Journal of Optimization Theory and Applications (2022) 192:335–359
https://doi.org/10.1007/s10957-021-01975-z

Exploiting Sparsity in Complex Polynomial Optimization

Jie Wang[1] · **Victor Magron**[2]

Received: 23 March 2021 / Accepted: 6 November 2021 / Published online: 26 November 2021

Abstract
In this paper, we study the sparsity-adapted complex moment-Hermitian sum of squares (moment-HSOS) hierarchy for complex polynomial optimization problems, where the *sparsity* includes correlative sparsity and term sparsity. We compare the strengths of the sparsity-adapted complex moment-HSOS hierarchy with the sparsity-adapted real moment-SOS hierarchy on either randomly generated complex polynomial optimization problems or the AC optimal power flow problem. The results of numerical experiments show that the sparsity-adapted complex moment-HSOS hierarchy provides a trade-off between the computational cost and the quality of obtained bounds for large-scale complex polynomial optimization problems.

Keywords Complex moment-HSOS hierarchy · Correlative sparsity · Term sparsity · Complex polynomial optimization · Optimal power flow

Mathematics Subject Classification 90C23 · 14P10 · 90C22 · 90C26 · 12D15

Communicated by Vaithilingam Jeyakumar.

✉ Victor Magron
vmagron@laas.fr
https://homepages.laas.fr/vmagron

Jie Wang
wangjie212@amss.ac.cn
https://wangjie212.github.io/jiewang

1 Academy of Mathematics and Systems Science, CAS, Beijing, China

2 Laboratory for Analysis and Architecture of Systems, CNRS, Toulouse, France

1 Introduction

In this paper, we consider the following complex polynomial optimization problem (CPOP):

$$(\mathrm{Q}): \quad \begin{cases} \inf_{\mathbf{z}\in\mathbb{C}^n} & f(\mathbf{z},\bar{\mathbf{z}}) := \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}} f_{\boldsymbol{\alpha},\boldsymbol{\beta}}\mathbf{z}^{\boldsymbol{\alpha}}\bar{\mathbf{z}}^{\boldsymbol{\beta}} \\ \text{s.t.} & g_j(\mathbf{z},\bar{\mathbf{z}}) := \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}} g_{j,\boldsymbol{\alpha},\boldsymbol{\beta}}\mathbf{z}^{\boldsymbol{\alpha}}\bar{\mathbf{z}}^{\boldsymbol{\beta}} \geq 0, \quad j = 1,\ldots,m, \\ & h_i(\mathbf{z},\bar{\mathbf{z}}) := \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}} h_{i,\boldsymbol{\alpha},\boldsymbol{\beta}}\mathbf{z}^{\boldsymbol{\alpha}}\bar{\mathbf{z}}^{\boldsymbol{\beta}} = 0, \quad i = 1,\ldots,t, \end{cases} \tag{1.1}$$

where n, m, and t are positive integers, $\bar{\mathbf{z}} := (\bar{z}_1,\ldots,\bar{z}_n)$ stands for the conjugate of complex variables $\mathbf{z} := (z_1,\ldots,z_n)$. The functions $f, g_1,\ldots,g_m, h_1,\ldots,h_t$ are real-valued polynomials, and their coefficients satisfy $f_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \bar{f}_{\boldsymbol{\beta},\boldsymbol{\alpha}}$, $g_{j,\boldsymbol{\alpha},\boldsymbol{\beta}} = \bar{g}_{j,\boldsymbol{\beta},\boldsymbol{\alpha}}$, and $h_{i,\boldsymbol{\alpha},\boldsymbol{\beta}} = \bar{h}_{i,\boldsymbol{\beta},\boldsymbol{\alpha}}$. The feasible set is defined as $\{\mathbf{z}\in\mathbb{C}^n \mid g_j(\mathbf{z},\bar{\mathbf{z}}) \geq 0, j = 1,\ldots,m, h_i(\mathbf{z},\bar{\mathbf{z}}) = 0, i = 1,\ldots,t\}$. For the sake of brevity, we assume that there are only inequality constraints in (1.1) in the rest of this paper. CPOP (1.1) arises naturally from diverse areas, such as imaging science [13], signal processing [2,3,23], automatic control [26], quantum mechanics [15], optimal power flow [5]. By introducing real variables for the real part and the imaginary part of each complex variables, respectively, CPOP (1.1) can be converted into a polynomial optimization problem (POP) involving only real variables.

The moment-sum of squares (SOS) hierarchy (also known as Lasserre's hierarchy) [19], which consists of a sequence of increasingly tight semidefinite relaxations, has become a popular tool to retrieve global optimal values of POPs involving real variables. For convenience, we refer to the moment-SOS hierarchy as the "real hierarchy" in this paper. The real hierarchy can be further adapted as the moment-Hermitian sum of squares (HSOS) hierarchy to handle CPOPs [17], which thereby we refer to as the "complex hierarchy" in this paper. Notice that for a CPOP, one can either apply the real hierarchy after converting it into a real POP or apply the complex hierarchy directly. It is known that the complex hierarchy never produces tighter bounds than the real hierarchy when using the same relaxation order, which is, however, still of interest because of its lower computational complexity as recently shown in [17].

On the other hand, due to the rapidly growing size of semidefinite relaxations, the standard real and complex hierarchies are typically solvable only for modest-size problems on a personal computer. To improve their scalability, it is then crucial to exploit the structure, e.g., sparsity, encoded in the problem data. For the real hierarchy, one can make use of the well-known correlative sparsity [29], or the newly proposed term sparsity [33,34], or the combination of both [35]. Exploiting sparsity in the real hierarchy has been successfully done for many practical applications, including computer arithmetic [20,21], control [28,31], machine learning [10], noncommutative optimization [18,32,36,37], rational function optimization [9], just to name a few. For the complex hierarchy, one can also make use of correlative sparsity; see [17] for an application to the AC optimal power flow problem. The main purpose of this paper is to develop sparsity-adapted complex hierarchies by taking into account term sparsity as for the real hierarchy.

Contribution Our contributions are threefold:

1. We propose sparsity-adapted complex hierarchies based on either term sparsity or correlative-term sparsity for CPOPs, which are indexed by two parameters: the relaxation order d and the sparse order k. The optima of the sparsity-adapted complex hierarchies for a fixed d are proved to converge to the optimum of the dense relaxation or the relaxation exploiting only correlative sparsity with the same relaxation order when the maximal chordal extension is chosen. We also prove that the block structure arising in the sparsity-adapted complex hierarchies is always a refinement of the block structure determined by the sign symmetries of the problem.
2. We propose a minimum initial relaxation step of the sparsity-adapted complex hierarchy. For the AC optimal power flow problem, this new relaxation is able to provide a tighter lower bound than Shor's relaxation and is, depending on the input, possibly much less expensive than the second-order relaxation of the complex hierarchy.
3. We provide a comprehensive comparison on the strengths of the sparsity-adapted real hierarchy and the sparsity-adapted complex hierarchy for CPOPs via numerical experiments. The largest numerical example is an instance of the AC optimal power flow problem involving 2869 complex variables (or 5738 real variables).

Our sparsity-adapted complex hierarchies can be viewed as complex variants of the ones obtained for the real case [33–35]. Throughout the paper, we emphasize in several places the subtle differences between the complex and real settings, in particular to define sparsity-adapted moment/localizing matrices and the connection with sign symmetries when (finite) convergence occurs. We also hope that it is of interest for researchers solving large-scale CPOPs (e.g., AC optimal power flow problems) to have a self-contained paper explaining in detail the construction of term sparsity pattern graphs, as well as the sparsity-adapted semidefinite formulations. Last but not least, we do not pretend that the sparsity-adapted complex hierarchy systematically provides better results than the real hierarchy. It rather provides a trade-off between efficiency and accuracy for large-scale CPOPs.

The rest of the paper is organized as follows. In Sect. 2, we recall some preliminary background. In Sect. 3, we establish the sparsity-adapted complex hierarchies and prove some of their properties. In Sect. 4, a minimum initial relaxation step for the sparsity-adapted complex hierarchy is introduced. Numerical experiments are provided in Sect. 5, and conclusions are provided in Sect. 6.

2 Notation and Preliminaries

Let $\mathbb{N}$ be the set of nonnegative integers. For $n \in \mathbb{N}\backslash\{0\}$, let $[n] := \{1, 2, \dots, n\}$. Let $\mathbf{i}$ be the imaginary unit, satisfying $\mathbf{i}^2 = -1$. Let $\mathbf{z} = (z_1, \dots, z_n)$ (resp. $\mathbf{x} = (x_1, \dots, x_n)$) be a tuple of complex (resp. real) variables and $\bar{\mathbf{z}} = (\bar{z}_1, \dots, \bar{z}_n)$ its conjugate. We denote by $\mathbb{C}[\mathbf{z}] := \mathbb{C}[z_1, \dots, z_n]$, $\mathbb{C}[\mathbf{z}, \bar{\mathbf{z}}] := \mathbb{C}[z_1, \dots, z_n, \bar{z}_1, \dots, \bar{z}_n]$, $\mathbb{R}[\mathbf{x}] := \mathbb{R}[x_1, \dots, x_n]$ the complex polynomial ring in $\mathbf{z}$, the complex polynomial ring in $\mathbf{z}, \bar{\mathbf{z}}$, the real polynomial ring in $\mathbf{x}$, respectively. For $d \in \mathbb{N}$, let $\mathbb{C}_d[\mathbf{z}]$ (resp.

$\mathbb{C}_d[\mathbf{z}, \bar{\mathbf{z}}]$) denote the set of polynomials in $\mathbb{C}[\mathbf{z}]$ (resp. $\mathbb{C}[\mathbf{z}, \bar{\mathbf{z}}]$) of degree no greater than d. A polynomial $f \in \mathbb{C}[\mathbf{z}, \bar{\mathbf{z}}]$ can be written as $f = \sum_{(\boldsymbol{\beta},\boldsymbol{\gamma})\in\mathscr{A}} f_{\boldsymbol{\beta},\boldsymbol{\gamma}} \mathbf{z}^{\boldsymbol{\beta}} \bar{\mathbf{z}}^{\boldsymbol{\gamma}}$ with $\mathscr{A} \subseteq \mathbb{N}^n \times \mathbb{N}^n$ and $f_{\boldsymbol{\beta},\boldsymbol{\gamma}} \in \mathbb{C}$, $\mathbf{z}^{\boldsymbol{\beta}} = z_1^{\beta_1} \cdots z_n^{\beta_n}$, $\bar{\mathbf{z}}^{\boldsymbol{\gamma}} = \bar{z}_1^{\gamma_1} \cdots \bar{z}_n^{\gamma_n}$. The support of f is defined by $\operatorname{supp}(f) = \{(\boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathscr{A} \mid f_{\boldsymbol{\beta},\boldsymbol{\gamma}} \neq 0\}$. The conjugate of f is $\bar{f} = \sum_{(\boldsymbol{\beta},\boldsymbol{\gamma})\in\mathscr{A}} \bar{f}_{\boldsymbol{\beta},\boldsymbol{\gamma}} \mathbf{z}^{\boldsymbol{\gamma}} \bar{\mathbf{z}}^{\boldsymbol{\beta}}$. A polynomial $\sigma = \sum_{(\boldsymbol{\beta},\boldsymbol{\gamma})} \sigma_{\boldsymbol{\beta},\boldsymbol{\gamma}} \mathbf{z}^{\boldsymbol{\beta}} \bar{\mathbf{z}}^{\boldsymbol{\gamma}} \in \mathbb{C}_{2d}[\mathbf{z}, \bar{\mathbf{z}}]$ is called a *Hermitian sum of squares* or an *HSOS* for short if there exist polynomials $f_i \in \mathbb{C}_d[\mathbf{z}]$, $i \in [t]$ such that $\sigma = \sum_{i=1}^t f_i \bar{f}_i$. We will use $|\cdot|$ to denote the cardinality of a set. For $(\boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathbb{N}^n \times \mathbb{N}^n$, $\mathscr{A} \subseteq \mathbb{N}^n \times \mathbb{N}^n$, let $(\boldsymbol{\beta}, \boldsymbol{\gamma}) + \mathscr{A} := \{(\boldsymbol{\beta} + \boldsymbol{\beta}', \boldsymbol{\gamma} + \boldsymbol{\gamma}') \mid (\boldsymbol{\beta}', \boldsymbol{\gamma}') \in \mathscr{A}\}$.

For a positive integer r, the set of $r \times r$ Hermitian matrices is denoted by $\mathbf{H}^r$ and the set of $r \times r$ positive semidefinite (PSD) Hermitian matrices is denoted by $\mathbf{H}^r_+$. Let $A \circ B \in \mathbf{H}^r$ denote the Hadamard product of $A, B \in \mathbf{H}^r$, defined by $[A \circ B]_{ij} = A_{ij} B_{ij}$. For $d \in \mathbb{N}$, let $\mathbb{N}^n_d := \{(\alpha_i)_i \in \mathbb{N}^n \mid \sum_{i=1}^n \alpha_i \leq d\}$ (sorted with respect to the lexicographic order). The set $\{\mathbf{z}^{\boldsymbol{\beta}} \mid \boldsymbol{\beta} \in \mathbb{N}^n_d\}$ is called the standard (complex) *monomial basis* up to degree d. For the sake of convenience, we abuse notation slightly in this paper and use the exponent set $\mathbb{N}^n_d$ to denote the monomial basis.

2.1 The Complex Moment-HSOS Hierarchy

Let $\mathbf{y} = (y_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}\in\mathbb{N}^n} \subseteq \mathbb{R}$ be a sequence indexed by $\boldsymbol{\alpha} \in \mathbb{N}^n$. Let $L^{\mathrm{r}}_{\mathbf{y}} : \mathbb{R}[\mathbf{x}] \to \mathbb{R}$ be the linear functional

$$f = \sum_{\boldsymbol{\alpha}} f_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} \mapsto L^{\mathrm{r}}_{\mathbf{y}}(f) = \sum_{\boldsymbol{\alpha}} f_{\boldsymbol{\alpha}} y_{\boldsymbol{\alpha}}.$$

The *real moment* matrix $\mathbf{M}^{\mathrm{r}}_d(\mathbf{y})$ ($d \in \mathbb{N}$) associated with $\mathbf{y}$ is the matrix with rows and columns indexed by $\mathbb{N}^n_d$ such that

$$\mathbf{M}^{\mathrm{r}}_d(\mathbf{y})_{\boldsymbol{\beta}\boldsymbol{\gamma}} := L^{\mathrm{r}}_{\mathbf{y}}(\mathbf{x}^{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\gamma}}) = y_{\boldsymbol{\beta}+\boldsymbol{\gamma}}, \quad \forall \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}^n_d.$$

Suppose $g = \sum_{\boldsymbol{\alpha}} g_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} \in \mathbb{R}[\mathbf{x}]$. The *real localizing* matrix $\mathbf{M}^{\mathrm{r}}_d(g\mathbf{y})$ associated with g and $\mathbf{y}$ is the matrix with rows and columns indexed by $\mathbb{N}^n_d$ such that

$$\mathbf{M}^{\mathrm{r}}_d(g\,\mathbf{y})_{\boldsymbol{\beta}\boldsymbol{\gamma}} := L^{\mathrm{r}}_{\mathbf{y}}(g\, \mathbf{x}^{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\gamma}}) = \sum_{\boldsymbol{\alpha}} g_{\boldsymbol{\alpha}} y_{\boldsymbol{\alpha}+\boldsymbol{\beta}+\boldsymbol{\gamma}}, \quad \forall \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}^n_d.$$

Now, let $\mathbf{y} = (y_{\boldsymbol{\beta},\boldsymbol{\gamma}})_{(\boldsymbol{\beta},\boldsymbol{\gamma})\in\mathbb{N}^n\times\mathbb{N}^n} \subseteq \mathbb{C}$ be a sequence indexed by $(\boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathbb{N}^n \times \mathbb{N}^n$ and satisfies $y_{\boldsymbol{\beta},\boldsymbol{\gamma}} = \bar{y}_{\boldsymbol{\gamma},\boldsymbol{\beta}}$. Let $L^{\mathrm{c}}_{\mathbf{y}} : \mathbb{C}[\mathbf{z}, \bar{\mathbf{z}}] \to \mathbb{R}$ be the linear functional

$$f = \sum_{(\boldsymbol{\beta},\boldsymbol{\gamma})} f_{\boldsymbol{\beta},\boldsymbol{\gamma}} \mathbf{z}^{\boldsymbol{\beta}} \bar{\mathbf{z}}^{\boldsymbol{\gamma}} \mapsto L^{\mathrm{c}}_{\mathbf{y}}(f) = \sum_{(\boldsymbol{\beta},\boldsymbol{\gamma})} f_{\boldsymbol{\beta},\boldsymbol{\gamma}} y_{\boldsymbol{\beta},\boldsymbol{\gamma}}.$$

The *complex moment* matrix $\mathbf{M}^{\mathrm{c}}_d(\mathbf{y})$ ($d \in \mathbb{N}$) associated with $\mathbf{y}$ is the matrix with rows and columns indexed by $\mathbb{N}^n_d$ such that

$$\mathbf{M}^{\mathrm{c}}_d(\mathbf{y})_{\boldsymbol{\beta}\boldsymbol{\gamma}} := L^{\mathrm{c}}_{\mathbf{y}}(\mathbf{z}^{\boldsymbol{\beta}} \bar{\mathbf{z}}^{\boldsymbol{\gamma}}) = y_{\boldsymbol{\beta},\boldsymbol{\gamma}}, \quad \forall \boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}^n_d.$$

Suppose that $g = \sum_{(\boldsymbol{\beta}',\boldsymbol{\gamma}')} g_{\boldsymbol{\beta}',\boldsymbol{\gamma}'}\mathbf{z}^{\boldsymbol{\beta}'}\bar{\mathbf{z}}^{\boldsymbol{\gamma}'} \in \mathbb{C}[\mathbf{z}, \bar{\mathbf{z}}]$ is a Hermitian polynomial, i.e., $\bar{g} = g$. The *complex localizing* matrix $\mathbf{M}_d^{\mathrm{c}}(g\mathbf{y})$ associated with g and $\mathbf{y}$ is the matrix with rows and columns indexed by $\mathbb{N}_d^n$ such that

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

Both the complex moment matrix and the complex localizing matrix are Hermitian matrices.

Note that a distinguished difference between the real moment matrix and the complex moment matrix is that the former has the Hankel property (i.e., $\mathbf{M}_d^{\mathrm{r}}(\mathbf{y})_{\boldsymbol{\beta}\boldsymbol{\gamma}}$ is a function of $\boldsymbol{\beta} + \boldsymbol{\gamma}$), whereas the latter does not have.

There are two ways to construct a "moment-SOS" hierarchy for CPOP (1.1). The first way is introducing real variables for both real and imaginary parts of each complex variable in (1.1), i.e., letting $z_i = x_i + x_{i+n}\mathbf{i}$ for $i \in [n]$. Then, one can convert CPOP (1.1) to a POP involving only real variables at the price of doubling the number of variables. Therefore, the usual real moment-SOS hierarchy applies to the resulting real POP. In order to improve scalability, correlative and term sparsity can be exploited to yield sparsity-adapted hierarchies [29,33–35].

On the other hand, as the second way, it might be advantageous to handle CPOP (1.1) directly with the complex moment-HSOS hierarchy introduced in [17]. Let $d_j := \lceil \deg(g_j)/2 \rceil$, $j \in [m]$ and let $d_{\min} := \max\{\lceil \deg(f)/2 \rceil, d_1, \ldots, d_m\}$. Then, the complex moment hierarchy indexed by $d \geq d_{\min}$ (called the relaxation order) for CPOP (1.1) is given by

$$(\mathrm{Q}_d): \quad \begin{cases} \inf & L_{\mathbf{y}}^{\mathrm{c}}(f) \\ \text{s.t.} & \mathbf{M}_d^{\mathrm{c}}(\mathbf{y}) \succeq 0, \\ & \mathbf{M}_{d-d_j}^{\mathrm{c}}(g_j\mathbf{y}) \succeq 0, \quad j \in [m], \\ & y_{\mathbf{0},\mathbf{0}} = 1, \end{cases} \tag{2.1}$$

which is a semidefinite program (SDP) with optimum denoted by ρ_d. The dual of (Q_d) (2.1) can be formulized as the following HSOS relaxation:

$$(\mathrm{Q}_d)^*: \quad \begin{cases} \sup & \rho \\ \text{s.t.} & f - \rho = \sigma_0 + \sigma_1 g_1 + \ldots + \sigma_m g_m, \\ & \sigma_j \text{ is an HSOS}, \quad j = 0, \ldots, m, \\ & \deg(\sigma_0) \leq 2d, \deg(\sigma_j g_j) \leq 2d, \quad j \in [m]. \end{cases} \tag{2.2}$$

Remark 2.1 In (2.1), the expression "$X \succeq 0$" means an Hermitian matrix X to be positive semidefinite. Since popular SDP solvers deal with only real SDPs, it is necessary to convert this condition to a condition involving only real matrices. This can be done by introducing the real part A and the imaginary part B of X, respectively, such that

$X = A + B\mathbf{i}$. Then,

$$X \succeq 0 \quad \Longleftrightarrow \quad \begin{bmatrix} A & -B \\ B & A \end{bmatrix} \succeq 0.$$

Remark 2.2 The first-order moment-(H)SOS relaxation for quadratically constrained quadratic programs (QCQP) is also known as Shor's relaxation. It was proved in [16] that the real Shor's relaxation and the complex Shor's relaxation for homogeneous QCQPs yield the same bound. However, generally the complex hierarchy is weaker (i.e., producing looser bounds) than the real hierarchy at the same relaxation order $d > 1$ as Hermitian sums of squares are a special case of real sums of squares; see [17].

Remark 2.3 By the complex Positivstellensatz theorem due to D'Angelo and Putinar [12], global convergence of the complex hierarchy is guaranteed when a sphere constraint is present.

Remark 2.4 If CPOP (1.1) is feasible, then the program (2.1) is always feasible for any $d \geq d_{\min}$ as one can take the Dirac measure centering a feasible point of (1.1) which leads to a feasible point of (2.1). Moreover, if we further assume that there is a ball/sphere constraint (alternatively, multi-ball/sphere constraints) in terms of all variables in (1.1), then by a similar argument as for Proposition 5.8 of [24], we can show that the feasible set of (2.2) is nonempty. So the optimum of (2.1) is bounded from below by weak duality. In addition, by Lemma 3.6 of [16], the strong duality also holds in this case.

2.2 Sparse Matrices and Chordal Graphs

In this subsection, we briefly revisit the relationship between sparse matrices and chordal graphs, which is crucial for the sparsity-exploitation of this paper. For more details on sparse matrices and chordal graphs, the reader is referred to the survey [27].

An (undirected) *graph* $G(V, E)$ or simply G consists of a set of nodes V and a set of edges $E \subseteq \{\{v_i, v_j\} \mid (v_i, v_j) \in V \times V\}$. When G is a graph, we also use $V(G)$ and $E(G)$ to indicate the node set of G and the edge set of G, respectively. The *adjacency matrix* of G is denoted by B_G for which we put ones on positions corresponding to edges of G as well as its diagonal and put zeros otherwise. For two graphs G, H, we say that G is a *subgraph* of H if $V(G) \subseteq V(H)$ and $E(G) \subseteq E(H)$, denoted by $G \subseteq H$. A graph is called a *chordal graph* if all its cycles of length at least four have a chord[1]. Any non-chordal graph $G(V, E)$ can always be extended to a chordal graph $\overline{G}(V, \overline{E})$ by adding appropriate edges to E, which is called a *chordal extension* of $G(V, E)$. A *clique* $C \subseteq V$ of G is a subset of nodes where $\{v_i, v_j\} \in E$ for any $v_i, v_j \in C$. If a clique C is not a subset of any other clique, then it is called a *maximal clique*. It is known that maximal cliques of a chordal graph can be enumerated efficiently in linear time in the number of nodes and edges of the graph [6].

[1] By a chord, we means an edge that joins two nonconsecutive nodes in a cycle.

Remark 2.5 For a graph G, we denote any specific chordal extension of G by $\overline{G}$. The chordal extension of G is generally not unique. In this paper, we will consider two particular types of chordal extensions: *the maximal chordal extension* and *approximately smallest chordal extensions*. By the maximal chordal extension, we refer to the chordal extension that completes every connected component of G. A chordal extension with the smallest clique number is called a *smallest chordal extension*. Computing a smallest chordal extension of a graph is an NP-complete problem in general. Fortunately, several heuristic algorithms, e.g., the greedy minimum degree and the greedy minimum fill-ins, are known to efficiently produce a good approximation; see [7] for more detailed discussions. Throughout the paper, we assume that for graphs G, H,

$$G \subseteq H \Longrightarrow \overline{G} \subseteq \overline{H}. \tag{2.3}$$

This assumption is reasonable since any chordal extension of H restricting to G is also a chordal extension of G.

Given a graph $G(V, E)$, a Hermitian matrix Q indexed by $V = [n]$ is said to have sparsity pattern G if $Q_{ij} = Q_{ji} = 0$ whenever $i \neq j$ and $\{i, j\} \notin E$, i.e., $B_G \circ Q = Q$. Let $\mathbf{H}_G$ be the set of Hermitian matrices with sparsity pattern G. A matrix in $\mathbf{H}_G$ exhibits a *block structure*. Each block corresponds to a maximal clique of G. The maximal block size is the maximal size of maximal cliques of G, namely, the *clique number* of G. Note that there might be overlaps between blocks because different maximal cliques may share nodes.

Given a maximal clique C of $G(V, E)$, we define a matrix $P_C \in \mathbb{R}^{|C|\times|V|}$ by

$$[P_C]_{ij} = \begin{cases} 1, & \text{if } C(i) = j, \\ 0, & \text{otherwise}, \end{cases} \tag{2.4}$$

where $C(i)$ denotes the i-th node in C, sorted in the ordering compatible with V. Note that $P_C Q P_C^T \in \mathbf{H}^{|C|}$ extracts a principal submatrix defined by the indices in the clique C from a Hermitian matrix Q, and $P_C^T Q_C P_C$ inflates a $|C| \times |C|$ matrix Q_C into a sparse $|V| \times |V|$ matrix.

The PSD matrices with sparsity pattern G form a convex cone

$$\mathbf{H}_+^{|V|} \cap \mathbf{H}_G = \{Q \in \mathbf{H}_G \mid Q \succeq 0\}. \tag{2.5}$$

When the sparsity pattern graph G is chordal, the cone $\mathbf{H}_+^{|V|} \cap \mathbf{H}_G$ can be decomposed as a sum of simple convex cones, as stated in the following theorem.

Theorem 2.6 ([1], Theorem 2.3) *Let $G(V, E)$ be a chordal graph and assume that $\{C_1, \ldots, C_t\}$ is the list of maximal cliques of $G(V, E)$. Then, a matrix $Q \in \mathbf{H}_+^{|V|} \cap \mathbf{H}_G$ if and only if there exist $Q_k \in \mathbf{H}_+^{|C_k|}$ for $k = 1, \ldots, t$ such that $Q = \sum_{k=1}^t P_{C_k}^T Q_k P_{C_k}$.*

Given a graph $G(V, E)$, let Π_G be the projection from $\mathbf{H}^{|V|}$ to the subspace $\mathbf{H}_G$, i.e., for $Q \in \mathbf{H}^{|V|}$,

$$\Pi_G(Q)_{ij} = \begin{cases} Q_{ij}, & \text{if } \{i, j\} \in E \text{ or } i = j, \\ 0, & \text{otherwise.} \end{cases} \tag{2.6}$$

We denote by $\Pi_G(\mathbf{H}_+^{|V|})$ the set of matrices in $\mathbf{H}_G$ that have a PSD completion, i.e.,

$$\Pi_G(\mathbf{H}_+^{|V|}) = \{\Pi_G(Q) \mid Q \in \mathbf{H}_+^{|V|}\}. \tag{2.7}$$

One can check that the PSD completable cone $\Pi_G(\mathbf{H}_+^{|V|})$ and the PSD cone $\mathbf{H}_+^{|V|} \cap \mathbf{H}_G$ form a pair of dual cones in $\mathbf{H}_G$; see [27, Section 10.1] for a proof. Moreover, for a chordal graph G, the decomposition result for the cone $\mathbf{H}_+^{|V|} \cap \mathbf{H}_G$ in Theorem 2.6 leads to the following characterization of the PSD completable cone $\Pi_G(\mathbf{H}_+^{|V|})$.

Theorem 2.7 ([14], Theorem 7) *Let $G(V, E)$ be a chordal graph and assume that $\{C_1, \ldots, C_t\}$ is the list of maximal cliques of $G(V, E)$. Then, a matrix $Q \in \Pi_G(\mathbf{H}_+^{|V|})$ if and only if $Q_k = P_{C_k} Q P_{C_k}^T \succeq 0$ for $k = 1, \ldots, t$.*

3 Sparsity-Adapted Complex Moment-HSOS Hierarchies

In this section, we adapt the real moment-SOS hierarchies that exploit sparsity developed in [29,33–35] to the complex case.

3.1 Correlative Sparsity

The procedure to exploit correlative sparsity for the complex hierarchy consists of two steps: (1) partition the set of variables into subsets according to the correlations between variables emerging in the problem data, and (2) construct a sparse complex hierarchy with respect to the former partition of variables [17,29].

Let us discuss in more details. Consider the CPOP defined by (1.1). Recall $d_j = \lceil \deg(g_j)/2 \rceil$, $j \in [m]$ and $d_{\min} = \max\{\lceil \deg(f)/2 \rceil, d_1, \ldots, d_m\}$. Fix a relaxation order $d \geq d_{min}$. Let $J' := \{j \in [m] \mid d_j = d\}$. For $\boldsymbol{\beta} = (\beta_i)_i \in \mathbb{N}^n$, let $\mathrm{supp}(\boldsymbol{\beta}) := \{i \in [n] \mid \beta_i \neq 0\}$. We define the *correlative sparsity pattern (csp) graph* associated with CPOP (1.1) to be the graph G^{csp} with nodes $V = [n]$ and edges E satisfying $\{i, j\} \in E$ if one of the following holds:

(i) there exists $(\boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathrm{supp}(f) \cup \bigcup_{j \in J'} \mathrm{supp}(g_j)$ such that $\{i, j\} \subseteq \mathrm{supp}(\boldsymbol{\beta}) \cup \mathrm{supp}(\boldsymbol{\gamma})$;
(ii) there exists $k \in [m] \backslash J'$ such that $\{i, j\} \subseteq \bigcup_{(\boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathrm{supp}(g_k)} (\mathrm{supp}(\boldsymbol{\beta}) \cup \mathrm{supp}(\boldsymbol{\gamma}))$.

Remark 3.1 We adopt the idea of "monomial sparsity" proposed in [17] in the definition of csp graphs, which thus is slightly different from the original definition in [29].

Let $\overline{G}^{\text{csp}}$ be a chordal extension of G^{csp} and $\{I_l\}_{l\in[p]}$ be the list of maximal cliques of $\overline{G}^{\text{csp}}$ with $n_l := |I_l|$. Let $\mathbb{C}[\mathbf{z}(I_l)]$ denote the ring of complex polynomials in the n_l variables $\mathbf{z}(I_l) = \{z_i \mid i \in I_l\}$. We then partition the constraint polynomials $g_j, j \in [m]\backslash J'$ into groups $\{g_j \mid j \in J_l\}, l \in [p]$ which satisfy:

(i) $J_1, \dots, J_p \subseteq [m]\backslash J'$ are pairwise disjoint and $\cup_{l=1}^p J_l = [m]\backslash J'$;
(ii) for any $j \in J_l$, $\bigcup_{(\boldsymbol{\beta},\boldsymbol{\gamma})\in\text{supp}(g_j)}(\text{supp}(\boldsymbol{\beta}) \cup \text{supp}(\boldsymbol{\gamma})) \subseteq I_l$, $l \in [p]$.

Example 3.2 Consider the following CPOP

$$\begin{cases}\inf_{\mathbf{z}\in\mathbb{C}^3} & z_1\bar{z}_2 + \bar{z}_1 z_2 + |z_3|^2\\ \text{s.t.} & g_1 = 1 - |z_1|^2 - |z_2|^2 \geq 0,\\ & g_2 = 1 - |z_2|^2 - |z_3|^2 \geq 0,\\ & g_3 = |z_1|^4 + z_2\bar{z}_3 + \bar{z}_2 z_3 \geq 0.\end{cases}$$

Taking $d = d_{min} = 2$, we have two variable cliques $I_1 = \{1, 2\}$, $I_2 = \{2, 3\}$, and $J' = \{3\}$, $J_1 = \{1\}$, $J_2 = \{2\}$; taking $d = 3$, we have one variable clique $I_1 = \{1, 2, 3\}$, and $J' = \emptyset$, $J_1 = \{1, 2, 3\}$.

Next, with $l \in [p]$ and $g \in \mathbb{C}[\mathbf{z}(I_l)]$, let $\mathbf{M}_d^{\text{c}}(\mathbf{y}, I_l)$ (resp. $\mathbf{M}_d^{\text{c}}(g\mathbf{y}, I_l)$) be the complex moment (resp. complex localizing) submatrix obtained from $\mathbf{M}_d^{\text{c}}(\mathbf{y})$ (resp. $\mathbf{M}_d^{\text{c}}(g\mathbf{y})$) by retaining only those rows and columns indexed by $\boldsymbol{\beta} \in \mathbb{N}_d^n$ of $\mathbf{M}_d^{\text{c}}(\mathbf{y})$ (resp. $\mathbf{M}_d^{\text{c}}(g\mathbf{y})$) with $\text{supp}(\boldsymbol{\beta}) \subseteq I_l$.

Then, the complex (moment) hierarchy based on correlative sparsity for CPOP (1.1) is defined as

$$(\mathrm{Q}_d^{\text{cs}}): \quad \begin{cases}\inf & L_{\mathbf{y}}^{\text{c}}(f)\\ \text{s.t.} & \mathbf{M}_d^{\text{c}}(\mathbf{y}, I_l) \succeq 0, \quad l \in [p],\\ & \mathbf{M}_{d-d_j}^{\text{c}}(g_j\mathbf{y}, I_l) \succeq 0, \quad j \in J_l, l \in [p],\\ & L_{\mathbf{y}}^{\text{c}}(g_j) \geq 0, \quad j \in J',\\ & y_{\mathbf{0},\mathbf{0}} = 1.\end{cases} \tag{3.1}$$

We denote the optimum of $(\mathrm{Q}_d^{\text{cs}})$ by ρ_d^{cs}.

Proposition 3.3 *If CPOP* (1.1) *is a QCQP, then* (Q_1^{cs}) *and* (Q_1) *yield the same lower bound for* (1.1)*, i.e.,* $\rho_1^{cs} = \rho_1$.

Proof By construction, the objective function and the affine constraints of (Q_1) involve only the decision variables $\{y_{\boldsymbol{\beta},\boldsymbol{\gamma}}\}_{(\boldsymbol{\beta},\boldsymbol{\gamma})}$ with $\text{supp}(\boldsymbol{\beta}) \cup \text{supp}(\boldsymbol{\gamma}) \subseteq I_l$ for some $l \in [p]$. Therefore, we can replace $\mathbf{M}_1^{\text{c}}(\mathbf{y}) \succeq 0$ by $B_G \circ \mathbf{M}_1^{\text{c}}(\mathbf{y}) \in \Pi_G(\mathbf{H}_+^{n+1})$ without changing the optimum, where G is the graph obtained from $\overline{G}^{\text{csp}}$ by adding a node 0 (corresponding to $\mathbf{0} \in \mathbb{N}^n$) and adding edges $\{0, i\}, i \in [n]$. Note that G is again a chordal graph and so the equality of optima of (Q_1) and $(\mathrm{Q}_1^{\text{cs}})$ follows from Theorem 2.7. □

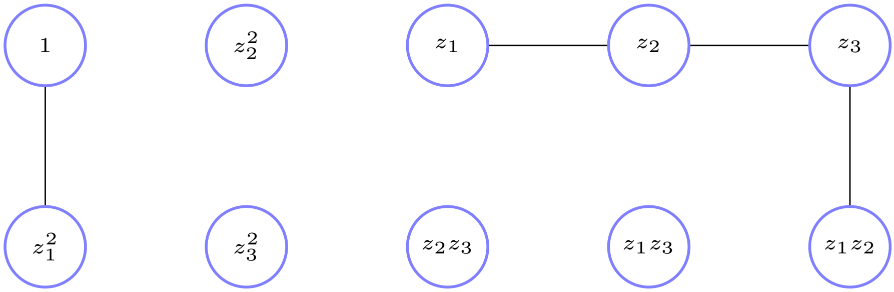

Fig. 1 The tsp graph with $d = 2$ for Example 3.5

3.2 Term Sparsity

Besides correlative sparsity, one can also exploit term sparsity for the complex hierarchy, which was recently developed for real POPs in [33–35]. The intuition behind this procedure is the following: starting with a minimal initial support set, one expands the support set that is taken into account by iteratively performing chordal extensions to the related sparsity pattern graphs inspired by Theorem 2.7. We next adapt it to the complex case.

Let $\mathscr{A} = \operatorname{supp}(f) \cup \bigcup_{j=1}^m \operatorname{supp}(g_j)$. We define the *term sparsity pattern (tsp) graph* at relaxation order d associated with CPOP (1.1) or the set $\mathscr{A}$, to be the graph G_d^{tsp} with nodes $V = \mathbb{N}_d^n$ and edges

$$E := \{\{\boldsymbol{\beta}, \boldsymbol{\gamma}\} \subseteq \mathbb{N}_d^n \mid (\boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathscr{A}\}. \tag{3.2}$$

Remark 3.4 There is a difference on the definitions of tsp graphs between the complex and real cases. In the real case, we use $\mathscr{A} \cup 2\mathbb{N}_d^n$ rather than $\mathscr{A}$ in (3.2) due to the Hankel structure of real moment matrices.

Example 3.5 Consider the following CPOP

$$\begin{cases} \inf_{\mathbf{z}\in\mathbb{C}^3} & z_1^2 + \bar{z}_1^2 + z_1\bar{z}_2 + \bar{z}_1z_2 + z_2\bar{z}_3 + \bar{z}_2z_3 + z_1z_2\bar{z}_3 + \bar{z}_1\bar{z}_2z_3 \\ \text{s.t.} & g_1 = 1 - |z_1|^2 - |z_2|^2 - |z_3|^2 \geq 0. \end{cases}$$

Figure 1 illustrates the tsp graph G_2^{tsp} for this CPOP, where the nodes are labeled by $\mathbf{z}^{\boldsymbol{\beta}}$ instead of $\boldsymbol{\beta}$ for better visualization.

For any graph G with $V \subseteq \mathbb{N}^n$ and $g = \sum_{(\boldsymbol{\beta}',\boldsymbol{\gamma}')} g_{\boldsymbol{\beta}',\boldsymbol{\gamma}'}\mathbf{z}^{\boldsymbol{\beta}'}\bar{\mathbf{z}}^{\boldsymbol{\gamma}'} \in \mathbb{C}[\mathbf{z}, \bar{\mathbf{z}}]$, we define the *g-support* of G by

$$\operatorname{supp}_g(G) := \{(\boldsymbol{\beta} + \boldsymbol{\beta}', \boldsymbol{\gamma} + \boldsymbol{\gamma}') \mid \boldsymbol{\beta} = \boldsymbol{\gamma} \in V(G) \text{ or } \{\boldsymbol{\beta}, \boldsymbol{\gamma}\} \in E(G), (\boldsymbol{\beta}', \boldsymbol{\gamma}') \in \operatorname{supp}(g)\}.$$

Let us set $d_0 := 0$ and $g_0 := 1$. Now, assume that $G_{d,0}^{(0)} = G_d^{\text{tsp}}$ and $G_{d,j}^{(0)}, j \in [m]$ are empty graphs. Then, we iteratively define an ascending chain of graphs $(G_{d,j}^{(k)}(V_{d,j}, E_{d,j}^{(k)}))_{k\geq 1}$ with $V_{d,j} = \mathbb{N}_{d-d_j}^n$ for each $j \in \{0\} \cup [m]$ by

$$G_{d,j}^{(k)} := \overline{F_{d,j}^{(k)}}, \tag{3.3}$$

where $F_{d,j}^{(k)}$ is the graph with $V(F_{d,j}^{(k)}) = \mathbb{N}_{d-d_j}^n$ and

$$E(F_{d,j}^{(k)}) = \{\{\boldsymbol{\beta}, \boldsymbol{\gamma}\} \subseteq \mathbb{N}_{d-d_j}^n \mid ((\boldsymbol{\beta}, \boldsymbol{\gamma}) + \text{supp}(g_j)) \cap (\bigcup_{i=0}^{m} \text{supp}_{g_i}(G_{d,i}^{(k-1)})) \neq \emptyset\}. \tag{3.4}$$

Let $r_j := \binom{n+d-d_j}{d-d_j}$ for $j \in \{0\} \cup [m]$. Then, with $d \geq d_{\min}$ and $k \geq 1$, the complex (moment) hierarchy based on term sparsity for CPOP (1.1) is defined as

$$(\text{Q}_{d,k}^{\text{ts}}): \quad \begin{cases} \inf & L_{\mathbf{y}}^{\text{c}}(f) \\ \text{s.t.} & B_{G_{d,0}^{(k)}} \circ \mathbf{M}_d^{\text{c}}(\mathbf{y}) \in \Pi_{G_{d,0}^{(k)}}(\mathbf{H}_+^{r_0}), \\ & B_{G_{d,j}^{(k)}} \circ \mathbf{M}_{d-d_j}^{\text{c}}(g_j\mathbf{y}) \in \Pi_{G_{d,j}^{(k)}}(\mathbf{H}_+^{r_j}), \quad j \in [m], \\ & y_{\mathbf{0},\mathbf{0}} = 1, \end{cases} \tag{3.5}$$

with optimum denoted by $\rho_{d,k}^{\text{ts}}$. The above hierarchy is called the (complex) *TSSOS* hierarchy, which is indexed by two parameters: the relaxation order d and the *sparse order* k.

Theorem 3.6 *Consider CPOP* (1.1)*. The following holds:*

(i) Fixing a relaxation order $d \geq d_{\min}$, the sequence $(\rho_{d,k}^{ts})_{k\geq 1}$ is monotonically nondecreasing and $\rho_{d,k}^{ts} \leq \rho_d$ for all k (with ρ_d defined in Sect. 2.1).

(ii) Fixing a sparse order $k \geq 1$, the sequence $(\rho_{d,k}^{ts})_{d\geq d_{\min}}$ is monotonically nondecreasing.

Proof (i). For all j, k, by construction we have $G_{d,j}^{(k)} \subseteq G_{d,j}^{(k+1)}$, which implies that $B_{G_{d,j}^{(k)}} \circ \mathbf{M}_{d-d_j}^{\text{c}}(g_j\mathbf{y}) \in \Pi_{G_{d,j}^{(k)}}(\mathbf{H}_+^{r_j})$ is less restrictive than $B_{G_{d,j}^{(k+1)}} \circ \mathbf{M}_{d-d_j}^{\text{c}}(g_j\mathbf{y}) \in \Pi_{G_{d,j}^{(k+1)}}(\mathbf{H}_+^{r_j})$. Hence, $(\text{Q}_{d,k}^{\text{ts}})$ is a relaxation of $(\text{Q}_{d,k+1}^{\text{ts}})$ and is clearly also a relaxation of (Q_d). As a result, $(\rho_{d,k}^{\text{ts}})_{k\geq 1}$ is monotonically nondecreasing and $\rho_{d,k}^{\text{ts}} \leq \rho_d$ for all k.

(ii). The conclusion follows if we can show that the inclusion $G_{d,j}^{(k)} \subseteq G_{d+1,j}^{(k)}$ holds for all d, j since this implies that $(\text{Q}_{d,k}^{\text{ts}})$ is a relaxation of $(\text{Q}_{d+1,k}^{\text{ts}})$. Let us prove $G_{d,j}^{(k)} \subseteq G_{d+1,j}^{(k)}$ by induction on k. For $k = 1$, we have $E(G_d^{\text{tsp}}) \subseteq E(G_{d+1}^{\text{tsp}})$ by (3.2), which implies $G_{d,j}^{(1)} \subseteq G_{d+1,j}^{(1)}$ for all d, j. Now, assume that $G_{d,j}^{(k)} \subseteq G_{d+1,j}^{(k)}$ holds for all d, j for a given $k \geq 1$. Then, by (2.3), (3.3), (3.4) and by the induction hypothesis, we deduce that $G_{d,j}^{(k+1)} \subseteq G_{d+1,j}^{(k+1)}$ holds for all d, j, which completes the induction. □

When building $(\mathrm{Q}^{\mathrm{ts}}_{d,k})$, we have the freedom to choose a specific chordal extension for any involved graph $G^{(k)}_{d,j}$, which offers a trade-off between the quality of obtained bounds and the computational cost. We show that if the maximal chordal extension is chosen, then with d fixed, the resulting sequence of optima of the hierarchy (as k increases) converges in finitely many steps to the optimum of the corresponding dense relaxation.

Theorem 3.7 *Consider CPOP (1.1). If the maximal chordal extension is used in (3.3), then for $d \ge d_{\min}$, $(\rho^{ts}_{d,k})_{k\ge 1}$ converges to ρ_d in finitely many steps.*

Proof Let d be fixed. It is clear that for all $j \in \{0\} \cup [m]$, the graph sequence $(G^{(k)}_{d,j})_{k\ge 1}$ stabilizes after finitely many steps and we denote the stabilized graph by $G^{(\circ)}_{d,j}$. Let $(\mathrm{Q}^{\mathrm{ts}}_{d,\circ})$ be the moment relaxation corresponding to the stabilized graphs and let $\mathbf{y}^* = (y^*_{\boldsymbol{\beta},\boldsymbol{\gamma}})$ be an arbitrary feasible solution. Notice that $\{y_{\boldsymbol{\beta},\boldsymbol{\gamma}} \mid (\boldsymbol{\beta},\boldsymbol{\gamma}) \in \bigcup_{i=0}^m \mathrm{supp}_{g_j}(G^{(\circ)}_{d,j})\}$ is the set of decision variables involved in $(\mathrm{Q}^{\mathrm{ts}}_{d,\circ})$ and $\{y_{\boldsymbol{\beta},\boldsymbol{\gamma}} \mid (\boldsymbol{\beta},\boldsymbol{\gamma}) \in \mathbb{N}^n_d \times \mathbb{N}^n_d\}$ is the set of decision variables involved in (Q_d). Define $\overline{\mathbf{y}}^* = (\overline{y}^*_{\boldsymbol{\beta},\boldsymbol{\gamma}})_{(\boldsymbol{\beta},\boldsymbol{\gamma})\in\mathbb{N}^n_d\times\mathbb{N}^n_d}$ as follows:

$$\overline{y}^*_{\boldsymbol{\beta},\boldsymbol{\gamma}} = \begin{cases} y^*_{\boldsymbol{\beta},\boldsymbol{\gamma}}, & \text{if } (\boldsymbol{\beta},\boldsymbol{\gamma}) \in \bigcup_{i=0}^m \mathrm{supp}_{g_j}(G^{(\circ)}_{d,j}), \\ 0, & \text{otherwise.} \end{cases}$$

If the maximal chordal extension is used in (3.3), then we have that the matrices in $\Pi_{G^{(k)}_{d,j}}(\mathbf{H}^{r_j}_+)$ are block-diagonal (up to permutation on rows and columns) for all j, k. As a consequence, $B_{G^{(k)}_{d,j}} \circ \mathbf{M}^{\mathrm{c}}_{d-d_j}(g_j\mathbf{y}) \in \Pi_{G^{(k)}_{d,j}}(\mathbf{H}^{r_j}_+)$ implies $B_{G^{(k)}_{d,j}} \circ \mathbf{M}^{\mathrm{c}}_{d-d_j}(g_j\mathbf{y}) \succeq 0$. By construction, we have $\mathbf{M}^{\mathrm{c}}_{d-d_j}(g_j\overline{\mathbf{y}}^*) = B_{G^{(\circ)}_{d,j}} \circ \mathbf{M}^{\mathrm{c}}_{d-d_j}(g_j\mathbf{y}^*) \succeq 0$ for all $j \in \{0\} \cup [m]$. Therefore, $\overline{\mathbf{y}}^*$ is a feasible solution of (Q_d) and hence $L^{\mathrm{c}}_{\mathbf{y}^*}(f) = L^{\mathrm{c}}_{\overline{\mathbf{y}}^*}(f) \ge \rho_d$, which yields $\rho^{\mathrm{ts}}_{d,\circ} \ge \rho_d$ since $\mathbf{y}^*$ is an arbitrary feasible solution of $(\mathrm{Q}^{\mathrm{ts}}_{d,\circ})$. By (i) of Theorem 3.6, we already have $\rho^{\mathrm{ts}}_{d,\circ} \le \rho_d$. So $\rho^{\mathrm{ts}}_{d,\circ} = \rho_d$ as desired. □

Proposition 3.8 *If CPOP (1.1) is a QCQP, then $(Q^{ts}_{1,1})$ and (Q_1) yield the same lower bound for CPOP (1.1), i.e., $\rho^{ts}_{1,1} = \rho_1$.*

Proof For a QCQP, (Q_1) reads as

$$(\mathrm{Q}_1): \quad \begin{cases} \inf & L^{\mathrm{c}}_{\mathbf{y}}(f) \\ \text{s.t.} & \mathbf{M}^{\mathrm{c}}_1(\mathbf{y}) \succeq 0, \\ & L^{\mathrm{c}}_{\mathbf{y}}(g_j) \ge 0, \quad j \in [m], \\ & y_{\mathbf{0},\mathbf{0}} = 1. \end{cases}$$

Note that the objective function and the affine constraints of (Q_1) involve only the decision variables $\{y_{\mathbf{0},\mathbf{0}}\}\cup\{y_{\boldsymbol{\beta},\boldsymbol{\gamma}}\}_{(\boldsymbol{\beta},\boldsymbol{\gamma})\in\mathscr{A}}$ with $\mathscr{A} = \mathrm{supp}(f)\cup\bigcup_{j=1}^m \mathrm{supp}(g_j)$. Hence, there is no discrepancy of optima in replacing (Q_1) with $(\mathrm{Q}^{\mathrm{ts}}_{1,1})$ by construction. □

Sign symmetry. Let $\mathscr{A} \subseteq \mathbb{N}^n \times \mathbb{N}^n$. The *sign symmetries* of $\mathscr{A}$ consist of all binary vectors $\mathbf{r} \in \mathbb{Z}_2^n := \{0, 1\}^n$ such that $\mathbf{r}^T(\boldsymbol{\beta} + \boldsymbol{\gamma}) \equiv 0 \pmod 2$ for all $(\boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathscr{A}$. For the monomial basis $\mathbb{N}_d^n$, a set of sign symmetries $R = [\mathbf{r}_1, \ldots, \mathbf{r}_s]$ (regarded as a matrix with columns $\mathbf{r}_i \in \mathbb{Z}_2^n$) induces a partition on $\mathbb{N}_d^n$: $\boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}_d^n$ belong to the same block in this partition if and only if $R^T(\boldsymbol{\beta} + \boldsymbol{\gamma}) \equiv \mathbf{0} \pmod 2$.

We now prove that the block structure at each step of the TSSOS hierarchy is a refinement of the one induced by the sign symmetries of the system.

Theorem 3.9 *Consider CPOP (1.1). Let* $\mathscr{A} = \mathrm{supp}(f) \cup \bigcup_{j=1}^m \mathrm{supp}(g_j)$ *and R be its sign symmetries. Assume* $d \geq d_{\min}$ *and* $k \geq 1$*. For any* $j \in \{0\} \cup [m]$ *and* $\boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}_{d-d_j}^n$*, if* $\{\boldsymbol{\beta}, \boldsymbol{\gamma}\} \in E(G_{d,j}^{(k)})$*, then* $R^T(\boldsymbol{\beta} + \boldsymbol{\gamma}) \equiv \mathbf{0} \pmod 2$.

Proof We prove the conclusion by induction on k. For $k = 1$, suppose $\{\boldsymbol{\beta}, \boldsymbol{\gamma}\} \in E(F_{d,j}^{(1)})$. By (3.4), there exists $(\boldsymbol{\beta}', \boldsymbol{\gamma}') \in \mathrm{supp}(g_j)$ such that $(\boldsymbol{\beta}, \boldsymbol{\gamma}) + (\boldsymbol{\beta}', \boldsymbol{\gamma}') \in \mathscr{A}$. Hence, $R^T(\boldsymbol{\beta} + \boldsymbol{\gamma}) + R^T(\boldsymbol{\beta}' + \boldsymbol{\gamma}') \equiv \mathbf{0} \pmod 2$ and since $R^T(\boldsymbol{\beta}' + \boldsymbol{\gamma}') \equiv \mathbf{0} \pmod 2$, it follows that $R^T(\boldsymbol{\beta} + \boldsymbol{\gamma}) \equiv \mathbf{0} \pmod 2$. Note that $G_{d,j}^{(1)}$ is a chordal extension of $F_{d,j}^{(1)}$ such that the additional edges connect two nodes belonging to the same connected component. Thus, it is not hard to show that if $\{\boldsymbol{\beta}, \boldsymbol{\gamma}\} \in E(G_{d,j}^{(k)})$, then $R^T(\boldsymbol{\beta} + \boldsymbol{\gamma}) \equiv \mathbf{0} \pmod 2$. Now, assume that the conclusion holds for a given $k \geq 1$, which implies that for all i and for any $(\boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathrm{supp}_{g_i}(G_{d,i}^{(k)})$, $R^T(\boldsymbol{\beta} + \boldsymbol{\gamma}) \equiv \mathbf{0} \pmod 2$. For $\{\boldsymbol{\beta}, \boldsymbol{\gamma}\} \in E(F_{d,j}^{(k+1)})$, by (3.4), there exists $(\boldsymbol{\beta}', \boldsymbol{\gamma}') \in \mathrm{supp}(g_j)$ such that $(\boldsymbol{\beta}, \boldsymbol{\gamma}) + (\boldsymbol{\beta}', \boldsymbol{\gamma}') \in \bigcup_{i=0}^m \mathrm{supp}_{g_i}(G_{d,i}^{(k)})$. Hence, by the induction hypothesis, $R^T(\boldsymbol{\beta} + \boldsymbol{\gamma}) + R^T(\boldsymbol{\beta}' + \boldsymbol{\gamma}') \equiv \mathbf{0} \pmod 2$ and so $R^T(\boldsymbol{\beta} + \boldsymbol{\gamma}) \equiv \mathbf{0} \pmod 2$. Because $G_{d,j}^{(k+1)}$ is a chordal extension of $F_{d,j}^{(k+1)}$ such that the additional edges connect two nodes belonging to the same connected component, we obtain that the conclusion holds for $k + 1$ and complete the induction. □

In contrast with the real case in which it was shown that the partition on $\mathbb{N}_{d-d_j}^n$, $j \in \{0\} \cup [m]$ at the final step of the TSSOS hierarchy (i.e., when it stabilizes as k increases) using the maximal chordal extension is exactly the one induced by the sign symmetries of the system [34], in the complex case this is not necessarily true as the following example illustrates.

Example 3.10 Consider the following CPOP

$$\inf\{z_1 + \bar{z}_1 : 1 - z_1\bar{z}_1 - z_2\bar{z}_2 \geq 0\}.$$

Let us take the relaxation order $d = 2$. One can easily check by hand that the partition on $\mathbb{N}_2^2$ at the final step of the TSSOS hierarchy using the maximal chordal extension is $\{1, z_1, z_1^2\}$, $\{z_2^2\}$ and $\{z_2, z_1 z_2\}$, while the partition induced by sign symmetries is $\{1, z_1, z_1^2, z_2^2\}$ and $\{z_2, z_1 z_2\}$. Clearly, the former is a strict refinement of the latter.

3.3 Correlative-Term Sparsity

We are now prepared to exploit correlative sparsity and term sparsity simultaneously in the complex hierarchy for CPOP (1.1).

Let $\{I_l\}_{l\in[p]}$, $\{n_l\}_{l\in[p]}$, J', $\{J_l\}_{l\in[p]}$ be defined as in Sect. 3.1. We apply the iterative procedure of exploiting term sparsity to each subsystem involving variables $\mathbf{z}(I_l)$ for $l \in [p]$ as follows. Let

$$\mathscr{A} := \operatorname{supp}(f) \cup \bigcup_{j=1}^{m} \operatorname{supp}(g_j) \tag{3.6}$$

and

$$\mathscr{A}_l := \{(\boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathscr{A} \mid \operatorname{supp}(\boldsymbol{\beta}) \cup \operatorname{supp}(\boldsymbol{\gamma}) \subseteq I_l\} \tag{3.7}$$

for $l \in [p]$. As before, $d_{\min} := \max\{\lceil \deg(f)/2 \rceil, d_1, \ldots, d_m\}$, $d_0 := 0$ and $g_0 := 1$. Fix a relaxation order $d \geq d_{\min}$. Let $G_{d,l}^{\mathrm{tsp}}$ be the tsp graph with nodes $\mathbb{N}_{d-d_j}^{n_l}$ associated with $\mathscr{A}_l$ defined as in Sect. 3.2. Note that here we embed $\mathbb{N}_{d-d_j}^{n_l}$ into $\mathbb{N}_{d-d_j}^{n}$ via the map $\boldsymbol{\alpha} = (\alpha_i)_{i\in I_l} \in \mathbb{N}_{d-d_j}^{n_l} \mapsto \boldsymbol{\alpha}' = (\alpha_i')_{i\in[n]} \in \mathbb{N}_{d-d_j}^{n}$ which satisfies

$$\alpha_i' = \begin{cases} \alpha_i, & \text{if } i \in I_l, \\ 0, & \text{otherwise.} \end{cases}$$

Assume that $G_{d,l,0}^{(0)} = G_{d,l}^{\mathrm{tsp}}$ and $G_{d,l,j}^{(0)}$, $j \in J_l$, $l \in [p]$ are empty graphs. Letting

$$\mathscr{C}_d^{(k-1)} := \bigcup_{l=1}^{p} \bigcup_{j\in\{0\}\cup J_l} \operatorname{supp}_{g_j}(G_{d,l,j}^{(k-1)}), \quad k \geq 1, \tag{3.8}$$

we iteratively define an ascending chain of graphs $(G_{d,l,j}^{(k)}(V_{d,l,j}, E_{d,l,j}^{(k)}))_{k\geq 1}$ with $V_{d,l,j} = \mathbb{N}_{d-d_j}^{n_l}$ for each $j \in \{0\} \cup J_l$ and each $l \in [p]$ by

$$G_{d,l,j}^{(k)} := \overline{F_{d,l,j}^{(k)}}, \tag{3.9}$$

where $F_{d,l,j}^{(k)}$ is the graph with $V(F_{d,l,j}^{(k)}) = \mathbb{N}_{d-d_j}^{n_l}$ and

$$E(F_{d,l,j}^{(k)}) = \{\{\boldsymbol{\beta}, \boldsymbol{\gamma}\} \subseteq \mathbb{N}_{d-d_j}^{n_l} \mid ((\boldsymbol{\beta}, \boldsymbol{\gamma}) + \operatorname{supp}(g_j)) \cap \mathscr{C}_d^{(k-1)} \neq \emptyset\}. \tag{3.10}$$

Let $r_{d,l,j} := \binom{n_l+d-d_j}{d-d_j}$ for all l, j. Then, with $d \geq d_{\min}$ and $k \geq 1$, the complex (moment) hierarchy based on correlative-term sparsity for CPOP (1.1) is defined as

$$(\mathrm{Q}_{d,k}^{\text{cs-ts}}) : \begin{cases} \inf & L_{\mathbf{y}}^{\mathrm{c}}(f) \\ \text{s.t.} & B_{G_{d,l,0}^{(k)}} \circ \mathbf{M}_d^{\mathrm{c}}(\mathbf{y}, I_l) \in \Pi_{G_{d,l,0}^{(k)}}(\mathbf{H}_+^{r_{d,l,0}}), \quad l \in [p], \\ & B_{G_{d,l,j}^{(k)}} \circ \mathbf{M}_{d-d_j}^{\mathrm{c}}(g_j\mathbf{y}, I_l) \in \Pi_{G_{d,l,j}^{(k)}}(\mathbf{H}_+^{r_{d,l,j}}), \quad j \in J_l, l \in [p], \\ & L_{\mathbf{y}}^{\mathrm{c}}(g_j) \geq 0, \quad j \in J', \\ & y_{\mathbf{0},\mathbf{0}} = 1, \end{cases} \tag{3.11}$$

with optimum denoted by $\rho_{d,k}^{\text{cs-ts}}$. The above hierarchy is called the (complex) *CS-TSSOS* hierarchy indexed by the relaxation order d and the sparse order k.

By similar arguments as for Theorem 3.6, we can prove the following theorem.

Theorem 3.11 *Consider CPOP (1.1). The following holds:*

(i) Fixing a relaxation order $d \geq d_{\min}$, the sequence $(\rho_{d,k}^{\text{cs-ts}})_{k\geq 1}$ is monotonically nondecreasing and $\rho_{d,k}^{\text{cs-ts}} \leq \rho_d^{\text{cs}}$ for all k (with ρ_d^{cs} defined in Sect. 3.1).
(ii) Fixing a sparse order $k \geq 1$, the sequence $(\rho_{d,k}^{\text{cs-ts}})_{d\geq d_{\min}}$ is monotonically nondecreasing.

From Theorem 3.11, we have the following two-level hierarchy of lower bounds for the optimum of CPOP (1.1):

$$\begin{array}{ccccccc} \rho_{d_{\min},1}^{\text{cs-ts}} & \leq & \rho_{d_{\min},2}^{\text{cs-ts}} & \leq \cdots \leq & \rho_{d_{\min}}^{\text{cs}} \\ \wedge| & & \wedge| & & \wedge| \\ \rho_{d_{\min}+1,1}^{\text{cs-ts}} & \leq & \rho_{d_{\min}+1,2}^{\text{cs-ts}} & \leq \cdots \leq & \rho_{d_{\min}+1}^{\text{cs}} \\ \wedge| & & \wedge| & & \wedge| \\ \vdots & & \vdots & \vdots & \vdots \\ \wedge| & & \wedge| & & \wedge| \\ \rho_{d,1}^{\text{cs-ts}} & \leq & \rho_{d,2}^{\text{cs-ts}} & \leq \cdots \leq & \rho_d^{\text{cs}} \\ \wedge| & & \wedge| & & \wedge| \\ \vdots & & \vdots & \vdots & \vdots \end{array} \tag{3.12}$$

By similar arguments as for Theorem 3.7, we can prove the convergence of the CS-TSSOS hierarchy at each relaxation order when the maximal chordal extension is chosen.

Theorem 3.12 *Consider CPOP (1.1). If the maximal chordal extension is used in (3.9), then for $d \geq d_{\min}$, $(\rho_{d,k}^{\text{cs-ts}})_{k\geq 1}$ converges to ρ_d^{cs} in finitely many steps.*

By slightly adapting the proof of Theorem 3.9, one can prove that the block structure at each step of the CS-TSSOS hierarchy is also "governed" by the sign symmetries of the system.

Theorem 3.13 *Consider CPOP (1.1). Let* $\mathscr{A} = \mathrm{supp}(f) \cup \bigcup_{j=1}^{m} \mathrm{supp}(g_j)$ *and* R *be its sign symmetries. Assume* $d \geq d_{\min}$ *and* $k \geq 1$*. For any* $j \in J_l$*,* $l \in [p]$ *and* $\boldsymbol{\beta}, \boldsymbol{\gamma} \in \mathbb{N}^{n_l}_{d-d_j}$*, if* $\{\boldsymbol{\beta}, \boldsymbol{\gamma}\} \in E(G^{(k)}_{d,l,j})$*, then* $R^T(\boldsymbol{\beta} + \boldsymbol{\gamma}) \equiv \mathbf{0}$ *(mod 2).*

If CPOP (1.1) is a QCQP, then by Proposition 3.3, we have $\rho_1^{\mathrm{cs}} = \rho_1$. To ensure any higher order relaxation $(\mathrm{Q}^{\text{cs-ts}}_{d,k})$ $(d > 1)$ achieves a better lower bound than Shor's relaxation, we may add an extra first-order moment matrix for each variable clique[2]:

$$(\mathrm{Q}^{\text{cs-ts}}_{d,k})' : \quad \begin{cases} \inf & L^{\mathrm{c}}_{\mathbf{y}}(f) \\ \text{s.t.} & B_{G^{(k)}_{d,l,0}} \circ \mathbf{M}^{\mathrm{c}}_{d}(\mathbf{y}, I_l) \in \Pi_{G^{(k)}_{d,l,0}}(\mathbf{H}_{+}^{r_{d,l,0}}), \quad l \in [p], \\ & \mathbf{M}^{\mathrm{c}}_{1}(\mathbf{y}, I_l) \succeq 0, \quad l \in [p], \\ & B_{G^{(k)}_{d,l,j}} \circ \mathbf{M}^{\mathrm{c}}_{d-d_j}(g_j\mathbf{y}, I_l) \in \Pi_{G^{(k)}_{d,l,j}}(\mathbf{H}_{+}^{r_{d,l,j}}), \quad j \in J_l, l \in [p], \\ & L^{\mathrm{c}}_{\mathbf{y}}(g_j) \geq 0, \quad j \in J', \\ & y_{\mathbf{0},\mathbf{0}} = 1. \end{cases} \tag{3.13}$$

4 The Minimum Initial Relaxation Step of the Complex Hierarchy

For CPOP (1.1), suppose that f is not homogeneous or the constraint polynomials $g_j, j \in [m]$ have different degrees. Then, it might be beneficial to assign different relaxation orders to different subsystems obtained from the correlative sparsity pattern for the initial relaxation step of the complex hierarchy instead of using the uniform minimum relaxation order $d_{\min}$. More specifically, we redefine the csp graph $G^{\mathrm{icsp}}(V, E)$ as follows: let $V = [n]$ and $\{i, j\} \in E$ if there exists $(\boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathrm{supp}(f) \cup \bigcup_{j \in [m]} \mathrm{supp}(g_j)$ such that $\{i, j\} \subseteq \mathrm{supp}(\boldsymbol{\beta}) \cup \mathrm{supp}(\boldsymbol{\gamma})$. This is clearly a subgraph of G^{csp} defined in Sect. 3.1 and hence typically has a smaller chordal extension. Let $\overline{G}^{\mathrm{icsp}}$ be a chordal extension of G^{icsp} and $\{I_l\}_{l \in [p]}$ be the list of maximal cliques of $\overline{G}^{\mathrm{icsp}}$ with $n_l := |I_l|$. Now, we partition the constraint polynomials $g_j, j \in [m]$ into groups $\{g_j \mid j \in J_l\}_{l \in [p]}$ and $\{g_j \mid j \in J'\}$ which satisfy:

(i) $J_1, \ldots, J_p, J' \subseteq [m]$ are pairwise disjoint and $\bigcup_{l=1}^{p} J_l \cup J' = [m]$;
(ii) for any $j \in J_l$, $\bigcup_{(\boldsymbol{\beta},\boldsymbol{\gamma}) \in \mathrm{supp}(g_j)} (\mathrm{supp}(\boldsymbol{\beta}) \cup \mathrm{supp}(\boldsymbol{\gamma})) \subseteq I_l$, $l \in [p]$;
(iii) for any $j \in J'$, $\bigcup_{(\boldsymbol{\beta},\boldsymbol{\gamma}) \in \mathrm{supp}(g_j)} (\mathrm{supp}(\boldsymbol{\beta}) \cup \mathrm{supp}(\boldsymbol{\gamma})) \nsubseteq I_l$ for all $l \in [p]$.

Assume that f decomposes as $f = \sum_{l \in [p]} f_l$ such that $\bigcup_{(\boldsymbol{\beta},\boldsymbol{\gamma}) \in \mathrm{supp}(f_l)} (\mathrm{supp}(\boldsymbol{\beta}) \cup \mathrm{supp}(\boldsymbol{\gamma})) \subseteq I_l$ for $l \in [p]$. We define the vector of minimum relaxation orders $\mathbf{o} = (o_l)_l \in \mathbb{N}^p$ with $o_l := \max(\{d_j : j \in J_l\} \cup \{\lceil \deg(f_l)/2 \rceil\})$. Then, with $k \geq 1$, we consider the following initial relaxation step of the complex hierarchy:

[2] Even if CPOP (1.1) is not a QCQP, this operation could also strengthen the relaxation.

Table 1 The notation

n	number of complex variables
d	relaxation order
k	sparse order
mb	maximal size of PSD blocks
opt	optimum
time	running time in seconds
gap	optimality gap
-	an out of memory error

$$
(\mathrm{Q}^{\text{cs-ts}}_{\min,k}):\quad \begin{cases} \inf & L^{\mathrm{c}}_{\mathbf{y}}(f) \\ \text{s.t.} & B_{G^{(k)}_{o_l,l,0}} \circ \mathbf{M}^{\mathrm{c}}_{o_l}(\mathbf{y}, I_l) \in \Pi_{G^{(k)}_{o_l,l,0}}(\mathbf{H}^{s_{l,0}}_{+}), \quad l \in [p], \\ & \mathbf{M}^{\mathrm{c}}_{1}(\mathbf{y}, I_l) \succeq 0, \quad l \in [p], \\ & B_{G^{(k)}_{o_l,l,j}} \circ \mathbf{M}^{\mathrm{c}}_{o_l-d_j}(g_j\mathbf{y}, I_l) \in \Pi_{G^{(k)}_{o_l,l,j}}(\mathbf{H}^{s_{l,j}}_{+}), \quad j \in J_l, l \in [p], \\ & L^{\mathrm{c}}_{\mathbf{y}}(g_j) \geq 0, \quad j \in J', \\ & y_{\mathbf{0},\mathbf{0}} = 1, \end{cases} \tag{4.1}
$$

where the sparsity pattern graphs $G^{(k)}_{o_l,l,j}$, $j \in J_l, l \in [p]$ are defined in the same manner as in Sect. 3.3 with $V(G^{(k)}_{o_l,l,j}) = \mathbb{N}^{n_l}_{o_l-d_j}$ and $s_{l,j} := \binom{n_l+o_l-d_j}{o_l-d_j}$ for all l, j.

Remark 4.1 Similarly, we can also define the minimum initial relaxation step for the real hierarchy.

5 Numerical Experiments

In this section, we present numerical results of the proposed sparsity-adapted complex hierarchies for complex polynomial optimization problems. The hierarchies were implemented in the Julia package `TSSOS` [22], which utilizes the Julia packages `LightGraphs` [8] to handle graphs, `ChordalGraph` [30] to generate approximately smallest chordal extensions, and `JuMP` [11] to model the SDP. For the numerical experiments in this paper, `Mosek` [25] is used as an SDP solver. `TSSOS` is freely available at https://github.com/wangjie212/TSSOS.

All numerical examples were computed on an Intel Core i5-8265U@1.60GHz CPU with 8GB RAM memory. We list the notations that are used in this section in Table 1. The running times include the time for pre-processing (to get the block structure), the time for modeling the SDP and the time for solving the SDP.

Table 2 The complex hierarchy versus the real hierarchy with $d = 2, k = 1$ for minimizing quartic objective functions on multi-balls

n	CE	Complex			Real		
		mb	opt	Time	mb	opt	Time
55	min	6	−24.6965	0.95	21	-21.2240	9.21
	max	36	−24.4543	5.82	–	–	–
105	min	6	−48.9783	2.45	21	−40.4650	36.5
	max	46	−48.8367	16.9	–	–	–
155	min	6	−69.4493	4.10	21	-57.9840	44.3
	max	52	−69.2345	29.3	–	–	–
205	min	6	−100.132	6.92	21	-82.7737	85.5
	max	54	−99.4924	65.4	–	–	–
255	min	8	−122.621	10.9	21	-102.728	107
	max	46	−121.754	81.7	–	–	–
305	min	8	−151.096	13.7	21	-126.094	134
	max	52	−149.406	54.7	–	–	–
355	min	8	−172.275	18.3	21	-144.936	182
	max	52	−170.655	65.2	–	–	–
405	min	8	−197.036	24.7	21	-163.360	229
	max	48	−195.287	79.6	–	–	–
455	min	8	−224.471	29.3	21	-184.701	278
	max	52	−222.837	85.5	–	–	–
505	min	8	−238.760	35.2	21	-199.774	357
	max	46	−237.437	108	–	–	–

5.1 Randomly Generated Examples

Given an integer $p \in \mathbb{N}$, we randomly generate a CPOP as follows:

(1) Let $f = \sum_{l=1}^{p} f_l \in \mathbb{C}_4[z_1, \dots, z_{5(p+1)}, \bar{z}_1, \dots, \bar{z}_{5(p+1)}]$, where for all $l \in [p]$, $f_l = \bar{f}_l \in \mathbb{C}_4[z_{5(l-1)+1}, \dots, z_{5(l-1)+10}, \bar{z}_{5(l-1)+1}, \dots, \bar{z}_{5(l-1)+10}]$ is a polynomial with 40 terms whose real/imaginary parts of coefficients are selected with respect to the uniform probability distribution on $[-1, 1]$;
(2) Let us encode multi-ball constraints with $g_l = 1 - \sum_{i=1}^{10} z_{5(l-1)+i}\bar{z}_{5(l-1)+i}$ for $l \in [p]$;
(3) The CPOP is defined as $\inf_{\mathbf{z}\in\mathbb{C}^{5(p+1)}}\{f(\mathbf{z}, \bar{\mathbf{z}}) : g_l(\mathbf{z}, \bar{\mathbf{z}}) \geq 0, l = 1, \dots, p\}$.

In such a way, we generate 10 random CPOPs with $p = 10, 20, \dots, 100$, respectively. We compare the performance of the complex hierarchy with that of the real hierarchy in solving these instances with $d = 2, k = 1$ and with approximately smallest chordal extensions or the maximal chordal extension. The results are displayed in Table 2. The column "CE" indicates which type of chordal extensions we use: "min" represents approximately smallest chordal extensions and "max" represents the maximal chordal extension.

As we can see from Table 2, for the complex hierarchy, the bound with the maximal chordal extension is slightly better than the bound with approximately smallest chordal

Table 3 The complex hierarchy with different sparse orders. Here, "gap" is the relative gap with respect to the bound given by the real hierarchy in Table 2

n	k	CE	mb	opt	Time	Gap
55	1	min	6	−24.6965	0.95	16.36%
55	2	min	38	−21.4174	13.3	0.91%

Table 4 The results for AC-OPF instances: typical operating conditions

Case	Order	Complex				Real			
		mb	opt	Time	Gap	mb	opt	Time	Gap
14_ieee	1st	6	2.1781e3	0.07	0.00%	6	2.1781e3	0.09	0.00%
30_ieee	1st	8	7.5472e3	0.12	8.06%	8	7.5472e3	0.15	8.06%
	1.5th	12	8.2073e3	0.66	0.02%	22	8.2085e3	0.97	0.00%
39_epri	1st	8	1.3565e4	0.17	2.00%	8	1.3565e4	0.22	2.00%
	1.5th	14	1.3765e4	1.08	0.55%	25	1.3842e4	1.12	0.00%
57_ieee	1st	12	3.7588e4	0.27	0.00%	12	3.7588e4	0.32	0.00%
89_pegase	1st	24	1.0670e5	0.72	0.55%	24	1.0670e5	0.74	0.55%
	1.5th	96	1.0709e5	263	0.19%	184	1.0715e5	1232	0.13%
118_ieee	1st	10	9.6900e4	0.49	0.32%	10	9.6901e4	0.57	0.32%
	1.5th	20	9.7199e4	5.22	0.02%	37	9.7214e4	8.78	0.00%
162_ieee_dtc	1st	28	1.0164e5	1.49	5.96%	28	1.0164e5	1.51	5.96%
	1.5th	40	1.0249e5	17.1	5.17%	74	1.0645e5	87.5	1.51%
	2nd	146	–	–	–	282	–	–	–
179_goc	1st	10	7.5016e5	0.72	0.55%	10	7.5016e5	0.77	0.55%
	1.5th	20	7.5078e5	6.77	0.46%	37	7.5382e5	10.6	0.06%
300_ieee	1st	14	5.5424e5	1.41	1.94%	16	5.5424e5	1.49	1.94%
	1.5th	22	5.6455e5	19.1	0.12%	40	5.6522e5	27.3	0.00%
1354_pegase	1st	26	1.2172e6	10.9	3.30%	26	1.2172e6	13.1	3.30%
	1.5th	26	1.2304e6	255	2.29%	49	1.2514e6	392	0.59%
2383wp_k	1st	48	1.8620e6	36.2	0.33%	50	1.8617e6	43.4	0.35%
2869_pegase	1st	26	2.4387e6	47.2	0.98%	26	2.4388e6	67.3	0.97%
	1.5th	98	2.4586e6	1666	0.17%	191	–	–	–

extensions. Both of them have relative gaps around 20% with respect to the bound given by the real hierarchy with approximately smallest chordal extensions. The gap can be reduced if we increase the sparse order of the complex hierarchy. For instance, for the random CPOP with $p = 10$, the complex hierarchy with $d = 2, k = 1$ has a relative gap of 16.36% and the complex hierarchy with $d = 2, k = 2$ has a relative gap of 0.91%; see Table 3.

On the other hand, when using approximately smallest chordal extensions and with $k = 1$, the complex hierarchy is over 9 times faster then the real hierarchy for each instance. Due to the limitation of memory, the real hierarchy with the maximal chordal extension is unsolvable.

Table 5 The results for AC-OPF instances: congested operating conditions

Case	Order	Complex				Real			
		mb	opt	Time	Gap	mb	opt	Time	Gap
14_ieee	1st	6	5.6886e3	0.06	5.18%	6	5.6886e3	0.07	5.18%
	1.5th	14	5.9981e3	0.39	0.02%	22	5.9994e3	0.52	0.00%
30_ieee	1st	8	1.7253e4	0.13	4.40%	8	1.7253e4	0.14	4.40%
	1.5th	12	1.7941e4	0.78	0.59%	22	1.8044e4	1.09	0.00%
39_epri	1st	8	2.4523e5	0.19	1.78%	8	2.4523e5	0.21	1.78%
	1.5th	14	2.4707e5	1.08	1.04%	25	2.4966e5	1.70	0.00%
57_ieee	1st	12	4.9289e5	0.22	0.00%	12	4.9289e5	0.29	0.00%
89_pegase	1st	24	1.0052e5	0.68	22.78%	24	1.0052e5	0.75	22.78%
	1.5th	96	1.0145e5	196	22.06%	184	1.0322e5	1448	20.71%
	2nd	224	–	–	–	426	–	–	–
118_ieee	1st	10	1.9375e5	0.58	20.02%	10	1.9375e5	0.65	20.02%
	1.5th	20	2.0193e5	6.15	16.64%	37	2.2286e5	8.58	9.00%
	2nd	48	2.2318e5	125	7.87%	92	–	–	–
162_ieee_dtc	1st	28	1.1206e5	1.67	7.38%	28	1.1206e5	1.71	7.38%
	1.5th	40	1.1284e5	22.5	6.74%	74	1.1955e5	81.4	1.19%
	2nd	146	–	–	–	282	–	–	–
179_goc	1st	10	1.7224e6	0.77	10.85%	10	1.7224e6	0.78	10.85%
	1.5th	20	1.8438e6	7.75	4.57%	37	1.9226e6	8.57	0.48%
300_ieee	1st	14	6.7932e5	1.22	0.83%	16	6.7932e5	1.69	0.83%
	1.5th	22	6.8411e5	19.4	0.13%	40	6.8493e5	25.2	0.01%
1354_pegase	1st	26	1.4871e6	9.55	0.75%	26	1.4872e6	12.7	0.74%
	1.5th	26	1.4929e6	265	0.36%	49	1.4929e6	379	0.36%
2383wp_k	1st	48	2.7913e5	36.0	0.00%	50	2.7913e5	39.3	0.00%
2869_pegase	1st	26	2.9059e6	44.8	0.81%	26	2.9059e6	67.8	0.81%
	1.5th	98	2.9262e6	1918	0.12%	191	–	–	–

5.2 AC-OPF Instances

The AC optimal power flow (AC-OPF) is a central problem in power systems, which aims to minimize the generation cost of an alternating current transmission network under the physical constraints (Kirchhoff's laws, Ohm's law, and power balance equations) as well as operational constraints. Mathematically, it can be formulated as the following CPOP:

Table 6 The results for AC-OPF instances: small angle difference conditions

Case	Order	Complex				Real			
		mb	opt	Time	Gap	mb	opt	Time	Gap
14_ieee	1st	6	2.7743e3	0.10	0.09%	6	2.7743e3	0.10	0.09%
30_ieee	1st	8	7.5472e3	0.16	8.06%	8	7.5472e3	0.17	8.06%
	1.5th	12	8.2072e3	0.83	0.02%	22	8.2085e3	0.94	0.00%
39_epri	1st	8	1.4791e4	0.14	0.29%	8	1.4791e4	0.18	0.29%
	1.5th	14	1.4831e4	0.91	0.02%	25	1.4832e4	1.16	0.01%
57_ieee	1st	12	3.8646e4	0.23	0.04%	12	3.8646e4	0.32	0.04%
89_pegase	1st	24	1.0672e5	0.61	0.54%	24	1.0672e5	0.75	0.54%
	1.5th	96	1.0700e5	159	0.27%	184	1.0713e5	1201	0.15%
118_ieee	1st	10	1.0191e5	0.48	3.10%	10	1.0191e5	0.56	3.10%
	1.5th	20	1.0239e5	5.52	2.64%	37	1.0336e5	8.56	1.71%
	2nd	48	1.0401e5	120	1.10%	92	–	–	–
162_ieee_dtc	1st	28	1.0283e5	1.39	5.39%	28	1.0283e5	1.57	5.39%
	1.5th	40	1.0434e5	20.6	4.00%	74	1.0740e5	69.1	1.19%
	2nd	146	–	–	–	282	–	–	–
179_goc	1st	10	7.5261e5	0.75	1.30%	10	7.5261e5	1.12	1.30%
	1.5th	20	7.5361e5	7.43	1.17%	37	7.5583e5	7.49	0.88%
300_ieee	1st	14	5.6162e5	1.25	0.72%	16	5.6168e5	1.81	0.71%
	1.5th	22	5.6557e5	20.9	0.02%	40	5.6572e5	22.8	0.00%
1354_pegase	1st	26	1.2172e6	10.4	3.30%	26	1.2172e6	13.1	3.30%
	1.5th	26	1.2358e6	259	1.83%	49	1.2586e6	260	0.02%
2383wp_k	1st	48	1.9060e6	35.5	0.27%	50	1.9061e6	48.6	0.27%
2869_pegase	1st	26	2.4488e6	40.2	0.81%	26	2.4490e6	57.1	0.80%
	1.5th	96	2.4495e6	1879	0.78%	191	–	–	–

$$
\begin{cases}
\inf\limits_{V_i, S_s^g} & \sum_{s\in G}(\mathbf{c}_{2s}(\Re(S_s^g))^2 + \mathbf{c}_{1s}\Re(S_s^g) + \mathbf{c}_{0s}) \\
\text{s.t.} & \angle V_r = 0, \\
& \mathbf{S}_s^{gl} \le S_s^g \le \mathbf{S}_s^{gu}, \quad \forall s \in G, \\
& \boldsymbol{v}_i^l \le |V_i| \le \boldsymbol{v}_i^u, \quad \forall i \in N, \\
& \sum_{s\in G_i} S_s^g - \mathbf{S}_i^d - \mathbf{Y}_i^{sh}|V_i|^2 = \sum_{(i,j)\in E_i\cup E_i^R} S_{ij}, \quad \forall i \in N, \\
& S_{ij} = (\mathbf{Y}_{ij}^* - \mathbf{i}\frac{\mathbf{b}_{ij}^c}{2})\frac{|V_i|^2}{|\mathbf{T}_{ij}|^2} - \mathbf{Y}_{ij}^*\frac{V_iV_j^*}{\mathbf{T}_{ij}}, \quad \forall (i,j) \in E, \\
& S_{ji} = (\mathbf{Y}_{ij}^* - \mathbf{i}\frac{\mathbf{b}_{ij}^c}{2})|V_j|^2 - \mathbf{Y}_{ij}^*\frac{V_i^*V_j}{\mathbf{T}_{ij}^*}, \quad \forall (i,j) \in E, \\
& |S_{ij}| \le \mathbf{s}_{ij}^u, \quad \forall (i,j) \in E \cup E^R, \\
& \boldsymbol{\theta}_{ij}^{\Delta l} \le \angle(V_iV_j^*) \le \boldsymbol{\theta}_{ij}^{\Delta u}, \quad \forall (i,j) \in E,
\end{cases}
\tag{5.1}
$$

where V_i is the voltage, S_s^g is the power generation, S_{ij} is the power flow (all are complex variables; $\Re(\cdot)$ and $\angle\cdot$ stand for the real part and the angle of a complex number, respectively) and all symbols in boldface are constants ($\mathbf{Y}$: admittance, $\boldsymbol{\theta}^{\Delta}$: voltage angle difference limit). Notice that G is the collection of generators and N is the collection of buses. For a full description on the AC-OPF problem, the reader may refer to [4] as well as [5]. We will consider AC-OPF instances for which the following assumption holds.

AC-OPF Assumption There is at most one generator attached to each bus, i.e., $|G_i| \le 1$ for all $i \in N$. In this case, we assume that a generator $s \in G$ is attached to the bus $i_s \in N$.

If **AC-OPF Assumption** holds, then we can eliminate the variables S_s^g from (5.1) to obtain a CPOP involving only the voltage V_i:

$$\begin{cases} \inf\limits_{V_i} & \sum_{s\in G}(\mathbf{c}_{2s}\Re(\mathbf{S}_{i_s}^d + \mathbf{Y}_{i_s}^{sh}|V_{i_s}|^2 + \sum_{(i_s,j)\in E_{i_s}\cup E_{i_s}^R} S_{i_sj})^2 \\ & \qquad +\mathbf{c}_{1s}\Re(\mathbf{S}_{i_s}^d + \mathbf{Y}_{i_s}^{sh}|V_{i_s}|^2 + \sum_{(i_s,j)\in E_{i_s}\cup E_{i_s}^R} S_{i_sj}) + \mathbf{c}_{0s}) \\ \text{s.t.} & \angle V_r = 0, \\ & \mathbf{S}_s^{gl} \le \mathbf{S}_{i_s}^d + \mathbf{Y}_{i_s}^{sh}|V_{i_s}|^2 + \sum_{(i_s,j)\in E_{i_s}\cup E_{i_s}^R} S_{i_sj} \le \mathbf{S}_s^{gu}, \quad \forall s \in G, \\ & \boldsymbol{v}_i^l \le |V_i| \le \boldsymbol{v}_i^u, \quad \forall i \in N, \\ & S_{ij} = (\mathbf{Y}_{ij}^* - \mathbf{i}\frac{\mathbf{b}_{ij}^c}{2})\frac{|V_i|^2}{|\mathbf{T}_{ij}|^2} - \mathbf{Y}_{ij}^*\frac{V_iV_j^*}{\mathbf{T}_{ij}}, \quad \forall (i,j) \in E, \\ & S_{ji} = (\mathbf{Y}_{ij}^* - \mathbf{i}\frac{\mathbf{b}_{ij}^c}{2})|V_j|^2 - \mathbf{Y}_{ij}^*\frac{V_i^*V_j}{\mathbf{T}_{ij}^*}, \quad \forall (i,j) \in E, \\ & |S_{ij}| \le \mathbf{s}_{ij}^u, \quad \forall (i,j) \in E \cup E^R, \\ & \boldsymbol{\theta}_{ij}^{\Delta l} \le \angle(V_iV_j^*) \le \boldsymbol{\theta}_{ij}^{\Delta u}, \quad \forall (i,j) \in E. \end{cases} \tag{5.2}$$

Note that in (5.2), if we substitute the expression of S_{ij} for S_{ij} in $|S_{ij}| \le \mathbf{s}_{ij}^u$, i.e., $S_{ij}\bar{S}_{ij} \le (\mathbf{s}_{ij}^u)^2$, we actually get a quartic constraint. To implement Shor's relaxation for QCQPs, we then relax it to a quadratic constraint by using the trick described in [5]. The minimum initial relaxation step of the complex hierarchy ($\mathrm{Q}_{\min,1}^{\text{cs-ts}}$) for (5.2) is able to provide a tighter lower bound than Shor's relaxation, which we thereby refer to as the 1.5th-order relaxation.

To tackle an AC-OPF instance, we first compute a locally optimal solution with a local solver and then rely on lower bounds obtained from SDP relaxations to certify global optimality. Suppose that the optimum reported by the local solver is AC and the optimum of a certain SDP relaxation is opt. The *optimality gap* between the locally optimal solution and the SDP relaxation is defined by

$$\text{gap} := \frac{\text{AC} - \text{opt}}{\text{AC}} \times 100\%.$$

If the optimality gap is less than 1%, then we accept the locally optimal solution to be globally optimal.

We select instances satisfying **AC-OPF Assumption** from the AC-OPF library *PGLiB* [4]. The number appearing in each instance name stands for the number of buses, which is equal to the number of complex variables involved in (5.2). For these instances, we compute the complex hierarchy as well as the real hierarchy with $k = 1$

and with the maximal chordal extension. The results are displayed in Tables 4, 5 and 6 in which the column "Order" indicates the relaxation order.

As we can see from the tables, for Shor's relaxation (the first-order relaxation), the complex hierarchy and the real hierarchy give the same lower bound (up to a given precision), while the complex hierarchy is slightly faster. For the 1.5th-order relaxation, the complex hierarchy typically gives a looser bound than the real hierarchy, whereas it is faster than the real hierarchy by a factor of $1 \sim 8$. Shor's relaxation is able to certify global optimality for 19 out of all 36 instances. For the remaining 17 instances, with the 1.5th-order relaxation the complex hierarchy is able to certify global optimality for 6 instances and the real hierarchy is able to certify global optimality for 11 instances.

6 Conclusions

In this paper, we have studied the sparsity-adapted complex hierarchy for complex polynomial optimization problems by taking into account both correlative and term sparsity. Numerical experiments demonstrate that the complex hierarchy offers a trade-off between the computational efficiency and the quality of obtained bounds relative to the real hierarchy. We hope that the sparsity-adapted complex hierarchy could help to tackle large-scale CPOPs arising either from the academic literature or from practical industrial problems.

Acknowledgements Both authors were supported by the Tremplin ERC Stg Grant ANR-18-ERC2-0004-01 (T-COPS project). The second author was supported by the FMJH Program PGMO (EPICS project), as well as the PEPS2 Program (FastOPF project) funded by AMIES and RTE. This work has benefited from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie Actions, Grant Agreement 813211 (POEMA) as well as from the AI Interdisciplinary Institute ANITI funding, through the French "Investing for the Future PIA3" program under the Grant agreement n°ANR-19-PI3A-0004.

References

1. Agler, J., Helton, W., McCullough, S., Rodman, L.: Positive semidefinite matrices with a given sparsity pattern. Linear Algebra Appl. **107**, 101–149 (1988)
2. Aittomaki, T., Koivunen, V.: Beampattern optimization by minimization of quartic polynomial. In: Piscataway, N.J. (ed.) 2009 IEEE/SP 15th Workshop on Statistical Signal Processing, pp. 437–440. IEEE (2009)
3. Aubry, A., De Maio, A., Jiang, B., Zhang, S.: Ambiguity function shaping for cognitive radar via complex quartic optimization. IEEE Trans. Signal Process. **61**(22), 5603–5619 (2013)
4. Babaeinejadsarookolaee, S., Birchfield, A., Christie, R.D., Coffrin, C., DeMarco, C., Diao, R., Ferris, M., Fliscounakis, S., Greene, S., Huang, R. et al.: The power grid library for benchmarking AC optimal power flow algorithms. (2019). arXiv preprint arXiv:1908.02788
5. Bienstock, D., Escobar, M., Gentile, C., Liberti, L.: Mathematical programming formulations for the alternating current optimal power flow problem. 4OR **18**(3), 249–292 (2020)
6. Blair, J.R., Peyton, B.: An introduction to chordal graphs and clique trees. In: George, A., Gilbert, J.R., Liu, J.W.H. (eds.) Graph Theory and Sparse Matrix Computation. The IMA Volumes in Mathematics and its Applications, vol. 56, pp. 1–29. Springer, New York, NY (1996)
7. Bodlaender, H.L., Koster, A.M.: Treewidth computations I. Upper bounds. Inf. Comput. **208**(3), 259–275 (2010)

8. Bromberger, S., Fairbanks, J.: and other contributors. JuliaGraphs/LightGraphs.jl: an optimized graphs package for the Julia programming language (2017)
9. Bugarin, F., Henrion, D., Lasserre, J.B.: Minimizing the sum of many rational functions. Math. Program. Comput. **8**(1), 83–111 (2016)
10. Chen, T., Lasserre, J.-B., Magron, V., Pauwels, E.: Semialgebraic optimization for bounding Lipschitz constants of Relu networks. In: Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M.F., Lin, H. (eds.) Proceeding of Advances in Neural Information Processing Systems, vol. 33 (2020)
11. Dunning, I., Huchette, J., Lubin, M.: JuMP: a modeling language for mathematical optimization. SIAM Rev. **59**(2), 295–320 (2017)
12. D'Angelo, J.P., Putinar, M.: Polynomial optimization on odd-dimensional spheres. In: Emerging Applications of Algebraic Geometry, pp. 1–15. Springer (2009)
13. Fogel, F., Waldspurger, I., d'Aspremont, A.: Phase retrieval for imaging problems. Math. Program. Comput. **8**(3), 311–335 (2016)
14. Grone, R., Johnson, C.R., Sá, E.M., Wolkowicz, H.: Positive definite completions of partial Hermitian matrices. Linear Algebra Appl. **58**, 109–124 (1984)
15. Hilling, J.J., Sudbery, A.: The geometric measure of multipartite entanglement and the singular values of a hypermatrix. J. Math. Phys. **51**(7), 072102 (2010)
16. Josz, C., Molzahn, D.K.: Moment/sum-of-squares hierarchy for complex polynomial optimization. (2015). arXiv preprint arXiv:1508.02068
17. Josz, C., Molzahn, D.K.: Lasserre hierarchy for large scale polynomial optimization in real and complex variables. SIAM J. Optim. **28**(2), 1017–1048 (2018)
18. Klep, I., Magron, V., Povh, J.: Sparse noncommutative polynomial optimization. Math. Program. **2021**, 1–41 (2021)
19. Lasserre, J.-B.: Global optimization with polynomials and the problem of moments. SIAM J. Optim. **11**(3), 796–817 (2001)
20. Magron, V.: Interval enclosures of upper bounds of roundoff errors using semidefinite programming. ACM Trans. Math. Softw. **44**(4), 1–18 (2018)
21. Magron, V., Constantinides, G., Donaldson, A.: Certified roundoff error bounds using semidefinite programming. ACM Trans. Math. Softw. **43**(4), 1–34 (2017)
22. Magron, V., Wang, J.: TSSOS: a Julia library to exploit sparsity for large-scale polynomial optimization. In: The 16th Effective Methods in Algebraic Geometry Conference (2021). https://puremath.no/Contributed%20MEGA/papers/MEGA_2021_paper_17.pdf
23. Mariere, B., Luo, Z.-Q., Davidson, T.N.: Blind constant modulus equalization via convex optimization. IEEE Trans. Signal Process. **51**(3), 805–818 (2003)
24. Marshall, M.: Representations of non-negative polynomials, degree bounds and applications to optimization. Can. J. Math. **61**(1), 205–221 (2009)
25. Mosek, A.: The MOSEK optimization Suite. Version 9.0 (2019)
26. Toker, O., Ozbay, H.: On the complexity of purely complex μ computation and related problems in multidimensional systems. IEEE Trans. Autom. Control **43**(3), 409–414 (1998)
27. Vandenberghe, L., Andersen, M..S., et al.: Chordal graphs and semidefinite optimization. Found. Trends® Optim. **1**(4), 241–433 (2015)
28. Vreman, N., Pazzaglia, P., Wang, J., Magron, V., Maggio, M.: Stability of control systems under extended weakly-hard constraints. (2021). arXiv preprint arXiv:2101.11312
29. Waki, H., Kim, S., Kojima, M., Muramatsu, M.: Sums of squares and semidefinite programming relaxations for polynomial optimization problems with structured sparsity. SIAM J. Optim. **17**(1), 218–242 (2006)
30. Wang, J.: ChordalGraph: A Julia Package to Handle Chordal Graphs (2020). https://github.com/wangjie212/ChordalGraph
31. Wang, J., Maggio, M., Magron, V.: SparseJSR: A fast algorithm to compute joint spectral radius via sparse SOS decompositions. In: 2021 American Control Conference (ACC), pp. 2254–2259. IEEE (2021)
32. Wang, J., Magron, V.: Exploiting term sparsity in noncommutative polynomial optimization. Comput. Optim. Appl. **80**(2), 483–521 (2021)
33. Wang, J., Magron, V., Lasserre, J.-B.: Chordal-TSSOS: a moment-SOS hierarchy that exploits term sparsity with chordal extension. SIAM J. Optim. **31**(1), 114–141 (2021)
34. Wang, J., Magron, V., Lasserre, J.-B.: TSSOS: A moment-SOS hierarchy that exploits term sparsity. SIAM J. Optim. **31**(1), 30–58 (2021)

35. Wang, J., Magron, V., Lasserre, J.-B., Mai, N.H.A.: CS-TSSOS: Correlative and term sparsity for large-scale polynomial optimization. (2020). arXiv:2005.02828
36. Zhou, Q., Marecek, J.: Proper learning of linear dynamical systems as a non-commutative polynomial optimisation problem (2020). arXiv:2002.01444
37. Zhou, Q., Marecek, J., Shorten, R.N.: Fairness in forecasting and learning linear dynamical systems. (2020). arXiv:2006.07315