SIAM J. OPTIM.
Vol. 35, No. 2, pp. 593–621

SPARSE POLYNOMIAL OPTIMIZATION WITH UNBOUNDED SETS*

LEI HUANG†, SHUCHENG KANG‡, JIE WANG§, AND HENG YANG‡

Abstract. This paper considers sparse polynomial optimization with unbounded sets. When the problem possesses correlative sparsity, we propose a sparse homogenized moment-SOS hierarchy with perturbations to solve it. The new hierarchy introduces one extra auxiliary variable for each variable clique according to the correlative sparsity pattern. Under the running intersection property, we prove that this hierarchy asymptotically converges to a value close to the optimum. Furthermore, we also provide two sparse homogenized moment-SOS hierarchies without perturbations, each having asymptotic convergence to the exact optimum. As by-products, new Positivstellensätze are obtained for sparse positive polynomials on unbounded sets. Extensive numerical experiments demonstrate the power of our approaches in solving sparse polynomial optimization problems on unbounded sets with up to thousands of variables. Finally, we apply our approaches to tackle two trajectory optimization problems: block-moving with minimum work, and optimal control of Van der Pol.

Key words. polynomial optimization, unbounded set, moment-SOS hierarchy, sparsity, semidefinite relaxation

MSC codes. 90C23, 90C17, 90C22, 90C26

DOI. 10.1137/24M1636010

1. Introduction.

In this paper, we consider the polynomial optimization problem (POP),

$$\begin{cases} \inf & f(\mathbf{x}) \\ \text{s.t.} & g_j(\mathbf{x}) \geq 0, \quad j = 1, \ldots, m, \end{cases} \tag{1.1}$$

where $f(\mathbf{x}), g_j(\mathbf{x})$ are polynomials in $\mathbf{x} := (x_1, \ldots, x_n) \in \mathbb{R}^n$. Let K denote the feasible set of (1.1) and let $f_{\min}$ denote the optimal value of (1.1). Throughout the paper, we assume that (1.1) has a finite minimum, i.e., $f_{\min} > -\infty$. The moment-SOS (sum-of-squares) hierarchy proposed by Lasserre [21] is efficient for solving (1.1). Under the Archimedeanness of the constraining polynomials (the feasible set K must be compact in this case; see [21, 24]), it yields a sequence of semidefinite relaxations whose optimal values converge to $f_{\min}$. Furthermore, it was shown in [14, 34] that the moment-SOS hierarchy converges in finitely many steps if standard optimality conditions hold at every global minimizer. We refer the reader to the books and surveys [23, 25, 32, 36] for more general introductions to polynomial optimization.

---

*Received by the editors January 31, 2024; accepted for publication (in revised form) December 2, 2024; published electronically April 2, 2025.
https://doi.org/10.1137/24M1636010
Funding: This work was jointly funded by National Key R&D Program of China under grants 2023YFA1009401 and 2022YFA1005102, by the Strategic Priority Research Program of the Chinese Academy of Sciences under grants XDB0640000 and XDB0640200, and by the National Natural Science Foundation of China under grants 12201618 and 12171324.
†Department of Mathematics, University of California San Diego, La Jolla, CA 92093 USA (leh010@ucsd.edu).
‡School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138 USA (skang1@g.harvard.edu, hankyang@seas.harvard.edu).
§Academy of Mathematics and Systems Science, Chinese Academy of Sciences (AMSS-CAS), Beijing, China (wangjie212@amss.ac.cn).

When the feasible set K is unbounded, the classical moment-SOS hierarchy typically does not converge. There exist some works on solving polynomial optimization with unbounded sets. Based on the Karush–Kuhn–Tucker (KKT) conditions and Lagrange multipliers, Nie proposed tight moment-SOS relaxations for solving (1.1) [35]. The paper [15] proposed moment-SOS relaxations by adding sublevel set constraints. The resulting hierarchy of relaxations also converges under the Archimedeanness for the new constraints. Based on Putinar andVasilescu's Positivstellensätze [39, 40], Mai, Lasserre, and Magron [29] proposed a new hierarchy of moment-SOS relaxations by adding a small perturbation to the objective. Convergence to a value close to $f_{\min}$ was proved for the case when the optimal value is achievable.[1] The complexity of this new hierarchy was studied in [30]. Recently, a homogenized moment-SOS hierarchy was proposed in [13] to solve polynomial optimization with unbounded sets using homogenization techniques. Finite convergence was proved if standard optimality conditions hold at every global minimizer, including those at infinity. A theoretically interesting problem in polynomial optimization with unbounded sets is the case when the optimal value is not achievable. We refer the reader to [10, 13, 37, 43, 46] for related works.

A drawback of the moment-SOS hierarchy is its limited scalability. This is because the size of the matrices involved at the kth order relaxation is $\binom{n+k}{k}$, which increases rapidly as n and k grow, and current semidefinite program (SDP) solvers based on interior-point methods can typically solve SDPs involving matrices of moderate size (e.g., ≤ 2000) in reasonable time on a standard laptop [45]. An effective way to improve scalability is by exploiting the sparsity of input polynomials. There are two types of sparsity patterns in the literature to reduce the size of semidefinite relaxations: *correlative sparsity* and *term sparsity.* Correlative sparsity [47] considers the sparsity pattern of the variables. The resulting sparse moment-SOS hierarchy is obtained by building blocks of SDP matrices based on subsets of the input variables. Under the so-called running intersection property (RIP) and Archimedeanness, this sparse hierarchy was shown to have asymptotic convergence in [7, 19, 22]. In contrast, term sparsity, proposed by Wang, Magron, and Lasserre [49, 50], considers the sparsity of monomials or terms. A bilevel block moment-SOS hierarchy can be obtained by using a two-step iterative procedure (a support extension operation followed by a block closure or chordal extension operation) to exploit term sparsity. For both types of sparsity, if the size of the obtained SDP blocks is relatively small, the resulting SDP relaxations are more tractable, and computational costs can be significantly reduced. These methods have been successfully applied to solve problems in optimal power flow [16, 52], round-off error bound analysis [26], noncommutative polynomial optimization [18, 48], neural network verification [33], dynamical systems analysis [53], and other areas.

However, the sparse moment-SOS hierarchies discussed above may not converge when K is unbounded, as in the dense case. For the unbounded case, Mai, Lasserre, and Magron [31] recently provided a sparse version of Putinar and Vasilescu's Positivstellensätze. More specifically, it was proved that if the problem (1.1) admits a correlative sparsity pattern $(\mathbf{x}(1), \ldots, \mathbf{x}(p))$ satisfying the RIP and $f \geq 0$ on K, then for every $\epsilon > 0$, there exist sums of squares $\sigma_{0,\ell}$, $\sigma_{j,\ell}$ $(j \in J_\ell)$ in variables $\mathbf{x}(\ell)$ $(\ell = 1, \ldots, p)$ such that

$$f + \varepsilon \sum_{\ell=1}^{p} \left(1 + \sum_{x_i \in \mathbf{x}(\ell)} x_i^2\right)^d = \sum_{\ell=1}^{p} \frac{\sigma_{0,\ell} + \sum_{j \in J_\ell} \sigma_{j,\ell} g_j}{\Theta_\ell^k},$$

[1] In this paper, a value close to $f_{\min}$ means one that lies within a narrow interval around $f_{\min}$, and the width of this interval depends on the perturbation parameter.

where $d \geq 1 + \lfloor \deg(f)/2 \rfloor$ and Θ_ℓ^k, $\ell = 1, \ldots, p$, are typically high-degree denominators (see section 2.3 for related notation and concepts). Based on this, a sparse moment-SOS hierarchy with perturbations is proposed to solve sparse polynomial optimization with unbounded sets. However, due to the presence of high-degree denominators, it is limited to solving problems with up to 10 variables. As a result, the computational benefit of this sparse hierarchy is rather limited, making it essentially a theoretical result, as stated in [31].

Contributions. This paper studies sparse polynomial optimization with unbounded sets using homogenization techniques. Our new contributions are as follows.

I. When the problem (1.1) admits correlative sparsity, we propose a sparse homogenized reformulation that inherits the correlative sparsity pattern of the original problem. This sparse reformulation introduces two new types of variables: the homogenization variable, and the auxiliary variables associated with each variable clique. We then apply the sparse moment-SOS hierarchy to solve the new reformulation with a small perturbation. Under the RIP, we prove that the sequence of lower bounds produced by this hierarchy converges to a value close to $f_{\min}$.

II. To remove undesired perturbations, we also propose two alternative sparse homogenized reformulations of (1.1), at the cost of possibly increasing the maximal clique size. When (1.1) admits a chain-like correlative sparsity pattern, we show that one of these sparse homogenized reformulations inherits the correlative sparsity pattern, with the maximal clique size increased by three. We establish the asymptotic convergence of the resulting sparse moment-SOS hierarchies to $f_{\min}$.

III. Based on the sparse homogenized reformulations, novel Positivstellensätze are provided for sparse positive polynomials on unbounded sets.

IV. Diverse numerical experiments demonstrate that our approach significantly outperforms the standard sparse moment-SOS hierarchy in solving sparse polynomial optimization problems on unbounded sets. In fact, it enables us to handle such problems with up to thousands of variables!

V. To further illustrate its power, we apply our approach to tackle two trajectory optimization problems (block-moving with minimum work, and optimal control of Van der Pol) from the fields of robotics and control. Our approaches can achieve globally optimal solutions for those problems with high accuracy.

This paper is organized as follows. Section 2 reviews some basics about polynomial optimization. Section 3 introduces the sparse homogenized moment-SOS hierarchy with perturbations and presents its asymptotic convergence result. Positivstellensätze with perturbations are provided. In section 4, we introduce two alternative sparse homogenized moment-SOS hierarchies without perturbations and prove their asymptotic convergence. Positivstellensätze without perturbations are provided. Numerical experiments are presented in section 5, and applications to trajectory optimization are discussed in section 6. Section 7 draws conclusions and presents some discussions.

2. Notation and preliminaries.

Notation. The symbol $\mathbb{N}$ (resp., $\mathbb{R}$) denotes the set of nonnegative integers (resp., real numbers). For $n \in \mathbb{N}$, let $[n] := \{1, \ldots, n\}$. Let $\mathbf{x} := (x_1, \ldots, x_n)$ denote a tuple of variables, and let $\mathbf{x}^{\circ 2} := (x_1^2, \ldots, x_n^2)$. By slight abuse of notation, we also view $\mathbf{x}$ as a set, i.e., $\mathbf{x} = \{x_1, \ldots, x_n\}$. For $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, let

$$\mathbf{x}^{\boldsymbol{\alpha}} := x_1^{\alpha_1} \cdots x_n^{\alpha_n}, \ \ |\boldsymbol{\alpha}| := \alpha_1 + \cdots + \alpha_n.$$

For $k \in \mathbb{N}$, let $\mathbb{N}^n_k := \{\boldsymbol{\alpha} \in \mathbb{N}^n \mid |\boldsymbol{\alpha}| \leq k\}$. Denote by $[\mathbf{x}]_k$ the vector of all monomials in $\mathbf{x}$ with degrees $\leq k$, i.e.,

$$[\mathbf{x}]_k := [1, x_1, x_2, \ldots, x_1^2, x_1x_2, \ldots, x_1^k, x_1^{k-1}x_2, \ldots, x_n^k]^\intercal.$$

Let $\mathbb{R}[\mathbf{x}] := \mathbb{R}[x_1, \ldots, x_n]$ be the ring of polynomials in $\mathbf{x}$ with real coefficients, and let $\mathbb{R}[\mathbf{x}]_k \subseteq \mathbb{R}[\mathbf{x}]$ be the subset of polynomials with degrees $\leq k$. For a polynomial $p \in \mathbb{R}[\mathbf{x}]$, denote by $\deg(p)$, $p^{(\infty)}$, $\tilde{p}$ its total degree, highest degree part, and homogenization with respect to the homogenization variable x_0 (i.e., $\tilde{p}(\tilde{\mathbf{x}}) = x_0^{\deg(p)} p(\mathbf{x}/x_0)$ with $\tilde{\mathbf{x}} := (x_0, x_1, \ldots, x_n)$), respectively. A homogeneous polynomial is said to be a *form.* A form p is *positive definite* if $p(\mathbf{x}) > 0$ for all nonzero $\mathbf{x} \in \mathbb{R}^n$. We write $A \succeq 0$ to indicate that a symmetric matrix A is positive semidefinite. For a vector $\mathbf{v} \in \mathbb{R}^n$, $\|\mathbf{v}\|$ denotes its standard Euclidean norm. We write $\mathbf{0}$ (resp., $\mathbf{1}$) for the zero (resp., all-one) vector whose dimension is clear from the context. For $t \in \mathbb{R}$, $\lceil t \rceil$ denotes the smallest integer greater than or equal to t.

2.1. Some basics for polynomial optimization. We review some basics in real algebraic geometry and polynomial optimization, referring the reader to [23, 25, 36] for more details.

A subset $I \subseteq \mathbb{R}[\mathbf{x}]$ is called an *ideal* if $I \cdot \mathbb{R}[\mathbf{x}] \subseteq I$, $I + I \subseteq I$. For a polynomial tuple $h := (h_1, \ldots, h_l)$, Ideal$[h]$ denotes the ideal generated by h, i.e.,

$$\mathrm{Ideal}[h] := h_1 \cdot \mathbb{R}[\mathbf{x}] + \cdots + h_l \cdot \mathbb{R}[\mathbf{x}].$$

For $k \in \mathbb{N}$, the kth degree truncation of Ideal$[h]$ is

$$\mathrm{Ideal}[h]_k := h_1 \cdot \mathbb{R}[\mathbf{x}]_{k-\deg(h_1)} + \cdots + h_l \cdot \mathbb{R}[\mathbf{x}]_{k-\deg(h_l)}.$$

Given a subset of variables $\mathbf{x}' \subseteq \mathbf{x}$, and a polynomial tuple $h \in \mathbb{R}[\mathbf{x}']^l$, we denote the ideal generated by h in $\mathbb{R}[\mathbf{x}']$ by

$$\mathrm{Ideal}[h, \mathbf{x}'] := h_1 \cdot \mathbb{R}[\mathbf{x}'] + \cdots + h_l \cdot \mathbb{R}[\mathbf{x}'].$$

Its kth degree truncation is defined as

$$\mathrm{Ideal}[h, \mathbf{x}']_k := h_1 \cdot \mathbb{R}[\mathbf{x}']_{k-\deg(h_1)} + \cdots + h_l \cdot \mathbb{R}[\mathbf{x}']_{k-\deg(h_l)}.$$

A polynomial $p \in \mathbb{R}[\mathbf{x}]$ is said to be a *sum of squares* (SOS) if $p = p_1^2 + \cdots + p_t^2$ for some $p_1, \ldots, p_t \in \mathbb{R}[\mathbf{x}]$. The set of all SOS polynomials in $\mathbb{R}[\mathbf{x}]$ is denoted by $\Sigma[\mathbf{x}]$. For $k \in \mathbb{N}$, let $\Sigma[\mathbf{x}]_k := \Sigma[\mathbf{x}] \cap \mathbb{R}[\mathbf{x}]_k$. For a polynomial tuple $g = (g_1, \ldots, g_m)$, the *quadratic module* generated by g is defined as

$$\mathrm{QM}[g] := \Sigma[\mathbf{x}] + g_1 \cdot \Sigma[\mathbf{x}] + \cdots + g_m \cdot \Sigma[\mathbf{x}]. \tag{2.1}$$

For $k \in \mathbb{N}$, the kth degree truncation of QM$[g]$ is

$$\mathrm{QM}[g]_k := \Sigma[\mathbf{x}]_k + g_1 \cdot \Sigma[\mathbf{x}]_{k-\deg(g_1)} + \cdots + g_m \cdot \Sigma[\mathbf{x}]_{k-\deg(g_m)}. \tag{2.2}$$

Similarly, if $g \in \mathbb{R}[\mathbf{x}']^m$ for $\mathbf{x}' \subseteq \mathbf{x}$, its quadratic module generated by g in $\mathbb{R}[\mathbf{x}']$ and kth degree truncation are denoted as

$$\begin{aligned} \mathrm{QM}[g, \mathbf{x}'] &:= \Sigma[\mathbf{x}'] + g_1 \cdot \Sigma[\mathbf{x}'] + \cdots + g_m \cdot \Sigma[\mathbf{x}'], \\ \mathrm{QM}[g, \mathbf{x}']_k &:= \Sigma[\mathbf{x}']_k + g_1 \cdot \Sigma[\mathbf{x}']_{k-\deg(g_1)} + \cdots + g_m \cdot \Sigma[\mathbf{x}']_{k-\deg(g_m)}. \end{aligned}$$

The set $\mathrm{Ideal}[h]+\mathrm{QM}[g]$ is said to be *Archimedean* if there exists $R>0$ such that $R-\|\mathbf{x}\|^2\in\mathrm{Ideal}[h]+\mathrm{QM}[g]$. Clearly, if $p\in\mathrm{Ideal}[h]+\mathrm{QM}[g]$, then $p\geq 0$ on the semialgebraic set $S:=\{\mathbf{x}\in\mathbb{R}^n\mid h(\mathbf{x})=\mathbf{0},g(\mathbf{x})\geq\mathbf{0}\}$, while the converse does not always hold. However, if p is positive on S, and $\mathrm{Ideal}[h]+\mathrm{QM}[g]$ is Archimedean, we have $p\in\mathrm{Ideal}[h]+\mathrm{QM}[g]$. This conclusion is known as Putinar's Positivstellensätze [38].

For $k\in\mathbb{N}$, let $\mathbb{R}^{\mathbb{N}^n_{2k}}$ be the set of all real vectors indexed by $\mathbb{N}^n_{2k}$. Given $\mathbf{y}\in\mathbb{R}^{\mathbb{N}^n_{2k}}$, define the following Riesz linear functional:

$$\langle p,\mathbf{y}\rangle:=\sum_{|\boldsymbol{\alpha}|\leq 2k}p_{\boldsymbol{\alpha}}y_{\boldsymbol{\alpha}}\quad\forall p=\sum_{|\boldsymbol{\alpha}|\leq 2k}p_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}}\in\mathbb{R}[\mathbf{x}]_{2k}. \tag{2.3}$$

For a polynomial $p\in\mathbb{R}[\mathbf{x}]$ and $\mathbf{y}\in\mathbb{R}^{\mathbb{N}^n_{2k+\deg(p)}}$, the kth *localizing matrix* $M_k[p\mathbf{y}]$ associated with p is the symmetric matrix $M_k[p\mathbf{y}]$ indexed by $\mathbb{N}^n_k$ such that

$$q^{\intercal}(M_k[p\mathbf{y}])q=\left\langle p(q^{\intercal}[\mathbf{x}]_k)^2,\mathbf{y}\right\rangle \tag{2.4}$$

for all $q\in\mathbb{R}^{\mathbb{N}^n_k}$. In particular, if $p=1$, then $M_k[\mathbf{y}]$ is called the kth *moment matrix.* For $\mathbf{x}'\subseteq\mathbf{x}$ and $p\in\mathbb{R}[\mathbf{x}']$, let $M_k[p\mathbf{y},\mathbf{x}']$ be the localizing submatrix obtained by retaining only those rows and columns of $M_k[p\mathbf{y}]$ that are indexed by $\boldsymbol{\alpha}\in\mathbb{N}^n$ with $\mathbf{x}^{\boldsymbol{\alpha}}\in\mathbb{R}[\mathbf{x}']$.

2.2. The homogenized moment-SOS hierarchy. When the feasible set K is unbounded, the standard moment-SOS hierarchy typically fails to converge. In this section, we present the homogenization approach introduced in [13] to solve polynomial optimization on unbounded sets.

Let $\tilde{\mathbf{x}}=(x_0,x)\in\mathbb{R}^{n+1}$. For the feasible set K given in (1.1), define the homogenized set

$$\widetilde{K}:=\left\{\tilde{\mathbf{x}}\in\mathbb{R}^{n+1}\;\middle|\;\begin{array}{ll}\tilde{g}_j(\tilde{\mathbf{x}})\geq 0, & j\in[m],\\ x_0\geq 0, & \|\tilde{\mathbf{x}}\|^2=1.\end{array}\right\} \tag{2.5}$$

DEFINITION 2.1. *The set K is said to be* closed at infinity (∞) *if*

$$\widetilde{K}=\mathrm{cl}(\widetilde{K}\cap\{\tilde{\mathbf{x}}\in\mathbb{R}^{n+1}\mid x_0>0\}),$$

where $\mathrm{cl}(\cdot)$ *is the closure operator.*

We remark that the closedness at infinity may depend on the description polynomials for K [13]. For instance, consider the two sets

$$S_1=\{(x_1,x_2)\in\mathbb{R}^2:x_1-x_2^2\geq 0\},\quad S_2=\{(x_1,x_2)\in\mathbb{R}^2:x_1-x_2^2\geq 0,x_1\geq 0\},$$

which are the same. However, for their description polynomials, the set S_2 is closed at ∞ while S_1 is not, since one can check that $(0,-1,0)\notin\widetilde{S}_1\backslash\mathrm{cl}(\widetilde{S}_1\cap\{\tilde{\mathbf{x}}\in\mathbb{R}^3\mid x_0>0\})$. On the other hand, being closed at infinity is a generic property and is satisfied for generic semialgebraic sets (see [9]).

Recall that for a polynomial p, $p^{(\infty)}$ denotes the highest degree terms of p. Define

$$K^{(\infty)}:=\left\{\mathbf{x}\in\mathbb{R}^n\;\middle|\;\begin{array}{ll}g_j^{(\infty)}(\mathbf{x})\geq 0, & j\in[m],\\ \|\mathbf{x}\|^2-1=0.\end{array}\right\}$$

The following is the basic property for closedness at infinity. We give a short proof to keep the paper self-contained.

PROPOSITION 2.2. *Suppose that the set K is closed at infinity and f is bounded from below on K, i.e., $f_{\min} > -\infty$. Then, we have that $\tilde{f} - f_{\min} x_0^{\deg(f)} \geq 0$ on $\widetilde{K}$, and $f^{(\infty)} \geq 0$ on $K^{(\infty)}$.*

Proof. Suppose that $\tilde{\mathbf{u}} = (u_0, \mathbf{u}) \in \widetilde{K}$. If $u_0 \neq 0$, then $\mathbf{u}/u_0 \in K$ and

$$\tilde{f}(\tilde{\mathbf{u}}) - f_{\min} u_0^{\deg(f)} = u_0^{\deg(f)} (f(\mathbf{u}/u_0) - f_{\min}) \geq 0.$$

If $u_0 = 0$, we have $\tilde{\mathbf{u}} := (0, \mathbf{u}) \in \widetilde{K}$ and $\mathbf{u} \in K^{(\infty)}$. Since K is closed at infinity, there exists a sequence of $\tilde{\mathbf{u}}^{(k)} = (u_0^{(k)}, \mathbf{u}^{(k)}) \in \widetilde{K} \cap \{\tilde{\mathbf{x}} \in \mathbb{R}^{n+1} \mid x_0 > 0\}$ such that $\tilde{\mathbf{u}}^{(k)} \to \tilde{\mathbf{u}}$ and each $u_0^{(k)} > 0$. Note that $\mathbf{u}^{(k)}/u_0^{(k)} \in K$ and

$$\tilde{f}(\tilde{\mathbf{u}}^{(k)}) - f_{\min} \cdot (u_0^{(k)})^{\deg(f)} = (u_0^{(k)})^{\deg(f)} (f(\mathbf{u}^{(k)}/u_0^{(k)}) - f_{\min}) \geq 0.$$

As $k \to \infty$, we get $f^{(\infty)}(\mathbf{u}) \geq 0$ for every $\mathbf{u} \in K^{(\infty)}$, so $f^{(\infty)} \geq 0$ on $K^{(\infty)}$. This is equivalent to $\tilde{f}(\tilde{\mathbf{u}}^{(k)}) - f_{\min}$ for $\tilde{\mathbf{u}} := (0, \mathbf{u}) \in \widetilde{K}$. ∎

When K is closed at infinity, it follows from Proposition 2.2 that $f - \gamma \geq 0$ on K if and only if $\tilde{f}(\tilde{\mathbf{x}}) - \gamma x_0^d \geq 0$ on $\widetilde{K}$ with $d := \deg(f)$. Therefore, in this case, (1.1) is equivalent to the following homogenized optimization problem:

$$\begin{cases} \sup & \gamma \\ \text{s.t.} & \tilde{f}(\tilde{\mathbf{x}}) - \gamma x_0^d \geq 0 \quad \text{on} \quad \widetilde{K}. \end{cases} \tag{2.6}$$

By applying the standard moment-SOS relaxations to solve the homogenized reformulation (2.6), a homogenized moment-SOS hierarchy was proposed in [12, 13] to solve (1.1) with unbounded sets. Asymptotic and finite convergence results were established under some generic assumptions.

2.3. Polynomial optimization with correlative sparsity. The Moment-SOS hierarchy with correlative sparsity was first studied in [47]. Suppose that the subsets of variables $\mathbf{x}(1), \ldots, \mathbf{x}(p) \subseteq \mathbf{x}$. We say that POP (1.1) admits a *correlative sparsity pattern* (csp) $(\mathbf{x}(1), \ldots, \mathbf{x}(p))$ if the following conditions hold:

(i) The objective function $f \in \mathbb{R}[\mathbf{x}]$ can be written as

$$f = \sum_{\ell=1}^{p} f_\ell \text{ with } f_\ell \in \mathbb{R}[\mathbf{x}(\ell)] \text{ for } \ell \in [p].$$

(ii) There exist subsets $J_1, \ldots, J_p$ of $[m]$ such that $\cup_{\ell=1}^{p} J_\ell = [m]$, and for any $\ell \in [p]$ and $j \in J_\ell$, one has $g_j \in \mathbb{R}[\mathbf{x}(\ell)]$.

Let

$$d_{\min} := \max\{\lceil \deg(f)/2 \rceil, \lceil \deg(g_1)/2 \rceil, \ldots, \lceil \deg(g_m)/2 \rceil\}.$$

Given $k \geq d_{\min}$, the kth order sparse SOS relaxation for (1.1) is

$$\begin{cases} \sup & \gamma \\ \text{s.t.} & f - \gamma \in \mathrm{QM}[(g_j)_{j \in J_1}, \mathbf{x}(1)]_{2k} + \cdots + \mathrm{QM}[(g_j)_{j \in J_p}, \mathbf{x}(p)]_{2k}. \end{cases} \tag{2.7}$$

The dual of (2.7) is the following kth order sparse moment relaxation:

$$\begin{cases} \inf & \langle f, \mathbf{y} \rangle \\ \text{s.t.} & y_{\mathbf{0}} = 1, \quad M_k[\mathbf{y}, \mathbf{x}(\ell)] \succeq 0, \quad \ell \in [p], \\ & M_{k - \lceil \deg(g_j)/2 \rceil}[g_j \mathbf{y}, \mathbf{x}(\ell)] \succeq 0, \quad j \in J_\ell, \ell \in [p]. \end{cases} \tag{2.8}$$

The csp $(\mathbf{x}(1),\dots,\mathbf{x}(p))$ is said to satisfy the *running intersection property* (RIP) if for every $\ell \in [p-1]$, there exists some $s \in [\ell]$ such that

$$\mathbf{x}(\ell+1) \cap \bigcup_{j=1}^{\ell} \mathbf{x}(j) \subseteq \mathbf{x}(s). \tag{2.9}$$

The RIP is closely related to the notion of chordal graphs. A *graph* $G(V,E)$ consists of a set of nodes V and a set of edges $E \subseteq \{\{v_i,v_j\} \mid v_i \neq v_j, (v_i,v_j) \in V \times V\}$. A *clique* of a graph is a subset of nodes that induces a complete subgraph. A *maximal clique* is a clique that is not contained in any other clique. A graph is called a *chordal graph* if all its cycles of length at least four have a chord.[2] Any nonchordal graph $G(V,E)$ can always be extended to a chordal graph $G'(V,E')$ by adding appropriate edges to E, which is called a *chordal extension* of $G(V,E)$. We refer the reader to [2, 5] for more details. The RIP actually gives an equivalent characterization of chordal graphs; see Theorem 3.4 or Corollary 1 of [5].

THEOREM 2.3 (see [5]). *A connected graph is chordal if and only if its maximal cliques after an appropriate ordering satisfy the RIP.*

The csp graph of POP (1.1) is defined as the graph $G(V,E)$ with $V = \{x_1,\dots,x_n\}$ and $\{x_i,x_j\} \in E$ if x_i,x_j appear in the same term of the objective f or in the same constraint polynomial. By Theorem 2.3, the set of maximal cliques of some chordal extension of $G(V,E)$ gives a csp of POP (1.1) that satisfies the RIP. Under the RIP, it was shown in [7, 19, 22] that the sparse moment-SOS hierarchy (2.7)–(2.8) has asymptotic convergence.

THEOREM 2.4 (see [7, 19, 22]). *Suppose that* (1.1) *admits the csp* $(\mathbf{x}(1),\dots,\mathbf{x}(p))$ *satisfying the RIP, and the quadratic module* $QM[(g_j)_{j\in J_\ell},\mathbf{x}(\ell)]$ *is Archimedean for each* $\ell \in [p]$. *If* $f > 0$ *on* K, *then*

$$f \in QM[(g_j)_{j\in J_1},\mathbf{x}(1)] + \cdots + QM[(g_j)_{j\in J_p},\mathbf{x}(p)].$$

3. The sparse homogenized moment-SOS hierarchy with perturbations. In this section, we give a hierarchy of sparse homogenized moment-SOS relaxations to solve polynomial optimization with unbounded sets. Suppose that POP (1.1) admits a csp $(\mathbf{x}(1),\dots,\mathbf{x}(p))$, and K is unbounded. Note that we cannot directly apply the sparse relaxations (2.7)–(2.8) to solve the homogenized reformation (2.6), since the spherical constraint $\|\tilde{\mathbf{x}}\|^2 = 1$ destroys the csp of (1.1). To overcome this difficulty, we give a sparse homogenized reformulation for (1.1) by introducing a tuple of auxiliary variables.

Let $\{J_1,\dots,J_p\}$ be subsets of $[m]$ such that $\cup_{\ell=1}^{p} J_\ell = [m]$, and for any $\ell \in [p]$ and $j \in J_\ell$, we have $g_j \in \mathbb{R}[\mathbf{x}(\ell)]$. Define the set

$$\widetilde{K}_s^1 := \left\{ (\tilde{\mathbf{x}},\mathbf{w}) \in \mathbb{R}^{n+1+p} \;\middle|\; \begin{array}{ll} x_0 \geq 0, \quad \tilde{g}_j(\tilde{\mathbf{x}}) \geq 0, & j \in [m], \\ \|\tilde{\mathbf{x}}(\ell)\|^2 + w_\ell^2 = 1, & \ell \in [p], \end{array} \right\} \tag{3.1}$$

where $\tilde{\mathbf{x}}(\ell) := (x_0,\mathbf{x}(\ell))$, and $\mathbf{w} := (w_1,\dots,w_p)$ is a tuple of auxiliary variables. The difference between $\widetilde{K}_s^1$ and $\widetilde{K}$ is that we replace the single nonsparse spherical constraint $\|\tilde{\mathbf{x}}\|^2 = 1$ by multiple sparse spherical constraints $\|\tilde{\mathbf{x}}(\ell)\|^2 + w_\ell^2 = 1$, $\ell \in [p]$. Consequently, $\widetilde{K}_s^1$ retains the csp of POP (1.1).

Let $d := \deg(f)$ and $d_0 := 2\lceil \frac{d}{2} \rceil$. Consider the sparse homogenized reformulation for (1.1) with perturbations,

[2] A chord is an edge that joins two nonconsecutive nodes in a cycle.

$$\text{(3.2)} \qquad \begin{cases} \sup & \gamma \\ \text{s.t.} & \tilde{f}(\tilde{\mathbf{x}}) + \epsilon \cdot \left(\sum_{i=0}^n x_i^{d_0} + \sum_{\ell=1}^p w_\ell^{d_0}\right) - \gamma x_0^d \geq 0 \quad \forall (\tilde{\mathbf{x}}, \mathbf{w}) \in \widetilde{K}_s^1, \end{cases}$$

where $\epsilon \geq 0$ is a tunable parameter. Denote by $f^{(\epsilon)}$ the optimal value of (3.2), and let

$$h_\ell := \|\tilde{\mathbf{x}}(\ell)\|^2 + w_\ell^2 - 1, \quad \ell \in [p].$$

For $k \geq d_{\min}$, the kth order sparse homogenized SOS relaxation for (3.2) is

(3.3)

$$\begin{cases} \sup & \gamma \\ \text{s.t.} & \tilde{f}(\tilde{\mathbf{x}}) + \epsilon \cdot \left(\sum_{i=0}^n x_i^{d_0} + \sum_{\ell=1}^p w_\ell^{d_0}\right) - \gamma x_0^d \\ & \quad \in \sum_{\ell=1}^p \left(\mathrm{Ideal}[h_\ell, \tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}]_{2k} + \mathrm{QM}[\{x_0\} \cup \{\tilde{g}_j\}_{j \in J_\ell}, \tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}]_{2k}\right). \end{cases}$$

The dual of (3.3) is the following kth order sparse moment relaxation:

$$\text{(3.4)} \qquad \begin{cases} \inf & \left\langle \tilde{f}(\tilde{\mathbf{x}}) + \epsilon \cdot \left(\sum_{i=0}^n x_i^{d_0} + \sum_{\ell=1}^p w_\ell^{d_0}\right), \mathbf{y} \right\rangle \\ \text{s.t.} & \langle x_0^d, \mathbf{y} \rangle = 1, \quad M_{k-1}[h_\ell \mathbf{y}, \tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}] = 0, \quad \ell \in [p], \\ & M_{k-1}[x_0, \tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}] \succeq 0, \quad M_k[\mathbf{y}, \tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}] \succeq 0, \quad \ell \in [p], \\ & M_{k - \lceil \deg(g_j)/2 \rceil}[\tilde{g}_j \mathbf{y}, \tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}] \succeq 0, \quad j \in J_\ell, \ell \in [p]. \end{cases}$$

The hierarchy of relaxations (3.3)–(3.4) is called the sparse homogenized moment-SOS hierarchy for (1.1). Let ρ_k denote the optimum of (3.3).

3.1. Convergence analysis. When K is closed at infinity, we establish the relationship between the optimal values of (1.1) and (3.2).

THEOREM 3.1. *Suppose that K is closed at infinity,* (1.1) *has a finite minimum which is achievable, and $\mathbf{x}^*$ is a minimizer of* (1.1)*. Let $f^{(\epsilon)}$ be the optimal value of* (3.2)*. For every $\epsilon > 0$, the following inequalities hold:*

$$f_{\min} < f^{(\epsilon)} \leq f_{\min} + \epsilon \cdot p \cdot \left(1 + \|\mathbf{x}^*\|^2\right)^{\frac{d}{2}}.$$

Moreover, one has $f^{(\epsilon)} = f_{\min}$ when $\epsilon = 0$.

Proof. Since K is closed at infinity and $f - f_{\min} \geq 0$ on K, we have $\tilde{f} - f_{\min} x_0^d \geq 0$ on $\widetilde{K}$. Note that $\mathbf{0} \notin \widetilde{K}_s^1$. Hence, for every $(\tilde{\mathbf{x}}, \mathbf{w}) \in \widetilde{K}_s^1$, we have

$$\tilde{f}(\tilde{\mathbf{x}}) - f_{\min} x_0^d + \epsilon \cdot \left(\sum_{i=0}^n x_i^{d_0} + \sum_{\ell=1}^p w_\ell^{d_0}\right) > 0,$$

which implies $f_{\min} < f^{(\epsilon)}$. On the other hand, let

$$\tilde{\mathbf{x}}^* := \frac{(1, \mathbf{x}^*)}{\sqrt{1 + \|\mathbf{x}^*\|^2}}, \; w_\ell^* := \sqrt{1 - \|\tilde{\mathbf{x}}^*(\ell)\|^2} \; (\ell \in [p]).$$

Then it holds that

$$\tilde{g}_j(\tilde{\mathbf{x}}^*) = g_j(\mathbf{x}^*) / (\sqrt{1 + \|\mathbf{x}^*\|^2})^{\deg(g_j)} \geq 0, \quad j \in [m],$$

and so $(\tilde{\mathbf{x}}^*, \mathbf{w}^*) \in \widetilde{K}_s^1$. If γ is feasible for (3.2), then we have (noting that $d_0 \geq 2$)

$$\gamma \cdot (\tilde{x}_0^*)^d \leq \tilde{f}(\tilde{\mathbf{x}}^*) + \epsilon \cdot \left(\sum_{i=0}^{n} (\tilde{x}_i^*)^{d_0} + \sum_{\ell=1}^{p} (w_\ell^*)^{d_0} \right) \leq (\tilde{x}_0^*)^d f_{\min} + \epsilon \cdot p.$$

It follows that $f^{(\epsilon)} \leq f_{\min} + \epsilon \cdot p/(\tilde{x}_0^*)^d = f_{\min} + \epsilon \cdot p \cdot \left(1 + \|\mathbf{x}^*\|^2\right)^{\frac{d}{2}}$. When $\epsilon = 0$, the above implies that $\gamma \leq f_{\min}$ for every feasible γ of (3.2). Thus, we have $f^{(0)} = f_{\min}$. □

Remark 3.2. The auxiliary variables $w_1, \ldots, w_p$ are necessary for Theorem 3.1 to hold. For instance, let $\mathbf{x}(1) = \{x_1, x_2\}$, $\mathbf{x}(2) = \{x_2, x_3\}$, and consider the unconstrained optimization problem of minimizing $f = (x_1 - x_2)^2 + (x_2 - 1)^2 + (x_2 - 2x_3)^2$. Clearly, we have $f_{\min} = 0$, and the unique minimizer is $(1, 1, \frac{1}{2})$. In this case, the sparse homogenized reformulation (3.2) without auxiliary variables reads as

$$\begin{cases} \sup & \gamma \\ \text{s.t.} & \tilde{f}(\tilde{\mathbf{x}}) + \epsilon \cdot (x_0^2 + x_1^2 + x_2^2 + x_3^2) - \gamma x_0^2 \geq 0 \, \forall \tilde{\mathbf{x}} \in \widetilde{K}_s^1, \end{cases} \tag{3.5}$$

where

$$\widetilde{K}_s^1 = \{(x_0, x_1, x_2, x_3) \in \mathbb{R}^4 : x_0^2 + x_1^2 + x_2^2 = 1, x_0^2 + x_2^2 + x_3^2 = 1\}.$$

Suppose that $\tilde{\mathbf{x}} = (x_0, \mathbf{x}) \in \widetilde{K}_s^1$. If $x_0 = 0$, we have

$$\begin{aligned} &\tilde{f}(\tilde{\mathbf{x}}) + \epsilon \cdot (x_0^2 + x_1^2 + x_2^2 + x_3^2) - \gamma x_0^2 \\ &= (x_1 - x_2)^2 + x_2^2 + (x_2 - 2x_3)^2 + \epsilon \cdot (x_1^2 + x_2^2 + x_3^2) \geq 0. \end{aligned}$$

If $x_0 \neq 0$, it holds that

$$\tilde{f}(\tilde{\mathbf{x}}) + \epsilon \cdot (x_0^2 + x_1^2 + x_2^2 + x_3^2) - \gamma x_0^2 = x_0^2 \left(\tilde{f}\left(\frac{x}{x_0}\right) + \epsilon \cdot \left(1 + \frac{x_1^2 + x_2^2 + x_3^2}{x_0^2}\right) - \gamma \right).$$

Note that for any $\tilde{\mathbf{x}} \in \widetilde{K}_s^1$, we have that $x_1^2 = x_3^2$. Let $t_1 = \frac{x_1}{x_0}, t_2 = \frac{x_2}{x_0}, t_3 = \frac{x_3}{x_0}$. Then, (3.5) is equivalent to

$$\begin{cases} \sup & \gamma \\ \text{s.t.} & f(\mathbf{t}) + \epsilon \cdot (1 + t_1^2 + t_2^2 + t_3^2) - \gamma \geq 0 \, \forall \mathbf{t} \in T, \end{cases} \tag{3.6}$$

where

$$T = \{\mathbf{t} := (t_1, t_2, t_3) \in \mathbb{R}^3 : t_1^2 = t_3^2\}.$$

If $t_1 = t_3$, we have

$$\begin{aligned} f(\mathbf{t}) &= (t_1 - t_2)^2 + (t_2 - 1)^2 + (t_2 - 2t_1)^2 \\ &= 5t_1^2 + 3t_2^2 - 6t_1 t_2 - 2t_2 + 1 \\ &= \left(3\sqrt{\frac{6}{11}} t_1 - \sqrt{\frac{11}{6}} t_2\right)^2 + \frac{1}{11} t_1^2 + \frac{7}{6}\left(t_2 - \frac{6}{7}\right)^2 + \frac{1}{7} \\ &\geq \frac{1}{7}. \end{aligned}$$

Similarly, if $t_1 = -t_3$, we have

$$f(\mathbf{t}) = (t_1+t_2)^2 + 4t_1^2 + 2\left(t_2 - \frac{1}{2}\right)^2 + \frac{1}{2} \geq \frac{1}{2}.$$

Thus, it follows that the optimal value of (3.5) is no less than $\frac{1}{7}$ for any $\epsilon \geq 0$. Since $f_{\min} = 0$, Theorem 3.1 fails.

The next lemma shows that (3.2) indeed inherits the csp of the original problem (1.1).

LEMMA 3.3. *Suppose that* (1.1) *admits the csp* $(\mathbf{x}(1),\ldots,\mathbf{x}(p))$. *Then,* (3.2) *admits the csp* $(\tilde{\mathbf{x}}(1)\cup\{w_1\},\ldots,\tilde{\mathbf{x}}(p)\cup\{w_p\})$. *Furthermore, if* $(\mathbf{x}(1),\ldots,\mathbf{x}(p))$ *satisfies the RIP, so does* $(\tilde{\mathbf{x}}(1)\cup\{w_1\},\ldots,\tilde{\mathbf{x}}(p)\cup\{w_p\})$.

Proof. The first assertion is immediate from the construction. The second assertion follows from the fact that for $\ell \in [p-1]$, it holds that

$$(\tilde{\mathbf{x}}(\ell+1)\cup\{w_{\ell+1}\}) \cap \bigcup_{j=1}^{\ell}(\tilde{\mathbf{x}}(j)\cup\{w_j\}) = \{x_0\} \cup \left(\mathbf{x}(\ell+1)\cap\bigcup_{j=1}^{\ell}\mathbf{x}(j)\right). \qquad \square$$

Now we prove that the sparse homogenized moment-SOS hierarchy (3.3)–(3.4) has asymptotic convergence to a value close to $f_{\min}$.

THEOREM 3.4. *Suppose that* (1.1) *admits the csp* $(\mathbf{x}(1),\ldots,\mathbf{x}(p))$ *satisfying the RIP, K is closed at infinity, and* (1.1) *has a finite minimum. Then, for every $\epsilon > 0$, we have $\rho_k \to f^{(\epsilon)}$ as $k \to \infty$.*

Proof. For any $\gamma < f^{(\epsilon)}$, we show that for each $(\tilde{\mathbf{x}},\mathbf{w}) \in \widetilde{K}_s^1$,

$$\tilde{f}(\tilde{\mathbf{x}}) + \epsilon\cdot\left(\sum_{i=0}^{n} x_i^{d_0} + \sum_{\ell=1}^{p} w_\ell^{d_0}\right) - \gamma x_0^d > 0. \tag{3.7}$$

If $x_0 = 0$, it follows from Proposition 2.2 that

$$\tilde{f}(\tilde{\mathbf{x}}) + \epsilon\cdot\left(\sum_{i=0}^{n} x_i^{d_0} + \sum_{\ell=1}^{p} w_\ell^{d_0}\right) - \gamma x_0^d \geq \epsilon\cdot\left(\sum_{i=1}^{n} x_i^{d_0} + \sum_{\ell=1}^{p} w_\ell^{d_0}\right) > 0.$$

If $x_0 \neq 0$, we have

$$\begin{aligned}&\tilde{f}(\tilde{\mathbf{x}}) + \epsilon\cdot\left(\sum_{i=0}^{n} x_i^{d_0} + \sum_{\ell=1}^{p} w_\ell^{d_0}\right) - \gamma x_0^d\\ &= \tilde{f}(\tilde{\mathbf{x}}) + \epsilon\cdot\left(\sum_{i=0}^{n} x_i^{d_0} + \sum_{\ell=1}^{p} w_\ell^{d_0}\right) - f^{(\epsilon)}x_0^d + (f^{(\epsilon)}-\gamma)x_0^d > 0.\end{aligned}$$

Thus, we have $\tilde{f}(\tilde{\mathbf{x}}) + \epsilon\cdot(\sum_{i=0}^{n} x_i^{d_0} + \sum_{\ell=1}^{p} w_\ell^{d_0}) - \gamma x_0^d > 0$ on $\widetilde{K}_s^1$ for any $\gamma < f^{(\epsilon)}$. The spherical constraint $h_\ell = 0$ implies that $\mathrm{Ideal}[h_\ell, \tilde{\mathbf{x}}(\ell)\cup\{w_\ell\}] + \mathrm{QM}[(\tilde{g}_j)_{j\in J_\ell}, \tilde{\mathbf{x}}(\ell)\cup\{w_\ell\}]$ is Archimedean in $\mathbb{R}[\tilde{\mathbf{x}}(\ell)\cup\{w_\ell\}]$ for each $\ell \in [p]$. By Lemma 3.3, POP (3.2) admits the csp $(\tilde{\mathbf{x}}(1)\cup\{w_1\},\ldots,\tilde{\mathbf{x}}(p)\cup\{w_p\})$ satisfying the RIP. It follows from Theorem 2.4 that

$$\begin{aligned}&\tilde{f}(\tilde{\mathbf{x}}) + \epsilon\cdot\left(\sum_{i=0}^{n} x_i^{d_0} + \sum_{\ell=1}^{p} w_\ell^{d_0}\right) - \gamma x_0^d\\ &\in \sum_{\ell=1}^{p}\left(\mathrm{Ideal}[h_\ell, \tilde{\mathbf{x}}(\ell)\cup\{w_\ell\}] + \mathrm{QM}[\{x_0\}\cup\{\tilde{g}_j\}_{j\in J_\ell}, \tilde{\mathbf{x}}(\ell)\cup\{w_\ell\}]\right).\end{aligned}$$

Thus, we obtain $\rho_k \to f^{(\epsilon)}$ as $k \to \infty$. $\square$

Remark 3.5. If there is no perturbation, i.e., $\epsilon = 0$, the hierarchy (3.3)–(3.4) still provides valid lower bounds on $f_{\min}$. However, the convergence to $f_{\min}$ may not occur. Below, we present two examples to illustrate this.

(1) Let us consider the unconstrained optimization problem with Delzell's polynomial $f = x_4^2(x_1^4x_2^2 + x_2^4x_3^2 + x_1^2x_3^4 - 3x_1^2x_2^2x_3^2) + x_3^8$, where there is only one clique $\mathbf{x} = \{x_1, x_2, x_3, x_4\}$. In the following, we prove that for arbitrary $\gamma < f_{\min} = 0$, it holds that

$$f - \gamma x_0^8 \notin \mathrm{Ideal}[\|\tilde{\mathbf{x}}\|^2 + w_1^2 - 1, \tilde{\mathbf{x}} \cup \{w_1\}] + \mathrm{QM}[x_0, \tilde{\mathbf{x}} \cup \{w_1\}].$$

Suppose that otherwise, there exist $h \in \mathbb{R}[\tilde{\mathbf{x}} \cup \{w_1\}]$, $\sigma \in \mathrm{QM}[x_0, \tilde{\mathbf{x}} \cup \{w_1\}]$ such that

$$f - \gamma x_0^8 = h \cdot (\|\tilde{\mathbf{x}}\|^2 + w_1^2 - 1) + \sigma.$$

Substituting $(0, 0)$ for (x_0, w_1) in the above identity, we get that $f \in \mathrm{Ideal}[\|\mathbf{x}\|^2 - 1] + \Sigma[\mathbf{x}]$, which is a contradiction as shown in [42].

(2) Let

$$\mathbf{x}(1) = \{x_1, x_2, x_3, x_4, x_5\}, \quad \mathbf{x}(2) = \{x_5, x_6, x_7\},$$

and consider the unconstrained optimization problem $\inf_{\mathbf{x}\in\mathbb{R}^7} f(\mathbf{x})$ with $f = f_1 + f_2$, where

$$\begin{aligned} f_1 &= (x_4^2 + x_5^2 + 1)\left(x_1^4x_2^2 + x_2^4x_3^2 + x_1^2x_3^4 - 3x_1^2x_2^2x_3^2\right) + x_3^8, \\ f_2 &= x_5^2x_6^2x_7^2. \end{aligned}$$

We show that for arbitrary $\gamma < f_{\min} = 0$, it holds that

$$\tilde{f} - \gamma x_0^8 \notin \sum_{\ell=1}^{2}(\mathrm{Ideal}[\|\tilde{\mathbf{x}}(\ell)\|^2 + w_\ell^2 - 1, \tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}] + \mathrm{QM}[x_0, \tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}]).$$

Suppose that otherwise, there exist $h_\ell \in \mathbb{R}[\tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}]$, $\sigma_\ell \in \mathrm{QM}[x_0, \tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}]$ such that

$$\tilde{f} - \gamma x_0^8 = \sum_{\ell=1}^{2}(h_\ell \cdot (\|\tilde{\mathbf{x}}(\ell)\|^2 + w_\ell^2 - 1) + \sigma_\ell). \tag{3.8}$$

Note that the homogenization of f_1 is

$$\tilde{f}_1 = (x_4^2 + x_5^2 + x_0^2)\left(x_1^4x_2^2 + x_2^4x_3^2 + x_1^2x_3^4 - 3x_1^2x_2^2x_3^2\right) + x_3^8.$$

Substituting $(0, 0, 1, 0, 0, 0)$ for $(x_0, x_5, x_6, x_7, w_1, w_2)$ in (3.8), we obtain

$$x_4^2\left(x_1^4x_2^2 + x_2^4x_3^2 + x_1^2x_3^4 - 3x_1^2x_2^2x_3^2\right) + x_3^8 = \sigma_0 + h_0 \cdot (x_1^2 + \cdots + x_4^2 - 1) \tag{3.9}$$

for some $\sigma_0 \in \Sigma[x_1, x_2, x_3, x_4]$, $h_0 \in \mathbb{R}[x_1, x_2, x_3, x_4]$. However, the left-hand side of (3.9) is Delzell's polynomial, and the above representation does not exist as shown in [42].

Remark 3.6. Actually, in the perturbed objective function, one can set d_0 to be any even positive integer. Our motivation for choosing $d_0 = 2\lceil d/2 \rceil$ is that it is the smallest positive integer such that the highest degree terms of $\tilde{f} - \gamma x_0^d + \epsilon \cdot (\sum_{i=0}^{n} x_i^d + \sum_{\ell=1}^{p} w_\ell^d)$

are positive definite. This typically leads to better numerical performance than other choices.

Remark 3.7. In theory, one needs to check whether $\widetilde{K} = \mathrm{cl}(\widetilde{K} \cap \{x_0 > 0\})$ to verify if the set K is closed at infinity, which is typically difficult. However, it is not necessary to do this in practice because of the following:

(1) It was shown in [9] that being closed at infinity is a generic property, i.e., the set K is closed at infinity, except on a set of zero measure in the input space of the constraining polynomials g_j $(j = 1, \ldots, m)$.

(2) In fact, the assumption that K is closed at infinity can be replaced by $f^{(\infty)} \geq 0$ on $K^{(\infty)}$, while all convergence results still hold, which can be seen from our proofs. Note that if K is closed at infinity and $f_{\min} > -\infty$, then $f^{(\infty)} \geq 0$ on $K^{(\infty)}$, as shown in Proposition 2.2. In addition, the condition $f^{(\infty)} \geq 0$ on $K^{(\infty)}$ holds as long as SDP relaxations in our proposed hierarchies are feasible at some relaxation order.

3.2. Sparse Positivstellensätze with perturbations. In this subsection, we provide new sparse Positivstellensätze for nonnegative polynomials on unbounded semialgebraic sets, based on the sparse homogenized reformulation (3.1). Unlike the sparse versions of Reznick's Positivstellensätze and Putinar and Vasilescu's Positivstellensätze in [31], our Positivstellensätze do not involve any denominators.

First, we consider the homogeneous case with $K = \mathbb{R}^n$.

THEOREM 3.8. *Let $f \in \mathbb{R}[\mathbf{x}]$ be a form of degree d. Suppose that f admits the csp $(\mathbf{x}(1), \ldots, \mathbf{x}(p))$ satisfying the RIP. Let $\{w_1, \ldots, w_p\}$ be a tuple of auxiliary variables. Then, we have the following:*

(1) *If $f \geq 0$ on $\mathbb{R}^n$, then for any $\epsilon > 0$, there exist $\sigma_\ell \in \Sigma[\mathbf{x}(\ell) \cup \{w_\ell\}]$, $\tau_\ell \in \mathbb{R}[\mathbf{x}(\ell) \cup \{w_\ell\}]$ $(\ell \in [p])$ such that*

$$f + \epsilon \cdot \left(\sum_{i=1}^{n} x_i^d + \sum_{\ell=1}^{p} w_\ell^d \right) = \sum_{\ell=1}^{p} \left(\sigma_\ell + \tau_\ell \left(\|\mathbf{x}(\ell)\|^2 + w_\ell^2 - 1 \right) \right). \tag{3.10}$$

(2) *If f is positive definite, then for any $\epsilon > 0$, there exist $\sigma_\ell \in \Sigma[\mathbf{x}(\ell) \cup \{w_\ell\}]$, $\tau_\ell \in \mathbb{R}[\mathbf{x}(\ell) \cup \{w_\ell\}]$ $(\ell \in [p])$ such that*

$$f + \epsilon \cdot \sum_{\ell=1}^{p} w_\ell^d = \sum_{\ell=1}^{p} \left(\sigma_\ell + \tau_\ell \left(\|\mathbf{x}(\ell)\|^2 + w_\ell^2 - 1 \right) \right). \tag{3.11}$$

Proof. Note that $f \geq 0$ on $\mathbb{R}^n$ is equivalent to $f \geq 0$ on $S := \{(\mathbf{x}, \mathbf{w}) \in \mathbb{R}^{n+p} \mid \|\mathbf{x}(\ell)\|^2 + w_\ell^2 = 1, \ell \in [p]\}$. Since $\mathbf{0} \notin S$, we have $f + \epsilon \cdot (\sum_{i=1}^n x_i^d + \sum_{\ell=1}^p w_\ell^d) > 0$ on S. If f is positive definite, and there exists $(\mathbf{x}, \mathbf{w}) \in \mathbb{R}^{n+p}$ such that $f(\mathbf{x}) + \epsilon \cdot \sum_{\ell=1}^p w_\ell^d = 0$, then we must have $(\mathbf{x}, \mathbf{w}) = \mathbf{0}$, implying that $f + \epsilon \cdot \sum_{\ell=1}^p w_\ell^d > 0$ on S. Under the given assumptions, items (1) and (2) follow from Theorem 2.4. ∎

Note that if the polynomial $f \geq 0$, f must have an even degree. Then, we know that $f \geq 0$ on $\mathbb{R}^n$ is equivalent to $\tilde{f} \geq 0$ on $\mathbb{R}^{n+1}$. Theorem 3.8 can be generalized to nonhomogeneous polynomials directly. We omit the proof for brevity.

THEOREM 3.9. *Let $f \in \mathbb{R}[\mathbf{x}]$ with $\deg(f) = d$. Suppose that f admits the csp $(\mathbf{x}(1), \ldots, \mathbf{x}(p))$ satisfying the RIP. Let $\{w_1, \ldots, w_p\}$ be a tuple of auxiliary variables. Then, we have the following:*

(1) *If $f \geq 0$ on $\mathbb{R}^n$, then for any $\epsilon > 0$, there exist $\sigma_\ell \in \Sigma[\tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}], \tau_\ell \in \mathbb{R}[\tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}]$ $(\ell \in [p])$ such that*

$$\tilde{f} + \epsilon \left(\sum_{i=0}^{n} x_i^d + \sum_{\ell=1}^{p} w_\ell^d \right) = \sum_{\ell=1}^{p} \left(\sigma_\ell + \tau_\ell \left(\|\tilde{\mathbf{x}}(\ell)\|^2 + w_\ell^2 - 1 \right) \right). \tag{3.12}$$

(2) *If $f > 0$ on $\mathbb{R}^n$, and the form $f^{(\infty)}$ is positive definite, then for any $\epsilon > 0$, there exist $\sigma_\ell \in \Sigma[\tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}]$, $\tau_\ell \in \mathbb{R}[\tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}]$ $(\ell \in [p])$ such that*

$$\tilde{f} + \epsilon \sum_{\ell=1}^{p} w_\ell^d = \sum_{\ell=1}^{p} \left(\sigma_i + \tau_\ell \left(\|\tilde{\mathbf{x}}(\ell)\|^2 + w_\ell^2 - 1 \right) \right). \tag{3.13}$$

Let K be defined as in (1.1). Recall that

$$K^{(\infty)} := \left\{ \mathbf{x} \in \mathbb{R}^n \;\middle|\; \begin{array}{ll} g_j^{(\infty)}(\mathbf{x}) \geq 0, & j \in [m], \\ \|\mathbf{x}\|^2 - 1 = 0. & \end{array} \right\}$$

If K is closed at infinity and f is bounded from below on K, then $f^{(\infty)} \geq 0$ on $K^{(\infty)}$, as shown in Proposition 2.2.

DEFINITION 3.10. *A polynomial f is said to be* positive at infinity *on K if $f^{(\infty)} > 0$ on $K^{(\infty)}$.*

If the polynomial f is positive at infinity on K, then f is coercive on K, i.e., the sublevel set $\{x \in K : f(x) \leq \vartheta\}$ is compact for every $\vartheta \in \mathbb{R}$, as shown in [13]. However, the converse does not necessarily hold. For instance, the polynomial $f = x_1^2 + x_2^4$ is not positive at infinity on $K = \mathbb{R}^2$, but it is coercive.

When K is closed at infinity, $f \geq 0$ on K if and only if $\tilde{f} \geq 0$ on $\widetilde{K}$. Consequently, we can derive the following sparse homogenized version of Putinar and Vasilescu's Positivstellensätze from Theorem 2.4.

THEOREM 3.11. *Suppose that K is closed at infinity, and POP (1.1) admits the csp $(\mathbf{x}(1), \ldots, \mathbf{x}(p))$ satisfying the RIP. Let $\{w_1, \ldots, w_p\}$ be a tuple of auxiliary variables. Then, we have the following:*

(1) *If $f \geq 0$ on K, then for any $\epsilon > 0$, there exist $\sigma_\ell \in QM[\{x_0\} \cup \{\tilde{g}_j\}_{j \in J_\ell}, \tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}], \tau_\ell \in \mathbb{R}[\tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}]$ $(\ell \in [p])$ such that*

$$\tilde{f} + \epsilon \left(\sum_{i=0}^{n} x_i^{d_0} + \sum_{\ell=1}^{p} w_\ell^{d_0} \right) = \sum_{\ell=1}^{p} \left(\sigma_\ell + \tau_\ell \left(\|\tilde{\mathbf{x}}(\ell)\|^2 + w_\ell^2 - 1 \right) \right). \tag{3.14}$$

(2) *If $f > 0$ on K, and f is positive at infinity on K, then for any $\epsilon > 0$, there exist $\sigma_\ell \in QM[\{x_0\} \cup \{\tilde{g}_j\}_{j \in J_\ell}, \tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}], \tau_\ell \in \mathbb{R}[\tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}]$ $(\ell \in [p])$ such that*

$$\tilde{f} + \epsilon \sum_{\ell=1}^{p} w_\ell^{d_0} = \sum_{\ell=1}^{p} \left(\sigma_\ell + \tau_\ell \left(\|\tilde{\mathbf{x}}(\ell)\|^2 + w_\ell^2 - 1 \right) \right). \tag{3.15}$$

4. Sparse homogenized moment-SOS hierarchy without perturbations. For the sparse homogenized hierarchy (3.3)–(3.4) to converge, perturbations are typically required, as illustrated in Remark 3.5. This raises the following natural question: Can we design a sparse homogenized moment-SOS hierarchy without perturbations that converges to the optimal value $f_{\min}$ rather than a value close to $f_{\min}$? In the following, we provide such a hierarchy by introducing a new sparse reformulation of (1.1).

Suppose that (1.1) admits the csp $(\mathbf{x}(1),\dots,\mathbf{x}(p))$. For $i\in[n]$, let p_i denote the frequency of the variable x_i occurring in $\mathbf{x}(1),\dots,\mathbf{x}(p)$. Define the set

$$\widetilde{K}_s^2 := \left\{ (\tilde{\mathbf{x}},\mathbf{w})\in\mathbb{R}^{n+p} \;\middle|\; \begin{array}{l} x_0\ge 0,\quad \tilde{g}_j(\tilde{\mathbf{x}})\ge 0,\quad j\in[m],\\ \sum_{x_i\in\mathbf{x}(1)}\frac{1}{p_i}x_i^2+\frac{1}{p}x_0^2=w_1^2,\\ \sum_{x_i\in\mathbf{x}(2)}\frac{1}{p_i}x_i^2+\frac{1}{p}x_0^2+w_1^2=w_2^2,\\ \vdots\\ \sum_{x_i\in\mathbf{x}(p)}\frac{1}{p_i}x_i^2+\frac{1}{p}x_0^2+w_{p-1}^2=1,\\ \mathbf{1}-\tilde{\mathbf{x}}^{\circ 2}\ge\mathbf{0},\quad \mathbf{1}-\mathbf{w}^{\circ 2}\ge\mathbf{0}, \end{array} \right\} \tag{4.1}$$

where $\mathbf{w}:=(w_1,\dots,w_{p-1})$, and the notation $\mathbf{1}-\tilde{\mathbf{x}}^{\circ 2}\ge\mathbf{0}$ means that $1-x_i^2\ge 0$ for $i=0,\dots,n$ (with a similar meaning for $\mathbf{1}-\mathbf{w}^{\circ 2}\ge\mathbf{0}$). Consider the following new sparse reformulation for (1.1) ($d:=\deg(f)$):

$$\begin{cases} \sup & \gamma\\ \text{s.t.} & \tilde{f}(\tilde{\mathbf{x}})-\gamma x_0^d\ge 0 \quad \forall(\tilde{\mathbf{x}},\mathbf{w})\in\widetilde{K}_s^2. \end{cases} \tag{4.2}$$

We set $\mathbf{u}:=(\tilde{\mathbf{x}},\mathbf{w})$ and assume that $(\mathbf{u}(1),\dots,\mathbf{u}(q))$ is the list of maximal cliques of some chordal extension of the csp graph associated with POP (4.2). Let

$$h_1' := \sum_{x_i\in\mathbf{x}(1)}\frac{1}{p_i}x_i^2+\frac{1}{p}x_0^2-w_1^2,\; h_p' := \sum_{x_i\in\mathbf{x}(p)}\frac{1}{p_i}x_i^2+\frac{1}{p}x_0^2+w_{p-1}^2-1,$$
$$h_\ell' := \sum_{x_i\in\mathbf{x}(\ell)}\frac{1}{p_i}x_i^2+\frac{1}{p}x_0^2+w_{\ell-1}^2-w_\ell^2 \;(\ell=2,\dots,p-1).$$

Let $\{J_1,\dots,J_p\}$ be subsets of $[m]$ such that $\cup_{\ell=1}^p J_\ell=[m]$, and for any $\ell\in[p]$ and $j\in J_\ell$, we have $g_j\in\mathbb{R}[\mathbf{x}(\ell)]$. Moreover, let $\{I_1,\dots,I_q\}$ be subsets of $[p]$ such that $\cup_{\ell=1}^q I_\ell=[p]$, and for any $\ell\in[q]$ and $j\in I_\ell$, we have $h_j'\in\mathbb{R}[\mathbf{u}(\ell)]$.

For $k\ge d_{\min}$, the kth order sparse SOS relaxation of (4.2) is

$$\begin{cases} \sup & \gamma\\ \text{s.t.} & \tilde{f}(\tilde{\mathbf{x}})-\gamma x_0^d\in \quad \sum_{\ell=1}^q \mathrm{Ideal}[\{h_j'\}_{j\in I_\ell},\mathbf{u}(\ell)]_{2k}\\ & \qquad\qquad +\sum_{\ell=1}^q \mathrm{QM}[\{x_0\}\cup\{1-u_i^2\}_{u_i\in\mathbf{u}(\ell)}\cup\{\tilde{g}_j\}_{j\in J_\ell},\mathbf{u}(\ell)]_{2k}. \end{cases} \tag{4.3}$$

The dual of (4.3) is the following kth order sparse moment relaxation:

$$\begin{cases} \inf & \langle\tilde{f},\mathbf{y}\rangle\\ \text{s.t.} & \langle x_0^d,\mathbf{y}\rangle=1,\quad M_k[\mathbf{y},\mathbf{u}(\ell)]\succeq 0,\quad \ell\in[q],\\ & M_{k-1}[h_j'\mathbf{y},\mathbf{u}(\ell)]=0,\quad j\in I_\ell,\ell\in[q],\\ & M_{k-1}[x_0\mathbf{y},\mathbf{u}(\ell)]\succeq 0,\quad \ell\in[q],\\ & M_{k-\lceil\deg(g_j)/2\rceil}[\tilde{g}_j\mathbf{y},\mathbf{u}(\ell)]\succeq 0,\quad j\in J_\ell,\ell\in[q],\\ & M_{k-1}[(1-u_i^2)\mathbf{y},\mathbf{u}(\ell)]=0,\quad u_i\in\mathbf{u}(\ell),\ell\in[q]. \end{cases} \tag{4.4}$$

Let $\bar{\rho}_k$ denote the optimum of (4.3).

4.1. Sparse Positivstellensätze without perturbations. In this subsection, we provide new sparse Positivstellensätze for positive polynomials on general (possibly unbounded) semialgebraic sets for which no perturbations are required.

First, we consider the unconstrained case, i.e., $K=\mathbb{R}^n$.

THEOREM 4.1. *Suppose that f admits the csp $(\mathbf{x}(1),\dots,\mathbf{x}(p))$ satisfying the RIP. If $f>0$ on $\mathbb{R}^n$ and $f^{(\infty)}$ is positive definite, then there exist $\sigma_{\ell,i}\in\Sigma[\mathbf{u}(\ell)]$ for each $u_i\in\mathbf{u}(\ell),\ell\in[q]$, and $\tau_{\ell,j}\in\mathbb{R}[\mathbf{u}(\ell)]$ for each $j\in I_\ell,\ell\in[q]$ such that*

$$\tilde{f}=\sum_{\ell=1}^{q}\left(\sum_{u_i\in\mathbf{u}(\ell)}\sigma_{\ell,i}(1-u_i^2)+\sum_{j\in I_\ell}\tau_{\ell,j}h'_j\right). \tag{4.5}$$

Proof. Let

$$S:=\left\{(\tilde{\mathbf{x}},\mathbf{w})\in\mathbb{R}^{n+p}\ \middle|\ \begin{array}{l}\sum_{x_i\in\mathbf{x}(1)}\frac{1}{p_i}x_i^2+\frac{1}{p}x_0^2=w_1^2,\\ \sum_{x_i\in\mathbf{x}(2)}\frac{1}{p_i}x_i^2+\frac{1}{p}x_0^2+w_1^2=w_2^2,\\ \qquad\vdots\\ \sum_{x_i\in\mathbf{x}(p)}\frac{1}{p_i}x_i^2+\frac{1}{p}x_0^2+w_{p-1}^2=1,\\ \mathbf{1}-\tilde{\mathbf{x}}^{\circ 2}\geq\mathbf{0},\quad \mathbf{1}-\mathbf{w}^{\circ 2}\geq\mathbf{0}.\end{array}\right\}$$

We show that $\tilde{f}>0$ on S. Take any $(\tilde{\mathbf{x}},\mathbf{w})\in S$. If $x_0=0$, then we must have $\mathbf{x}\neq\mathbf{0}$. Suppose that otherwise, $\mathbf{x}=\mathbf{0}$. Then $w_1^2=w_2^2=\cdots=w_{p-1}^2=0$, which contradicts the constraint $\sum_{x_i\in\mathbf{x}(p)}\frac{1}{p_i}x_i^2+\frac{1}{p}x_0^2+w_{p-1}^2=1$ in the definition of S. Since the form $f^{(\infty)}$ is positive definite, we have

$$\tilde{f}(0,\mathbf{x})=f^{(\infty)}(\mathbf{x})>0.$$

If $x_0\neq 0$, then we have $\tilde{f}(\tilde{\mathbf{x}})=x_0^d f(\mathbf{x}/x_0)>0$. Moreover, for each $\ell\in[q]$, the quadratic module $\sum_{\ell=1}^{q}\mathrm{QM}[\{1-u_i^2\}_{u_i\in\mathbf{u}(\ell)},\mathbf{u}(\ell)]$ is Archimedean in $\mathbb{R}[\mathbf{u}(\ell)]$. Thus, the conclusion follows from Theorem 2.4. ∎

The following theorem addresses the constrained case. Since the proof is quite similar to that of Theorem 4.1, we omit it.

THEOREM 4.2. *Suppose that K is closed at infinity, and POP* (1.1) *admits the csp $(\mathbf{x}(1),\dots,\mathbf{x}(p))$ satisfying the RIP. If $f>0$ on K and f is positive at infinity on K, then there exist $\sigma_\ell\in QM[\{x_0\}\cup\{1-u_i^2\}_{u_i\in\mathbf{u}(\ell)}\cup\{\tilde{g}_j\}_{j\in J_\ell},\mathbf{u}(\ell)]$ for each $\ell\in[q]$, and $\tau_{\ell,j}\in\mathbb{R}[\mathbf{u}(\ell)]$ for each $j\in I_\ell,\ell\in[q]$ such that*

$$\tilde{f}=\sum_{\ell=1}^{q}\left(\sigma_\ell+\sum_{j\in I_\ell}\tau_{\ell,j}h'_j\right). \tag{4.6}$$

4.2. Convergence analysis. Next, we show that (1.1) and (4.2) have the same optimal values.

THEOREM 4.3. *Suppose that K is closed at infinity and* (1.1) *has a finite minimum. Then, the optimal value of* (4.2) *is equal to $f_{\min}$.*

Proof. Since K is closed at infinity and $f_{\min}>-\infty$, we know that $f^{(\infty)}\geq 0$ on $K^{(\infty)}$ by Proposition 2.2. Take any $(\tilde{\mathbf{x}},\mathbf{w})\in\widetilde{K}_s^2$. If $x_0=0$, then we have $g_j^{(\infty)}(\mathbf{x})=\tilde{g}_j(\tilde{\mathbf{x}})\geq 0$ for $j\in[m]$, and thus $\tilde{f}(\tilde{\mathbf{x}})-f_{\min}x_0^d=f^{(\infty)}(\mathbf{x})\geq 0$. If $x_0\neq 0$, we have $\mathbf{x}/x_0\in K$, and

$$\tilde{f}(\tilde{\mathbf{x}})-f_{\min}x_0^d=x_0^d(f(\mathbf{x}/x_0)-f_{\min})\geq 0.$$

It follows that $f_{\min}$ is not greater than the optimal value of (4.2). For the converse, let $\{\mathbf{x}^{(k)}\}\subseteq\mathbb{R}^n$ be a sequence such that $f(\mathbf{x}^{(k)})\to f_{\min}>-\infty$. Let

$$\tilde{\mathbf{x}}^{(k)} = \frac{(1, \mathbf{x}^{(k)})}{\sqrt{1 + \|\mathbf{x}^{(k)}\|^2}},\ w_\ell^{(k)} = \sqrt{\sum_{s=1}^{\ell} \left(\sum_{x_i \in \mathbf{x}(s)} \frac{1}{p_i} \left(x_i^{(k)}\right)^2 + \frac{1}{p} \left(x_0^{(k)}\right)^2 \right)}\ (\ell \in [p-1]).$$

By computations, one can verify that $(\tilde{\mathbf{x}}^{(k)}, \mathbf{w}^{(k)}) \in \widetilde{K}_s^2$. If γ is feasible for (4.2), then $\gamma \le \tilde{f}(\tilde{\mathbf{x}}^{(k)})/(x_0^{(k)})^d = f(\mathbf{x}^{(k)})$. Letting k go to infinity, we obtain

$$\gamma \le \lim_{k \to \infty} f(\mathbf{x}^{(k)}) = f_{\min},$$

which completes the proof. ∎

We now establish the asymptotic convergence of the sparse homogenized moment-SOS hierarchy (4.3)–(4.4).

THEOREM 4.4. *Suppose that* (1.1) *admits the csp* $(\mathbf{x}(1), \ldots, \mathbf{x}(p))$*,* K *is closed at infinity, and* $f^{(\infty)}$ *is positive at infinity on* K*. Then,* $\bar{\rho}_k \to f_{\min}$ *as* $k \to \infty$*.*

Proof. Since $f^{(\infty)}$ is positive at infinity on K, (1.1) has a finite minimum, i.e., $f_{\min} > -\infty$ [13]. For any $\gamma < f_{\min}$, we know that $f - \gamma > 0$ on K. Note that for each $\ell \in [q]$, the quadratic module

$$\mathrm{QM}[\{x_0\} \cup \{1 - u_i^2\}_{u_i \in \mathbf{u}(\ell)} \cup \{\tilde{g}_j\}_{j \in J_\ell}, \mathbf{u}(\ell)]$$

is Archimedean in $\mathbb{R}[\mathbf{u}(\ell)]$. Since the optimal value of (4.2) is $f_{\min}$ (see Theorem 4.3), it follows from Theorem 4.2 that for any $\gamma < f_{\min}$,

$$\tilde{f}(\tilde{\mathbf{x}}) - \gamma x_0^d \in \sum_{\ell=1}^{q} \left(\mathrm{Ideal}[\{h_j'\}_{j \in I_\ell}, \mathbf{u}(\ell)] + \mathrm{QM}[\{x_0\} \cup \{1 - u_i^2\}_{u_i \in \mathbf{u}(\ell)} \cup \{\tilde{g}_j\}_{j \in J_\ell}, \mathbf{u}(\ell)] \right).$$

Then, we obtain that $\bar{\rho}_k \to f_{\min}$ as $k \to \infty$. ∎

4.3. An alternative sparse homogenized moment-SOS hierarchy without perturbations. There is also an alternative way to remove perturbations. Suppose that $(\mathbf{x}(1), \ldots, \mathbf{x}(p))$ is the csp of (1.1), and let p_i denote the frequency of the variable x_i occurring in $\mathbf{x}(1), \ldots, \mathbf{x}(p)$ for $i \in [n]$. Define the set

$$\widetilde{K}_s^3 := \left\{ (\tilde{\mathbf{x}}, \mathbf{w}) \in \mathbb{R}^{n+1+p} \;\middle|\; \begin{array}{l} x_0 \ge 0, \quad \tilde{g}_j(\tilde{\mathbf{x}}) \ge 0, \quad j \in [m], \\ \sum_{x_i \in \mathbf{x}(\ell)} \frac{1}{p_i} x_i^2 + \frac{1}{p} x_0^2 + w_\ell^2 = 1, \quad \ell \in [p], \\ \|\mathbf{w}\|^2 = p - 1, \quad \mathbf{1} - \tilde{\mathbf{x}}^{\circ 2} \ge \mathbf{0}, \quad \mathbf{1} - \mathbf{w}^{\circ 2} \ge \mathbf{0}, \end{array} \right\} \tag{4.7}$$

where $\mathbf{w} := (w_1, \ldots, w_p)$. Consider the following sparse homogenized reformulation for (1.1):

$$\begin{cases} \sup & \gamma \\ \text{s.t.} & \tilde{f}(\tilde{\mathbf{x}}) - \gamma x_0^d \ge 0 \quad \forall (\tilde{\mathbf{x}}, \mathbf{w}) \in \widetilde{K}_s^3. \end{cases} \tag{4.8}$$

In the following, we show that (1.1) and (4.8) have the same optimal values.

THEOREM 4.5. *Suppose that* K *is closed at infinity and* (1.1) *has a finite minimum. Then, the optimal value of* (4.8) *is equal to* $f_{\min}$*.*

Proof. Since K is closed at infinity and $f_{\min} > -\infty$, we know that $f^{(\infty)} \ge 0$ on $K^{(\infty)}$ by Proposition 2.2. Take any $(\tilde{\mathbf{x}}, \mathbf{w}) \in \widetilde{K}_s^3$. If $x_0 = 0$, then we have $g_j^{(\infty)}(\mathbf{x}) = \tilde{g}_j(\tilde{\mathbf{x}}) \ge 0$ for $j \in [m]$, and thus $\tilde{f}(\tilde{\mathbf{x}}) - f_{\min} x_0^d = f^{(\infty)}(\mathbf{x}) \ge 0$. If $x_0 \ne 0$, we have $\mathbf{x}/x_0 \in K$, and

$$\tilde{f}(\tilde{\mathbf{x}}) - f_{\min} x_0^d = x_0^d (f(\mathbf{x}/x_0) - f_{\min}) \geq 0.$$

Thus, we know that the optimal value of (4.8) is at least $f_{\min}$. For the converse, let $\{\mathbf{x}^{(k)}\} \subseteq \mathbb{R}^n$ be a sequence such that $f(\mathbf{x}^{(k)}) \to f_{\min} > -\infty$. Let

$$\tilde{\mathbf{x}}^{(k)} = \frac{(1, \mathbf{x}^{(k)})}{\sqrt{1 + \|\mathbf{x}^{(k)}\|^2}},\ w_\ell^{(k)} = \sqrt{1 - \sum_{x_i \in \mathbf{x}(\ell)} \frac{1}{p_i} \left(\tilde{x}_i^{(k)}\right)^2 - \frac{1}{p}\left(\tilde{x}_0^{(k)}\right)^2}\ (\ell \in [p]).$$

One can verify that $(\tilde{\mathbf{x}}^{(k)}, \mathbf{w}^{(k)}) \in \widetilde{K}_s^3$. If γ is feasible for (4.8), then

$$\gamma \leq \lim_{k \to \infty} \tilde{f}(\tilde{\mathbf{x}}^{(k)})/(x_0^{(k)})^d = f_{\min},$$

which completes the proof. ∎

Now, we set $\mathbf{u} := (\tilde{\mathbf{x}}, \mathbf{w})$ and assume that $(\mathbf{u}(1), \ldots, \mathbf{u}(q))$ is the list of maximal cliques of some chordal extension of the csp graph associated with POP (4.8). Let

$$h'_\ell := \sum_{x_i \in \mathbf{x}(\ell)} \frac{1}{p_i} x_i^2 + \frac{1}{p} x_0^2 + w_\ell^2 - 1\ (\ell \in [p]),\ h'_{p+1} := \|\mathbf{w}\|^2 - p + 1.$$

Let $\{J_1, \ldots, J_q\}$ and $\{I_1, \ldots, I_q\}$ be defined similarly as before.

For $k \geq d_{\min}$, the kth order sparse SOS relaxation of (4.8) is

$$\text{(4.9)} \quad \begin{cases} \sup & \gamma \\ \text{s.t.} & \tilde{f}(\tilde{\mathbf{x}}) - \gamma x_0^d \in \sum_{\ell=1}^q \mathrm{Ideal}[\{h'_j\}_{j \in I_\ell}, \mathbf{u}(\ell)]_{2k} \\ & \qquad + \sum_{\ell=1}^q \mathrm{QM}[\{x_0\} \cup \{1 - u_i^2\}_{u_i \in \mathbf{u}(\ell)} \cup \{\tilde{g}_j\}_{j \in J_\ell}, \mathbf{u}(\ell)]_{2k}. \end{cases}$$

The dual of (4.9) is the following kth order sparse moment relaxation:

$$\text{(4.10)} \quad \begin{cases} \inf & \langle \tilde{f}, \mathbf{y} \rangle \\ \text{s.t.} & \langle x_0^d, \mathbf{y} \rangle = 1, \quad M_k[\mathbf{y}, \mathbf{u}(\ell)] \succeq 0, \quad \ell \in [q], \\ & M_{k-1}[h'_j \mathbf{y}, \mathbf{u}(\ell)] = 0, \quad j \in I_\ell, \ell \in [q], \\ & M_{k-1}[x_0 \mathbf{y}, \mathbf{u}(\ell)] \succeq 0, \quad \ell \in [q], \\ & M_{k - \lceil \deg(g_j)/2 \rceil}[\tilde{g}_j \mathbf{y}, \mathbf{u}(\ell)] \succeq 0, \quad j \in J_\ell, \ell \in [q], \\ & M_{k-1}[(1 - u_i^2)\mathbf{y}, \mathbf{u}(\ell)] = 0, \quad u_i \in \mathbf{u}(\ell), \ell \in [q]. \end{cases}$$

The following theorem establishes the asymptotic convergence of the sparse homogenized moment-SOS hierarchy (4.9)–(4.10). Since the proof is quite similar to that of Theorem 4.4, we omit it for brevity.

THEOREM 4.6. *Suppose that* (1.1) *admits the csp* $(\mathbf{x}(1), \ldots, \mathbf{x}(p))$*,* K *is closed at infinity, and* $f^{(\infty)}$ *is positive at infinity on* K*. Then the optimal value of* (4.9) *converges to* $f_{\min}$ *as* $k \to \infty$*.*

Remark 4.7. In the case of the dense moment-SOS hierarchy, a convenient criterion for detecting global optimality is flat extension/truncation (see [6, 25, 36]), and a related procedure for extracting minimizers is given in [11]. There is a similar procedure adapted to the sparse setting in [22], which can be also used to extract minimizers for the sparse homogenized hierarchies.

So far, we have discussed how to exploit correlative sparsity for homogenized polynomial optimization but have not yet considered term sparsity. In fact, correlative sparsity and term sparsity can be exploited simultaneously to achieve further reductions on the size of SDP relaxations arising from moment-SOS hierarchies. We refer the reader to [28, 51] for more details.

4.4. Comparisons of different hierarchies. In this subsection, we make some comparisons and present discussions about the sparse homogenized hierarchies mentioned above.

The first hierarchy, (3.3)–(3.4), is based on the homogenized reformulation (3.2). It introduces a perturbation parameter ϵ and converges to a value close to $f_{\min}$ (see Theorem 3.1). When the csp $(\mathbf{x}(1),\dots,\mathbf{x}(p))$ satisfies the RIP, POP (3.2) admits the csp $(\tilde{\mathbf{x}}(1)\cup\{w_1\},\dots,\tilde{\mathbf{x}}(p)\cup\{w_p\})$, which also satisfies the RIP. Therefore, the maximal clique size of the first hierarchy always increases by two compared to the original csp.

The second hierarchy, (4.3)–(4.4), and the third hierarchy, (4.9)–(4.10), are based on the sparse reformulations (4.2) and (4.8), respectively. They have no perturbation parameters, and both converge to the optimal value $f_{\min}$, under some general assumptions. Note that both reformulations generally break up the original csp, and the maximal clique sizes depend on the related chordal extensions. Finding a minimum chordal extension of a graph is an NP-complete problem in general [2]. However, it is worth noting that when the number of initial cliques p is large, the size of the second hierarchy is usually smaller than that of the third hierarchy, making it more preferable from the computational aspect. This is because (4.8) involves a variable clique of size at least p (due to the spherical constraint $\|\mathbf{w}\|^2=p-1$), and a chordal extension must be performed to recover the RIP. On the other hand, the maximal clique size of the second hierarchy increases by three if (1.1) admits a "chain-like" csp, which frequently occurs in practice. To be precise, we say that a csp $(\mathbf{x}(1),\dots,\mathbf{x}(p))$ is *chain-like* if for every $\ell\in[p-1]$, it holds that $\mathbf{x}(\ell+1)\cap\bigcup_{j=1}^{\ell}\mathbf{x}(j)\subseteq\mathbf{x}(\ell)$. A chain-like csp clearly satisfies the RIP. In the following, we show that when (1.1) admits a chain-like csp, the reformulation (4.2) "inherits" that csp.

LEMMA 4.8. *Suppose that* (1.1) *admits a chain-like csp* $(\mathbf{x}(1),\dots,\mathbf{x}(p))$. *Then,* (4.2) *also admits a chain-like csp* $(\tilde{\mathbf{x}}(1)\cup\{w_1\},\tilde{\mathbf{x}}(2)\cup\{w_1,w_2\},\dots,\tilde{\mathbf{x}}(p)\cup\{w_{p-1}\})$.

Proof. It is clear that (4.2) admits the csp $(\tilde{\mathbf{x}}(1)\cup\{w_1\},\tilde{\mathbf{x}}(2)\cup\{w_1,w_2\},\dots,\tilde{\mathbf{x}}(p)\cup\{w_{p-1}\})$. For every $\ell\in[p-2]$, it holds that

$$\begin{aligned}&(\tilde{\mathbf{x}}(\ell+1)\cup\{w_\ell,w_{\ell+1}\})\cap\left(\tilde{\mathbf{x}}(1)\cup\{w_1\}\cup\bigcup_{j=1}^{\ell}\tilde{\mathbf{x}}(j)\cup\{w_{j-1},w_j\}\right)\\&\quad=\left(\tilde{\mathbf{x}}(\ell+1)\cap\bigcup_{j=1}^{\ell}\tilde{\mathbf{x}}(j)\right)\cup\{w_\ell\}\subseteq\tilde{\mathbf{x}}(\ell)\cup\{w_\ell\}.\end{aligned}$$

Moreover, we have

$$\begin{aligned}&(\tilde{\mathbf{x}}(p)\cup\{w_{p-1}\})\cap\left(\tilde{\mathbf{x}}(1)\cup\{w_1\}\cup\bigcup_{j=2}^{p-1}\tilde{\mathbf{x}}(j)\cup\{w_{j-1},w_j\}\right)\\&\quad=\left(\tilde{\mathbf{x}}(p)\cap\bigcup_{j=1}^{p-1}\tilde{\mathbf{x}}(j)\right)\cup\{w_{p-1}\}\subseteq\tilde{\mathbf{x}}(p-1)\cup\{w_{p-1}\}.\end{aligned}$$

Hence, the conclusion follows. ∎

By Lemma 4.8, if (1.1) admits a chain-like csp, then (4.2) (and thus the second hierarchy) also admits a chain-like csp, with the maximal clique size increased by three.

5. Numerical examples. In this section, we present numerical results on solving POPs with three sparse homogenized moment-SOS hierarchies proposed in this

TABLE 1
Notation.

n	Number of variables
k	relaxation order
opt	optimum
time	running time in seconds
SSOS	the sparse SOS relaxation (2.7)
HSOS	the dense homogenized SOS relaxation
HSSOS1	the sparse homogenized SOS relaxation (3.3)
HSSOS2	the sparse homogenized SOS relaxation (4.3)
HSSOS3	the alternative sparse homogenized SOS relaxation (4.9)
bold font	global optimality being certified
*	indicating unknown termination status
**	infeasible SDP
-	returning an out of memory error

paper. All numerical experiments are performed on a desktop computer with Intel Core i9-10900 CPU@2.80 GHz and 64G RAM. To model the homogenized hierarchies, we use the Julia package `TSSOS`[3] [27], which relies on `Mosek` 10.0 [1] as the SDP backend with default settings. Unless specified otherwise, we set $\epsilon = 10^{-4}$ for the relaxations (3.3)–(3.4). We do not implement and compare with the approach proposed in [31], as it is limited to problems of modest size. Notation is listed in Table 1. All feasible sets are unbounded and closed at infinity in our numerical examples.

5.1. Unconstrained polynomial optimization.

Example 5.1. Let $\mathbf{x}(1) = \{x_1, x_2, x_3\}$ and $\mathbf{x}(2) = \{x_1, x_2, x_4\}$. Consider POP (1.1) with the csp $(\mathbf{x}(1), \mathbf{x}(2))$, where

$$f = f_1 + f_2, \quad f_1 = x_3^2(x_1^2 + x_1^4x_2^2 + x_3^4 - 3x_1^2x_2^2) + x_2^8, \quad f_2 = x_1^2x_2^2x_4^2.$$

The polynomial f_1 is the dehomogenized Delzell's polynomial, which is nonnegative but not an SOS [42]. This example is a variation of Example 1 in [31]. As shown in [31], f is nonnegative, and $f \notin \Sigma[\mathbf{x}(1)] + \Sigma[\mathbf{x}(2)]$. By solving (3.3) with $\epsilon = 0$ at the order $k = 5$, we obtain $\rho_5 \approx -1.6 \times 10^{-7}$, which confirms $\rho_5 = f_{\min} = 0$ (up to numerical round-off errors).

Example 5.2. Let

$$\mathbf{x}(1) = \{x_1, x_2, x_3, x_4\}, \quad \mathbf{x}(2) = \{x_4, x_5, x_6, x_7\}, \quad \mathbf{x}(3) = \{x_7, x_8, x_9, x_{10}\}.$$

Consider POP (1.1) with the csp $(\mathbf{x}(1), \mathbf{x}(2), \mathbf{x}(3))$, where $f = f_1 + f_2 + f_3$ $(x_0 := 1)$ with

$$\begin{aligned}
f_1 &= \sum_{i=1}^{4} x_i^4 + \sum_{i=0}^{4} \prod_{j \neq i} (x_i - x_j)\,, \\
f_2 &= \sum_{i=4}^{7} x_i^4 + \sum_{i=0,4,\ldots,7} \prod_{j \neq i} (x_i - x_j)\,, \\
f_3 &= \sum_{i=7}^{10} x_i^4 + \sum_{i=0,7,\ldots,10} \prod_{j \neq i} (x_i - x_j)\,.
\end{aligned}$$

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

TABLE 2
Results of Example 5.2.

k	SSOS		HSOS		HSSOS1		HSSOS2		HSSOS3	
	opt	time	opt	time	opt	time	opt	time	opt	time
2	0.5497	0.02	0.5497	0.05	0.5497	0.03	0.5497	0.03	0.5497	0.02
3	0.5497	0.21	**0.6927**	13.3	**0.6927**	0.37	**0.6927**	0.20	**0.6927**	0.15
4	0.5864*	0.73	**0.6927**	683	**0.6927**	3.27	**0.6927**	1.77	**0.6927**	1.38

TABLE 3
Results of Example 5.3.

k	SSOS		HSOS		HSSOS1		HSSOS2		HSSOS3	
	opt	time	opt	time	opt	time	opt	time	opt	time
2	1.1804	0.01	1.1804	0.54	1.1804	0.04	1.1804	0.11	1.1804	0.09
3	1.1804	0.07	-	-	1.1895	0.34	**1.1900**	0.96	1.1969*	1.18
4	1.1809	0.40	-	-	**1.1900**	1.48	1.1901	5.94	1.4871*	17.4

Here, we set $\epsilon = 0$ for HSSOS1. The numerical results for this problem are presented in Table 2. From the table, we make the following observations: (1) Without homogenization, the sparse hierarchy converges slowly; (2) by exploiting sparsity, we gain a significant speed-up, especially when the relaxation orders are high and (3) all three sparse homogenized moment-SOS hierarchies achieve the optimum $f_{\min} \approx 0.6927$ at the third order relaxation.

Example 5.3. Let

$$\begin{aligned} \mathbf{x}(1) &= \{x_1, x_2, x_3, x_4, x_5\}, & \mathbf{x}(2) &= \{x_1, x_2, x_6, x_7, x_8\}, \\ \mathbf{x}(3) &= \{x_1, x_2, x_9, x_{10}, x_{11}\}, & \mathbf{x}(4) &= \{x_1, x_2, x_{12}, x_{13}, x_{14}\}, \\ \mathbf{x}(5) &= \{x_1, x_2, x_{15}, x_{16}, x_{17}\}, & \mathbf{x}(6) &= \{x_1, x_2, x_{18}, x_{19}, x_{20}\}. \end{aligned}$$

Consider POP (1.1) with csp $(\mathbf{x}(1), \mathbf{x}(2), \mathbf{x}(3), \mathbf{x}(4), \mathbf{x}(5), \mathbf{x}(6))$, where $f = \sum_{i=1}^{6} f_i$, and for $i = 1, \ldots, 6$,

$$\begin{aligned} f_i = {} & x_1^2 (x_1 - 1)^2 + x_2^2 (x_2 - 1)^2 + x_{3i}^2 (x_{3i} - 1)^2 + 2x_1 x_2 x_{3i} (x_1 + x_2 + x_{3i} - 2) \\ & + \frac{1}{4}((x_1 - 1)^2 + (x_2 - 1)^2 + (x_{3i} - 1)^2 + (x_{3i+1} - 1)^2) + (x_{3i+1} x_{3i+2} - 1)^2. \end{aligned}$$

Here, we set $\epsilon = 0$ for HSSOS1. The numerical results for this problem are presented in Table 3. From the table, we make the following observations: (1) Without homogenization, the sparse hierarchy converges slowly; (2) by exploiting sparsity, we gain a significant speed-up for higher order relaxations; and (3) HSSOS1 achieves the optimum at $k = 4$, and HSSOS2 achieves the optimum at $k = 3$. However, HSSOS3 gives incorrect results due to numerical issues when $k \geq 3$.

5.2. Constrained polynomial optimization.

Example 5.4. Let

$$\mathbf{x}(1) = \{x_1, x_2\}, \quad \mathbf{x}(2) = \{x_2, x_3\}, \quad \mathbf{x}(3) = \{x_2, x_4, x_5\}.$$

Consider POP (1.1) with the csp $(\mathbf{x}(1), \mathbf{x}(2), \mathbf{x}(3))$ as follows:

$$\begin{cases} \inf & x_1^2 + 3x_2^2 - 2x_2x_3^2 + x_3^4 - x_2(x_4^2 + x_5^2) \\ \text{s.t.} & x_1^2 - 2x_1x_2 - 1 \geq 0,\, x_1^2 + 2x_1x_2 - 1 \geq 0, \\ & x_2 - 1 \geq 0,\, x_2 - x_4^2 - x_5^2 \geq 0. \end{cases} \tag{5.1}$$

For this problem, the optimal value is $4+2\sqrt{2} \approx 6.8284$. The numerical results of this problem are presented in Table 4. From the table, we draw the following conclusions: (1) Without homogenization, the sparse hierarchy converges slowly; and (2) HSOS, HSSOS2, and HSSOS3 all achieve the optimum at $k = 4$, while HSSOS1 converges more slowly.

Example 5.5. Let

$$\mathbf{x}(1) = \{x_1, x_2, x_3, x_7\}, \quad \mathbf{x}(2) = \{x_4, x_5, x_6, x_7\}.$$

Consider POP (1.1) with the csp $(\mathbf{x}(1), \mathbf{x}(2))$ as follows:

$$\begin{cases} \inf & f_1 + f_2 \\ \text{s.t.} & x_1 - x_2x_3 \geq 0, -x_2 + x_3^2 \geq 0, 1 - x_4^2 - x_5^2 - x_6^2 \geq 0, \end{cases}$$

where

$$f_1 = x_1^4x_2^2 + x_2^4x_3^2 + x_3^4x_1^2 - 3(x_1x_2x_3)^2 + x_2^2 + x_7^2(x_1^2 + x_2^2 + x_3^2),$$
$$f_2 = x_4^2x_5^2(10 - x_6^2) + x_7^2(x_4^2 + 2x_5^2 + 3x_6^2).$$

For this problem, the optimal value is 0 [41]. The numerical results are presented in Table 5. From the table, we make the following observations: (1) Without homogenization, the sparse hierarchy either yields infeasible SDPs or gives very loose bounds; (2) by exploiting sparsity, we gain some speed-up; and (3) HSOS achieves the optimum at $k = 4$, while HSSOS2 and HSSOS3 achieve the optimum at $k = 5$. However, HSSOS1 converges to a value close to $f_{\min}$ at $k = 4$.

Example 5.6. For an integer $p \geq 2$, let

$$\mathbf{x}(i) = \{x_{8i-7}, x_{8i-6}, \ldots, x_{8i+2}\}, \quad i \in [p].$$

TABLE 4
Results of Example 5.4.

k	SSOS		HSOS		HSSOS1		HSSOS2		HSSOS3	
	opt	time	opt	time	opt	time	opt	time	opt	time
2	2.0000	0.03	1.0499*	0.02	1.1712	0.02	1.0555*	0.02	1.1053*	0.02
3	2.0236*	0.05	5.3774*	0.09	3.1243	0.08	5.4444*	0.12	5.4477*	0.13
4	2.1899*	0.06	**6.8284**	0.39	4.5613	0.36	**6.8284**	0.41	**6.8284**	0.33

TABLE 5
Results of Example 5.5.

k	SSOS		HSOS		HSSOS1		HSSOS2		HSSOS3	
	opt	time	opt	time	opt	time	opt	time	opt	time
3	**	0.04	-4532	0.28	−1756*	0.16	−1106*	0.20	−1065*	0.24
4	**	0.19	**-1.6e-8**	2.71	0.0001	0.82	-0.0002	1.77	-0.0002	1.37
5	-4.0e5	0.89	**-9.8e-9**	33.4	0.0001	5.33	**1.4e-7**	6.38	**1.1e-7**	6.98

TABLE 6
Results of Example 5.6.

p	SSOS		HSOS		HSSOS1		HSSOS2		HSSOS3	
	opt	time	opt	time	opt	time	opt	time	opt	time
2	-11053^*	2.72	**6.1488**	156	6.0984	14.6	**6.1488**	12.8	**6.1488**	15.5
3	-18999^*	4.14	**9.2232**	2763	9.1475	20.1	9.2228	31.3	9.2227	20.6
4	-26984^*	5.14	-	-	12.196	29.4	12.295	55.1	12.294	30.4
5	-31198^*	6.54	-	-	15.246	39.0	15.364	69.4	15.365	39.4
10	-80847^*	12.8	-	-	30.491	101	30.504	170	30.543	122

For $i \in [p]$, let

$$f_i = \left(\sum_{j=1}^{10}\left(x_j^{(i)}\right)^2+1\right)^2 - 4\left(\left(x_1^{(i)}x_2^{(i)}\right)^2+\cdots+\left(x_4^{(i)}x_5^{(i)}\right)^2+\left(x_5^{(i)}x_1^{(i)}\right)^2\right)$$
$$-4\left(\left(x_6^{(i)}x_7^{(i)}\right)^2+\cdots+\left(x_9^{(i)}x_{10}^{(i)}\right)^2+\left(x_{10}^{(i)}x_6^{(i)}\right)^2\right)+\frac{1}{5}\sum_{j=1}^{10}\left(x_j^{(i)}\right)^4.$$

Consider POP (1.1) with the csp $(\mathbf{x}(1),\mathbf{x}(2),\ldots,\mathbf{x}(p))$ as follows:

$$\begin{cases}\inf & \sum_{i=1}^p f_i\\ \text{s.t.} & \|\mathbf{x}(i)^{\circ 2}\|^2-1\geq 0, \quad i=1,\ldots,p.\end{cases}$$

We solve the fourth order relaxations for different p. The numerical results for this problem are presented in Table 6. From the table, we can draw the following conclusions: (1) Without homogenization, the sparse relaxation yields very loose bounds; (2) for $p=2,3$, HSOS achieves the optimum, whereas for $p\geq 4$, HSOS runs out of memory; and (3) by exploiting sparsity, we improve the scalability of the homogenization approach and still obtain good bounds.

Example 5.7. We generate random instances of quadratic optimization on unbounded sets as follows. For $n\in\{20,40,100,200,400,800,2000\}$, let $p=\lceil\frac{n}{3}\rceil$. Let

$$\mathbf{x}(1)=\{x_1,x_2,x_3\},\ \mathbf{x}(i)=\{x_{3(i-1)},\ldots,x_{3i}\}\ (i=2,\ldots,p-1),\ \mathbf{x}(p)=\{x_{3(p-1)},\ldots,x_n\}.$$

Let $A_1\in\mathbb{R}^{3\times 3}$, $b_1\in\mathbb{R}^3$, $A_i\in\mathbb{R}^{4\times 4}, b_i\in\mathbb{R}^4$ $(i=2,\ldots,p-1)$, $b_p\in\mathbb{R}^{n-3p+4}$, $A_p\in\mathbb{R}^{(n-3p+4)\times(n-3p+4)}$ be randomly generated with entries being uniformly taken from $[0,1]$. For $i=1,\ldots,p$, let $f_i=\|A_i\mathbf{x}(i)^{\circ 2}\|^2+b_i^\intercal\mathbf{x}(i)^{\circ 2}$. Consider the POP with the csp $(\mathbf{x}(1),\mathbf{x}(2),\ldots,\mathbf{x}(p))$ as follows:

$$\begin{cases}\inf & \sum_{i=1}^p f_i\\ \text{s.t.} & \|\mathbf{x}(i)^{\circ 2}\|^2-1\geq 0, \quad i=1,\ldots,p.\end{cases}$$

We solve the fourth order relaxations for different n. The numerical results for this problem are presented in Table 7. From the table, we can draw the following conclusions: (1) Without homogenization, the sparse relaxation yields much looser bounds; (2) by exploiting sparsity, we gain some sizable speed-up; (3) HSOS and HSOSS3 do not scale well with the problem size (HSOS runs out of memory when $n\geq 40$, and HSOSS3 runs out of memory when $n\geq 100$); and (4) HSSOS1 and HSOSS2 scale well with the problem size (up to $n=2000$).

TABLE 7
Results of Example 5.7.

n	SSOS		HSOS		HSSOS1		HSSOS2		HSSOS3	
	opt	time	opt	time	opt	time	opt	time	opt	time
20	5.5065	0.33	**8.8328**	240	8.4216	0.93	**8.8328**	1.52	**8.8328**	1.21
40	11.813	0.35	-	-	17.481	1.51	17.856	3.59	18.059	29.7
100	27.976	1.29	-	-	42.273	6.94	41.336	16.3	-	-
200	60.178	2.62	-	-	87.726	19.7	82.240	52.9	-	-
400	111.35	6.52	-	-	164.06	55.4	146.66*	190	-	-
800	228.42	18.7	-	-	337.01	229	296.70*	702	-	-
2000	577.88	88.0	-	-	854.14*	1736	768.31*	6424	-	-

Before closing this section, we summarize the conclusions that can be drawn from the numerical experiments. (1) The homogenization approaches significantly improve the numerical performance of the moment-SOS hierarchy when solving sparse POPs with unbounded sets. (2) Exploiting sparsity greatly improves the scalability of the homogenized moment-SOS hierarchy. (3) Among the three sparse homogenized hierarchies, HSSOS1 is usually the most efficient, though it only provides approximate optima. (4) HSOSS2 is more scalable than HSOSS3 and is preferable for handling large-scale problems, especially when the POP admits a chain-like csp.

6. Applications to trajectory optimization.

Trajectory optimization plays an essential role in the fields of robotics and control [4]. In [44], it is demonstrated that applying the sparse moment-SOS hierarchy to specific trajectory optimization problems with compact feasible sets tends to yield tight solutions. However, assuming predefined bounds over all physical quantities can be unrealistic. For instance, it is particularly hard to bound the generalized momentum of highly nonlinear systems or the contact forces in contact-rich scenarios [3]. In such contexts, the ability to relax the compactness assumption while still achieving tight solutions is desirable. In this section, we explore the following two trajectory optimization problems with unbounded feasible sets: (1) block-moving with minimum work using direct collocation, and (2) optimal control of Van der Pol oscillator with direct multiple shooting. We compare the performance of SSOS, HSSOS1, and HSSOS2 (noting that HSOS and HSSOS3 do not scale well with large numbers of cliques).

6.1. Block-moving with minimum work.

The continuous time version of block-moving with minimum work is shown as follows [17]:

$$
\begin{cases}
\displaystyle\min_{u(\tau),x_1(\tau),x_2(\tau)} & \displaystyle\int_{\tau=0}^{1} |u(\tau)x_2(\tau)|\,\mathrm{d}\tau \\
\text{s.t.} & \dfrac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{f}(\mathbf{x},u) = \begin{bmatrix} x_2 \\ u \end{bmatrix}, \\
& x_1(0)=0, x_2(0)=0, x_1(1)=1, x_2(1)=0,
\end{cases}
\tag{6.1}
$$

where u is the control input, and x_1, x_2 are the block's position and velocity, respectively. Starting from the origin $\mathbf{x}(0)=[0,0]^\intercal$ in the state space, our goal is to push the block to a terminal state $\mathbf{x}(1)=[1,0]^\intercal$ at time $t=1$ while minimizing the work done. To achieve this, slack variables are introduced, and direct collocation is applied to discretize (6.1). This process results in the following POP:

$$
\begin{cases}
\min\limits_{u_k,\mathbf{x}_k,\mathbf{s}_k} & \sum\limits_{k=0}^{N}(s_{k,1}+s_{k,2})\cdot h \\
\text{s.t.} & s_{k,1}\geq 0, s_{k,2}\geq 0, \quad k=0,\ldots,N, \\
& s_{k,1}-s_{k,2}=u_k\cdot x_{k,2}, \quad k=0,\ldots,N, \\
& \dot{\bar{\mathbf{x}}}_k=\mathbf{f}(\bar{\mathbf{x}}_k,\bar{u}_k), \quad k=0,\ldots,N-1, \\
& x_{0,1}=0, x_{0,2}=0, x_{N,1}=1, x_{N,2}=0,
\end{cases} \tag{6.2}
$$

where N are the total time steps and h is the time step. With slight abuse of notation, we have used k to denote the index of the discretized time step in (6.2) (recall k also denotes the relaxation order). Since the terminal time is fixed to be 1, $N\cdot h=1$ should hold. The equation $\dot{\bar{\mathbf{x}}}_k=\mathbf{f}(\bar{\mathbf{x}}_k,\bar{u}_k)$ in (6.2) is called the collocation constraints, which iare given by

$$
\begin{aligned}
\dot{\bar{\mathbf{x}}}_k &= -\frac{3}{2h}(\mathbf{x}_k-\mathbf{x}_{k+1})-\frac{1}{4}(\mathbf{f}(\mathbf{x}_k,u_k)+\mathbf{f}(\mathbf{x}_{k+1},u_{k+1})), \\
\bar{\mathbf{x}}_k &= \frac{1}{2}(\mathbf{x}_k+\mathbf{x}_{k+1})+\frac{h}{8}(\mathbf{f}(\mathbf{x}_k,u_k)-\mathbf{f}(\mathbf{x}_{k+1},u_{k+1})), \\
\bar{u}_k &= \frac{1}{2}(u_k+u_{k+1}).
\end{aligned}
$$

It should be noted that (6.2) is a nonconvex problem due to the inclusion of quadratic equality constraints. Moreover, (6.2) exhibits a chain-like csp. Specifically, if we consider the kth variable clique as $(u_{k-1},\mathbf{x}_{k-1},\mathbf{s}_{k-1},u_k,\mathbf{x}_k,\mathbf{s}_k), k=1,\ldots,N$, the RIP is satisfied due to the Markov property [44], since the system's state at time step k only depends on the state and control input at time step $k-1$. Setting the relaxation order k to 2, we test the three approaches' performance on multiple values of N and $u_{\max}$. For HSSOS1, the perturbation parameter ϵ is set to 10^{-4}. The results are shown in Table 8. Note that we also reported the suboptimality gap η between the SDP's solution and the solution refined by nonlinear programming solvers (here we use the MATLAB method fmincon). In all our experiments, fmincon can always round a feasible solution (with maximum constraint violation below 10^{-10}). The rounding time of our nonlinear programming solver is negligible compared to the SDP solving time. Denote the SDP's optimal value as f_{lower} and the fmincon local minimum as f_{upper}. Then η is defined as

$$
\eta=\frac{|f_{\text{upper}}-f_{\text{lower}}|}{1+|f_{\text{upper}}|+|f_{\text{lower}}|}. \tag{6.3}
$$

This gap is shown in logarithmic form, i.e., $\log_{10}\eta$. From Table 8, we see that HSSOS2 achieves tight solutions in all parameter settings, with the suboptimality gap always lower than 10^{-4}. However, both SSOS and HSSOS1 suffer from numerical issues. Further trajectory visualizations for $u(t)$ are given in Figure 1.

6.2. Optimal control of Van der Pol. Now we consider the optimal control problem for a Van der Pol oscillator [8], a highly nonlinear and potentially unstable system. Its continuous time dynamics is

$$
\mathbf{f}(\mathbf{x},u)=\frac{d}{dt}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}(1-x_2^2)x_1-x_2+u\\x_1\end{bmatrix}. \tag{6.4}
$$

Here $\mathbf{x}=[x_1,x_2]^\intercal$ is the system state, and u is the control input. Utilizing the direct multiple shooting technique leads to the trajectory optimization problem as follows:

TABLE 8
Results of the block-moving example.

N	$u_{\max}$	SSOS			HSSOS1			HSSOS2		
		opt	time	gap	opt	time	gap	opt	time	gap
10	10	1.1164	2.90	−2.62	0.5044*	25.9	−0.43	1.1111	10.0	−12.0
	12	0.9660	3.26	−2.74	0.4391*	26.5	−0.43	0.9625	12.5	−10.5
	14	0.8100*	6.30	−2.01	0.3774*	24.2	−0.45	0.7944	10.8	−10.2
	16	0.6370*	4.72	−1.61	0.3101*	26.4	−0.49	0.6069	13.9	−12.9
	18	0.6448*	5.43	−0.63	0.2565*	23.4	−0.55	0.4000	10.6	−7.89
	20	0.3651*	3.41	−0.45	0.1370	28.4	−0.92	0.1741	11.9	−8.17
20	10	1.2301*	7.35	−2.76	0.6873*	51.0	−0.55	1.2266	29.3	−10.9
	12	1.1725*	10.4	−1.97	0.6663*	59.6	−0.57	1.1476	27.7	−11.1
	14	1.1410*	12.7	−1.46	0.6600*	84.2	−0.63	1.0646	34.3	−7.40
	16	1.0924*	9.23	−1.40	0.6181*	56.3	−0.62	1.0096	26.6	−8.56
	18	1.0854*	8.31	−1.21	0.5918*	59.6	−0.63	0.9591	27.1	−8.04
	20	1.0278*	8.70	−1.19	0.5680*	74.7	−0.64	0.9036	29.7	−8.66
30	10	1.2724*	9.50	−2.02	0.8184*	73.5	−0.68	1.2483	42.7	−9.22
	12	1.2107*	11.5	−1.92	0.8024*	23.2	−0.72	1.1822	63.0	−8.10
	14	1.1985*	10.1	−1.52	0.7787*	23.4	−0.74	1.1294	44.3	−7.91
	16	1.1804*	10.1	−1.41	0.7649*	23.5	−0.75	1.0923	47.4	−8.27
	18	1.1718*	8.38	−1.27	0.7436*	22.6	−0.76	1.0532	54.9	−7.06
	20	1.2133*	6.74	−1.04	0.7331*	27.1	−0.80	1.0119	67.7	−4.07

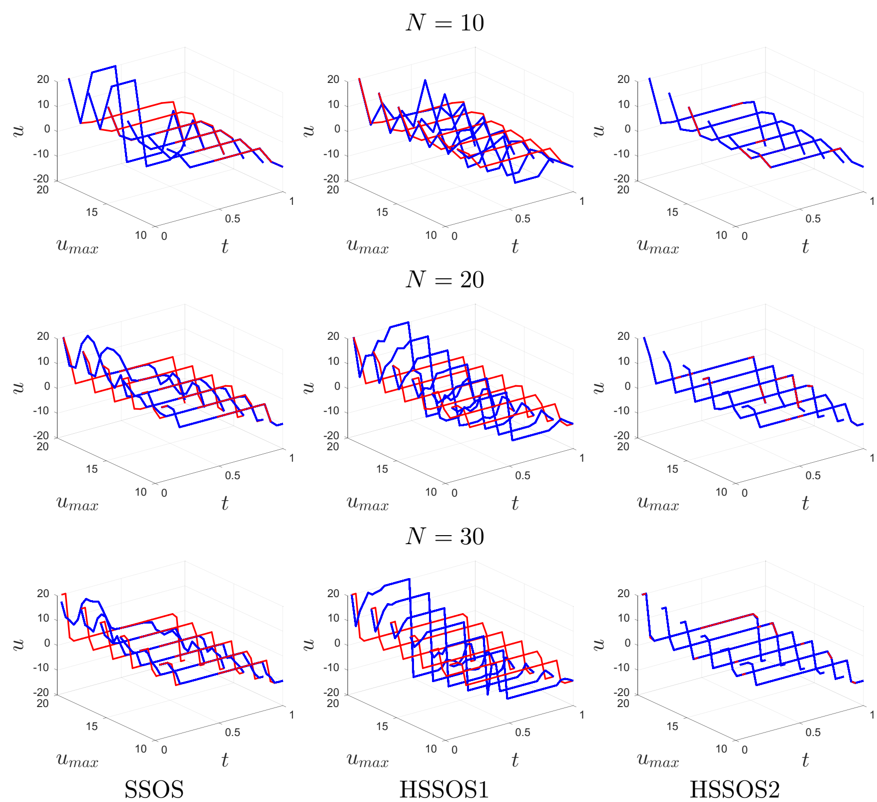

FIG. 1. *Comparison between the SDP's solutions (blue lines) and solutions refined by fmincon (red lines) in the block-moving example. In HSSOS2, red lines and blue lines are nearly indistinguishable, indicating the attainment of tight solutions. (See online version for color.)*

Table 9
Results of the Van der Pol example.

N	SSOS			HSSOS1			HSSOS2		
	opt	time	gap	opt	time	gap	opt	time	gap
10	11.559	0.17	−5.93	11.454	0.48	−2.94	11.559	0.51	−7.55
20	18.457	0.33	−5.05	18.230	0.97	−2.57	18.534	1.32	−7.15
30	23.485	0.55	−3.98	23.012	2.50	−1.93	23.734	5.11	−6.28
40	26.728	0.78	−2.06	25.760	3.89	−1.54	27.419	9.68	−5.99
50	28.122	1.64	−1.57	27.418	12.6	−1.44	29.780	21.1	−5.62
60	28.655	2.07	−1.44	28.434	25.5	−1.45	31.058	42.5	−5.00
70	28.782	1.12	−1.37	29.118	5.07	−1.49	31.768	18.3	−4.94
80	28.874	1.44	−1.35	29.582	5.63	−1.55	32.131	10.7	−4.34
90	28.978	1.52	−1.34	29.918	9.79	−1.65	32.235	33.6	−4.22
100	29.033	1.74	−1.34	30.198	10.2	−1.71	32.257	13.5	−4.03

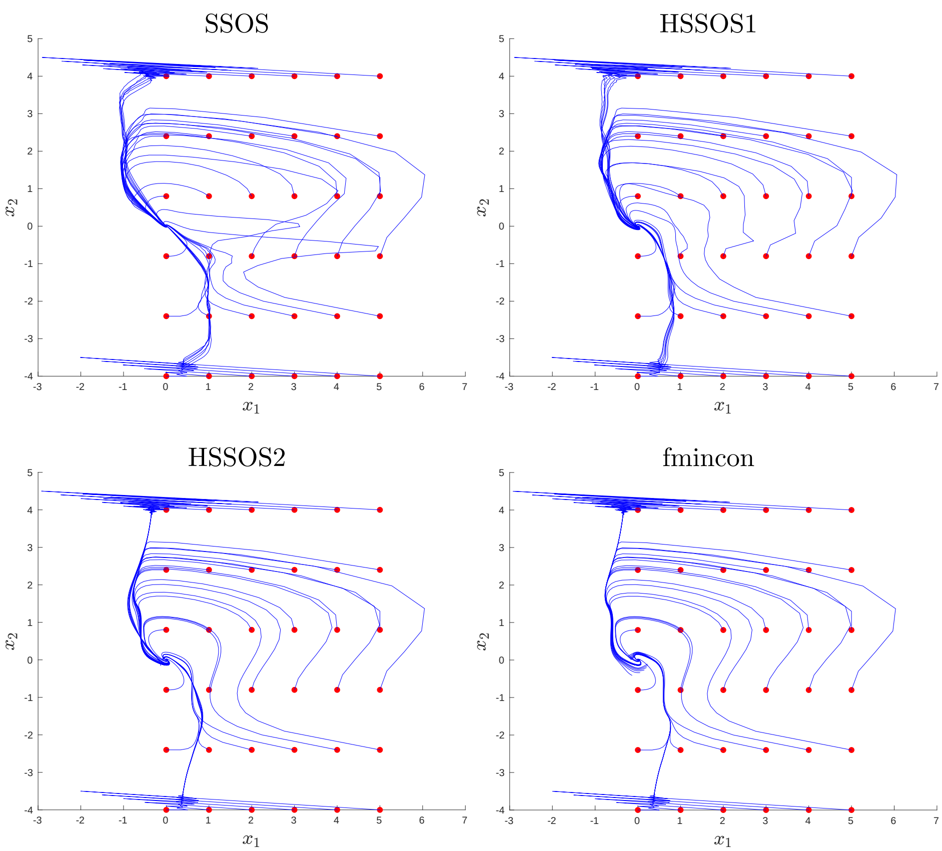

Fig. 2. *Comparison between the SDP's solutions and solutions refined by fmincon in the Van der Pol example. The red dots represent different initial states sampled with* 2*-D grids on the* x_1*-*x_2 *plane. Notably, all three approaches' initial guesses lead to the same refined trajectories by fmincon. Among these initial guesses, the one offered by HSSOS*2 *is of the best quality, since its trajectories are the closest to the fminconrounding results.*

$$
\begin{cases}
\min\limits_{u_k,\mathbf{x}_k} & \displaystyle\sum_{k=0}^{N-1} (u_k^2 + \|\mathbf{x}_k\|^2) \cdot h + \|\mathbf{x}_N\|^2 \cdot h \\
\text{s.t.} & \mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{f}(\mathbf{x}_k, u_k) \cdot h, \quad k = 0, \dots, N-1, \\
& u_{\max}^2 - u_k^2 \geq 0, \quad k = 0, \dots, N-1, \\
& \mathbf{x}_0 = \mathbf{x}_{\text{init}},
\end{cases} \tag{6.5}
$$

where N are the total time steps and h is the step length. Like (6.2), POP (6.5) also exhibits a chain-like csp by assigning the N sequential cliques as $\{(\mathbf{x}_{k-1}, u_{k-1}, \mathbf{x}_k)\}_{k=1}^N$. However, (6.5) does not fulfill the Archimedeanness assumption since the variables $\{\mathbf{x}_k\}_{k=1}^N$ are not subject to any bound. With the relaxation order $k = 2$, we incrementally vary N from 10 to 100 with step length 10. At each N, the performance of the three approaches is assessed using 36 predetermined initial states. Table 9 presents the average results across these states. From the table, we can draw the conclusion that the extracted solutions of HSSOS2 are better than those of SSOS and HSSOS1 in terms of achieving one or two orders of magnitude lower suboptimality gap η. Further comparison for solutions extracted from SDP relaxations and refined by fmincon are shown in Figure 2. Interestingly, despite the varying initial guesses supplied by the three approaches, they all converge to identical refined solutions.

7. Conclusions and discussions.

We have presented the sparse homogenized moment-SOS hierarchies to solve sparse polynomial optimization with unbounded sets. Asymptotic convergence has been established under the RIP, and extensive numerical experiments demonstrate the power of our approaches in solving problems with up to thousands of variables. Additionally, we provide applications to two trajectory optimization problems and obtain global solutions of high accuracy.

Recently, polynomial upper bounds on the convergence rate of the moment-SOS hierarchy with correlative sparsity (2.7)–(2.8) were obtained in [20]. It is promising that we will get similar convergence rates for our sparse homogenized hierarchies, with additional considerations on the behavior of f at infinity of K.

Acknowledgments. The authors would like to thank the handling editor and the anonymous referees for their valuable comments and suggestions which helped to greatly improve the quality and presentation of this paper.

REFERENCES

[1] E. Andersen and K. Andersen, *The Mosek interior point optimizer for linear programming: An implementation of the homogeneous algorithm*, in High Performance Optimization, Appl. Optim. 33, H. Frenk, K. Roos, T. Terlaky, and S. Zhang, eds., Springer, 2000, pp. 197–232.

[2] S. Arnborg, D. G. Corneil, and A. Proskurowski, *Complexity of finding embeddings in a k-tree*, SIAM J. Algebraic Discrete Methods, 8 (1987), pp. 277–284, https://doi.org/10.1137/0608024.

[3] A. Aydinoglu, P. Sieg, V. M. Preciado, and M. Posa, *Stabilization of complementarity systems via contact-aware controllers*, IEEE Trans. Robot., 38 (2022), pp. 1735–1754, https://doi.org/10.1109/TRO.2021.3120931.

[4] J. T. Betts, *Survey of numerical methods for trajectory optimization*, J. Guidance Control Dynam., 21 (1998), pp. 193–207, https://doi.org/10.2514/2.4231.

[5] J. R. Blair and B. Peyton, *An introduction to chordal graphs and clique trees*, in Graph Theory and Sparse Matrix Computation, Springer, 1993, pp. 1–29.

[6] R. Curto and L. Fialkow, *Truncated K-moment problems in several variables*, J. Operat. Theory, 54 (2005), pp. 189–226, https://www.jstor.org/stable/24715679.

[7] 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, https://doi.org/10.1007/s00013-007-2234-z.

[8] J. Guckenheimer, K. Hoffman, and W. Weckesser, *The forced van Der Pol equation* I*: The slow flow and its bifurcations*, SIAM J. Appl. Dyn. Syst., 2 (2003), pp. 1–35, https://doi.org/10.1137/S1111111102404738.

[9] F. Guo, L. Wang, and G. Zhou, *Minimizing rational functions by exact Jacobian SDP relaxation applicable to finite singularities*, J. Global Optim., 58 (2014), pp. 261–284, https://doi.org/10.1007/s10898-013-0047-0.

[10] H. V. Hà and T. S. Pham, *Solving polynomial optimization problems via the truncated tangency variety and sums of squares*, J. Pure Appl. Algebra, 213 (2009), pp. 2167–2176, https://doi.org/10.1016/j.jpaa.2009.03.014.

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

[12] L. Huang, J. Nie, and Y.-X. Yuan, *Generalized truncated moment problems with unbounded sets*, J. Sci. Comput., 95 (2023), 15, https://doi.org/10.1007/s10915-023-02139-z.

[13] L. Huang, J. Nie, and Y.-X. Yuan, *Homogenization for polynomial optimization with unbounded sets*, Math. Program., 200 (2023), pp. 105–145, https://doi.org/10.1007/s10107-022-01878-5.

[14] L. Huang, J. Nie, and Y.-X. Yuan, *Finite convergence of moment-SOS relaxations with nonreal radical ideals*, SIAM J. Optim., 34 (2024), pp. 3399–3428, https://doi.org/10.1137/23M1605430.

[15] V. Jeyakumar, J. Lasserre, and G. Li, *On polynomial optimization over non-compact semialgebraic sets*, J. Optim. Theory Appl., 163 (2014), pp. 707–718, https://doi.org/10.1007/s10957-014-0545-3.

[16] C. Josz and D. K. Molzahn, *Lasserre hierarchy for large scale polynomial optimization in real and complex variables*, SIAM J. Optim., 28 (2018), pp. 1017–1048, https://doi.org/10.1137/15M1034386.

[17] M. Kelly, *An introduction to trajectory optimization: How to do your own direct collocation*, SIAM Rev., 59 (2017), pp. 849–904, https://doi.org/10.1137/16M1062569.

[18] I. Klep, V. Magron, and J. Povh, *Sparse noncommutative polynomial optimization*, Math. Program., 193 (2022), pp. 789–829, https://doi.org/10.1007/s10107-020-01610-1.

[19] M. Kojima and M. Muramatsu, *A note on sparse SOS and SDP relaxations for polynomial optimization problems over symmetric cones*, Comput. Optim. Appl., 42 (2009), pp. 31–41, https://doi.org/10.1007/s10589-007-9112-2.

[20] M. Korda, V. Magron, and R. Ríos-Zertuche, *Convergence rates for sums-of-squares hierarchies with correlative sparsity*, Math. Program., 209 (2025), 209, pp. 435–473, https://doi.org/10.1007/s10107-024-02071-6.

[21] 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.

[22] J. B. Lasserre, *Convergent SDP-relaxations in polynomial optimization with sparsity*, SIAM J. Optim., 17 (2006), pp. 822–843, https://doi.org/10.1137/05064504X.

[23] J. B. Lasserre, *An Introduction to Polynomial and Semi-algebraic Optimization*, Cambridge Texts Appl. Math. 52, Cambridge University Press, 2015.

[24] M. Laurent, *Semidefinite representations for finite varieties*, Math. Program., 109 (2007), pp. 1–26, https://doi.org/10.1007/s10107-004-0561-4.

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

[26] V. Magron, *Interval enclosures of upper bounds of roundoff errors using semidefinite programming*, ACM Trans. Math. Software (TOMS), 44 (2018), 41, https://doi.org/10.1145/3206430.

[27] V. Magron and J. Wang, *TSSOS: A Julia Library to Exploit Sparsity for Large-Scale Polynomial Optimization*, preprint, https://arxiv.org/abs/2103.00915, 2021.

[28] V. Magron and J. Wang, *Sparse Polynomial Optimization: Theory and Practice*, World Scientific, 2023.

[29] N. Mai, J. Lasserre, and V. Magron, *Positivity certificates and polynomial optimization on non-compact semialgebraic sets*, Math. Program., 194 (2022), pp 443–485, https://doi.org/10.1007/s10107-021-01634-1.

[30] N. Mai and V. Magron, *On the complexity of Putinar—Vasilescu's Positivstellensatz*, J. Complexity, 72 (2022), 101663, https://doi.org/10.1016/j.jco.2022.101663.

[31] N. Mai, V. Magron, and J. Lasserre, *A sparse version of Reznick's Positivstellensatz*, Math. Oper. Res., 48 (2023), pp. 812–833, https://doi.org/10.1287/moor.2022.1284.

[32] M. Marshall, *Positive Polynomials and Sums of Squares*, Math. Surveys Monogr. 146, American Mathematical Soc., 2008.

[33] M. Newton and A. Papachristodoulou, *Sparse polynomial optimisation for neural network verification*, Automatica, 157 (2023), 111233, https://doi.org/10.1016/j.automatica.2023.111233.

[34] J. Nie, *Optimality conditions and finite convergence of Lasserre's hierarchy*, Math. Program., 146 (2014), pp. 97–121, https://doi.org/10.1007/s10107-013-0680-x.

[35] J. Nie, *Tight relaxations for polynomial optimization and Lagrange multiplier expressions*, Math. Program. Ser. A, 178 (2019), pp. 1–37, https://doi.org/10.1007/s10107-018-1276-2.

[36] J. Nie, *Moment and Polynomial Optimization*, MOS-SIAM Ser. Optim. 31, SIAM, 2023, https://doi.org/10.1137/1.9781611977608.

[37] T. Pham, *Tangencies and polynomial optimization*, Math. Program. Ser. A, 199 (2023), pp. 1239–1272, https://doi.org/10.1007/s10107-022-01869-6.

[38] M. Putinar, *Positive polynomials on compact semi-algebraic sets*, Indiana Univ. Math. J., 42 (1993), pp. 969–984, https://doi.org/10.1512/iumj.1993.42.42045.

[39] M. Putinar and F.-H. Vasilescu, *Positive polynomials on semi-algebraic sets*, C. R. Acad. Sci. Ser. I Math., 328 (1999), pp. 585–589, https://doi.org/10.1016/S0764-4442(99)80251-1.

[40] M. Putinar and F.-H. Vasilescu, *Solving moment problems by dimensional extension*, C. R. Acad. Sci. Ser. I Math., 328 (1999), pp. 495–499, https://doi.org/10.1016/S0764-4442(99)80197-9.

[41] Z. Qu and X. Tang, *A correlatively sparse Lagrange multiplier expression relaxation for polynomial optimization*, SIAM J. Optim., 34 (2024), pp. 127–162, https://doi.org/10.1137/22M1515689.

[42] B. Reznick, *Some concrete aspects of Hilbert's* 17*th problem*, in Real Algebraic Geometry and Ordered Structures Contemp. Math. 253, C. N. Delzell and J. J. Madden, eds., Amer. Math. Soc., 2000, pp. 251–272, https://doi.org/10.1090/conm/253/03936.

[43] M. Schweighofer, *Global optimization of polynomials using gradient tentacles and sums of squares*, SIAM J. Optim., 17 (2006), pp. 920–942, https://doi.org/10.1137/050647098.

[44] S. Teng, A. Jasour, R. Vasudevan, and M. Ghaffari, *Convex Geometric Motion Planning on Lie Groups via Moment Relaxation*, preprint, https://arxiv.org/abs/2305.13565, 2023.

[45] K. Toh, *Some numerical issues in the development of SDP algorithms*, Informs OS Today, 8 (2018), pp. 7–20.

[46] H. H. Vui and P. T. So'n, *Global optimization of polynomials using the truncated tangency variety and sums of squares*, SIAM J. Optim., 19 (2008), pp. 941–951, https://doi.org/10.1137/080719212.

[47] 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.

[48] J. Wang and V. Magron, *Exploiting term sparsity in noncommutative polynomial optimization*, Comput. Optim. Appl., 80 (2021), pp. 483–521, https://doi.org/10.1007/s10589-021-00301-7.

[49] J. Wang, V. Magron, and J.-B. Lasserre, *Chordal-TSSOS: A moment-SOS hierarchy that exploits term sparsity with chordal extension*, SIAM J. Optim., 31 (2021), pp. 114–141, https://doi.org/10.1137/20M1323564.

[50] J. Wang, V. Magron, and J.-B. Lasserre, *TSSOS: A moment-SOS hierarchy that exploits term sparsity*, SIAM J. Optim., 31 (2021), pp. 30–58, https://doi.org/10.1137/19M1307871.

[51] J. Wang, V. Magron, J. B. Lasserre, and N. Mai, *CS-TSSOS: Correlative and term sparsity for large-scale polynomial optimization*, ACM Trans. Math. Software, 48 (2022), 42, https://doi.org/10.1145/3569709.

[52] J. Wang, V. Magron, and J. B. Lasserre, *Certifying global optimality of AC-OPF solutions via sparse polynomial optimization*, Electr. Pow. Syst. Res., 213 (2022), 108683, https://doi.org/10.1016/j.epsr.2022.108683.

[53] J. Wang, C. Schlosser, M. Korda, and V. Magron, *Exploiting term sparsity in moment-SOS hierarchy for dynamical systems*, IEEE Trans. Automat. Control, 68 (2023), pp. 8232–8237, https://doi.org/10.1109/TAC.2023.3293014.