Illustrative Example for Exploiting Sparsity. To illustrate the selective application of second-order constraints, consider the five-bus optimal power flow problem in [14] which is an instance of QCQP-$\mathbb{C}$. Let ind$(\cdot)$ denote the set of indices corresponding to monomials of either the objective f or constraint functions $(g_i)_{1\leqslant 20}$. We have

$$\begin{aligned}
&\mathrm{ind}(f) = \{(1,1),(1,2),(1,3),(3,5),(4,5),(5,5)\}, \\
&\mathrm{ind}(g_1) = \mathrm{ind}(g_2) = \{(1,1),(1,2),(1,3)\} && \left[P_1^{\min}, Q_1^{\min}\right], \\
&\mathrm{ind}(g_3) = \mathrm{ind}(g_4) = \{(1,2),(2,2),(2,3),(2,4)\} && [P_2, Q_2], \\
&\mathrm{ind}(g_5) = \mathrm{ind}(g_6) = \{(1,3),(2,3),(3,3),(3,5)\} && [P_3, Q_3], \\
&\mathrm{ind}(g_7) = \mathrm{ind}(g_8) = \{(2,4),(4,4),(4,5)\} && [P_4, Q_4], \\
&\mathrm{ind}(g_9) = \mathrm{ind}(g_{10}) = \{(3,5),(4,5),(5,5)\} && \left[P_5^{\min}, Q_5^{\min}\right], \\
&\mathrm{ind}(g_{11}) = \mathrm{ind}(g_{12}) = \{(1,1)\} && \left[V_1^{\min}, V_1^{\max}\right], \\
&\mathrm{ind}(g_{13}) = \mathrm{ind}(g_{14}) = \{(2,2)\} && \left[V_2^{\min}, V_2^{\max}\right], \\
&\mathrm{ind}(g_{15}) = \mathrm{ind}(g_{16}) = \{(3,3)\} && \left[V_3^{\min}, V_3^{\max}\right], \\
&\mathrm{ind}(g_{17}) = \mathrm{ind}(g_{18}) = \{(4,4)\} && \left[V_4^{\min}, V_4^{\max}\right], \\
&\mathrm{ind}(g_{19}) = \mathrm{ind}(g_{20}) = \{(5,5)\} && \left[V_5^{\min}, V_5^{\max}\right],
\end{aligned} \tag{F.1}$$

where the text in brackets indicates the origin of the constraint: P_i and Q_i for active and reactive power injection equality constraints, $P_i^{\min}$ and $Q_i^{\min}$ for lower limits on active and reactive power injections, and $V_i^{\min}$ and $V_i^{\max}$ for squared voltage magnitude limits at bus i.

For brevity, the sphere constraints discussed in Section 3.3 are not enforced in this example. Regardless, the complex hierarchy with $d_i = 1$, $\forall i \in \{1,2,3,4,5,6,11,12,13,14,15,16\}$, $d_i = 2$, $\forall i \in \{7,8,9,10,17,18,19,20\}$ converges to the global solution. The second-order constraints are identified using the maximum power injection mismatch heuristic in [58].

The graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ corresponding to (F.1) is shown in Fig. 2. The nodes correspond to the complex variables $\mathcal{N} = \{1, \ldots, 5\}$. Edges $\mathcal{E}$, which are denoted by solid lines in Fig. 2, connect variables that appear in the same monomial in any of the constraint equations or objective function. The supergraph $\hat{\mathcal{G}} = \left(\mathcal{N}, \hat{\mathcal{E}}\right)$ has edges $\hat{\mathcal{E}}$ comprised of $\mathcal{E}$ (solid lines in Fig. 2) augmented with edges connecting all variables within each constraint with $d_i > 1$ (dashed lines in Fig. 2). In this case, $\hat{\mathcal{G}}$ is already chordal, so there is no need to form a chordal extension $\mathcal{G}^{\mathrm{ch}}$.

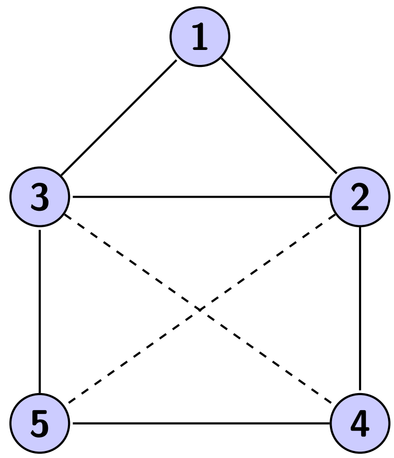

Fig. 2. *Graph Corresponding to Equations* (F.1) *from Five-Bus System in [14]*

The maximal cliques of $\hat{\mathcal{G}}$ are $\mathcal{C}_1 = \{1,2,3\}$ and $\mathcal{C}_2 = \{2,3,4,5\}$. Clique $\mathcal{C}_2$ is the minimal covering clique for all second-order constraints $g_i(z)$, $\forall i \in \{7,8,9,10,17,18,19,20\}$. The order associated with $\mathcal{C}_2$ is two ($\tilde{d}_2 = 2$) since the highest order d_i among all constraints for which $\mathcal{C}_2$ is the minimal covering clique is two. Clique $\mathcal{C}_1$ is not the minimal covering clique for any constraints with $d_i > 1$, so $\tilde{d}_1 = 1$.

The globally optimal objective value obtained from the complex hierarchy specified above is 946.8 with corresponding decision variables

$$z = \begin{pmatrix} 1.0467 + 0.0000\mathbf{i} \\ 0.9550 - 0.0578\mathbf{i} \\ 0.9485 - 0.0533\mathbf{i} \\ 0.7791 + 0.6011\mathbf{i} \\ 0.7362 + 0.7487\mathbf{i} \end{pmatrix}. \tag{F.2}$$

REFERENCES

[1] T. Aittomaki and V. Koivunen, *Beampattern Optimization by Minimization of Quartic Polynomial*, IEEE/SP 15th W. Stat. Signal Process., 51 (2009), pp. 437–440.

[2] N.I. Akhiezer, *The Classical Moment Problem and Some Related Questions in Analysis*, Hafner Publ. Co., New York, 1965.

[3] C.D. Aliprantis and K. Border, *Infinite Dimensional Analysis*, A Hitchhiker's guide, Second Edition, Springer-Verlag Berlin Heidelberg, 1999.

[4] M.S. Andersen, A. Hansson, and L. Vandenberghe, *Reduced-Complexity Semidefinite Relaxations of Optimal Power Flow Problems*, IEEE Trans. Power Syst., 29 (2014), pp. 1855–1863.

[5] E.J. Anderson and P. Nash, *Linear Programming in Infinite-Dimensional Spaces, Theory and Applications*, Wiley Int. Ser. Disc. Math. Optim., 1987.

[6] E. Artin, *Über die Zerlegung Definiter Funktionen in Quadrate*, Abhandlungen aus dem Math. Sem. der Univ. Hamburg, 5 (1927), pp. 100–115.

[7] A. Atzmon, *A Moment Problem for Positive Measures on the Unit Disc*, Pacific J. Math., 59 (1975), pp. 317–325.
[8] A. Aubry, A. De Maio, B. Jiang, and S. Zhang, *Ambiguity Function Shaping for Cognitive Radar via Complex Quartic Optimization*, IEEE Trans. Signal Process., 61 (2013), pp. 5603–5619.
[9] X. Bai, H. Wei, K. Fujisawa, and Y. Wang, *Semidefinite Programming for Optimal Power Flow Problems*, Int. J. Elec. Power, 30 (2008), pp. 383–392.
[10] A.S. Bandeira, N. Boumal, and A. Singer, *Tightness of the Maximum Likelihood Semidefinite Relaxation for Angular Synchronization*, Preprint, available at: `http://arxiv.org/abs/1411.3272`, (2014).
[11] S. Bose, D.F. Gayme, K.M. Chandy, and S.H. Low, *Quadratically Constrained Quadratic Programs on Acyclic Graphs with Application to Power*, IEEE Trans. Contr. Network Syst., (2015).
[12] S. Bose, S.H. Low, T. Teeraratkul, and B. Hassibi, *Equivalent Relaxations of Optimal Power Flow*, IEEE Trans. Automat. Control, (2014), p. 99.
[13] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2009.
[14] W.A Bukhsh, A. Grothey, K.I. McKinnon, and P.A. Trodden, *Local Solutions of the Optimal Power Flow Problem*, IEEE Trans. Power Syst., 28 (2013), pp. 4780–4788.
[15] E.J. Candès, Y. C. Eldar, T. Strohmer, and V. Voroninski, *Phase Retrieval for Imaging Problems*, SIAM J. Imaging Sci., 6 (2013), pp. 199–225.
[16] M.J. Carpentier, *Contribution à l'Étude du Dispatching Économique*, Bull. de la Soc. Fran. des Élec., 8 (1962), pp. 431–447.
[17] A. Castillo and R.P. O'Neill, *Computational Performance of Solution Techniques Applied to the ACOPF (OPF Paper 5)*, tech. report, US FERC, Jan. 2013.
[18] ———, *Survey of Approaches to Solving the ACOPF (OPF Paper 4)*, tech. report, US FERC, Mar. 2013.
[19] D.W. Catlin and J.P. DAngelo, *A Stabilization Theorem for Hermitian Forms and Applications to Holomorphic Mappings*, Math. Res. Lett., 3 (1996), pp. 149–166.
[20] C. Chen and P.P. Vaidyanathan, *MIMO Radar Waveform Optimization With Prior Information of the Extended Target and Clutter*, IEEE Trans. Signal Process., 57 (2009), pp. 3533–3544.
[21] C. Coffrin, H.L. Hijazi, and P. Van Hentenryck, *The QC Relaxation: Theoretical and Computational Results on Optimal Power Flow*, Preprint: `http://arxiv.org/abs/1502.07847`, (2015).
[22] R. Curto and L. Fialkow, *Solution of the Truncated Complex Moment Problem for Flat Data*, Memoirs Amer. Math. Soc., 568 (1996).
[23] ———, *The Quadratic Moment Problem for the Unit Circle and Unit Disk*, Integral Equations Operator Theory, 38 (2000), pp. 377–409.
[24] ———, *The Truncated Complex K-Moment Problem*, Trans. Amer. Math. Soc., 353 (2000), pp. 2825–2855.
[25] ———, *Truncated K-Moment Problems in Several Variables*, J. Operator Theory, 54 (2005), pp. 189–226.
[26] J.P. D'Angelo, *Inequalities from Complex Analysis*, Carus Math. Monogr., MAA, 2002.
[27] ———, *Hermitian Analogues of Hilbert's 17th Problem*, Adv. Math., 226 (2011), pp. 4607–4637.
[28] J.P. D'Angelo and M. Putinar, *Polynomial Optimization on Odd-Dimensional Spheres*, in Emerging Applications of Algebraic Geometry, Springer New York, 2008.
[29] ———, *Hermitian Complexity of Real Polynomial Ideals*, Int. J. Math., 23 (2012).
[30] K. Roos E. de Klerk, T. Terlaky, *Self-dual embeddings*, in Handbook of Semidefinite Programming – Theory, Algorithms, and Applications, H. Wolkowicz, R. Saigal, and L. Vandenberghe, eds., Kluwer Acad. Publ., Boston, 2000.
[31] S. Fliscounakis, P. Panciatici, F. Capitanescu, and L. Wehenkel, *Contingency Ranking with Respect to Overloads in Very Large Power Systems Taking into Account Uncertainty, Preventive and Corrective Actions*, IEEE Trans. Power Syst., 28 (2013), pp. 4909–4917.
[32] F. Fogel, I. Waldspurger, and A. d'Aspremont, *Phase Retrieval for Imaging Problems*, Preprint, available at: `http://arxiv.org/abs/1304.7735`, (2014).
[33] B. Ghaddar, J. Marecek, and M. Mevissen, *Optimal Power Flow as a Polynomial Optimization Problem*, To appear in *IEEE Trans. Power Syst.*
[34] R. Grone, C.R. Johnson, E.M. Sá, and H. Wolkowicz, *Positive Definite Completions of Partial Hermitan Matrices*, Linear Algebra Appl., 58 (1984), pp. 109–124.
[35] J.J. Hilling and A. Sudbery, *The Geometric Measure of Multipartite Entanglement and the Singular Values of a Hypermatrix*, J. Math. Phys., 51 (2010).
[36] M. Huneault and F. Galiana, *A Survey of the Optimal Power Flow Literature*, IEEE Trans.

Power Syst., 6 (1991), pp. 762–770.
[37] R.A. Jabr, *Radial Distribution Load Flow using Conic Programming*, IEEE Trans. Power Syst., 21 (2006), pp. 1458–1459.
[38] B. Jiang, Z. Li, and S. Zhang, *Approximation Methods for Complex Polynomial Optimization*, Springer Comput. Optim. Appl., 59 (2014), pp. 219–248.
[39] ———, *Characterizing Real-Valued Multivariate Complex Polynomials and Their Symmetric Tensor Representations*, Preprint, available at: `http://arxiv.org/abs/1501.01058`, (2015).
[40] C. Josz and D. Henrion, *Strong Duality in Lasserre's Hierarchy for Polynomial Optimization*, Springer Optim. Lett., (2015).
[41] C. Josz, J. Maeght, P. Panciatici, and J.C. Gilbert, *Application of the Moment-SOS Approach to Global Optimization of the OPF Problem*, IEEE Trans. Power Syst., 30 (2015), pp. 463–470.
[42] J.L. Krivine, *Anneaux préordonnés*, J. d'Anal. Math., 12 (1964), pp. 307–326.
[43] J.B. Lasserre, *Optimisation Globale et Théorie des Moments*, C. R. Acad. Sci. Paris, Série I, 331 (2000), pp. 929–934.
[44] ———, *Global Optimization with Polynomials and the Problem of Moments*, SIAM J. Optim., 11 (2001), pp. 796–817.
[45] ———, *Moments, Positive Polynomials and Their Applications*, no. 1 in Imperial College Press Optimization Series, Imperial College Press, 2010.
[46] M. Laurent, *A Comparison of Sherali-Adams, Lovasz-Schrijver, and Lasserre Relaxations for 0-1 Programming*, Math. Oper. Res., 28 (2003), pp. 470–496.
[47] J. Lavaei and S.H. Low, *Zero Duality Gap in Optimal Power Flow Problem*, IEEE Trans. Power Syst., 27 (2012), pp. 92–107.
[48] B.C. Lesieutre, D.K. Molzahn, A.R. Borden, and C.L. DeMarco, *Examining the Limits of the Application of Semidefinite Programming to Power Flow Problems*, in 49th Annu. Allerton Conf. Commun., Control, Comput., 2011, pp. 28–30.
[49] Z. Li, S. He, and S. Zhang, *Approximation Methods for Polynomial Optimization: Models, Algorithms, and Applications*, Comput. Optim. Appl., Springer, New York, 2012.
[50] J. Löfberg, *YALMIP: A Toolbox for Modeling and Optimization in MATLAB*, in IEEE Int. Symp. Comput. Aided Contr. Syst. Des., 2004, pp. 284–289.
[51] S.H. Low, *Convex Relaxation of Optimal Power Flow: Parts I & II*, IEEE Trans. Control Network Syst., 1 (2014), pp. 15–27.
[52] Z. Luo, W.-K. Ma, A.M.-C. So, Y. Ye, and S. Zhang, *Semidefinite Relaxation of Quadratic Optimization Problems*, IEEE Signal Process. Mag., 27 (2010), pp. 20–34.
[53] R. Madani, M. Ashraphijuo, and J. Lavaei, *Promises of Conic Relaxation for Contingency-Constrained Optimal Power Flow Problem*, in 52nd Annu. Allerton Conf. Commun., Control, Comput., Sept. 2014, pp. 1064–1071.
[54] B. Maricic, Z.-Q. Luo, and T.N. Davidson, *Blind Constant Modulus Equalization via Convex Optimization*, IEEE Trans. Signal Process., 51 (2003), pp. 805–818.
[55] D.K. Molzahn, S.S. Baghsorkhi, and I.A. Hiskens, *Semidefinite Relaxations of Equivalent Optimal Power Flow Problems: An Illustrative Example*, in IEEE Int. Symp. Circ. Syst. (ISCAS), May 24-27 2015.
[56] D.K. Molzahn and I.A. Hiskens, *Moment-Based Relaxation of the Optimal Power Flow Problem*, 18th Power Syst. Comput. Conf. (PSCC), (2014).
[57] ———, *Mixed SDP/SOCP Moment Relaxations of the Optimal Power Flow Problem*, in IEEE Eindhoven PowerTech, 29 June–2 July 2015.
[58] ———, *Sparsity-Exploiting Moment-Based Relaxations of the Optimal Power Flow Problem*, IEEE Trans. Power Syst., 30 (2015), pp. 3168–3180.
[59] D.K. Molzahn, J.T. Holzer, B.C. Lesieutre, and C.L. DeMarco, *Implementation of a Large-Scale Optimal Power Flow Solver Based on Semidefinite Programming*, IEEE Trans. Power Syst., 28 (2013), pp. 3987–3998.
[60] D.K. Molzahn, C. Josz, I.A. Hiskens, and P. Panciatici, *A Laplacian-Based Approach for Finding Near Globally Optimal Solutions to OPF Problems*, Submitted. Preprint available: `http://arxiv.org/abs/1507.07212`.
[61] D.K. Molzahn, C. Josz, I.A. Hiskens, and P. Panciatici, *Solution of Optimal Power Flow Problems using Moment Relaxations Augmented with Objective Function Penalization*, To appear in IEEE 54th Ann. Conf. Decis. Contr. (CDC), (2015).
[62] D.K. Molzahn, B.C. Lesieutre, and C.L. DeMarco, *Investigation of Non-Zero Duality Gap Solutions to a Semidefinite Relaxation of the Power Flow Equations*, in 47th Hawaii Int. Conf. Syst. Sci. (HICSS), 6-9 Jan. 2014.
[63] J. Nie, *Optimality Conditions and Finite Convergence of Lasserre's Hierarchy*, Math. Pro-

gram., 146 (2014), pp. 97–121.
[64] K.S. Pandya and S.K. Joshi, *A Survey of Optimal Power Flow Methods*, J. Theor. Appl. Inf. Tech., 4 (2008), pp. 450–458.
[65] P.A. Parrilo, *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*, PhD thesis, Cal. Inst. of Tech., May 2000.
[66] ———, *Semidefinite Programming Relaxations for Semialgebraic Problems*, Math. Program., 96 (2003), pp. 293–320.
[67] A. Prestel and C.N. Delzell, *Positive Polynomials*, Springer Monogr. Math., 2001.
[68] M. Putinar, *Positive Polynomials on Compact Semi-Algebraic Sets*, Indiana Univ. Math. J., 42 (1993), pp. 969–984.
[69] ———, *A Two-Dimensional Moment Problem*, J. Funct. Anal., 80 (1998), pp. 1–8.
[70] ———, *On Hermitian Polynomial Optimization*, Arch. Math., 87 (2006), pp. 41–51.
[71] M. Putinar and C. Scheiderer, *Quillen Property of Real Algebraic Varieties*, To appear in Muenster J. Math.
[72] ———, *Hermitian Algebra on the Ellipse*, Illinois J. Math., 56 (2012), pp. 213–220.
[73] D.G. Quillen, *On the Representation of Hermitian Forms as Sums of Squares*, Invent. Math., 5 (1968), pp. 237–242.
[74] W. Rudin, *Real and Complex Analysis*, Math. Ser., Third Edition, McGraw Hill Int. Ed., 1987.
[75] C. Scheiderer, *Positivity and Sums of Squares: A Guide to Recent Results*, vol. 149 of IMA Vol. Math. Appl., Springer, New York, 2009.
[76] K. Schmüdgen, *The K-Moment Problem for Semi-Algebraic Sets*, Math. Ann., 289 (1991), pp. 203–206.
[77] ———, *Around Hilbert's 17th Problem*, Doc. Math., Extra Vol. ISMP, (2012), pp. 433–438.
[78] M. Schweighöfer, *Optimization of Polynomials on Compact Semialgebraic Sets*, SIAM J. Optim., 15 (2005), pp. 805–825.
[79] N.Z. Shor, *Quadratic Optimization Problems*, Sov. J. Comput. Syst. Sci., 25 (1987), pp. 1–11.
[80] A. Singer, *Angular Synchronization by Eigenvectors and Semidefinite Programming*, Appl. Comput. Harmon. Anal., 30 (2011), pp. 20–36.
[81] L. Sorber, M.V. Barel, and L. De Lathauwer, *Unconstrained Optimization of Real Functions in Complex Variables*, SIAM J. Optim., 22 (2012), pp. 879–898.
[82] G. Stengle, *A Nullstellensatz and a Positivstellensatz in Semialgebraic Geometry*, Math. Ann., 207 (1974), pp. 87–97.
[83] J. Stochel, *Solving the Truncated Moment Problem Solves the Full Moment Problem*, Glasg. Math. J., 43 (2001), pp. 335–341.
[84] R.E. Tarjan and M. Yannakakis, *Simple Linear-Time Algorithms to Test Chordality of Graphs, Test Acyclicity of Hypergraphs, and Selectively Reduce Acyclic Hypergraphs*, SIAM J. Comput., 13 (1984), p. 566.
[85] J.A. Taylor, *Convex Optimization of Power Systems*, Cambridge University Press, 2015.
[86] J.A. Taylor and F.S. Hover, *Conic AC Transmission System Planning*, IEEE Trans. Power Syst., 28 (2013), pp. 952–959.
[87] O. Toker and H. Ozbay, *On the Complexity of Purely Complex Mu Computation and Related Problems in Multidimensional Systems*, IEEE Trans. Automat. Control, 43 (1998), pp. 409–414.
[88] M. Trnovská, *Strong Duality Conditions in Semidefinite Programming*, J. Electr. Eng., 56 (2005), pp. 1–5.
[89] University of Edinburgh Power Systems Test Case Archive, *GB Network*.
[90] F.H. Vascilescu, *Spectral Measures and Moment Problems*, Spectral Theory and Its Applications, (2003), pp. 173–215.
[91] H. Waki, S. Kim, M. Kojima, and M. Muramatsu, *Sums of Squares and Semidefinite Program Relaxations for Polynomial Optimization Problems with Structured Sparsity*, SIAM J. Optim., 17 (2006), pp. 218–242.
[92] W. Wirtinger, *Zur Formalen Theorie der Funktionen von Mehr Komplexen Veränderlichen*, Math. Ann., 97 (1927), pp. 357–375.
[93] B. Zhang and D. Tse, *Geometry of Feasible Injection Region of Power Networks*, IEEE Trans. Power Syst., 28 (2013), pp. 788–797.
[94] R. Zimmerman, C. Murillo-Sánchez, and R. Thomas, *MATPOWER: Steady-State Operations, Planning, and Analysis Tools for Power Systems Research and Education*, IEEE Trans. Power Syst., 99 (2011), pp. 1–8.