SIAM J. OPTIM.
Vol. 28, No. 2, pp. 1017–1048

LASSERRE HIERARCHY FOR LARGE SCALE POLYNOMIAL OPTIMIZATION IN REAL AND COMPLEX VARIABLES*

CÉDRIC JOSZ† AND DANIEL K. MOLZAHN‡

Abstract. We propose general notions to deal with large scale polynomial optimization problems and demonstrate their efficiency on a key industrial problem of the 21st century, namely the optimal power flow problem. These notions enable us to find global minimizers on instances with up to 4,500 variables and 14,500 constraints. First, we generalize the Lasserre hierarchy from real to complex numbers in order to enhance its tractability when dealing with complex polynomial optimization. Complex numbers are typically used to represent oscillatory phenomena, which are omnipresent in physical systems. Using the notion of hyponormality in operator theory, we provide a finite convergence criterion which generalizes the Curto–Fialkow conditions of the real Lasserre hierarchy. Second, we introduce the multi-ordered Lasserre hierarchy in order to exploit sparsity in polynomial optimization problems (in real or complex variables) while preserving global convergence. It is based on two ideas: (1) to use a different relaxation order for each constraint, and (2) to iteratively seek a closest measure to the truncated moment data until a measure matches the truncated data. Third and last, we exhibit a block diagonal structure of the Lasserre hierarchy in the presence of commonly encountered symmetries. To the best of our knowledge, the Lasserre hierarchy was previously limited to small scale problems, while we solve a large scale industrial problem with thousands of variables and constraints to global optimality.

Key words. multi-ordered Lasserre hierarchy, Hermitian sum-of-squares, chordal sparsity, semidefinite programming, optimal power flow

AMS subject classifications. 90C22, 90C06, 90C26, 28A99, 14Q99, 47N10

DOI. 10.1137/15M1034386

1. Introduction.

Polynomial optimization encompasses NP-hard nonconvex problems that arise in various applications, and it includes, as special cases, integer programming and quadratically constrained quadratic programming. The Lasserre hierarchy [61, 79, 80], which draws on algebraic geometry [83], enables one to solve such problems to global optimality using semidefinite programming. A big challenge today is to make it applicable to large scale real-world problems. Recent approaches in this direction include the use of chordal sparsity [100], the BSOS hierarchy [59] and sparse-BSOS hierarchy [101], the DSOS and SDSOS hierarchies [3, 51, 57], and ADMM for sum-of-squares [102]. The Lasserre hierarchy has two dual facets, moments and sums-of-squares, and most approaches to reduce the computational burden can be viewed as a restriction on the sum-of-squares: [100] restricts the number of variables, [59] restricts the degree, and [3] restricts the number of terms inside the square. Following this line of research, we propose to restrict sum-of-squares to *Hermitian* sum-of-squares [34] for optimization problems with oscillatory phenomena (e.g., power

*Received by the editors August 10, 2015; accepted for publication (in revised form) January 12, 2018; published electronically April 5, 2018.
http://www.siam.org/journals/siopt/28-2/M103438.html
Funding: The first author's research was funded by the CIFRE ANRT contract 2013/0179 and by the ERC under the European Union's Horizon 2020 research and innovation program (grant agreement 666981 TAMING). The second author received support from Dow Sustainability Fellowship, ARPA-E grant DE-AR0000232 and Los Alamos National Laboratory subcontract 270958.
†INRIA, Paris-Rocquencourt, BP 105, F-78153 Le Chesnay, France, and RTE 9, rue de la Porte de Buc, BP 561, F-78000 Versailles, France, and University of Paris VI, 4 Place Jussieu, 75005 Paris, France (cedric.josz@gmail.com).
‡Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 (molzahn@umich.edu).

systems [11, 17, 67], imaging science [13, 16, 93], signal processing [4, 21, 71, 72], automatic control [98], and quantum mechanics [45]). In addition, we propose restraining the use of high degree sum-of-squares to only some constraints by using a different degree for each constraint. Finally, we show that if the polynomials defining the objective and the constraints are even (i.e., all the monomials have an even degree), then we can restrict the sum-of-squares to be even at no loss of bound quality. We show that a similar result holds for Hermitian sum-of-squares. The relevance of the restrictions to sum-of-squares that we propose is demonstrated on the optimal power flow problem in electrical engineering.

Optimal power flow is a central problem in power systems introduced half a century ago in [17]. It tries to find a steady state operating point of an alternating current transmission network that respects Kirchhoff's laws, Ohm's law, and power balance equations. In addition, the point has to be optimal under a criterion such as total power generation or generation costs. It must also satisfy operational constraints which include narrow voltage ranges around nominal values and line ratings to keep Joule heating to acceptable levels. While many nonlinear methods [19, 104] have been developed to solve this difficult problem, there is a strong motivation for producing more 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 (e.g., 10,000 buses in the synchronous grid of continental Europe) 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. Since 2006, semidefinite and second-order conic relaxations have been proposed [7, 24, 47, 70, 76, 96]. It has emerged that the only approach that systematically yields global minimizers is the Lasserre hierarchy [38, 54, 75], although so far only for medium sized problems [77]. We solve large scale instances within minutes thanks to the restrictions of sum-of-squares discussed above.

This paper is organized as follows. Section 2 generalizes the Lasserre hierarchy to complex numbers to deal with complex polynomial optimization. Asymptotic convergence is discussed in section 3, while finite convergence is studied in sections 4 and 5. Sparsity is exploited in real and complex numbers via the multi-ordered Lasserre hierarchy in section 6, and symmetry is exploited via the block diagonal Lasserre hierarchy in section 7.

2. Complex Lasserre hierarchy.

Consider the problem of finding global solutions to a complex polynomial optimization problem

$$\begin{aligned} &\inf_{z\in\mathbb{C}^n} && f(z,\bar z) := \sum_{\alpha,\beta} f_{\alpha,\beta} z^\alpha \bar z^\beta \\ &\text{s.t.} && g_i(z,\bar z) := \sum_{\alpha,\beta} g_{i,\alpha,\beta} z^\alpha \bar z^\beta \geqslant 0, \quad i=1,\dots,m. \end{aligned} \tag{2.1}$$

We use the multi-index notation $z^\alpha := z_1^{\alpha_1}\cdots z_n^{\alpha_n}$ for $z\in\mathbb{C}^n$, $\alpha\in\mathbb{N}^n$, and $\bar z$ stands for the conjugate of z. As usual, $\mathbb{C}$ denotes the set of complex numbers (with $\mathbf{i}$ the imaginary number), and $\mathbb{R}$ denotes the set of real numbers. The functions $f, g_1, \dots, g_m$ are real-valued polynomials so that in the above sums only a finite number of coefficients $f_{\alpha,\beta}$ and $g_{i,\alpha,\beta}$ are nonzero, and they satisfy $\overline{f_{\alpha,\beta}} = f_{\beta,\alpha}$ and $\overline{g_{i,\alpha,\beta}} = g_{i,\beta,\alpha}$. The feasible set is defined as $K := \{z\in\mathbb{C}^n \ : \ g_i(z,\bar z)\geqslant 0,\ i=1,\dots,m\}$.

Example 2.1. The optimal power flow problem is a complex polynomial optimization problem. It reads as follows:

$$\inf_{z\in\mathbb{C}^n} \quad \sum_{i=1}^{n} C_{i2}\left(\sum_{j=1}^{n}\frac{\overline{Y}_{ij}}{2}z_i\overline{z}_j+\frac{Y_{ij}}{2}z_j\overline{z}_i\right)^2 + C_{i1}\left(\sum_{j=1}^{n}\frac{\overline{Y}_{ij}}{2}z_i\overline{z}_j+\frac{Y_{ij}}{2}z_j\overline{z}_i\right) + C_{i0}$$

$$\text{s.t.} \quad \begin{cases} (V_i^{\min})^2 \leqslant |z_i|^2 \leqslant (V_i^{\max})^2, \quad i=1,\ldots,n, \\ P_i^{\min}-P_i^{\text{dem}} \leqslant \sum_{j=1}^{n}\frac{\overline{Y}_{ij}}{2}z_i\overline{z}_j+\frac{Y_{ij}}{2}z_j\overline{z}_i \leqslant P_i^{\max}-P_i^{\text{dem}}, \quad i=1,\ldots,n, \\ Q_i^{\min}-Q_i^{\text{dem}} \leqslant \sum_{j=1}^{n}\frac{\overline{Y}_{ij}}{2\mathbf{i}}z_i\overline{z}_j-\frac{Y_{ij}}{2\mathbf{i}}z_j\overline{z}_i \leqslant Q_i^{\max}-Q_i^{\text{dem}}, \quad i=1,\ldots,n, \\ |B_{ij}z_i\overline{z}_i+Y_{ij}z_j\overline{z}_i|^2 \leqslant (S_{ij}^{\max})^2, \quad i,j=1,\ldots,n, \text{ when } Y_{ij}\neq 0, \end{cases}$$

where all symbols in capital letters are physical constants. Figure 1 illustrates a global solution on an instance with 14 complex variables [1, IEEE 14 Bus]. Each variable corresponds to a node in the graph and represents the voltage at that node. Once the voltages are computed, one can deduce the power production at each generator and the power flows on the edges. In order to supply 261 MW of power to consumers (in red), the least expensive generation plan entails a total power production of 268 MW (in black). (Color available online.) The global solution was computed with the first-order relaxation of the real Lasserre hierarchy (after converting the problem to real numbers). This is in accordance with the seminal work of Lavaei and Low [67], who showed that the first-order relaxation solves many instances of the optimal power flow problem. It was later shown that there are also many instances that need higher-order relaxations [68]. Such instances can be found in Table 1.

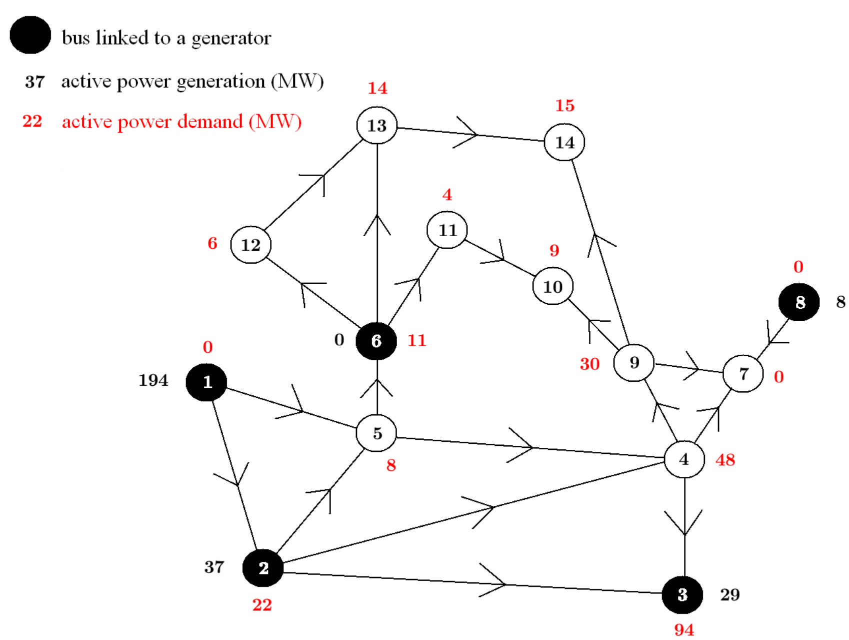

FIG. 1. *Globally optimal power flow in the midwestern United States.*

In order to solve complex polynomial optimization problems, approximative and iterative approaches have recently been proposed [48, 49, 94]. We pursue a different direction and follow the point of view of Lasserre [60, 61] in real numbers, that is, the reformulation

$$\inf_{\mu \in \mathcal{M}_+(K)} \int_K f d\mu \quad \text{subject to} \quad \int_K d\mu = 1, \tag{2.2}$$

where $\mathcal{M}_+(K)$ denotes the set of finite positive Borel measures on K. Lasserre observes that if the objective and constraints are real polynomials, then one may invoke the real moments $\int_K x^\alpha d\mu,\ \alpha \in \mathbb{N}^n$, of the measure μ. We remark that with complex polynomials, this leads instead to the complex moments of the measure μ, that is,

$$\int_K z^\alpha \bar{z}^\beta d\mu, \quad \forall \alpha, \beta \in \mathbb{N}^n. \tag{2.3}$$

Complex moments, like real moments, characterize the measure when K is compact, thanks to the Stone–Weierstrass theorem. Note that when K is compact, Borel measures are referred to as Radon measures and identify with the topological dual of the continuous functions from K to $\mathbb{R}$ equipped with the operator norm. This is due to the Riesz representation theorem (see standard textbooks, e.g., [91]).

In order to define the original Lasserre hierarchy, the sequence of moments is truncated $\left\{\int_K x^\alpha d\mu,\ |\alpha| \leqslant 2d\right\}$, where d is the truncation order and $|\alpha| := \sum_{k=1}^n \alpha_k$. In order to define the complex Lasserre hierarchy, we suggest truncating as follows: $\left\{\int_K z^\alpha \bar{z}^\beta d\mu,\ |\alpha|, |\beta| \leqslant d\right\}$. This naturally leads to a moment/sum-of-squares hierarchy in complex numbers:

$$\boxed{\begin{array}{llll} \inf_y & L_y(f) & \text{s.t.} & y_{0,0} = 1, \quad M_d(y) \succcurlyeq 0, \quad \text{and} \quad M_{d-k_i}(g_i y) \succcurlyeq 0,\ i = 1, \ldots, m, \\[2ex] \sup_{\lambda,\sigma} & \lambda & \text{s.t.} & f - \lambda = \sigma_0 + \sigma_1 g_1 + \cdots + \sigma_m g_m, \end{array}}$$

where $\succcurlyeq$ stands for positive semidefinite. It relies on the following key notions:

- The *complex moment matrix* is a Hermitian matrix defined by

  $$M_d(y) := (y_{\alpha,\beta})_{|\alpha|,|\beta| \leqslant d}. \tag{2.4}$$

  In contrast to the real moment matrix (in the original Lasserre hierarchy), it is *not* a Hankel matrix. In other words, $y_{\alpha,\beta}$ is not necessarily only a function of $\alpha + \beta$. In the real case, we have $x^\alpha x^\beta = x^{\alpha+\beta}$ for $x \in \mathbb{R}^n$, whereas in the complex case, no such relationship holds for $z^\alpha \bar{z}^\beta$ where $z \in \mathbb{C}^n$.[1]
- The *Riesz functional* is defined by

  $$L_y(f) \ := \ \sum_{\alpha,\beta} f_{\alpha,\beta}\ y_{\alpha,\beta}. \tag{2.5}$$

- The *localizing matrices* are defined by

  $$M_{d-k_i}(g_i y) := \left(\sum_{\gamma,\delta} g_{i,\gamma,\delta}\ y_{\alpha+\gamma,\beta+\delta} \right)_{|\alpha|,|\beta| \leqslant d-k_i}, \tag{2.6}$$

[1] In fact, if one were to enforce the Hankel property in the complex moment matrix, one obtains the real Lasserre hierarchy applied to the complex polynomial optimization problem where all the complex variables are restrained to the real line. But make no mistake: this is *not* the real polynomial optimization problem obtained by identifying real and imaginary parts of the complex variables.

where $k_i := \max\{|\alpha|, |\beta| \text{ s.t. } g_{i,\alpha,\beta} \neq 0\}$. Naturally, the truncation order d must be greater than or equal to $d^{\min} := \max\{k_0, k_1, \dots, k_m\}$, where $k_0 := \max\{|\alpha|, |\beta| \text{ s.t. } f_{\alpha,\beta} \neq 0\}$.

- A polynomial $\sigma(z,\bar z) = \sum_{|\alpha|,|\beta| \leqslant d} \sigma_{\alpha,\beta} z^\alpha \bar z^\beta$ is a *Hermitian sum-of-squares*, i.e., it belongs to $\Sigma_d[z,\bar z]$, if it is of the form[2]

$$\sigma(z,\bar z) \;=\; \sum_k \left| \sum_{|\alpha|\leqslant d} p_{k,\alpha} z^\alpha \right|^2 \quad \text{where} \quad p_{k,\alpha} \in \mathbb{C}. \tag{2.7}$$

  This is equivalent to $(\sigma_{\alpha,\beta})_{|\alpha|,|\beta|\leqslant d} \succcurlyeq 0$, where $\alpha, \beta \in \mathbb{N}^n$. In the complex Lasserre hierarchy, $\sigma_0 \in \Sigma_d[z,\bar z]$ and $\sigma_i \in \Sigma_{d-k_i}[z,\bar z]$, $i = 1,\dots,m$.
- A Hermitian sum-of-squares is a special case of a *real sum-of-squares* (used in the original Lasserre hierarchy), that is, a polynomial of the form

$$\sigma(z,\bar z) \;=\; \sum_k \left| \sum_{|\alpha+\beta|\leqslant d} p_{k,\alpha,\beta} z^\alpha \bar z^\beta \right|^2 \quad \text{where} \quad p_{k,\alpha,\beta} \in \mathbb{C}. \tag{2.8}$$

  This is equivalent to the existence of a real positive semidefinite matrix $(\varphi_{\alpha,\beta})_{|\alpha|,|\beta|\leqslant d}$, where $\alpha, \beta \in \mathbb{N}^{2n}$ such that $\sigma(z,\bar z) = \sum_{\alpha,\beta} \varphi_{\alpha,\beta} x^{\alpha+\beta}$. (We have identified real and imaginary parts $z_k := x_k + x_{k+n}\mathbf{i}$.)

Example 2.2. $x_1^2 + 2x_1 + 1 + x_2^2$ with $x_1, x_2 \in \mathbb{R}$ is a Hermitian sum-of-squares because it is equal to $|1 + x_1 + x_2\mathbf{i}|^2 = |1+z|^2$ where $z := x_1 + x_2\mathbf{i}$. In contrast, $x_1^2 + 2x_1 + 1$ is a real sum-of-squares but not a Hermitian sum-of-squares. Indeed, we have $x_1^2 + 2x_1 + 1 = \left|1 + \frac{1}{2}z + \frac{1}{2}\bar z\right|^2 = 1 + z + \bar z + \frac{1}{4}z^2 + \frac{1}{2}|z|^2 + \frac{1}{4}\bar z^2$. In other words,

$$x_1^2 + 2x_1 + 1 \;=\; \begin{pmatrix} 1 \\ z \\ z^2 \end{pmatrix}^* \begin{pmatrix} 1 & 1 & 1/4 \\ 1 & 1/2 & 0 \\ 1/4 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ z \\ z^2 \end{pmatrix}, \tag{2.9}$$

where $(\cdot)^*$ stands for conjugate transpose. The above matrix is unique and it is not positive semidefinite. Hence the polynomial is not a Hermitian sum-of-squares. The unicity in the Hermitian decomposition (which is true for any polynomial, not just in this example) contrasts with the nonunicity in the real decomposition $x_1^2 + 2x_1 + 1 = \dots$

$$\begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \end{pmatrix}^T \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \end{pmatrix} = \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \end{pmatrix}^T \begin{pmatrix} 1 & 1 & 0 & 1/2 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1x_2 \\ x_2^2 \end{pmatrix}$$

where $(\cdot)^T$ stands for transpose. One of the above matrices is positive semidefinite, making the polynomial a real sum-of-squares (which is otherwise obvious).

Example 2.3. Consider the following complex polynomial optimization problem:

$$\inf_{z\in\mathbb{C}} \quad z + \bar z \quad \text{s.t.} \quad |z|^2 = 1, \tag{2.10}$$

[2] We use $|\cdot|$ to denote the modulus of a complex number. In a Hermitian sum-of-squares, the dependence on both z and $\bar z$ can be seen upon developing the squares.

whose optimal value is -2. Letting $z =: x_1 + \mathbf{i}x_2$, it can be converted into real numbers:

$$\inf_{x_1,x_2\in\mathbb{R}} \ 2x_1 \quad \text{s.t.} \quad x_1^2 + x_2^2 = 1. \tag{2.11}$$

It can be solved to global optimality using real sum-of-squares since

$$2x_1 \ - \ (-2) \ = \ 1 + (x_1 + x_2)^2 \ + \ 1 \times (1 - x_1^2 - x_2^2). \tag{2.12}$$

But it can also be solved using Hermitian sum-of-squares since

$$z + \bar{z} \ - \ (-2) \ = \ |1 + z|^2 \ + \ 1 \times (1 - |z|^2). \tag{2.13}$$

Hermitian sum-of-squares entail a trade-off. At each truncation order d, they are cheaper to compute, but they potentially provide a relaxation bound of poorer quality. More precisely, as the number of variables grows, the moment matrix in the real Lasserre hierarchy is 2^d times bigger than the moment matrix in the complex hierarchy.[3] Likewise, the localizing matrix of constraint $g_i(z,\bar{z}) \geqslant 0$ is 2^{d-k_i} times bigger in the real hierarchy than in the complex hierarchy. Regarding the optimal power flow problem, the relaxation bounds for the real and complex hierarchies are the same at each order in all our numerical experiments.

Example 2.4. The advantage of Hermitian sum-of-squares for finding global minimizers to the optimal power flow problem can be seen in Table 1. In 17 out of the 18 instances, they are faster, sometimes up to an order of magnitude. The minimizers obtained are feasible up to 0.005 p.u. at voltage constraints and 1 MVA at all other

TABLE 1
Global value found by multi-ordered Lasserre hierarchy and solver time in seconds.

Case name	**Real vars.**	**Const-raints**	**Real Lasserre**			**Complex Lasserre**		
			Order 2	**Obj.**	**Time**	**Order 2**	**Obj.**	**Time**
case14Q	28	66	16	3,302	4.7	20	3,302	2.6
case14L	28	106	20	9,359	1.9	26	9,359	1.5
case39Q	78	268	268	11,221	741.7	268	11,211	58.7
case39L	78	268	18	41,921	2.3	18	41,921	1.4
case57Q	114	242	16	7,352	3.4	20	7,352	3.3
case57L	114	402	16	43,984	1.4	22	43,984	1.3
case118Q	236	580	16	81,515	15.7	20	81,515	3.7
case118L	236	952	40	134,907	10.5	52	134,907	5.5
case300	600	1,338	14	720,040	7.2	20	720,040	4.1
nesta_case24	48	538	181	6,421	246.1	203	6,421	61.7
nesta_case30	60	296	118	372	302.7	134	372	15.4
nesta_case73	146	1,642	284	20,125	506.9	290	20,124	52.6
PL-2383wp	4,354	14,702	14	24,990	583.4	80	24,991	53.9
PL-2746wop	4,378	14,828	34	19,210	2,662.4	34	19,212	124.3
PL-3012wp	4,584	16,474	114	27,642	318.7	90	27,644	141.0
PL-3120sp	4,628	16,000	96	21,512	386.6	112	21,512	193.9
PEGASE-1354	1,966	7,210	68	74,043	406.9	210	74,042	1,132.6
PEGASE-2869	4,240	14,468	84	133,944	921.3	148	133,939	700.8

[3]Indeed, the size of the real moment matrix is $\frac{(2n+d)\dots(2n+1)}{d!}$. The size of the complex moment matrix is $\frac{(n+d)\dots(n+1)}{d!}$. The ratio is $\frac{(2n+d)\dots(2n+1)}{(n+d)\dots(n+1)} = \frac{2n+d}{n+d}\cdots\frac{2n+1}{n+1}$, and each of the d terms converges toward 2 as n increases to infinity with d fixed.

constraints,[4] and the objective evaluated in the minimizers matches the relaxation bound with 0.05% relative to the bound. In order to obtain these results, sparsity is exploited using the multi-ordered Lasserre hierarchy (section 6), and symmetry is exploited using the block diagonal Lasserre hierarchy (section 7). As indicated by the "Order 2" columns, we increment the relaxation order at up to 268 constraints, and up to order 2. The largest maximal clique size is 19. Regarding the software, YALMIP 2015.06.26 [69] and MOSEK are used for the experiments. For test case descriptions, see [77] for case14Q–case300, [23] for the "nesta" cases with "active power increases" (API) loading scenarios, and [36, 52, 104] for PL-2383wp–PEGASE-2869. For the Polish (PL) and PEGASE cases, a preprocessing step was used to eliminate lines with impedances less than 1×10^{-3} and 3×10^{-3} *per unit, respectively, a* 1×10^{-4} per unit minimum resistance was enforced on each line (as in [67]), and the objective was active power loss minimization. Table 1 displays the number of variables and constraints after the preprocessing step.

Having motivated the introduction of the complex Lasserre hierarchy, we discuss asymptotic convergence and finite convergence in the next two sections. These two aspects are significantly different from the real hierarchy. Most other aspects of the real hierarchy carry over to the complex hierarchy in a straightforward fashion, including strong duality [53] and the generalized Lagrangian interpretation (see [50, section 7] for details). One aspect that is unresolved is the question of generic finite convergence [78], which is a subject for future research.

3. Asymptotic convergence. In 1968, Quillen [89] showed that a real-valued bihomogeneous complex polynomial that is positive away from the origin can be decomposed as a Hermitian sum-of-squares when it is multiplied by $(|z_1|^2 + \cdots + |z_n|^2)^r$ for some $r \in \mathbb{N}$. The result was rediscovered by Catlin and D'Angelo [20] and ignited a search for complex analogues of Hilbert's seventeenth problem [32, 33] and the ensuing Positivstellensätze [35, 84, 85, 86]. Notably, D'Angelo and Putinar proved the following powerful result in 2009.

THEOREM 3.1 (D'Angelo's and Putinar's Positivstellenstatz [34]). *Assume that one of the constraints of K is a sphere $|z_1|^2 + \cdots + |z_n|^2 = R^2$ for some radius $R > 0$. If $f > 0$ on K, then there exists Hermitian sum-of-squares $\sigma_0, \ldots, \sigma_m$ such that*

$$f = \sigma_0 + \sum_{i=1}^{m} \sigma_i g_i. \tag{3.1}$$

This theorem naturally admits a dual perspective.

THEOREM 3.2 (Putinar and Scheiderer [86]). *If one of the constraints of K is a sphere, then the following properties are equivalent:*

1. $\exists \mu \in \mathcal{M}_+(K): \ \forall \alpha, \beta \in \mathbb{N}^n, \ y_{\alpha,\beta} = \int_K z^\alpha \bar{z}^\beta d\mu$;
2. $\forall d \geqslant d^{\min}, \ M_d(y) \succcurlyeq 0, \ M_{d-k_i}(g_i y) \succcurlyeq 0$.

Global convergence in the complex hierarchy is thus guaranteed in the presence of a sphere constraint. This is in contrast to the real hierarchy where it is guaranteed in the presence of a ball constraint. A sphere may appear more restrictive than a ball. However, this is sufficient to solve complex polynomial optimization problems with compact feasible sets. Indeed, one can add a slack variable $z_{n+1} \in \mathbb{C}$ and a redundant

[4]Typical violations are smaller than 1 MVA. For instance, with the complex hierarchy PL-3012wp has over 99% of the buses with less than 0.02 MVA violation, and only 0.09% of the buses with greater than 0.1 MVA violation. Maximum line flow violation is 0.0006 MVA.

constraint $|z_1|^2 + \cdots + |z_{n+1}|^2 = R^2$ to the description of the feasible set when it is in a ball of radius R. This is similar to Lasserre who proposes to add a redundant ball constraint $x_1^2 + \cdots + x_n^2 \leqslant R^2$.

Example 3.1. Consider the optimization problem

$$\inf_{z\in\mathbb{C}} \quad 1 - \frac{4}{3}|z|^2 + \frac{7}{18}|z|^4 \qquad \text{s.t.} \qquad 1 - |z|^2 \geqslant 0, \tag{3.2}$$

whose optimal value is 1/18. D'Angelo and Putinar [34] have demonstrated that there do not exist Hermitian sum-of-squares σ_0 and σ_1 such that

$$1 - \frac{4}{3}|z|^2 + \frac{7}{18}|z|^4 \;=\; \sigma_0(z,\bar{z}) + \sigma_1(z,\bar{z})(1 - |z|^2). \tag{3.3}$$

As a result, the complex hierarchy cannot exceed the value 0. In fact, it finds $-1/3$ at all orders because

$$M(y) = \begin{array}{c|ccccc} & 1 & \bar{z} & \bar{z}^2 & \bar{z}^3 & \\ 1 & 1 & 0 & 0 & 0 & \cdots \\ z & 0 & 1 & 0 & 0 & \\ z^2 & 0 & 0 & 0 & 0 & \\ z^3 & 0 & 0 & 0 & 0 & \\ & \vdots & & & & \ddots \end{array} \tag{3.4}$$

is a primal optimal point. We propose to add a complex slack variable

$$\inf_{z_1,z_2\in\mathbb{C}} \quad 1 - \frac{4}{3}|z_1|^2 + \frac{7}{18}|z_1|^4 \qquad \text{s.t.} \qquad 1 - |z_1|^2 - |z_2|^2 = 0, \tag{3.5}$$

enabling the second order complex relaxation to find the global infimum

$$\begin{aligned} &1 - \tfrac{4}{3}|z_1|^2 + \tfrac{7}{18}|z_1|^4 - \tfrac{1}{18} \\ &= \tfrac{5}{18}|z_2|^2 + \tfrac{5}{18}|z_1 z_2|^2 + \tfrac{2}{3}|z_2^2|^2 \\ &+ \left(\tfrac{17}{18} - \tfrac{7}{18}|z_1|^2 + \tfrac{2}{3}|z_2|^2\right)(1 - |z_1|^2 - |z_2|^2). \end{aligned} \tag{3.6}$$

Note that the polynomial that multiplies the constraint is not a Hermitian sum-of-squares. This would be a contradiction when taking $z_2 = 0$.

We next discuss a weaker condition ensuring global convergence in the Lasserre hierarchy. In the real hierarchy, convergence is guaranteed if the Archimedean condition holds, that is to say, if there exists $R > 0$ and real sums-of-squares $\sigma_0, \ldots, \sigma_m$ such that $R^2 - x_1^2 - \cdots - x_n^2 = \sigma_0(x) + \sum_{i=1}^m \sigma_i(x) g_i(x) \, \forall x \in \mathbb{R}^n$. In the complex hierarchy, a similar condition can be deduced from the work of Putinar and Scheiderer [86, Propositions 6.6 and 3.2(iii)]. For notational convenience, suppose that some of the inequality constraints $g_i(z,\bar{z}) \geqslant 0$ are actually equality constraints $g_i(z,\bar{z}) = 0$. Let $E \subset \{1,\ldots,m\}$ denote the indices of the equality constraints. Global convergence in the complex hierarchy is guaranteed if there exist $R > 0$, a Hermitian sum-of-squares σ_0, and real-valued complex polynomials p_i's such that $R^2 - |z_1|^2 - \cdots - |z_n|^2 = \sigma_0(z,\bar{z}) + \sum_{i\in E} p_i(z,\bar{z}) g_i(z,\bar{z})$, $\forall z \in \mathbb{C}^n$. In particular, in the presence of the equalities $|z_k|^2 = 1, \; k = 1,\ldots,n$, convergence is guaranteed. In this case, there is no need to add a slack variable as suggested above. This applies, for instance, to the nonbipartite Grothendieck problem over the complex numbers [13]. Interestingly, in the optimal power flow problem, despite the absence of such equalities, global convergence is attained in all numerical experiments without adding a slack variable (as reported in Table 1). Of course, it is also attained when adding a slack variable.

When the weaker assumption presented above does not hold, there exists a way to quantify how far D'Angelo's and Putinar's Positivstellensatz is from being true. This is given by the Hermitian complexity [35] of the ideal associated with the equality constraints. This number is related to the greatest number of distinct points (possibly infinite) $z^{(i)} \in \mathbb{C}^n,\ 1 \leqslant i \leqslant p$, such that $g_k(z^{(i)}, z^{(j)}) = 0$ for all $k \in E$. Loosely speaking, the greater this number, the farther away the Positivstellensatz is from being true. In particular, when one of the equalities is $\sigma(z,\bar{z}) + |z_1|^2 + \cdots + |z_n|^2 = R^2$ with σ a Hermitian sum-of-squares and $R > 0$, then the Hermitian complexity is equal to 1. The Positivstellensatz is then true, in accordance with the weaker assumption presented above.

4. Finite convergence. The relaxation of order d of the complex Lasserre hierarchy yields a set of complex numbers $(y_{\alpha,\beta})_{|\alpha|,|\beta| \leqslant d}$. Global solutions may be extracted if there exists a measure that represents those $y_{\alpha,\beta}$ that appear in the objective and constraint functions. In particular, this is true if there exists positive Borel measure μ supported on the semialgebraic set K such that $y_{\alpha,\beta} = \int_{\mathbb{C}^n} z^\alpha \bar{z}^\beta d\mu$ for all $|\alpha|, |\beta| \leqslant d^{\min}$. In this case, the complex hierarchy has finite convergence and global optimality is attained. In this section, we will show that the conditions ensuring finite convergence in the complex hierarchy differ significantly from the real hierarchy. Additional conditions need to be satisfied, as shown in the next result. We will use the definition $d_K := \max\{2, k_1, \ldots, k_m\}$, where the number 2 will be explained in the next section. It is meant to guarantee that $d_K \geqslant 2$, in contrast to the real case when it is only required that $d_K \geqslant 1$ (see [65, equation (6.1)]).

PROPOSITION 4.1. *Assume that one of the constraints of the multivariate* $(n > 1)$ *optimization problem is a ball* $|z_1|^2 + \cdots + |z_n|^2 \leqslant R^2$ *for some radius* $R > 0$*. Consider an optimal solution* y *to the complex moment relaxation of order* d*.*

- *If there is an integer* t *such that* $d^{\min} \leqslant t \leqslant d$ *and* $\mathrm{rank} M_t(y) = 1$*, then global optimality is attained and there is at least one global solution.*
- *If there is an integer* t *such that* $\max\{d^{\min}, d_K\} \leqslant t \leqslant d$ *and if the following conditions hold:*

1. $\mathrm{rank}\ M_t(y) = \mathrm{rank}\ M_{t-d_K}(y)\ (=: S)$,
2. $\begin{pmatrix} M_{t-d_K}(y) & M_{t-d_K}(z_i y) & M_{t-d_K}(z_j y) \\ M_{t-d_K}(\bar{z}_i y) & M_{t-d_K}(|z_i|^2 y) & M_{t-d_K}(z_j \bar{z}_i y) \\ M_{t-d_K}(\bar{z}_j y) & M_{t-d_K}(z_i \bar{z}_j y) & M_{t-d_K}(|z_j|^2 y) \end{pmatrix} \succcurlyeq 0,\ \forall 1 \leqslant i < j \leqslant n,$

then global optimality is attained and there are at least S *global solutions.*

Proof. This is a consequence of Theorem 5.1. ◻

Under the assumptions of Proposition 4.1, if the rank is equal to one, then a global solution z can be read from the moment matrix, i.e., $z = (y_{\alpha,0})_{|\alpha|=1} \in \mathbb{C}^n$. This is just like in the real Lasserre hierarchy. Otherwise, if the rank is greater than one $(S > 1)$, then S global solutions can be extracted using [42, Algorithm 4.1]. In fact, this algorithm can also extract global solutions from the real Lasserre hierarchy. It appears to be the most efficient way to do so as it only requires one singular value decomposition followed by an eigendecomposition. Earlier approaches can be found

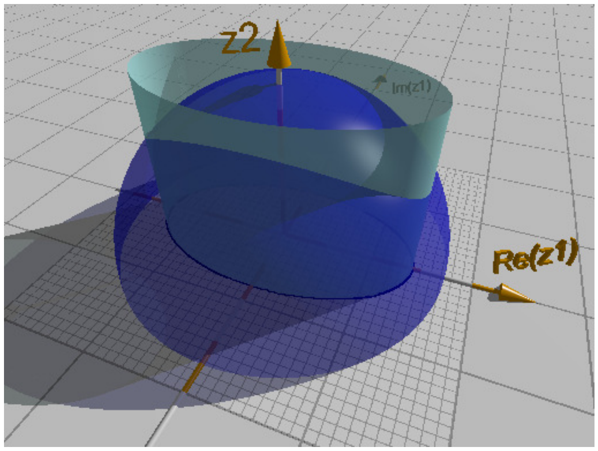

FIG. 2. *Feasible set at the intersection of the sphere and the elliptic cylinder.*

in [43, 44, 65], and recent work can found in [88].

Example 4.1. Consider the following problem whose elliptic constraint is taken from [85]:

$$\inf_{z_1,z_2\in\mathbb{C}} \quad 3-|z_1|^2-\frac{1}{2}\mathbf{i}z_1\bar{z}_2^2+\frac{1}{2}\mathbf{i}z_2^2\bar{z}_1, \tag{4.1}$$

$$\text{s.t.} \quad \begin{cases} |z_1|^2-\frac{1}{4}z_1^2-\frac{1}{4}\bar{z}_1^2=1, \\ |z_1|^2+|z_2|^2=3, \\ \mathbf{i}z_2-\mathbf{i}\bar{z}_2=0, \quad z_2+\bar{z}_2\geqslant 0. \end{cases} \tag{4.2}$$

The feasible set is represented in Figure 2, which we generated using POV-Ray 3.7.0 [81]. The hierarchy starts at the second order ($d^{\min}=d_K=2$), which yields[5] the lower bound 0.155089 and the optimal moment matrix

$$M_2(y)=\begin{array}{c|cccccc} & 1 & \bar{z}_1 & \bar{z}_2 & \bar{z}_1^2 & \bar{z}_1\bar{z}_2 & \bar{z}_2^2 \\ \hline 1 & 1.0000 & 0.3747\mathbf{i} & 0.8485 & 1.8272 & 0.5100\mathbf{i} & 1.0864 \\ z_1 & -0.3747\mathbf{i} & 1.9136 & -0.5100\mathbf{i} & 0.1929\mathbf{i} & 1.0505 & -0.9313\mathbf{i} \\ z_2 & 0.8485 & 0.5100\mathbf{i} & 1.0864 & 0.9245 & 0.9313\mathbf{i} & 1.4950 \\ z_1^2 & 1.8272 & -0.1929\mathbf{i} & 0.9245 & 4.5886 & -0.1162\mathbf{i} & 0.9324 \\ z_1z_2 & -0.5100\mathbf{i} & 1.0505 & -0.9313\mathbf{i} & 0.1162\mathbf{i} & 1.1523 & -1.4140\mathbf{i} \\ z_2^2 & 1.0864 & 0.9313\mathbf{i} & 1.4950 & 0.9324 & 1.4140\mathbf{i} & 2.1069 \end{array}$$

It holds that $\text{rank}M_0(y)=1$, $\text{rank}M_1(y)=3$, and $\text{rank}M_2(y)=3$. Since $\text{rank}M_0(y)\neq\text{rank}M_2(y)$, the rank condition in Proposition 4.1 does not hold. The positive semidef-

[5]MATLAB 2015b, CVX 2.0 [39], and SDPT3 4.0 [97] are used for the numerical experiments.

inite condition holds with $t = 2$ but not with $t = 3$:

$$\mathrm{sp}\left\{\begin{pmatrix} M_1(y) & M_1(z_1 y) & M_1(z_2 y) \\ M_1(\bar{z}_1 y) & M_1(|z_1|^2 y) & M_1(z_2\bar{z}_1 y) \\ M_1(\bar{z}_2 y) & M_1(z_1\bar{z}_2 y) & M_1(|z_2|^2 y) \end{pmatrix}\right\} = \begin{pmatrix} -1.5874 \\ -0.1295 \\ -0.0000 \\ 0.0000 \\ 0.1574 \\ 0.7711 \\ 3.5471 \\ 5.0544 \\ 8.1869 \end{pmatrix} \tag{4.3}$$

where sp$\{\cdot\}$ stands for spectrum. The third-order complex relaxation yields the value 0.428175 and the moment matrix satisfies $\mathrm{rank} M_3(y) = 1$. This yields a global solution $(z_1, z_2) = (-0.8165\mathbf{i}, 1.5275)$ which can be read from the third-order moment matrix. Interestingly, it is not necessary to go up to the third-order relaxation. Since the positive semidefinite condition is a convex property, it can be added to the second order relaxation, with $t = 3$, for instance:

$$\begin{pmatrix} M_1(y) & M_1(z_1 y) & M_1(z_2 y) \\ M_1(\bar{z}_1 y) & M_1(|z_1|^2 y) & M_1(z_2\bar{z}_1 y) \\ M_1(\bar{z}_2 y) & M_1(z_1\bar{z}_2 y) & M_1(|z_2|^2 y) \end{pmatrix} \succcurlyeq 0. \tag{4.4}$$

We then obtain the value 0.428175 and the following moment matrix:

$$M_2(y) = \begin{array}{c|cccccc} & 1 & \bar{z}_1 & \bar{z}_2 & \bar{z}_1^2 & \bar{z}_1\bar{z}_2 & \bar{z}_2^2 \\ 1 & 1.0000 & 0.8165\mathbf{i} & 1.5275 & -0.6667 & 1.2472\mathbf{i} & 2.3333 \\ z_1 & -0.8165\mathbf{i} & 0.6667 & -1.2472\mathbf{i} & 0.5443\mathbf{i} & 1.0184 & -1.9052\mathbf{i} \\ z_2 & 1.5275 & 1.2472\mathbf{i} & 2.3333 & -1.0184 & 1.9052\mathbf{i} & 3.5642 \\ z_1^2 & -0.6667 & -0.5443\mathbf{i} & -1.0184 & 0.4444 & -0.8315\mathbf{i} & -1.5556 \\ z_1 z_2 & -1.2472\mathbf{i} & 1.0184 & -1.9052\mathbf{i} & 0.8315\mathbf{i} & 1.5556 & -2.9102\mathbf{i} \\ z_2^2 & 2.3333 & 1.9052\mathbf{i} & 3.5642 & -1.5556 & 2.9102\mathbf{i} & 5.4444 \end{array}$$

which satisfies $\mathrm{rank} M_2(y) = 1$. A global solution can be read in the first column: $(z_1, z_2) = (-0.8165\mathbf{i}, 1.5275)$. We have just used a notion in operator theory to reduce the rank from 3 to 1 in a convex relaxation. For explanations, see the next section.

5. Truncated moment problem. The complex Lasserre hierarchy brings into the picture a truncated moment problem which has not been considered in past literature to the best of our knowledge. Given a set of complex numbers $(y_{\alpha,\beta})_{|\alpha|,|\beta|\leqslant d}$, it raises the question of whether there exists positive Borel measure μ supported on the semialgebraic set K such that

$$y_{\alpha,\beta} = \int_{\mathbb{C}^n} z^\alpha \bar{z}^\beta d\mu, \quad \forall \ |\alpha|, |\beta| \leqslant d. \tag{5.1}$$

In this section, we propose a solution to this problem. Precisely, we characterize when there exists a $\mathrm{rank} M_d(y)$-atomic representing measure for the data $(y_{\alpha,\beta})_{|\alpha|,|\beta|\leqslant d}$. We do so via the existence of an extension of the data, which must satisfy certain conditions, following in the footsteps of Curto and Fialkow [26, 27, 28]. An example is how they characterize [29, Theorem 5.1] atomic measures μ supported on the semialgebraic set K such that

$$y_{\alpha,\beta} = \int_{\mathbb{C}^n} z^\alpha \bar{z}^\beta d\mu, \quad \forall\, |\alpha| + |\beta| \leqslant 2d \tag{5.2}$$

given some complex numbers $(y_{\alpha,\beta})_{|\alpha|+|\beta|\leqslant 2d}$. However, this moment problem is not relevant for the complex Lasserre hierarchy since the truncation of the data is different. Below, we have represented the second-order truncation in the complex hierarchy in bullets and the second order truncation of Curto and Fialkow in bullets and circles:

$$\begin{array}{c|cccccc}
 & 1 & z & \bar{z}^2 & \bar{z}^3 & \bar{z}^4 & \\
\hline
1 & \bullet & \bullet & \bullet & \circ & \circ & \dots \\
z & \bullet & \bullet & \bullet & \circ & & \\
z^2 & \bullet & \bullet & \bullet & & & \\
z^3 & \circ & \circ & & \ddots & & \\
z^4 & \circ & & & & & \\
 & \vdots & & & & &
\end{array}$$

This leads to different notions of moment matrices. Below, we have represented the moment matrix of the complex hierarchy (on the left, in bullets) and the moment matrix of Curto and Fialkow (on the right, in circles):

$$\begin{array}{c|ccc}
 & 1 & \bar{z} & \bar{z}^2 \\
\hline
1 & \bullet & \bullet & \bullet \\
z & \bullet & \bullet & \bullet \\
z^2 & \bullet & \bullet & \bullet
\end{array}
\qquad\qquad
\begin{array}{c|cccccc}
 & 1 & \bar{z} & z & \bar{z}^2 & \bar{z}z & z^2 \\
\hline
1 & \circ & \circ & \circ & \circ & \circ & \circ \\
z & \circ & \circ & \circ & \circ & \circ & \circ \\
\bar{z} & \circ & \circ & \circ & \circ & \circ & \circ \\
z^2 & \circ & \circ & \circ & \circ & \circ & \circ \\
z\bar{z} & \circ & \circ & \circ & \circ & \circ & \circ \\
\bar{z}^2 & \circ & \circ & \circ & \circ & \circ & \circ
\end{array}$$

The moment matrix in the complex hierarchy is referred to as a *pruned* complex moment matrix in [63]. However, the associated moment problem is not considered. Despite the discrepancies between the moment matrices, like Curto and Fialkow, we will rely on the notion of *flat extension*, which is an extension of the moment matrix that preserves the positive semidefiniteness and the rank.

Note that the complex moment problem of Curto and Fialkow is equivalent in some sense (see [29, Theorem 5.2]) to the real moment problem, i.e., where we seek a measure on a real semialgebraic set such that

$$y_\alpha = \int_{\mathbb{R}^{2n}} x^\alpha d\mu, \quad \forall\, |\alpha| \leqslant 2d \tag{5.3}$$

given some real numbers $(y_\alpha)_{|\alpha|\leqslant 2d}$. In contrast, the truncated moment problem arising in the complex hierarchy captures the real truncated moment problem as a special case. It corresponds to the case where the moment data forms a Hankel matrix (see Theorem 5.2 below).

To provide a solution to the truncated moment problem arising in the complex hierarchy (Theorem 5.1 below), we rely on the notion of hyponormality in operator theory. Indeed, we are unable to adapt the algebraic arguments used by Curto and Fialkow. They consider the ideal generated by the monomials that are indexes of the rows of their moment matrix. We are unable to make use of this in our context. We thus pursue a different approach, based exclusively on operator theory. The relationship between this discipline and the moment problem was recognized early on, as described by Akhiezer [5, Chapter 4] in 1965. It has since been enriched by a vast literature including the works of Cassier [18], Schmüdgen [92], and Putinar [83]. In particular, Atzmon [10] used operator theory to solve the full moment problem on the unit disc in the complex plane. Later, Curto and Putinar [30, Theorem 3.1] extended this result to subalgebraic subsets of the complex plane defined by one inequality. For such sets, the real moment problem was shown to be reducible to a complex moment problem in [82]. In fact, Putinar employed this complexification of the real moment problem in his seminal result [83] which implies convergence of the (real) Lasserre hierarchy. Going back to the full complex moment problem, a solution was given in [95] when the support is the entire complex plane. An operator-valued moment problem is also considered in that work. In the more recent paper [56], a Hermitian-matrix-valued truncated moment problem is investigated in the multivariate setting.

We now outline our approach. We need a few notations: let $\mathcal{B}(\mathcal{H})$ denote the set of linear bounded operators acting on a Hilbert space $(\mathcal{H}, \langle \cdot, \cdot \rangle)$. For all $T \in \mathcal{B}(\mathcal{H})$, let $T \succcurlyeq 0$ denote $\langle Tu, u \rangle \geqslant 0$ for all $u \in \mathcal{H}$. In addition, the commutator of $A, B \in \mathcal{B}(\mathcal{H})$ is defined as $[A, B] := AB - BA$. Finally, let A^* denote the adjoint of $A \in \mathcal{B}(\mathcal{H})$. Following Halmos [41], an operator $T \in \mathcal{B}(\mathcal{H})$ is said to be

- *normal* if $[T^*, T] = T^*T - TT^* = 0$;
- *subnormal* if it can be extended to a normal operator N on a larger Hilbert space $\mathcal{K}$;
- *hyponormal* if $[T^*, T] = T^*T - TT^* \succcurlyeq 0$.

The notions of subnormality and hyponormality were introduced by Halmos in 1950 in order to extend the spectral theory of normal operators to a larger class of operators. They have since been used to shed light on the moment problem, as in [9, 99] and the works cited above. The following implications hold (for explanations, see, e.g., [31, 73]):

$$\text{normal} \quad \Longrightarrow \quad \text{subnormal} \quad \Longrightarrow \quad \text{hyponormal}$$

The gap between subnormality and hyponormality has been the subject of much investigation, such as in [25, 74]. It was later discovered in [30] that there is in fact a significant gap: even polynomially hyponormal operators (i.e., such that $p(T)$ is hyponormal for all $p \in \mathbb{C}[z]$) are not necessarily subnormal. The key ingredient for our proof is that in finite dimension, normality, subnormality, and hyponormality are all equivalent. Indeed, if $\mathcal{H}$ is finite dimensional, then the trace of $[T^*, T]$ is equal to zero. If in addition $[T^*, T] \succcurlyeq 0$, then it must be that $[T^*, T] = 0$. We next show how this observation is relevant for a tuple of operators.

Following the definition of Athavale [8], operators $T_1, \ldots, T_n \in \mathcal{B}(\mathcal{H})$ are *jointly* *hyponormal* if

$$
\begin{pmatrix}
[T_1^*, T_1] & [T_2^*, T_1] & \dots & [T_n^*, T_1] \\
[T_1^*, T_2] & [T_2^*, T_2] & \dots & [T_n^*, T_2] \\
\vdots & \vdots & & \vdots \\
[T_1^*, T_n] & [T_2^*, T_n] & \dots & [T_n^*, T_n]
\end{pmatrix} \succcurlyeq 0 \tag{5.4}
$$

in the sense that for all $u_1, \dots u_n \in \mathcal{H}$, there holds $\sum_{i,j=1}^n \langle u_i, [T_j^*, T_i] u_j \rangle \geqslant 0$. Thanks to our previous observation, in finite dimension, this is equivalent to[6]

$$
\begin{pmatrix}
[T_1^*, T_1] & [T_2^*, T_1] & \dots & [T_n^*, T_1] \\
[T_1^*, T_2] & [T_2^*, T_2] & \dots & [T_n^*, T_2] \\
\vdots & \vdots & & \vdots \\
[T_1^*, T_n] & [T_2^*, T_n] & \dots & [T_n^*, T_n]
\end{pmatrix} = 0. \tag{5.5}
$$

This is itself trivially equivalent to

$$
\begin{pmatrix}
[T_i^*, T_i] & [T_j^*, T_i] \\
[T_i^*, T_j] & [T_j^*, T_j]
\end{pmatrix} = 0, \quad \forall\, 1 \leqslant i < j \leqslant n, \tag{5.6}
$$

and to

$$
\begin{pmatrix}
I & T_i^* & T_j^* \\
T_i & T_i^* T_i & T_i^* T_j \\
T_j & T_i^* T_j & T_j^* T_j
\end{pmatrix} \succcurlyeq 0, \quad \forall\, 1 \leqslant i < j \leqslant n, \tag{5.7}
$$

thanks to a Schur complement.

The purpose of the third condition of Theorem 5.1 below is to guarantee that joint hyponormality holds for a certain set of operators acting on a finite dimensional space. The purpose of the second condition of Theorem 5.1 is to ensure that these operators commute (recall that $d_K := \max\{2, k_1, \dots, k_m\}$). The operators in question are *shift* operators, which are commonly used when dealing with the moment problem (e.g., [87, Proposition 8]). More explanations can be found in the proof.

THEOREM 5.1. *Consider a positive integer d and some complex numbers $(y_{\alpha,\beta})_{|\alpha|,|\beta| \leqslant d}$. Assume that K contains a ball constraint $|z_1|^2 + \cdots + |z_n|^2 \leqslant R^2$ for some radius $R > 0$ and that we are in the multivariate setting ($n > 1$). Then there exists a positive* $\operatorname{rank} M_d(y)$*-atomic measure μ supported on K such that*

$$
y_{\alpha,\beta} = \int_{\mathbb{C}^n} z^\alpha \bar{z}^\beta d\mu, \quad \forall\, |\alpha|, |\beta| \leqslant d \tag{5.8}
$$

if and only if there exists an extension $(y_{\alpha,\beta})_{d < |\alpha|, |\beta| \leqslant d + d_K}$ such that the following

[6]Make no mistake: this is not the definition of a normal tuple of operators, which is that $[T_i, T_j] = [T_i^*, T_i] = 0$ for all $i, j = 1, \dots, n$ (see [55, p. 1505]).

hold:

1. *Positivity of moment and localizing matrices:*
$M_{d+d_K}(y) \succcurlyeq 0$ and $M_{d+d_K-k_i}(g_i y) \succcurlyeq 0,\ i = 1,\dots,m$

2. *Commutativity of the shifts:*
$\text{rank} M_{d+d_K}(y) = \text{rank} M_d(y)$

3. *Joint hyponormality of the shifts:*
$$\begin{pmatrix} M_d(y) & M_d(z_i y) & M_d(z_j y) \\ M_d(\bar{z}_i y) & M_d(|z_i|^2 y) & M_d(z_j \bar{z}_i y) \\ M_d(\bar{z}_j y) & M_d(z_i \bar{z}_j y) & M_d(|z_j|^2 y) \end{pmatrix} \succcurlyeq 0,\ \forall 1 \leqslant i < j \leqslant n.$$

Proof. See Appendix A. ∎

In the univariate setting ($n = 1$), the "joint hyponormality of the shifts" condition must be replaced by

$$\begin{pmatrix} M_d(y) & M_d(zy) \\ M_d(\bar{z}y) & M_d(|z|^2 y) \end{pmatrix} \succcurlyeq 0. \tag{5.9}$$

Theorem 5.1 then holds with $d_K := \max\{1, k_1, \dots, k_m\}$, in contrast to the multivariate setting where $d_K := \max\{2, k_1, \dots, k_m\}$ (see proof for explanations).

In the next result, we consider two cases where the ball constraint and the "joint hyponormality of the shifts" condition can be removed. One case is when the moment data forms a Toeplitz matrix, that is to say when $y_{\alpha,\beta} = \overline{y_{\beta,\alpha}}$ only depends on $\alpha - \beta$. This is relevant when optimizing in the presence of the constraints $|z_k|^2 = 1,\ k = 1,\dots,n$. The other case is when the moment data forms a Hankel matrix, that is to say when $y_{\alpha,\beta}$ is real and only depends on $\alpha + \beta$. This is relevant for real polynomial optimization, which can be viewed as an instance of complex polynomial optimization with the constraints $\mathbf{i}z_k - \mathbf{i}\bar{z}_k = 0,\ k = 1,\dots,n$. It corresponds exactly to the moment data generated by the original (real) Lasserre hierarchy.

THEOREM 5.2. *Consider a positive integer d and some complex numbers $(y_{\alpha,\beta})_{|\alpha|,|\beta| \leqslant d}$. Assume that K contains either the constraints $|z_k|^2 = 1,\ k = 1,\dots,n$, or the constraints $\mathbf{i}z_k - \mathbf{i}\bar{z}_k = 0,\ k = 1,\dots,n$. Then there exists a positive* $\text{rank} M_d(y)$*-atomic measure μ supported on K such that*

$$y_{\alpha,\beta} = \int_{\mathbb{C}^n} z^\alpha \bar{z}^\beta d\mu, \quad \forall\, |\alpha|, |\beta| \leqslant d \tag{5.10}$$

if and only if there exists an extension $(y_{\alpha,\beta})_{d<|\alpha|,|\beta| \leqslant d+d_K}$ such that the following hold:

1. *Positivity of moment and localizing matrices:*
$$M_{d+d_K}(y) \succcurlyeq 0 \quad \textit{and} \quad M_{d+d_K-k_i}(g_i y) \succcurlyeq 0, \quad i = 1,\dots,m.$$

2. *Commutativity of the shifts:*
$$\text{rank} M_{d+d_K}(y) = \text{rank} M_d(y).$$

Proof. See Appendix B. ∎

Our solution to the moment problem in the Toeplitz case is new to the best of our knowledge. In the univariate case $n = 1$ with support equal to the full space

$K = \mathbb{C}$, it corresponds to the truncated trigonometric moment problem. A solution to this problem has been given by [46, p. 211], [6, Theorem I.I.12], and [26, Theorem 6.12]. It can be stated as follows. A Toeplitz matrix can be represented by a positive Borel measure if and only if it is positive semidefinite. In other words, there need not exist a flat extension for there to exist a measure. For some more recent work on the trigonometric moment problem, see [2, 37, 103]. See also [12] for its relevance in the context of matrix completions.

The Hankel case in Theorem 5.2 corresponds to the solution of real truncated moments found in [62, Theorem 3.11] due to Curto and Fialkow [29, Theorem 1.1]. However, their result is stronger because it only requires that $d_K \geqslant 1$, while we require that $d_K \geqslant 2$. The reason why we record this result is to underscore the link between the moment problems arising in the real and complex hierarchies. It also provides a new proof based solely on operator theory, in contrast to the proof of Curto and Fialkow, and the more recent proof of Laurent [64]. The latter relies partly on algebraic tools, while the former relies only on algebraic tools. Note that the result has been generalized to moment matrices indexed by arbitrary monomials in [66].

6. Multi-ordered Lasserre hierarchy.

In [77], a heuristic was proposed to exploit sparsity in the Lasserre hierarchy when applied to the optimal power flow problem. Inspired by that work, we propose a general approach to exploit sparsity in any polynomial optimization problem (in real or complex variables) which preserves global convergence. The approach associates a different relaxation order to each constraint, in contrast to the Lasserre hierarchy, which associates the same relaxation order to all constraints.

6.1. Defining a relaxation order at each constraint.

In order to define a relaxation order for each constraint, we build on the work of Waki et al. [100]. Those authors propose using chordal sparsity in the Lasserre hierarchy. They draw on the *correlative sparsity* graph whose vertices are the variables and whose edges signify that two variables appear simultaneously either in a constraint or in a monomial of the objective. The idea of Waki et al. is to restrain the variables appearing in the sum-of-squares (a priori all variables) to subsets of variables. Indeed, in the sum-of-squares decomposition, i.e.,

$$f - \lambda = \sigma_0 + \sum_{i=1}^{m} \sigma_i g_i, \tag{6.1}$$

one would like to restrain the variables appearing in σ_i as a function of the variables appearing in the constraint g_i. For instance, if the variables appearing in one constraint g_i are x_1, x_2, x_3, x_4 (among, say, $x_1, \ldots, x_{100}$), one could hope to restrain the variables appearing in σ_i to x_1, x_2, x_3, x_4 (or some slightly larger set). This hope becomes a reality when considering the maximal cliques of a chordal extension of the correlative sparsity graph. Then, to each constraint g_i, one can associate a maximal clique containing all the variables of g_i (preferably with the fewest number of variables if several cliques work). Next, one can restrain the variables in the sum-of-squares σ_i to that clique. The sum-of-squares σ_0 can be restricted to a sum of terms, where each term is a sum-of-squares with variables belonging to a clique. At a given order, the relaxation might be weaker, but global convergence is preserved, as was first shown by Lasserre [62, Theorems 2.28 and 4.7] and later confirmed in [40] and [58]. These results easily generalize to complex numbers: the proof is the same as in the real case [62, Lemma B.13 and 4.10.2, Proof of Theorem 4.7] once the real vector spaces on which measures are defined are replaced by complex vector spaces. Note that the assumption of a redundant ball constraint per clique must be replaced by a sphere and slack variable per clique. To sum up, if $C_1, \dots, C_p$ are the cliques, Waki et al. propose to restrain (6.1) to

$$f - \lambda \quad = \quad \sum_{k=1}^{p} \left(\sigma_0^k \quad + \sum_{\substack{\text{constraints } i \\ \text{associated to } C_k}} \sigma_i g_i \right), \tag{6.2}$$

where the variables of σ_0^k and σ_i are restrained to the clique C_k.

The approach of Waki et al. reduces the computational burden of the Lasserre hierarchy for sparse problems. Concerning the optimal power flow problem, it allows one to solve some hard instances to global optimality with up to 80 variables [38] (instead of 20 without exploiting sparsity [75]). However, by using the correlative sparsity graph discussed above, a lot of the sparsity is lost. We thus propose a finer notion of sparsity that takes advantage of the fact that the constraints are polynomials. To that effect, we define the *monomial sparsity* graph whose vertices are the variables and whose edges signify that two variables appear simultaneously in a monomial of either the objective or a constraint.[7] We can then define a relaxation order d_i for each constraint g_i. If we want a constraint g_i to have a high order, i.e., with a sum-of-squares σ_i of degree greater than zero, then we add the correlative sparsity induced by g_i to the monomial sparsity graph. We then consider the maximal cliques of a chordal extension of the resulting graph. To each constraint g_i that is of high order, one can associate a maximal clique containing all the variables of g_i. The variables in the sum-of-squares σ_i can then be restrained to the clique associated to g_i when it is of high order; if not of high order, the sum-of-squares is a nonnegative real number. The sum-of-squares σ_0 can be restricted to a sum of terms, where each term is a sum-of-squares with variables belonging to a clique. To sum up, we replace (6.2) by

$$f - \lambda \quad = \quad \sum_{k=1}^{p} \left(\sigma_0^k \quad + \sum_{\substack{\text{constraints } i \\ \text{of high order} \\ \text{associated to } C_k}} \sigma_i g_i \right) + \sum_{\substack{\text{constraints } i \\ \text{not of high order}}} \sigma_i g_i, \tag{6.3}$$

where the variables of σ_0^k are restrained to the clique C_k, and the same goes for σ_i in the case of high-order constraints associated to C_k; otherwise σ_i is a nonnegative real number. The polynomial σ_0^k is a sum of squares of polynomials of degree less than or equal to the maximal relaxation order d_i among all high order constraints associated to C_k. If no high-order constraints are associated to C_k, the degree is less than or equal to one. As can be seen in (6.3), if all the constraints have a high order, we are back to (6.2), i.e., the approach of Waki et al.,[8] but with different orders at each constraint. In that case, the relaxation is at least as tight as the approach of Waki et al., whose relaxation order is taken to be the minimal order among all constraints

[7] A formal definition is as follows. Given $\alpha \in \mathbb{N}^n$ with $n > 1$, let $\text{supp}(\alpha) := \{1 \leqslant s \leqslant n \mid \alpha_s \neq 0\}$. The set of edges of the monomial sparsity graph is $\{(l, m) \mid l \neq m \text{ s.t. } \exists \alpha, \beta \in \mathbb{N}^n \text{ s.t. } \{l, m\} \subset \text{supp}(\alpha + \beta) \text{ and } |f_{\alpha,\beta}| + |g_{1,\alpha,\beta}| + \cdots + |g_{m,\alpha,\beta}| \neq 0\}$.

[8] The monomial sparsity graph augmented with the correlative sparsity induced by each constraint is none other than the correlative sparsity graph.

in the multi-ordered hierarchy. Global convergence is thus preserved as the minimal order increases to infinity.

Example 6.1. Consider the following optimization problem:

$$\inf_{x_1,x_2,x_3,x_4 \in \mathbb{R}} \quad x_1x_2 + x_1x_4 \quad \text{s.t.} \quad \begin{cases} x_1x_2 + x_1x_3 \geqslant 0, \\ x_1x_3 + x_1x_4 + x_1x_2 \geqslant 0. \end{cases} \tag{6.4}$$

It is solely meant to illustrate the above notions; the next example is much more interesting from a numerical perspective. Figure 3 illustrates the correlative sparsity used by Waki et al. and the monomial sparsity advocated in this paper. Suppose one wants to impose order 2 at the first constraint (i.e., a high order) and order 1 at the second constraint (i.e., not a high order). The correlative sparsity induced by the first constraint is the triangle formed by the first three variables. When added to the monomial sparsity pattern, it yields the graph on the left of Figure 3 if the edge (2,3) is added. The variables in the sum-of-squares σ_1 can then be restrained to x_1, x_2, x_3, while the sum-of-squares σ_2 is a nonnegative real number. The reason we must add the correlative sparsity induced by the high-order constraint to the monomial sparsity pattern is that in the expression $\sigma_1(x_1,x_2,x_3)g_1(x_1,x_2,x_3)$, all the possible products x_1x_2, x_1x_3, x_2x_3 appear.

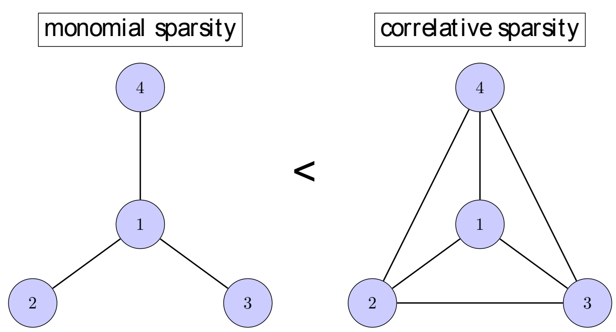

FIG. 3. *Two different notions of sparsity.*

Example 6.2. In [14, WB5, $Q_5^{\min} = -30.00$ MVAr] an instance of the optimal power flow problem is proposed. It can be viewed as a complex polynomial optimization problem with five variables $z_1, z_2, z_3, z_4, z_5 \in \mathbb{C}$. Let $\text{mon}(\cdot)$ denote the monomial sparsity induced either by the objective f or by one of the constraints $g_1, \ldots, g_{20}$:

$$\begin{aligned} \text{mon}(f) &= \{(1,2),(1,3),(3,5),(4,5)\}, \\ \text{mon}(g_1) = \text{mon}(g_2) &= \{(1,2),(1,3)\}, \\ \text{mon}(g_3) = \text{mon}(g_4) &= \{(1,2),(2,3),(2,4)\}, \\ \text{mon}(g_5) = \text{mon}(g_6) &= \{(1,3),(2,3),(3,5)\}, \\ \text{mon}(g_7) = \text{mon}(g_8) &= \{(2,4),(4,5)\}, \\ \text{mon}(g_9) = \text{mon}(g_{10}) &= \{(3,5),(4,5)\}, \\ \text{mon}(g_{11}) = \cdots = \text{mon}(g_{20}) &= \emptyset. \end{aligned} \tag{6.5}$$

The monomial sparsity is empty when no two distinct variables appear in one monomial, such as in $g_{11}(z,\bar{z}) = z_1\bar{z}_1 - 0.90$. Otherwise, it corresponds to all the couples of distinct variables that appear in one monomial, such as (z_1, z_2) and (z_1, z_3) in

$g_1(z,\bar{z}) = 8.12 z_1\bar{z}_1 - (2.06-4.64\mathbf{i})z_2\bar{z}_1 - (2.06+4.64\mathbf{i})z_1\bar{z}_2 - (2.00-4.00\mathbf{i})z_3\bar{z}_1 - (2.00+4.00\mathbf{i})z_1\bar{z}_3$. The complex hierarchy with $d_i = 1,\ \forall i \in \{1,2,3,4,5,6,11,12,13,14,15,16\}$, and $d_i = 2,\ \forall i \in \{7,8,9,10,17,18,19,20\}$, yields a global solution. (Second-order constraints are identified using the procedure described in the next section.) With this choice of high order constraints, the relevant graph is illustrated in Figure 4. It is already chordal, and its maximal cliques are {1,2,3} and {2,3,4,5}. We can associate the latter to all high-order constraints since it contains their variables. The globally optimal objective value thus obtained is 946.6 MW (in accordance with [15]) with corresponding decision variable $z = (1.0467+0.0000\mathbf{i}, 0.9550-0.0578\mathbf{i}, 0.9485-0.0533\mathbf{i}, 0.7791+0.6011\mathbf{i}, 0.7362+0.7487\mathbf{i})^T$.

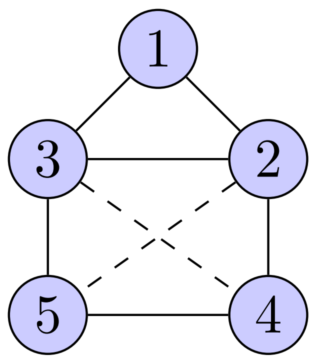

FIG. 4. *Monomial sparsity graph (solid lines) plus correlative sparsity of high order constraints (dashed lines).*

6.2. Updating the relaxation order at each constraint. Consider a polynomial optimization problem with 10,000 constraints, as encountered in Table 1. Say that one has computed the first-order Lasserre relaxation and that it does not yield a global solution. How does one choose the constraints at which to augment the relaxation order? Even if only two constraints require a high order, the combinatorial difficulty is tremendous: there are 49,995,000 combinations to choose from. This section provides one way to choose the high-order constraints. There could be other ways, of course. The proposed approach works well for the optimal power flow problem and is guaranteed to converge globally for any problem, but its performance on other applications is left for future study. We next present the approach when applied to polynomial optimization in real numbers (i.e., $\inf_{x\in\mathbb{R}^n} \sum_\alpha f_\alpha x^\alpha$ s.t. $\sum_\alpha g_{i,\alpha} x^\alpha \geqslant 0,\ i = 1,\dots,m$), but it also applies to complex numbers. We begin by computing a solution y to the moment relaxation with the lowest possible order at each constraint. Next, we do the following: Until a measure can be extracted from a solution y to the moment relaxation:

1. find a closest measure μ to y not necessarily supported on K:

$$\underset{\mu}{\arg\min} \quad \sum_\alpha \left(y_\alpha - \int_{\mathbb{R}^n} x^\alpha d\mu \right)^2; \tag{6.6}$$

2. increment $d_i = d_i + 1$ at the largest mismatch, that is to say,

$$\underset{1\leqslant i\leqslant m}{\arg\max} \left| \sum_\alpha g_{i,\alpha} \left(y_\alpha - \int_{\mathbb{R}^n} x^\alpha d\mu \right) \right|; \tag{6.7}$$

3. compute a solution y to the moment relaxation of order $(d_1, \ldots, d_m)$.

The first step is a priori challenging computationally, so we use a proxy for it. For each clique C_k, we have a set of pseudomoments $(y_{\alpha+\beta})_{|\alpha|,|\beta|=1}$ from which we can extract an eigenvector of highest eigenvalue. This eigenvector u_k is defined up to a sign change when dealing with real numbers since $(-u_k)(-u_k)^T = u_k u_k^T$ (respectively, up to phase shift when dealing with complex numbers since $(e^{\mathbf{i}\theta}u_k)(e^{\mathbf{i}\theta}u_k)^* = u_k u_k^*$). In order to synchronize the eigenvectors among the overlapping cliques, we must therefore choose the signs (respectively, phase shifts), which can be done approximately via convex optimization. Next, we use a least-squares optimization to find a vector $x \in \mathbb{R}^n$ (respectively, $z \in \mathbb{C}^n$) that best matches the signed eigenvectors on each clique (respectively, phased eigenvectors). This approximately provides a closest measure to the pseudomoments, namely the Dirac measure with atom equal to x (respectively, z) and weight equal to 1.

The second step depends on three parameters: a mismatch tolerance $\epsilon > 0$; the number h of largest mismatches considered at each iteration; and an upper bound $\Delta_{\min}^{\max}$ on the difference between maximum and minimum relaxation orders. (In the experiments of Table 1, we take $\epsilon = 1$ MVA, $h = 2$, and $\Delta_{\min}^{\max} = 2$.) First, assume that there are constraints with a mismatch greater than ϵ, i.e., $|L_y(g_i) - g_i(x)| > \epsilon$ (respectively, $|L_y(g_i) - g_i(z, \bar{z})| > \epsilon$), and whose relaxation orders have not yet been increased. Then increment the order at those that have the h largest mismatches and consider a set of cliques which contain all their variables; increment also the order of any constraint whose variables are included in those cliques. For all other constraints, keep the same relaxation order unless the bound $\Delta_{\min}^{\max}$ is violated, in which case increment all those with the smallest order. Second, assume that all the constraints with a mismatch greater than ϵ have already been augmented. Then, among those, consider the h largest mismatches, and repeat the above procedure. Third, if all the mismatches are below ϵ, then the point x (respectively, z) is feasible up to ϵ and globally optimal.

7. Block diagonal Lasserre hierarchy.

7. Block diagonal Lasserre hierarchy. Finally, we exhibit a block diagonal structure of the Lasserre hierarchy in the presence of symmetries. We begin with an illustrative example.

Example 7.1. In [14, WB2, $V_2^{\max} = 1.022$ p.u.], an instance of the optimal power flow is proposed. It yields the following complex polynomial optimization problem:

$$\inf_{z_1, z_2 \in \mathbb{C}} \quad 8|z_1 - z_2|^2$$

$$\text{s.t.} \quad \begin{cases} 0.9025 \leqslant |z_1|^2 \leqslant 1.1025, \\ 0.9025 \leqslant |z_2|^2 \leqslant 1.0568, \\ (2+10\mathbf{i})z_1\bar{z}_2 + (2-10\mathbf{i})z_2\bar{z}_1 - 4|z_2|^2 = 350, \\ (-10+2\mathbf{i})z_1\bar{z}_2 + (-10-2\mathbf{i})z_2\bar{z}_1 + 20|z_2|^2 = -350. \end{cases}$$

Notice that if (z_1, z_2) is a feasible point, then so is $(e^{\mathbf{i}\theta} z_1, e^{\mathbf{i}\theta} z_2)$ for all $\theta \in \mathbb{R}$. When converted to real numbers $z_1 := x_1 + x_3\mathbf{i}$ and $z_2 := x_2 + x_4\mathbf{i}$, it yields

$$
\inf_{x_1,x_2,x_3,x_4 \in \mathbb{R}} \quad 8(x_1 - x_2)^2 + 8(x_3 - x_4)^2
$$

$$
\text{s.t.} \quad \left\{ \begin{array}{rcr}
0.9025 \leqslant x_1^2 + x_3^2 \leqslant 1.1025, & & \\
0.9025 \leqslant x_2^2 + x_4^2 \leqslant 1.0568, & & \\
4x_1x_2 + 4x_3x_4 + 20x_1x_4 - 20x_3x_2 - 4x_2^2 + 4x_4^2 & = & 350, \\
-20x_1x_2 - 20x_3x_4 + 4x_1x_4 - 4x_3x_2 + 20x_2^2 + 20x_4^2 & = & -350.
\end{array} \right.
$$

Notice that if (x_1, x_2, x_3, x_4) is a feasible point, then so is $(-x_1, -x_2, -x_3, -x_4)$.

The above symmetries allow one to cancel many terms in the Lasserre hierarchy at no loss of bound quality. We next illustrate this. The real and complex hierarchies yield the same bounds at the first, second, and third orders (888.1, 894.3, and 905.7 MW, respectively). This is in accordance with [54, Table I]. The rank of the real and complex moment matrices guarantees that global convergence is reached at the third order. At that order, one can set to zero the following terms in the moment matrices (the bullets represent potentially nonzero terms):

Complex moment matrix:

$$
\begin{array}{r|cccccccccc}
 & 1 & \bar{z}_1 & \bar{z}_2 & \bar{z}_1\bar{z}_1 & \bar{z}_1\bar{z}_2 & \bar{z}_2\bar{z}_2 & \bar{z}_1\bar{z}_1\bar{z}_1 & \bar{z}_1\bar{z}_1\bar{z}_2 & \bar{z}_1\bar{z}_2\bar{z}_2 & \bar{z}_2\bar{z}_2\bar{z}_2 \\
1 & \bullet & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
z_1 & 0 & \bullet & \bullet & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
z_2 & 0 & \bullet & \bullet & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
z_1z_1 & 0 & 0 & 0 & \bullet & \bullet & \bullet & 0 & 0 & 0 & 0 \\
z_1z_2 & 0 & 0 & 0 & \bullet & \bullet & \bullet & 0 & 0 & 0 & 0 \\
z_2z_2 & 0 & 0 & 0 & \bullet & \bullet & \bullet & 0 & 0 & 0 & 0 \\
z_1z_1z_1 & 0 & 0 & 0 & 0 & 0 & 0 & \bullet & \bullet & \bullet & \bullet \\
z_1z_1z_2 & 0 & 0 & 0 & 0 & 0 & 0 & \bullet & \bullet & \bullet & \bullet \\
z_1z_2z_2 & 0 & 0 & 0 & 0 & 0 & 0 & \bullet & \bullet & \bullet & \bullet \\
z_2z_2z_2 & 0 & 0 & 0 & 0 & 0 & 0 & \bullet & \bullet & \bullet & \bullet
\end{array}
$$

Real moment matrix:

	1	x_1	x_2	x_3	x_4	x_1x_1	x_1x_2	x_1x_3	x_1x_4	x_2x_2	x_2x_3	x_2x_4	x_3x_3	x_3x_4	x_4x_4	$x_1x_1x_1$	$x_1x_1x_2$	$x_1x_1x_3$	$x_1x_1x_4$	$x_1x_2x_2$	$x_1x_2x_3$	$x_1x_2x_4$	$x_1x_3x_3$	$x_1x_3x_4$	$x_1x_4x_4$	$x_2x_2x_2$	$x_2x_2x_3$	$x_2x_2x_4$	$x_2x_3x_3$	$x_2x_3x_4$	$x_2x_4x_4$	$x_3x_3x_3$	$x_3x_3x_4$	$x_3x_4x_4$	$x_4x_4x_4$
1	•	0	0	0	0	•	•	•	•	•	•	•	•	•	•	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
x_1	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
x_2	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
x_3	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
x_4	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
x_1x_1	•	0	0	0	0	•	•	•	•	•	•	•	•	•	•	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
x_1x_2	•	0	0	0	0	•	•	•	•	•	•	•	•	•	•	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
x_1x_3	•	0	0	0	0	•	•	•	•	•	•	•	•	•	•	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
x_1x_4	•	0	0	0	0	•	•	•	•	•	•	•	•	•	•	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
x_2x_2	•	0	0	0	0	•	•	•	•	•	•	•	•	•	•	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
x_2x_3	•	0	0	0	0	•	•	•	•	•	•	•	•	•	•	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
x_2x_4	•	0	0	0	0	•	•	•	•	•	•	•	•	•	•	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
x_3x_3	•	0	0	0	0	•	•	•	•	•	•	•	•	•	•	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
x_3x_4	•	0	0	0	0	•	•	•	•	•	•	•	•	•	•	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
x_4x_4	•	0	0	0	0	•	•	•	•	•	•	•	•	•	•	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$x_1x_1x_1$	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
$x_1x_1x_2$	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
$x_1x_1x_3$	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
$x_1x_1x_4$	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
$x_1x_2x_2$	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
$x_1x_2x_3$	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
$x_1x_2x_4$	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
$x_1x_3x_3$	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
$x_1x_3x_4$	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
$x_1x_4x_4$	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
$x_2x_2x_2$	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
$x_2x_2x_3$	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
$x_2x_2x_4$	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
$x_2x_3x_3$	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
$x_2x_3x_4$	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
$x_2x_4x_4$	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
$x_3x_3x_3$	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
$x_3x_3x_4$	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
$x_3x_4x_4$	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
$x_4x_4x_4$	0	•	•	•	•	0	0	0	0	0	0	0	0	0	0	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•

Note that there are four diagonal blocks in the complex moment matrix, while there are two diagonal blocks in the real moment matrix (after permutation).

The above example illustrates a block structure in the Lasserre hierarchy which can be generalized. At order d, there are $d+1$ blocks in the complex moment matrix, as opposed to 2 blocks in the real moment matrix (regardless of the relaxation order). This can easily be deduced from the following considerations.

If $\sigma(z,\bar{z}) = \sum_{\alpha,\beta} \sigma_{\alpha,\beta} z^\alpha \bar{z}^\beta$ is a Hermitian sum-of-squares, then its associated *balanced form*, i.e., $\sum_{|\alpha|=|\beta|} \sigma_{\alpha,\beta} z^\alpha \bar{z}^\beta$, is also a Hermitian sum-of-squares. This follows readily from the *homogeneous decomposition*

$$\sum_{|\alpha|=|\beta|} \sigma_{\alpha,\beta} z^\alpha \bar{z}^\beta \quad = \quad \sum_{h=0}^{d} \; \sum_{k} \left| \sum_{|\alpha|=h} p_{k,\alpha} z^\alpha \right|^2, \tag{7.1}$$

where $\sigma(z,\bar{z}) = \sum_k |\sum_{|\alpha|\leqslant d} p_{k,\alpha} z^\alpha|^2$. The homogeneous decomposition becomes relevant when the objective f and constraints $g_1,\dots,g_m$ are themselves balanced forms. Indeed, we then get the following property. If $(\sigma_0,\dots,\sigma_m)$ is a feasible point of the relaxation of order d, that is to say $f - \lambda = \sigma_0 + \sigma_1 g_1 + \cdots + \sigma_m g_m$, then the associated balanced forms also constitute a feasible point. In other words, the nonbalanced terms $\sigma_{\alpha,\beta} z^\alpha \bar{z}^\beta$, $|\alpha| \neq |\beta|$, can be discarded. The homogeneous decomposition accounts for the $(d+1)$-block diagonal structure of the complex Lasserre hierarchy.

Likewise, if $\sigma(x) = \sum_\alpha \sigma_\alpha x^\alpha$ is a real sum-of-squares, then its associated *even form*, i.e., $\sum_{|\alpha| \text{ even}} \sigma_\alpha x^\alpha$, is also a real sum-of-squares. This follows readily from the *even/odd decomposition*

$$\sum_{|\alpha|\ \text{even}} \sigma_\alpha x^\alpha = \sum_k \left(\sum_{|\alpha|\ \text{even}} p_{k,\alpha} x^\alpha \right)^2 + \left(\sum_{|\alpha|\ \text{odd}} p_{k,\alpha} x^\alpha \right)^2, \tag{7.2}$$

where $\sigma(x) = \sum_k \left(\sum_\alpha p_{k,\alpha} x^\alpha\right)^2$. The even/odd decomposition becomes relevant when the objective f and constraints $g_1, \dots, g_m$ are themselves even forms. If $(\sigma_0, \dots, \sigma_m)$ is a feasible point of the relaxation of order d, then the associated even forms also constitute a feasible point. The even/odd decomposition accounts for the 2-block diagonal structure of the real Lasserre hierarchy.

We next analyse the above results via a dual perspective based on measure theory. We illustrate it with Figure 5, which we have generated with MATLAB and Paint.

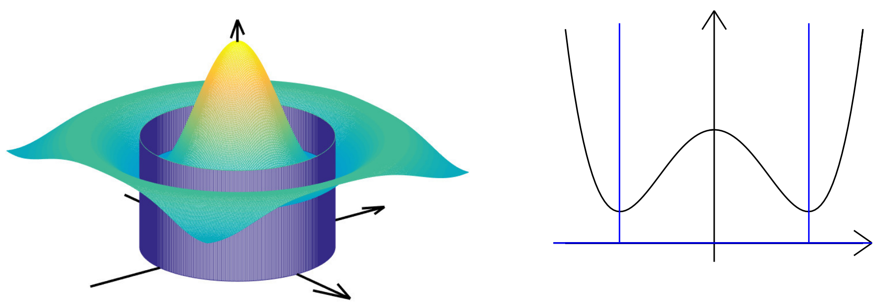

FIG. 5. *Invariant measure on the complex plane versus on the real line.*

On the left, we seek to minimize a function $f(z)$ of one unconstrained complex variable. The two bottom axes correspond to the real and imaginary parts of the variable, respectively. The objective function is invariant under the action of the torus $\mathbb{T}$ (i.e., $f(e^{i\theta} z) = f(z)$). Thus, one may seek a measure that is also invariant, instead of looking for a Dirac measure. The cylinder represents an invariant measure μ minimizing $\int f d\mu$. Such an invariant measure satisfies $\int z^\alpha \bar{z}^\beta d\mu = 0$ if $|\alpha| \neq |\beta|$. In order to seek such a measure, one may therefore set the corresponding pseudomoments $y_{\alpha,\beta}$ to zero in the complex moment matrix.

Seeking an invariant measure under the action of a finite group in the real Lasserre hierarchy was proposed in [90]. Above, we applied this idea to the complex Lasserre hierarchy for a compact group (the torus). We next apply it to the real Lasserre hierarchy regarding a symmetry not considered in [90], namely $\{-1, +1\}$. In particular, that work does not exhibit a block diagonal structure of the Lasserre hierarchy. Note also that a general theory of invariance in sums of squares was developed in [22]. However, the two symmetries that we consider are not studied.

On the right of Figure 5, we seek to minimize a function $f(x)$ of one unconstrained real variable. It is invariant under the action of the symmetry group $\{-1, +1\}$ (i.e.,

$f(-x) = f(x)$). Thus one may seek a measure that is also invariant. The vertical lines represent an invariant measure μ minimizing $\int f d\mu$. Such an invariant measure satisfies $\int x^\alpha d\mu = 0$ if $|\alpha|$ is odd. In order to seek such a measure, one may therefore set the corresponding pseudomoments y_α to zero in the real moment matrix.

Appendix A. Proof of Theorem 5.1. ($\Longleftarrow$) The positive semidefinite moment matrix of rank $r := \text{rank} M_{d+d_K}(y)$ can be factorized in Grammian form as $y_{\alpha,\beta} = x_\alpha^* x_\beta$, for all $|\alpha|, |\beta| \leqslant d + d_K$, where $x_\alpha \in \mathbb{C}^r$. This leads us to consider the finite-dimensional Hilbert space $\mathbb{C}^r = \text{span}(x_\alpha)_{|\alpha| \leqslant d+d_K} = \text{span}(x_\alpha)_{|\alpha|\leqslant d}$, the last equality being a consequence of $\text{rank} M_{d+d_K}(y) = \text{rank} M_d(y)$.[9] Since $d_K \geqslant 1$, it is true in particular that $\mathbb{C}^r = \text{span}(x_\alpha)_{|\alpha|\leqslant d+d_K-1}$. On this space, we define the *shift operators* $T_1, \ldots, T_n$ as

$$(\text{A.1}) \qquad \begin{array}{rccc} T_k : & \mathbb{C}^r & \longrightarrow & \mathbb{C}^r, \\ & \displaystyle\sum_{|\alpha| \leqslant d+d_K-1} u_\alpha x_\alpha & \longmapsto & \displaystyle\sum_{|\alpha| \leqslant d+d_K-1} u_\alpha x_{\alpha+e_k}, \end{array}$$

where e_k is the row vector of size n that contains only zeros apart from 1 in position k. In order to make sure that the shifts are well defined, we must check that each element of $\mathbb{C}^r$ has a unique image by T_k. In other words, given two sets of coefficients $(u_\alpha)_{|\alpha|\leqslant d+d_K-1}$ and $(v_\alpha)_{|\alpha|\leqslant d+d_K-1}$, if $\sum_{|\alpha|\leqslant d+d_K-1} u_\alpha x_\alpha = \sum_{|\alpha|\leqslant d+d_K-1} v_\alpha x_\alpha$, then it must be that $\sum_{|\alpha|\leqslant d+d_K-1} u_\alpha x_{\alpha+e_k} = \sum_{|\alpha|\leqslant d+d_K-1} v_\alpha x_{\alpha+e_k}$. Indeed, this is true because

$$(\text{A.2}) \qquad \left\| \sum_{|\alpha|\leqslant d+d_K-1} (u_\alpha - v_\alpha) x_{\alpha+e_k} \right\| \leqslant R \left\| \sum_{|\alpha|\leqslant d+d_K-1} (u_\alpha - v_\alpha) x_\alpha \right\|,$$

where $\|x\| := \sqrt{x^* x}$ denotes the 2-norm of a vector $x \in \mathbb{C}^r$. We now explain why the above inequality holds. Given some complex numbers $(w_\alpha)_{|\alpha|\leqslant d+d_K-1}$, the positivity of the localizing matrix associated to the ball constraint, i.e., $M_{d+d_K-1}[(R^2 - |z_1|^2 - \ldots - |z_n|^2)y] \succcurlyeq 0$, implies that

$$(\text{A.3}) \qquad \begin{aligned} \left\| \sum_{|\alpha|\leqslant d+d_K-1} w_\alpha x_{\alpha+e_k} \right\|^2 &= \sum_{|\alpha|,|\beta|\leqslant d+d_K-1} \overline{w}_\alpha w_\beta \; x_{\alpha+e_k}^* x_{\beta+e_k} \\ &= \sum_{|\alpha|,|\beta|\leqslant d+d_K-1} \overline{w}_\alpha w_\beta \; y_{\alpha+e_k,\beta+e_k} \\ &\leqslant R^2 \sum_{|\alpha|,|\beta|\leqslant d+d_K-1} \overline{w}_\alpha w_\beta \; y_{\alpha+e_k,\beta+e_k} \end{aligned}$$

$$(\text{A.4}) \qquad \leqslant R^2 \left\| \sum_{|\alpha|\leqslant d+d_K-1} w_\alpha x_{\alpha+e_k} \right\|^2.$$

We now proceed to show that $T_1, \ldots, T_n, T_1^*, \ldots, T_n^*$ commute pairwise. When $\text{rank} M_{d+d_K}(y) = \text{rank} M_d(y) = 1$, this is trivial since $T_1, \ldots, T_n$ are then a set of

[9] Indeed, consider $d < |\beta| \leqslant d + d_K$. The column of $M_{d+d_K}(y)$ indexed by β is a linear combination of the columns of $M_{d+d_K}(y)$ indexed by α with $|\alpha| \leqslant d$. In other words, there exist some complex numbers $(c_\alpha)_{|\alpha|\leqslant d}$ such that $y_{\gamma,\beta} = \sum_{|\alpha|\leqslant d} c_\alpha y_{\gamma,\alpha}$, $\forall |\gamma| \leqslant d + d_K$. As a result, $x_\gamma^* x_\beta = \sum_{|\alpha|\leqslant d} c_\alpha x_\gamma^* x_\alpha$, $\forall|\gamma| \leqslant d + d_K$. To conclude, $x_\beta - \sum_{|\alpha|\leqslant d} c_\alpha x_\alpha \in (\text{span}(x_\alpha)_{|\alpha|\leqslant d+d_K})^\perp \cap (\text{span}(x_\alpha)_{|\alpha|\leqslant d+d_K}) = \{0\}$, where $(\cdot)^\perp$ stands for orthogonal.

complex numbers. Otherwise, we use that $d_K \geqslant 2$ to prove that $T_1, \ldots, T_n$ commute pairwise. Indeed, for all $|\alpha| \leqslant d \leqslant d + d_K - 2$, it holds that $T_i T_j x_\alpha = T_i x_{\alpha+e_j} = x_{\alpha+e_j+e_i} = x_{\alpha+e_i+e_j} = T_j T_i x_\alpha$. As a result, given $u \in \mathbb{C}^r$, say with decomposition $u = \sum_{|\alpha|\leqslant d} u_\alpha x_\alpha$, we have that

$$T_i T_j u = T_i T_j \left(\sum_{|\alpha|\leqslant d} u_\alpha x_\alpha \right) = \sum_{|\alpha|\leqslant d} u_\alpha T_i T_j x_\alpha = \sum_{|\alpha|\leqslant d} u_\alpha T_j T_i x_\alpha = T_j T_i u. \tag{A.5}$$

We go on to prove the stronger property that $T_1, \ldots, T_n, T_1^*, \ldots, T_n^*$ commute pairwise. Consider $u, v, w \in \mathbb{C}^r$ admitting the following decompositions:

$$\begin{aligned} u &= \sum_{|\alpha|\leqslant d} u_\alpha x_\alpha, \quad \vec{u} := (u_\alpha)_{|\alpha|\leqslant d}, \\ v &= \sum_{|\alpha|\leqslant d} v_\alpha x_\alpha, \quad \vec{v} := (v_\alpha)_{|\alpha|\leqslant d}, \\ w &= \sum_{|\alpha|\leqslant d} w_\alpha x_\alpha, \quad \vec{w} := (w_\alpha)_{|\alpha|\leqslant d}. \end{aligned} \tag{A.6}$$

A simple computation (details below) yields that, for all $1 \leqslant i < j \leqslant n$,

$$\begin{aligned} \begin{pmatrix} u \\ v \\ w \end{pmatrix}^* \begin{pmatrix} I & T_i^* & T_j^* \\ T_i & T_i^* T_i & T_j^* T_i \\ T_j & T_i^* T_j & T_j^* T_j \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} \quad &= \quad \cdots \\ \begin{pmatrix} \vec{u} \\ \vec{v} \\ \vec{w} \end{pmatrix}^* \begin{pmatrix} M_d(y) & M_d(z_i y) & M_d(z_j y) \\ M_d(\bar{z}_i y) & M_d(|z_i|^2 y) & M_d(z_j \bar{z}_i y) \\ M_d(\bar{z}_j y) & M_d(z_i \bar{z}_j y) & M_d(|z_j|^2 y) \end{pmatrix} \begin{pmatrix} \vec{u} \\ \vec{v} \\ \vec{w} \end{pmatrix}&. \end{aligned} \tag{A.7}$$

The "joint hyponormality of the shifts" condition then implies that

$$\begin{pmatrix} I & T_i^* & T_j^* \\ T_i & T_i^* T_i & T_i^* T_j \\ T_j & T_i^* T_j & T_j^* T_j \end{pmatrix} \succcurlyeq 0. \tag{A.8}$$

We now dwell on the computational details. We will use the notation $T^\alpha := T_1^{\alpha_1} \ldots T_n^{\alpha_n}$ for convenience. For any complex polynomial $g \in \mathbb{C}[z, \bar{z}]$, it holds that $u^* g(T^*, T) v = \vec{u}^* M_{d-d_K}(gy) \vec{v}$ since

$$\begin{aligned} u^* g(T^*, T) v &= u^* \left(\sum_{\gamma,\delta} g_{\gamma,\delta} (T^*)^\gamma T^\delta \right) v \\ &= \sum_{\gamma,\delta} g_{\gamma,\delta} \; (T^\gamma u)^* T^\delta v \\ &= \sum_{\alpha,\beta,\gamma,\delta} \overline{u}_\alpha v_\beta \; g_{\gamma,\delta} \; (T^\gamma x_\alpha)^* T^\delta x_\beta \\ &= \sum_{\alpha,\beta,\gamma,\delta} \overline{u}_\alpha v_\beta \; g_{\gamma,\delta} \; x^*_{\alpha+\gamma} x_{\beta+\delta} \end{aligned} \tag{A.9}$$

$$\begin{aligned} &= \sum_{\alpha,\beta} \overline{u}_\alpha v_\beta \left(\sum_{\gamma,\delta} g_{\gamma,\delta} \; y_{\alpha+\gamma,\beta+\delta} \right) \\ &= \vec{u}^* M_d(gy) \vec{v}. \end{aligned} \tag{A.10}$$

Let us pursue the proof: in accordance with section 5, it holds that $T_1, \dots, T_n$ are jointly hyponormal and that $[T_i, T_j] = T_i^* T_j - T_i T_j^* = 0$. Together with the fact that $T_1, \dots, T_n$ commute pairwise, we deduce that $T_1, \dots, T_n, T_1^*, \dots, T_n^*$ commute pairwise. The operators must then be simultaneously diagonalizable. In other words, there exists a unitary matrix P such that $T_k = P D_k P^*,\ k = 1, \dots, n$, where $D_k = \mathrm{diag}(d_{k1}, \dots, d_{kr})$ is a diagonal matrix. For all $|\alpha|, |\beta| \leqslant d + d_K$, we thus have

$$\begin{aligned} y_{\alpha,\beta} &= x_\alpha^* x_\beta \\ &= (T^\alpha x_0)^* (T^\beta x_0) \\ &= x_0^* (T^\alpha)^* T^\beta x_0 \\ &= x_0^* (P D^\alpha P^*)^* P D^\beta P^* x_0 \\ &= x_0^* P \overline{D}^\alpha P^* P D^\beta P^* x_0 \\ &= x_0^* P \overline{D}^\alpha D^\beta P^* x_0 \\ &= x_0^* \left(\sum_{j=1}^r p_j \overline{d}_j^\alpha d_j^\beta p_j^* \right) x_0 \\ &= \sum_{j=1}^r x_0^* p_j p_j^* x_0 \ \overline{d}_j^\alpha d_j^\beta \\ &= \sum_{j=1}^r |x_0^* p_j|^2 \ \overline{d}_j^\alpha d_j^\beta, \end{aligned} \tag{A.11}$$

where $P =: (p_1, \dots, p_r)$ denote the columns of P and $d_j := (d_{1j}, \dots, d_{nj})$. As a result, eigenvalues of the shift operators yield the support of a measure, and their eigenvectors yield the weights of a measure. Precisely, the measure $\mu = \sum_{j=1}^r |x_0^* p_j|^2 \ \delta_{\overline{d}_j}$ satisfies $y_{\alpha,\beta} = \int_{\mathbb{C}^n} z^\alpha \bar{z}^\beta d\mu$ for all $|\alpha|, |\beta| \leqslant d + d_K$. In addition, the atoms are distinct and the weights are positive because $r = \mathrm{rank} M_d(y)$. Finally, the measure is supported on K because

$$\begin{aligned} g_i(\overline{d}_j, d_j) &= p_j^* p_j \sum_{\gamma,\delta} g_{i,\gamma,\delta} \ \overline{d}_j^\gamma d_j^\delta \qquad (p_j^* p_j = 1) \\ &= \sum_{\gamma,\delta} g_{i,\gamma,\delta} \ (d_j^\gamma p_j)^* (d_j^\delta p_j) \\ &= \sum_{\gamma,\delta} g_{i,\gamma,\delta} \ (T^\gamma p_j)^* (T^\delta p_j) \qquad \left(\text{let } p_j =: \sum_{|\alpha| \leqslant d} p_\alpha x_\alpha \right) \\ &= \sum_{|\alpha|,|\beta| \leqslant d} \overline{p}_\alpha p_\beta \left(\sum_{\gamma,\delta} g_{i,\gamma,\delta} \ (T^\gamma x_\alpha)^* (T^\delta x_\beta) \right) \\ &= \sum_{|\alpha|,|\beta| \leqslant d} \overline{p}_\alpha p_\beta \left(\sum_{\gamma,\delta} g_{i,\gamma,\delta} \ x_{\alpha+\gamma}^* x_{\beta+\delta} \right) \\ &= \sum_{|\alpha|,|\beta| \leqslant d} \overline{p}_\alpha p_\beta \left(\sum_{\gamma,\delta} g_{i,\gamma,\delta} \ y_{\alpha+\gamma,\beta+\delta} \right) \ \geqslant \ 0. \end{aligned}$$

The above inequality is a consequence of $M_{d+d_K-k_i}(g_i y) \succcurlyeq 0$ and $d \leqslant d + d_K - k_i$.

($\Longrightarrow$) Consider the natural extension given by $y_{\alpha,\beta} = \int_{\mathbb{C}^n} z^\alpha \bar{z}^\beta d\mu$ for all $d < |\alpha|, |\beta| \leqslant d + d_K$. The positivity of the moment matrix follows from the positivity of the weights of the atomic measure. The positivity of the localizing matrices follows from the inclusion of the support of the measure in K. The rank is preserved because the rank of the moment matrix cannot exceed the number of atoms. Finally, we have

$$\begin{pmatrix} \vec{u} \\ \vec{v} \\ \vec{w} \end{pmatrix}^* \begin{pmatrix} M_d(y) & M_d(z_i y) & M_d(z_j y) \\ M_d(\bar{z}_i y) & M_d(|z_i|^2 y) & M_d(z_j \bar{z}_i y) \\ M_d(\bar{z}_j y) & M_d(z_i \bar{z}_j y) & M_d(|z_j|^2 y) \end{pmatrix} \begin{pmatrix} \vec{u} \\ \vec{v} \\ \vec{w} \end{pmatrix} = \cdots \tag{A.12}$$

$$\int_{\mathbb{C}^n} |u(z) + z_i v(z) + z_j w(z)|^2 \, d\mu \geqslant 0,$$

where $u(z) := \sum_{|\alpha| \leqslant d} u_\alpha z^\alpha$, $v(z) := \sum_{|\alpha| \leqslant d} v_\alpha z^\alpha$, and $w(z) := \sum_{|\alpha| \leqslant d} w_\alpha z^\alpha$.

Appendix B. Proof of Theorem 5.2. ($\Longrightarrow$) This part is identical to the proof of Theorem 5.1.

($\Longleftarrow$) Just like in the proof of Theorem 5.1, it holds that $T_1, \ldots, T_n$ are pairwise commuting. There are two points that need to be addressed: (1) the existence of the shift operators and (2) the pairwise commutativity of the operators $T_1, \ldots, T_n, T_1^*, \ldots, T_n^*$. To address them, we make use of well-known properties on shift operators (namely unitary and self-adjoint; see [10, p. 319], for instance).

• *K contains the constraints* $|z_k|^2 = 1$, $k = 1, \ldots, n$: The localizing matrix associated to $|z_k|^2 = 1$ is equal to zero, that is, $M_{d+d_K-1}[(1 - |z_k|^2) y] = 0$. As a result, for all complex numbers $(w_\alpha)_{|\alpha| \leqslant d+d_K-1}$, it holds that

$$\left\| \sum_{|\alpha| \leqslant d+d_K-1} w_\alpha x_{\alpha+e_k} \right\| = \left\| \sum_{|\alpha| \leqslant d+d_K-1} w_\alpha x_\alpha \right\|. \tag{B.1}$$

The shifts are thus well defined. In addition, for all $|\alpha|, |\beta| \leqslant d$, we have that

$$x_\alpha^* T_k^* T_k x_\beta = (T_k x_\alpha)^* (T_k x_\beta) = x_{\alpha+e_k}^* x_{\beta+e_k} = y_{\alpha+e_k, \beta+e_k} = y_{\alpha,\beta} = x_\alpha^* x_\beta. \tag{B.2}$$

As a result, given $u \in \mathbb{C}^r$, say with decomposition $u = \sum_{|\alpha| \leqslant d} u_\alpha x_\alpha$, we have that

$$u^* T_k^* T_k u = \sum_{|\alpha|, |\beta| \leqslant d} \overline{u}_\alpha u_\beta x_\alpha^* T_k^* T_k x_\beta = \sum_{|\alpha|, |\beta| \leqslant d} \overline{u}_\alpha u_\beta x_\alpha^* x_\beta = u^* u. \tag{B.3}$$

Hence $T_k^* T_k$ is the identity matrix; in other words, the shift operators are unitary. This means that $(T_1, \ldots, T_n, T_1^*, \ldots, T_n^*) = (T_1, \ldots, T_n, T_1^{-1}, \ldots, T_n^{-1})$ is a pairwise commuting tuple of operators. Indeed, if two invertible square matrices A and B commute, so do A^{-1} and B^{-1} (since $A^{-1}B^{-1}ABB^{-1}A^{-1} = A^{-1}B^{-1}BAB^{-1}A^{-1}$), and so do A and B^{-1} (since $B^{-1}ABB^{-1} = B^{-1}BAB^{-1}$).

• *K contains the constraints* $\mathbf{i} z_k - \mathbf{i} \bar{z}_k = 0$, $k = 1, \ldots, n$: Consider two sets of complex numbers $(u_\alpha)_{|\alpha| \leqslant d+d_K-1}$ and $(v_\alpha)_{|\alpha| \leqslant d+d_K-1}$ and assume that $\sum_{|\alpha| \leqslant d+d_K-1} u_\alpha x_\alpha = \sum_{|\alpha| \leqslant d+d_K-1} v_\alpha x_\alpha$. We next demonstrate that $\sum_{|\alpha| \leqslant d+d_K-1} u_\alpha x_{\alpha+e_k} = \sum_{|\alpha| \leqslant d+d_K-1} v_\alpha x_{\alpha+e_k}$. To do so, define $w_\alpha := u_\alpha - v_\alpha$ for all $|\alpha| \leqslant d + d_K - 1$.

For all $|\beta| \leqslant d + d_K - 1$, it holds that

$$
\begin{aligned}
x_\beta^* \left(\sum_{|\alpha| \leqslant d+d_K-1} w_\alpha x_{\alpha+e_k} \right) &= \sum_{|\alpha| \leqslant d+d_K-1} w_\alpha \; x_\beta^* x_{\alpha+e_k} \\
&= \sum_{|\alpha| \leqslant d+d_K-1} w_\alpha \; y_{\beta,\alpha+e_k} \\
&= \sum_{|\alpha| \leqslant d+d_K-1} w_\alpha \; y_{\beta+e_k,\alpha} \\
&= x_{\beta+e_k}^* \left(\sum_{|\alpha| \leqslant d+d_K-1} w_\alpha x_\alpha \right) = 0.
\end{aligned} \tag{B.4}
$$

Since $\mathbb{C}^r = \mathrm{span}(x_\alpha)_{|\alpha| \leqslant d+d_K-1}$, we conclude that $\sum_{|\alpha| \leqslant d+d_K-1} w_\alpha x_{\alpha+e_k} = 0$.

Moving on to the latter part of the proof, for all $|\alpha|, |\beta| \leqslant d$, we have that

$$x_\alpha^* T_k^* x_\beta = (T_k x_\alpha)^* x_\beta = x_{\alpha+e_k}^* x_\beta = y_{\alpha+e_k,\beta} = y_{\alpha,\beta+e_k} = x_\alpha^* x_{\beta+e_k} = x_\alpha^* T_k x_\beta. \tag{B.5}$$

Hence $T_k^* = T_k$; in other words, the shift operators are self-adjoint. This means that $(T_1, \ldots, T_n, T_1^*, \ldots, T_n^*) = (T_1, \ldots, T_n, T_1, \ldots, T_n)$ is pairwise commuting.

Acknowledgments. We wish to thank the anonymous reviewers for their precious time and valuable feedback. Special thanks to Mihai Putinar for the fruitful discussions that helped us to improve this paper. We also wish to thank Jean-Bernard Baillon, Didier Henrion, Jean Bernard Lasserre, Bruno Nazaret, and Markus Schweighofer for their insightful comments.

REFERENCES

[1] *Power Systems Test Case Archive*, University of Washington Department of Electrical Engineering, http://www.ee.washington.edu/research/pstca/.

[2] X. L. A. and S. Ranga, *Szegö polynomials and the truncated trigonometric moment problem*, Ramanujan J., 12 (2006), pp. 461–472.

[3] A. A. Ahmadi and A. Majumdar, *DSOS and SDSOS Optimization: LP and SOCP-based alternatives to sum of squares optimization*, in 48th Annual Conference on Information Sciences and Systems (CISS), 2014.

[4] T. Aittomaki and V. Koivunen, *Beampattern Optimization by Minimization of Quartic Polynomial*, in the 15th IEEE/SP Workshop on Statistical Signal Processing, 2009, pp. 437–440.

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

[6] N. I. Akhiezer and M. Krein, *Some questions in the theory of moments*, Transl. Math. Monogr. 2, 58 (1962), pp. 164–168.

[7] M. Andersen, A. Hansson, and L. Vandenberghe, *Reduced-complexity semidefinite relaxations of optimal power flow problems*, IEEE Trans. Power Syst., 29 (2014), pp. 1855–1863.

[8] A. Athavale, *On joint hyponormality of operators*, Proc. Amer. Math. Soc., 103 (1988), pp. 417–423.

[9] A. Athavale and S. Pederson, *Moment problems and subnormality*, J. Math. Anal. Appl., 146 (1990), pp. 434–441.

[10] A. Atzmon, *A moment problem for positive measures on the unit disc*, Pacific J. Math., 59 (1975), pp. 317–325.

[11] 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.

[12] M. Bakonyi and H. J. Woerdeman, *Matrix Completions, Moments and Sums of Hermitian Squares*, Princeton University Press, Princeton, NJ, 2011.

[13] A. Bandeira, N. Boumal, and A. Singer, *Tightness of the Maximum Likelihood Semidefinite Relaxation for Angular Synchronization*, Math. Program. Ser. A, 163 (2017), pp. 145–167.

[14] W. Bukhsh, A. Grothey, K. McKinnon, and P. Trodden, *Local solutions of the optimal power flow problem*, IEEE Trans. Power Syst., 28 (2013), pp. 4780–4788.

[15] W. Bukhsh, A. Grothey, K. McKinnon, and P. Trodden, *Test Case Archive of Optimal Power Flow (OPF) Problems with Local Optima*, http://www.maths.ed.ac.uk/optenergy/LocalOpt/WB5.html, 2013.

[16] E. J. Candès, Y. C. Eldar, T. Strohmer, and V. Voroninski, *Phase retrieval via matrix completion*, SIAM J. Imaging Sci., 6 (2013), pp. 199–225, https://doi.org/10.1137/110848074.

[17] M. Carpentier, *Contribution à l'Étude du Dispatching Économique*, Bull. de la Soc. Fran. des Élec., 8 (1962), pp. 431–447.

[18] G. Cassier, *Problème des moments sur un compact de Rn et décomposition de polynômes à plusieurs variables*, J. Funct. Anal., 58 (1984), pp. 254–266.

[19] A. Castillo and R. O'Neill, *Survey of Approaches to Solving the ACOPF (OPF Paper* 4*)*, Tech. report, US FERC, 2013.

[20] D. Catlin and J. D'Angelo, *A stabilization theorem for Hermitian forms and applications to holomorphic mappings*, Math. Res. Lett., 3 (1996), pp. 149–166.

[21] C. Chen and P. Vaidyanathan, *MIMO Radar waveform optimization with prior information of the extended target and clutter*, IEEE Trans. Signal Process., 57 (2009), pp. 3533–3544.

[22] J. Cimpric, S. Kuhlmann, and C. Scheiderer, *Sums of squares and moment problems in equivariant situations*, Trans. Amer. Math. Soc., 361 (2009), pp. 735–765.

[23] C. Coffrin, D. Gordon, and P. Scott, *NESTA, the NICTA Energy System Test Case Archive*, preprint, https://arxiv.org/abs/1411.0359, 2016.

[24] C. Coffrin, H. Hijazi, and P. Van Hentenryck, *The QC relaxation: Theoretical and computational results on optimal power flow*, IEEE Trans. Power Syst., 31 (2016), pp. 3008–3018.

[25] R. E. Curto, *Joint hyponormality: A bridge between hyponormality and subnormality*, in Operator Theory: Operator Algebras and Applications, Part 2 (Durham, NH, 1988), Proc. Sympos. Pure Math. 51, Part 2, AMS, Providence, RI, 1990, pp. 69–91.

[26] R. E. Curto and L. A. Fialkow, *Recursiveness, positivity, and truncated moment problems*, Houston J. Math., 17 (1991), pp. 603–635.

[27] R. E. Curto and L. A. Fialkow, *Solution of the truncated complex moment problem for flat data*, Mem. Amer. Math. Soc., 119 (1996), 568.

[28] R. E. Curto and L. A. Fialkow, *The truncated complex K-moment problem*, Trans. Amer. Math. Soc., 352 (2000), pp. 2825–2855.

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

[30] R. E. Curto and M. Putinar, *Nearly subnormal operators and moment problems*, J. Funct. Anal., 2 (1993), pp. 480–497.

[31] R. E. Curto and M. Putinar, *Polynomially hyponormal operators*, in A Glimpse at Hilbert Space Operators, Oper. Theory Adv. Appl. 207, Birkhäuser, Basel, 2010, pp. 195–207.

[32] J. D'Angelo, *Inequalities from Complex Analysis*, Carus Math. Monogr. 28, MAA, Washington, DC, 2002.

[33] J. D'Angelo, *Hermitian analogues of Hilbert's* 17*th problem*, Adv. Math., 226 (2011), pp. 4607–4637.

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

[35] J. D'Angelo and M. Putinar, *Hermitian complexity of real polynomial ideals*, Internat. J. Math., 23 (2012), 1250026.

[36] 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.

[37] J.-P. Gabardo, *Truncated trigonometric moment problems and determinate measures*, J. Math. Anal. Appl., 239 (1999), pp. 349–370.

[38] B. Ghaddar, J. Marecek, and M. Mevissen, *Optimal power flow as a polynomial optimization problem*, IEEE Trans. Power Syst., 31 (2016), pp. 539–546.

[39] M. Grant and S. Boyd, *CVX: MATLAB Software for Disciplined Convex Programming, Version* 2.0. http://cvxr.com/cvx, 2012.

[40] D. Grimm, T. Netzer, and M. Schweighofer, *A note on the representation of positive polynomials with structured sparsity*, Arch. Math., 89 (2007), pp. 399–403.

[41] P. Halmos, *Normal dilations and extensions of operators*, Summa Brasil. Math., 2 (1950), pp. 125–134.

[42] J. Harmouch, H. Khalil, and B. Mourrain, *Structured Low Rank Decomposition of Multivariate Hankel Matrices*, Tech. report hal-01440063, https://hal.inria.fr/hal-01440063, 2017.

[43] D. Henrion and J.-B. Lasserre, *Detecting global optimality and extracting solutions in GloptiPoly*, in Positive Polynomials in Control, Lect. Notes Control Inf. Sci. 312, Springer, Berlin, 2005, pp. 293–310.

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

[45] J. Hilling and A. Sudbery, *The geometric measure of multipartite entanglement and the singular values of a hypermatrix*, J. Math. Phys., 51 (2010), 072102.

[46] I. S. Iohvidov, *Hankel and Toeplitz Matrices and Forms: Algebraic Theory*, Birkhäuser Boston, Boston, 1982.

[47] R. Jabr, *Radial distribution load flow using conic programming*, IEEE Trans. Power Syst., 21 (2006), pp. 1458–1459.

[48] B. Jiang, Z. Li, and S. Zhang, *Approximation methods for complex polynomial optimization*, Comput. Optim. Appl., 59 (2014), pp. 219–248.

[49] B. Jiang, Z. Li, and S. Zhang, *Characterizing real-valued multivariate complex polynomials and their symmetric tensor representations*, SIAM J. Matrix Anal. Appl., 37 (2016), pp. 381–408, https://doi.org/10.1137/141002256.

[50] C. Josz, *Application of Polynomial Optimization to Electricity Transmission Networks* Ph.D. thesis, https://arxiv.org/pdf/1608.03871v1.pdf, 2016.

[51] C. Josz, *Counterexample to Global Convergence of DSOS and SDSOS Hierarchies*, preprint, https://arxiv.org/pdf/1707.02964.pdf, 2017.

[52] C. Josz, S. Fliscounakis, J. Maeght, and P. Panciatici, *AC Power Flow Data in MATPOWER and QCQP Format: iTesla, RTE Snapshots, and PEGASE*, preprint, https://arxiv.org/abs/1603.01533, 2016.

[53] C. Josz and D. Henrion, *Strong duality in Lasserre's hierarchy for polynomial optimization*, Optim. Lett., 10 (2016), pp. 3–10.

[54] C. Josz, J. Maeght, P. Panciatici, and J. Gilbert, *Application of the moment-SOS approach to global optimization of the OPF problem*, IEEE Trans. Power Syst., 30 (2015), pp. 463–470, https://doi.org/10.1109/TPWRS.2014.2320819.

[55] D. P. Kimsey, *The subnormal completion problem in several variables*, J. Math. Anal. Appl., 434 (2016), pp. 1504–1532.

[56] D. P. Kimsey and H. J. Woerdeman, *The truncated matrix-valued K-moment problem on* $\mathbb{R}^d$, $\mathbb{C}^d$, *and* $\mathbb{T}^d$, Trans. Amer. Math. Soc., 365 (2013), pp. 5393–5430.

[57] X. Kuang, B. Ghaddar, J. Naoum-Sawaya, and L. F. Zuluaga, *Alternative LP and SOCP hierarchies for ACOPF problems*, IEEE Trans. Power Syst., 32 (2017), pp. 2828–2836.

[58] S. Kuhlmann and M. Putinar, *Positive polynomials on fibre products*, C. R. Math. Acad. Sci. Paris, 344 (2007), pp. 681–684.

[59] J. Lasserre, K. Toh, and S. Yang, *A bounded degree SOS hierarchy for polynomial optimization*, EURO J. Comput. Optim., 5 (2017), pp. 87–117.

[60] J. B. Lasserre, *Optimisation globale et théorie des moments*, C. R. Acad. Sci. Paris Sér. I Math., 331 (2000), pp. 929–934.

[61] J. B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim., 11 (2001), pp. 796–817, https://doi.org/10.1137/S1052623400366802.

[62] J. B. Lasserre, *Moments, Positive Polynomials and Their Applications*, Imperial College Press Optimization Series 1, Imperial College Press, London, 2010.

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

[64] M. Laurent, *Revisiting two theorems of Curto and Fialkow on moment matrices*, Proc. Amer. Math. Soc., 10 (2005), pp. 2965–2976.

[65] M. Laurent, *Sums of squares, moment matrices and optimization over polynomials*, in Emerging Applications of Algebraic Geometry, IMA Vol. Math. Appl. 149, Springer, New York, 2009, pp. 157–270.

[66] M. Laurent and B. Mourrain, *A generalized flat extension theorem for moment matrices*, Arch. Math. (Basel), 6 (2009), pp. 87–98.

[67] J. Lavaei and S. Low, *Zero duality gap in optimal power flow problem*, IEEE Trans. Power Syst., 27 (2012), pp. 92–107.

[68] B. Lesieutre, D. Molzahn, A. Borden, and C. DeMarco, *Examining the limits of the application of semidefinite programming to power flow problems*, in 49th Annual Allerton Conference on Communication, Control, and Computing, IEEE, Washington, DC, 2011, pp. 28–30.

[69] J. Löfberg, *YALMIP: A toolbox for modeling and optimization in MATLAB*, in IEEE International Symposium on Computer Aided Control System Design, IEEE, Washington, DC, 2004, pp. 284–289.

[70] S. Low, *Convex relaxation of optimal power flow: Parts* I *&* II, IEEE Trans. Control Network Syst., 1 (2014), pp. 15–27.

[71] Z. Luo, W.-K. Ma, A.-C. So, Y. Ye, and S. Zhang, *Semidefinite relaxation of quadratic optimization problems*, IEEE Signal Process. Mag., 27 (2010), pp. 20–34.

[72] B. Maricic, Z.-Q. Luo, and T. Davidson, *Blind constant modulus equalization via convex optimization*, IEEE Trans. Signal Process., 51 (2003), pp. 805–818.

[73] M. Martin and M. Putinar, *Lectures on Hyponormal Operators*, Oper. Theory Adv. Appl. 39, Birkhäuser Verlag, Basel, 1989.

[74] S. McCullough and V. Paulsen, *A note on joint hyponormality*, Proc. Amer. Math. Soc., 107 (1989), pp. 187–195.

[75] D. Molzahn and I. Hiskens, *Moment-based relaxation of the optimal power flow problem*, in 18th Power Systems Computation Conference (PSCC), IEEE, Washington, DC, 2014.

[76] D. Molzahn and I. Hiskens, *Mixed SDP/SOCP moment relaxations of the optimal power flow problem*, in IEEE Eindhoven PowerTech, 2015.

[77] D. Molzahn and I. Hiskens, *Sparsity-exploiting moment-based relaxations of the optimal power flow problem*, IEEE Trans. Power Syst., 30 (2015), pp. 3168–3180.

[78] J. Nie, *Optimality conditions and finite convergence of Lasserre's hierarchy*, Math. Program., 146 (2014), pp. 97–121.

[79] P. Parrilo, *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*, Ph.D. thesis, California Institute of Technology, Pasadena, CA, 2000.

[80] P. Parrilo, *Semidefinite programming relaxations for semialgebraic problems*, Math. Program., 96 (2003), pp. 293–320.

[81] POV-Ray, *Persistence of Vision Raytracer* 3.7.0, http://www.povray.org/, 2013.

[82] M. Putinar, *Sur la complexification du problème des moments*, C. R. Acad. Sci. Paris Sér. I Math., 314 (1992), pp. 743–745.

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

[84] M. Putinar, *On Hermitian polynomial optimization*, Arch. Math., 87 (2006), pp. 41–51.

[85] M. Putinar and C. Scheiderer, *Hermitian algebra on the ellipse*, Illinois J. Math., 56 (2012), pp. 213–220.

[86] M. Putinar and C. Scheiderer, *Quillen property of real algebraic varieties*, Muenster J. Math., 7 (2014), pp. 671–696.

[87] M. Putinar and K. Schmüdgen, *Multivariate determinateness* , Indiana Univ. Math. J., 57 (2008), pp. 2931–2968.

[88] M. L. Quijorna, *Detecting Optimality and Extracting Solutions in Polynomial Optimization with the Truncated GNS Construction*, preprint, https://arxiv.org/abs/1704.02034, 2017.

[89] D. Quillen, *On the representation of Hermitian forms as sums of squares*, Invent. Math., 5 (1968), pp. 237–242.

[90] C. Riener, T. Theobald, L. J. Andrén, and J. B. Lasserre, *Exploiting symmetries in SDP-relaxations for polynomial optimization*, Math. Oper. Res., 38 (2013), pp. 122–141.

[91] W. Rudin, *Real and Complex Analysis*, 3rd ed., McGraw–Hill, New York, 1987.

[92] K. Schmüdgen, *The K-moment problem for semi-algebraic sets*, Math. Ann., 289 (1991), pp. 203–206.

[93] A. Singer, *Angular synchronization by eigenvectors and semidefinite programming*, Appl. Comput. Harmon. Anal., 30 (2011), pp. 20–36.

[94] L. Sorber, M. Van Barel, and L. De Lathauwer, *Unconstrained optimization of real functions in complex variables*, SIAM J. Optim., 22 (2012), pp. 879–898, https://doi.org/10.1137/110832124.

[95] J. Stochel and F. H. Szafraniec, *The complex moment problem and subnormality: A polar decomposition approach*, J. Funct. Anal., 159 (1998), pp. 432–491.

[96] J. Taylor, *Convex Optimization of Power Systems*, Cambridge University Press, Cambridge, UK, 2015.

[97] K.-C. Toh, M. J. Todd, and R. H. Tütüncü, *On the Implementation and Usage of SDPT*3*—a MATLAB Software Package for Semidefinite-Quadratic-Linear Programming, Version* 4.0, Springer, Boston, 2012, pp. 715–754.

[98] O. Toker and H. Ozbay, *On the complexity of purely complex μ computation and related problems in multidimensional systems*, IEEE Trans. Automat. Control, 43 (1998), pp. 409–414.

[99] F. H. Vascilescu, *Subnormality and moment problems*, Extracta Math., 24 (2009), pp. 167–186.

[100] 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, https://doi.org/10.1137/050623802.

[101] T. Weisser, J. Lasserre, and K. Toh, *Sparse-BSOS: A bounded degree SOS hierarchy for large scale polynomial optimization with sparsity*, Math. Program. Comput., 5 (2017), pp. 1–32.

[102] A. P. Y. Zheng, G. Fantuzzi, and A. Papachristodoulou, *Fast ADMM for Sum-of-Squares Programs Using Partial Orthogonality*, preprint, https://arxiv.org/pdf/1708.04174.pdf, 2017.

[103] S. Zagorodnyuk, *On the truncated operator trigonometric moment problem*, Concr. Oper., 2 (2015), pp. 37–46.

[104] 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.