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 has asymptotic convergence. Furthermore, we provide two alternative sparse hierarchies to remove perturbations while preserving asymptotic convergence. As byproducts, new Positivstellensätze are obtained for sparse positive polynomials on unbounded sets. Extensive numerical experiments demonstrate the power of our approach in solving sparse polynomial optimization problems on unbounded sets with up to thousands of variables. Finally, we apply our approach 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

AMS subject classifications. 90C23, 90C17, 90C22, 90C26

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 $\mathbf{x} \coloneqq (x_1, \ldots, x_n) \in \mathbb{R}^n$. 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 $f_{\min} > -\infty$. The Moment-SOS hierarchy proposed by Lasserre [18] is efficient in solving (1.1). Under the Archimedeanness of constraining polynomials (the feasible set K must be compact in this case; see [18, 21]), it yields a sequence of semidefinite relaxations whose optimal values converge to $f_{\min}$. Furthermore, it was shown in [9, 31] that the Moment-SOS hierarchy converges in finitely many steps if standard optimality conditions hold at every global minimizer. We refer to books and surveys [20, 22, 29, 33] for more general introductions to polynomial optimization.

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 Karush-Kuhn-Tucker (KKT) conditions and Lagrange multipliers, Nie proposed tight Moment-SOS relaxations for solving (1.1) [32]. In [12], the authors proposed Moment-SOS relaxations by adding sublevel set constraints. The resulting hierarchy of relaxations is also convergent under the Archimedeanness for the new constraints. Based on Putinar-Vasilescu's Positivstellensatz [36, 37], Mai, Lasserre, and Magron [28] proposed a new hierarchy of Moment-SOS relaxations by adding a small perturbation to the objective, and convergence to a neighborhood of $f_{\min}$ was proved if the optimal value is achievable. The complexity of this new

*Submitted to the editors DATE.

Funding: This work was supported by the NSFC under grants 12201618 and 12171324.

†Department of Mathematics, University of California San Diego (leh010@ucsd.edu).

‡School of Engineering and Applied Sciences, Harvard University (skang1@g.harvard.edu).

§Academy of Mathematics and Systems Science, Chinese Academy of Sciences (AMSS-CAS) (wangjie212@amss.ac.cn).

¶School of Engineering and Applied Sciences, Harvard University (hankyang@seas.harvard.edu).

hierarchy was studied in [26]. Recently, a homogenized Moment-SOS hierarchy was proposed in [11] to solve polynomial optimization with unbounded sets employing homogenization techniques, and finite convergence was proved if standard optimality conditions hold at every global minimizer, including those at infinity. A theoretically interesting problem for polynomial optimization with unbounded sets is the case where the optimal value is not achievable. We refer to [7, 11, 34, 40, 43] for related works.

A drawback of the Moment-SOS hierarchy is its limited scalability. This is because the size of involved matrices at the kth order relaxation is $\binom{n+k}{k}$ which increases rapidly as n, k grow, and current semidefinite program (SDP) solvers based on interior-point methods can typically solve SDPs involving matrices of moderate size (say, $\leq 2,000$) in reasonable time on a standard laptop [42]. An important way to improve the scalability is exploiting sparsity of inputting polynomials. There are two types of sparsity patterns in the literature to reduce the size of SDP relaxations: *correlative sparsity* and *term sparsity*. Correlative sparsity [44] considers the sparsity pattern of variables. The resulting sparse Moment-SOS hierarchy is obtained by building blocks of SDP matrices with respect to 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 [5, 16, 19]. In contrast, term sparsity proposed by Wang et al. [46, 47] considers the sparsity of monomials or terms. One can obtain a two-level block Moment-SOS hierarchy 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 obtained SDP blocks is relatively small, then the resulting SDP relaxations are more tractable and computational costs can be significantly reduced. They have been successfully applied to solve optimal power flow problems [13, 49], round-off error bound analysis [23], noncommutative polynomial optimization [15, 45], neural network verification [30], dynamical systems analysis [50], etc.

However, the above sparse Moment-SOS hierarchies may not converge when K is unbounded as in the dense case. For the unbounded case, Mai, Lasserre, and Magron [27] have recently provided a sparse version of Putinar–Vasilescu's Positivstellensatz. To be more specific, 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 $\sigma_{0,\ell}$, $\sigma_{j,\ell}$, $j \in J_\ell$ of suitable degrees in variables $\mathbf{x}(\ell)$, $\ell = 1,\ldots,p$ such that

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

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 notations 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 occurrence of high-degree denominators, it is limited to solving problems with up to 10 variables. The computational benefit of this sparse hierarchy is hence rather limited, and it is essentially a theoretical result as stated in [27].

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 for (1.1) while preserving the correlative sparsity pattern of the original problem. The sparse reformulation introduces two new types of variables. One is the homogenization variable, and the other consists of auxiliary variables associated to each variable clique. Then we 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 near neighborhood of $f_{\min}$.

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. We establish 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 performs much better than the usual sparse Moment-SOS hierarchy when solving sparse polynomial optimization on unbounded sets. In fact, with it we are able to handle such problems with up to thousands of variables!

V. To further illustrate its power, we apply our approach to trajectory optimization problems arising from the fields of robotics and control. It turns out that our approach can achieve global solutions for those problems with high accuracy.

The rest of 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. Then 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. Applications to trajectory optimization are provided in Section 6. Section 7 draws conclusions and make some discussions.

2. Notations 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 $[n] \coloneqq \{1, \ldots, n\}$. Let $\mathbf{x} \coloneqq (x_1, \ldots, x_n)$ denote a tuple of variables and let $\mathbf{x}^2 \coloneqq (x_1^2, \ldots, x_n^2)$. 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

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

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

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

Let $\mathbb{R}[\mathbf{x}] \coloneqq \mathbb{R}[x_1, \ldots, x_n]$ be the ring of polynomials in $\mathbf{x}$ with real coefficients, and $\mathbb{R}[\mathbf{x}]_k \subseteq \mathbb{R}[\mathbf{x}]$ is 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 $\tilde{\mathbf{x}} \coloneqq (x_0, x_1, \ldots, x_n)$), 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 the 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 to [20, 22, 33] for more details.

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

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

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

$$\mathrm{QM}[g]_k := \Sigma[\mathbf{x}]_k + g_1 \cdot \Sigma[\mathbf{x}]_{k-\lceil \deg(g_1)/2 \rceil} + \cdots + g_m \cdot \Sigma[\mathbf{x}]_{k-\lceil \deg(g_m)/2 \rceil}. \tag{2.2}$$

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

$$\mathrm{QM}[g, \mathbf{x}'] := \Sigma[\mathbf{x}'] + g_1 \cdot \Sigma[\mathbf{x}'] + \cdots + g_m \cdot \Sigma[\mathbf{x}'],$$

$$\mathrm{QM}[g, \mathbf{x}']_k := \Sigma[\mathbf{x}']_k + g_1 \cdot \Sigma[\mathbf{x}']_{k-\lceil \deg(g_1)/2 \rceil} + \cdots + g_m \cdot \Sigma[\mathbf{x}']_{k-\lceil \deg(g_m)/2 \rceil}.$$

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 is not always true. 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 referred to as Putinar's Positivstellensatz [35].

For $k \in \mathbb{N}$, let $\mathbb{R}^{\mathbb{N}^n_{2k}}$ be the set of all real vectors that are 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 indexed by $\mathbb{N}^n_k$ such that

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

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 $M_k[p\mathbf{y}]$ 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 have convergence. In this section, we present the homogenization approach introduced [11] for solving polynomial optimization with 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

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

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. A basic property for closedness at infinity is 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 $d \coloneqq \deg(f)$ [11]. Therefore, when K is closed at infinity, (1.1) is equivalent to the following homogenized optimization problem:

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

By applying the standard Moment-SOS relaxations to solve the homogenized reformulation (2.6), a homogenized Moment-SOS hierarchy was proposed in [10, 11] to solve (1.1) with unbounded sets. Asymptotic and finite convergences were proved under some generic assumptions.

2.3. Polynomial optimization with correlative sparsity. The Moment-SOS hierarchy with correlative sparsity was first studied in [44]. Suppose that the subsets of variables $\mathbf{x}(1), \ldots, \mathbf{x}(p) \subseteq \mathbf{x}$ satisfy $\cup_{\ell=1}^p \mathbf{x}(\ell) = \mathbf{x}$. The POP (1.1) is said to have a *correlative sparsity pattern* (csp) $(\mathbf{x}(1), \ldots, \mathbf{x}(p))$ if

(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 exists a partition $\{J_1, \ldots, J_p\}$ of $[m]$ such that for every $\ell \in [p]$ and every $j \in J_\ell$, we have $g_j \in \mathbb{R}[\mathbf{x}(\ell)]$.

Let

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

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

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

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

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

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

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

Under the RIP, it was shown in [5, 16, 19] that the sparse Moment-SOS hierarchy (2.7)–(2.8) has asymptotic convergence.

THEOREM 2.1 ([5, 16, 19]). *Suppose that (1.1) has the csp $(\mathbf{x}(1), \ldots, \mathbf{x}(p))$ 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*

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

Remark 2.2. For a given POP, a csp $(\mathbf{x}(1), \ldots, \mathbf{x}(p))$ satisfying the RIP can be obtained by: (1) building the csp graph, (2) generating a chordal extension, (3) taking the list of maximal cliques; see [25, 44]. Note that chordal extensions of a graph are typically not unique and so are correlative sparsity patterns.

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 the POP (1.1) admits a correlative sparse pattern and K is unbounded. Note that we can not 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 $(\mathbf{x}(1), \ldots, \mathbf{x}(p))$ be a csp of (1.1) and $\{J_1, \ldots, J_p\}$ a partition of $[m]$ such that for every $\ell \in [p]$ and every $j \in J_\ell$, $g_j \in \mathbb{R}[\mathbf{x}(\ell)]$. Define the sparse set

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

where $\tilde{\mathbf{x}}(\ell) := (x_0, \mathbf{x}(\ell))$ and $\mathbf{w} := (w_1, \ldots, w_p)$ is a tuple of auxiliary variables. The difference between $\widetilde{K}_s^1$ and $\widetilde{K}$ is that we replace the single non-sparse spherical constraint $\|\tilde{\mathbf{x}}\|^2 = 1$ by multiple spherical constraints $\|\tilde{\mathbf{x}}(\ell)\|^2 + w_\ell^2 = 1$, $\ell \in [p]$. Interestingly, $\widetilde{K}_s^1$ retains the correlative sparse pattern of the feasible set K.

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

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

where $\epsilon \geq 0$ is a tunable parameter. Let $f^{(\epsilon)}$ be the optimal value of (3.2) and

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

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

$$\left\{\begin{array}{ll} \sup & \gamma \\ \text{s.t.} & \tilde{f}(\tilde{\mathbf{x}}) + \epsilon \cdot \left(\sum\limits_{i=0}^{n} x_i^{d_0} + \sum\limits_{\ell=1}^{p} w_\ell^{d_0}\right) - \gamma x_0^d \\ & \qquad \in \sum\limits_{\ell=1}^{p} \left(\text{Ideal}[h_\ell, \tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}]_{2k} + \text{QM}[\{x_0\} \cup \{\tilde{g}_j\}_{j\in J_\ell}, \tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}]_{2k}\right). \end{array}\right. \tag{3.3}$$

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

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

The hierarchy of relaxations (3.3)–(3.4) is called the sparse homogenized Moment-SOS hierarchy for solving (1.1). Let f_k, f_k' denote the optima of (3.3) and (3.4), respectively.

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 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 holds*

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

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

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

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

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

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

Remark 3.2. Auxiliary variables $\mathbf{w}$ are necessary for Theorem 3.1 to be true. For instance, let $\mathbf{x}(1) = \{x_1, x_2\}$, $\mathbf{x}(2) = \{x_2, x_3\}$. Consider the unconstrained optimization problem with $f = x_1^2 x_2^4 - x_2^4 x_3^2$. Clearly, we have $f_{\min} = -\infty$. In this case, the sparse homogenized reformulation (3.2) without auxiliary variables $\mathbf{w}$ reads as

$$\begin{cases} \sup & \gamma \\ \text{s.t.} & x_1^2 x_2^4 - x_2^4 x_3^2 + \epsilon \cdot (x_0^4 + x_1^4 + x_2^4) - \gamma x_0^4 \geq 0, \ \forall \mathbf{x} \in \widetilde{K}_s^1, \end{cases} \tag{3.5}$$

where

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

For arbitrary $\epsilon \geq 0$, the optimal value of (3.5) is nonnegative. So 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 a 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. By the assumption, we know that $f = f_1 + \cdots + f_p \in \mathbb{R}[\mathbf{x}]$ with $f_\ell \in \mathbb{R}[\mathbf{x}(\ell)]$ $(\ell \in [p])$. For $\ell = 1, \ldots, p$, we have that

$$\tilde{f}_\ell \in \mathbb{R}[\tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}], \ h_\ell \in \mathbb{R}[\tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}], \ \tilde{g}_j \in \mathbb{R}[\tilde{\mathbf{x}}(\ell) \cup \{w_\ell\}] \ (j \in J_\ell).$$

Thus, (3.2) admits the csp $(\tilde{\mathbf{x}}(1) \cup \{w_1\}, \ldots, \tilde{\mathbf{x}}(p) \cup \{w_p\})$. Then the conclusion follows. For every $\ell \in [p-1]$, it holds that

$$\begin{aligned} &(\tilde{\mathbf{x}}(\ell+1) \cup \{w_{\ell+1}\}) \cap (\textstyle\bigcup_{j=1}^{\ell} \tilde{\mathbf{x}}(j) \cup \{w_j\}) \\ &\qquad = (x(\ell+1) \cup \{x_0, w_{\ell+1}\}) \cap \textstyle\bigcup_{j=1}^{\ell} (x(j) \cup \{x_0, w_j\}), \\ &\qquad = \{x_0\} \cup (I_{\ell+1} \cap \textstyle\bigcup_{j=1}^{\ell} I_j). \end{aligned} \tag{3.6}$$

Hence, the conclusion follows. ∎

In the following, we prove that the sparse homogenized Moment-SOS hierarchy (3.3)-(3.4) has asymptotic convergence to a neighbourhood of $f_{\min}$.

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

Proof. For $\gamma < f^{(\epsilon)}$, 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 $x_0 = 0$, then

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

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 $\tilde{f}(\tilde{\mathbf{x}}) + \epsilon \cdot \left(\sum_{i=0}^{n} x_i^{d_0} + \sum_{\ell=1}^{p} w_\ell^{d_0}\right) - \gamma x_0^d > 0$ 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 $(\tilde{\mathbf{x}}(1) \cup \{w_1\}, \dots, \tilde{\mathbf{x}}(p) \cup \{w_p\})$ satisfying the RIP. It follows from Theorem 2.1 that

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

Thus we obtain $f_k \to f^{(\epsilon)}$ as $k \to \infty$. ☐

Remark 3.5. If there is no perturbation, i.e., $\epsilon = 0$, the hierarchy (3.3)–(3.4) still provides valid lower bounds to $f_{\min}$. However, the convergence to $f_{\min}$ may not happen. For instance, 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

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

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

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

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

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

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 the dehomogenized Delzell's polynomial and the above representation does not exist as shown in [39].

3.2. Sparse Positivstellensätze with perturbations. In this subsection, we provide new sparse Positivstellensätze for nonnegative polynomials on unbounded semialgebraic sets, based on sparse homogenized reformulations. Our Positivstellensätze do not need any denominator, and thus are quite different from the sparse versions of Reznick's Positivstellensatz and Putinar–Vasilescu's Positivstellensatz given in [27].

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

THEOREM 3.6. *Let $f \in \mathbb{R}[\mathbf{x}]$ be a form of degree d. Suppose f admits a csp $(\mathbf{x}(1), \ldots, \mathbf{x}(p))$ satisfying the RIP.*

(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

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

Proof. Note that $f \geq 0$ on $\mathbb{R}^n$ is equivalent to $f \geq 0$ on $S := \{(\mathbf{x}, \mathbf{w}) \in \mathbb{R}^{n+p} \mid \|\mathbf{x}(\ell)\|^2 + w_\ell^2 = 1, \ell \in [p]\}$. Sicen $0 \notin S$, we have $f + \epsilon(\sum_{i=1}^{n} x_i^d + \sum_{\ell=1}^{p} w_\ell^d) > 0$ on S. If f is positive definite and there exists $(\mathbf{x}, \mathbf{w}) \in \mathbb{R}^{n+p}$ such that $f(\mathbf{x}) + \epsilon \sum_{\ell=1}^{p} w_\ell^d = 0$, we must have $(\mathbf{x}, \mathbf{w}) = \mathbf{0}$. Hence, we have $f + \epsilon \sum_{\ell=1}^{p} w_\ell^d > 0$ on S. Under given assumptions, items (i), (ii) follow from Theorem 2.1. ∎

Note that a polynomial $f \geq 0$ on $\mathbb{R}^n$ is equivalent to $\tilde{f} \geq 0$ on $\mathbb{R}^{n+1}$. Theorem 3.6 can be generalized to non-homogeneous polynomials directly. We omit the proof for cleanness.

THEOREM 3.7. *Let $f \in \mathbb{R}[\mathbf{x}]$ with $\deg(f) = d$. Suppose that f admits a csp $(\mathbf{x}(1), \ldots, \mathbf{x}(p))$ satisfying the RIP.*

(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). Define

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

If K is closed at infinity and f is bounded from below on K, then $f^{(\infty)} \geq 0$ on $K^{(\infty)}$ [11, Theorem 3.6.]. The polynomial f is said to be *positive at infinity* on K if $f^{(\infty)} > 0$ on $K^{(\infty)}$. When K is closed at infinity, $f \geq 0$ on K if and only if $\tilde{f} \geq 0$ on $\widetilde{K}$. Thus we can derive the following sparse homogenized version of Putinar-Vasilescu's Positivstellensatz from Theorem 2.1.

THEOREM 3.8. *Suppose that K is closed at infinity, and POP (1.1) admits a csp $(\mathbf{x}(1),\ldots,\mathbf{x}(p))$ satisfying the RIP.*

(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 definite 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

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

3.3. Extraction of minimizers. In the case of the dense Moment-SOS hierarchy, a convenient criterion for detecting global optimality is flat extension/truncation (see [4, 22, 33]). A procedure for extracting minimizers is given in [8]. This procedure was generalized to polynomial optimization with correlative sparsity in [19]. We adapt it to extract minimizers from the sparse homogenized moment relaxtions (3.4).

Let

$$d_K\coloneqq\max\{\lceil\deg(g_j)/2\rceil,j\in[m]\}\text{ and }d_\ell\coloneqq\max\{\lceil\deg(g_j)/2\rceil,j\in J_\ell\},\ \ell\in[p]. \tag{3.16}$$

Suppose that $\mathbf{y}^*$ is an optimal solution of (3.4) at the kth order relaxation. If there exists an integer $t\in[d_K,k]$ such that

$$\begin{aligned}\operatorname{rank}M_t\left(\mathbf{y}^*,\tilde{\mathbf{x}}(\ell)\cup\{w_\ell\}\right)&=\operatorname{rank}M_{t-d_\ell}\left(\mathbf{y}^*,\tilde{\mathbf{x}}(\ell)\cup\{w_\ell\}\right),\ \text{for all }\ell\in[p],\\ \operatorname{rank}M_t\left(\mathbf{y}^*,\tilde{\mathbf{x}}(i)\cap\tilde{\mathbf{x}}(j)\right)&=1,\ \text{for all }i\neq j\in[p]\text{ with }\tilde{\mathbf{x}}(i)\cap\tilde{\mathbf{x}}(j)\neq\emptyset,\end{aligned} \tag{3.17}$$

then the moment relaxation (3.4) is exact, i.e., $f_k'=f^{(\epsilon)}$. Furthermore, by applying the extraction procedure in [8] to each moment matrix $M_t(\mathbf{y}^*,\tilde{\mathbf{x}}(\ell)\cup\{w_\ell\})$, we obtain a set of points

$$\Delta_\ell\coloneqq\{(\tilde{\mathbf{x}}^*(\ell),w_\ell^*)\}\subseteq\mathbb{R}^{|\mathbf{x}(\ell)|+2},\quad \ell\in[p],$$

where $|\mathbf{x}(\ell)|$ stands for the dimension of $\mathbf{x}(\ell)$. Next we show that approximate minimizers (exact minimizers if $\epsilon=0$) of (1.1) can be extracted from the sets Δ_ℓ ($\ell\in[p]$).

THEOREM 3.9. *Suppose that K is closed at infinity and POP (1.1) admits a csp $(\mathbf{x}(1),\ldots,\mathbf{x}(p))$, and $\mathbf{y}^*$ is an optimal solution of (3.4) satisfying (3.17). Let*

$$\Omega\coloneqq\left\{(x_0^*,\mathbf{x}^*)\in\mathbb{R}^{n+1}\mid \textit{there exists } w_\ell^*\in\mathbb{R}\textit{ such that }(\tilde{\mathbf{x}}^*(\ell),w_\ell^*)\in\Delta_\ell\textit{ for }\ell\in[p]\right\}.$$

If $\tilde{\mathbf{x}}^*\in\Omega$ with $x_0^*>0$, then $\mathbf{x}^*/x_0^*\in K$ and

$$f(\mathbf{x}^*/x_0^*)+\epsilon\left(\sum_{i=0}^{n}(x_i^*)^{d_0}+\sum_{\ell=1}^{p}(w_\ell^*)^{d_0}\right)/(x_0^*)^d=f^{(\epsilon)}, \tag{3.18}$$

where w_ℓ^* is from Δ_ℓ. In particular, when $\epsilon=0$, we have $f(\mathbf{x}^*/x_0^*)=f_{\min}$ and $\mathbf{x}^*/x_0^*$ is a minimizer of (1.1).

Proof. Consider the optimization problem:

$$
\left\{\begin{array}{ll} \inf & \tilde{f}(\tilde{\mathbf{x}}) + \epsilon\left(\sum\limits_{i=0}^{n} x_i^{d_0} + \sum\limits_{\ell=1}^{p} w_\ell^{d_0}\right) - f^{(\epsilon)} x_0^d \\ \text{s.t.} & x_0 \geq 0, \quad \tilde{g}_j(\tilde{\mathbf{x}}) \geq 0, \quad j \in [m], \\ & \|\tilde{\mathbf{x}}(\ell)\|^2 + w_\ell^2 = 1, \quad \ell \in [p], \end{array}\right. \tag{3.19}
$$

whose optimal value is clearly 0. Let $\tilde{\mathbf{x}}^* \in \Omega$ with $x_0^* > 0$. It follows from [19, Theorem 3.2] that $(\tilde{\mathbf{x}}^*, \mathbf{w}^*)$ is a minimizer of (3.19) and

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

Note that $\tilde{f}(\tilde{\mathbf{x}}^*) = (x_0^*)^d f(\mathbf{x}^*/x_0^*)$, $\tilde{g}_j(\tilde{\mathbf{x}}^*) = (x_0^*)^{\deg(g_j)} g(\mathbf{x}^*/x_0^*)$ $(j \in [m])$. Thus, we have that $\mathbf{x}^*/x_0^* \in K$ and (3.18) holds. ∎

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. A natural question is whether we can design a sparse homogenized Moment-SOS hierarchy without perturbations while having asymptotic convergence to the optimal value $f_{\min}$ rather than a neighborhood of $f_{\min}$. In the following, we provide such a hierarchy by introducing a new sparse reformulation of (1.1).

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

$$
\widetilde{K}_s^2 \coloneqq \left\{ (\tilde{\mathbf{x}}, \mathbf{w}) \in \mathbb{R}^{n+1+p} \;\middle|\; \begin{array}{l} x_0 \geq 0, \quad \tilde{g}_j(\tilde{\mathbf{x}}) \geq 0, \quad j \in [m], \\ \sum\limits_{x_i \in \mathbf{x}(\ell)} \frac{1}{p_i} x_i^2 + \frac{1}{p} x_0^2 + w_\ell^2 = 1, \quad \ell \in [p], \\ \|\mathbf{w}\|^2 = p - 1, \quad \mathbf{1} - \tilde{\mathbf{x}}^2 \geq \mathbf{0}, \quad \mathbf{1} - \mathbf{w}^2 \geq \mathbf{0}. \end{array} \right\} \tag{4.1}
$$

Consider the new sparse reformulation for (1.1) ($d \coloneqq \deg(f)$):

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

Let $\mathbf{u} \coloneqq (\tilde{\mathbf{x}}, \mathbf{w})$. Assume that $(\mathbf{u}(1), \ldots, \mathbf{u}(q))$ be the list of maximal cliques of some chordal extension of the csp graph associated with POP (4.2). Let

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

Let $\{J_1, \ldots, J_q\}$ be a partition of $[m]$ such that for every $\ell \in [q]$ and every $j \in J_\ell$, $g_j \in \mathbb{R}[\mathbf{u}(\ell)]$. Moreover, let $\{I_1, \ldots, I_q\}$ be a partition of $[p+1]$ such that for every $\ell \in [q]$ and every $j \in I_\ell$, $h'_j \in \mathbb{R}[\mathbf{u}(\ell)]$.

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

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

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

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

Let $\bar{f}_k$, $\bar{f}_k'$ be the optimal values of (4.3) and (4.4), respectively.

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

THEOREM 4.1. *Suppose that K is closed at infinity. Then the optimal value of (4.2) is $f_{\min}$.*

Proof. Since K is closed at infinity and $f_{\min} > -\infty$, we know that $f^{(\infty)} \geq 0$ on $K^{(\infty)}$. 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 no greater than the optimal value of (4.2). For the converse, let $\mathbf{x}^*$ be a minimizer of (1.1). Let

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

One can verify $(\tilde{\mathbf{x}}^*, \mathbf{w}^*) \in \widetilde{K}_s^2$. If γ is feasible for (4.2), then $\gamma \leq \tilde{f}(\tilde{\mathbf{x}}^*)/(x_0^*)^d = f_{\min}$. □

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

THEOREM 4.2. *Suppose that K is closed at infinity, $f^{(\infty)}$ is positive definite at ∞ on K, and (1.1) admits a csp $(\mathbf{x}(1), \ldots, \mathbf{x}(p))$. Then, $\bar{f}_k \to f_{\min}$ as $k \to \infty$.*

Proof. For any $\gamma < f_{\min}$, we show that $\tilde{f}(\tilde{\mathbf{x}}) - \gamma x_0^d > 0$ on $\widetilde{K}_s^2$. Take any $(\tilde{\mathbf{x}}, \mathbf{w}) \in \widetilde{K}_s^2$. If $x_0 = 0$, then we must have $\mathbf{x} \neq \mathbf{0}$. Suppose otherwise $\mathbf{x} = \mathbf{0}$. Then $w_1^2 = w_2^2 = \cdots = w_p^2 = 1$, which contradicts to the constraint $\|\mathbf{w}\|^2 - p + 1 = 0$ in the definition of $\widetilde{K}_s^2$. Since $f^{(\infty)}$ is positive definite at ∞ on K, we have

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

If $x_0 \neq 0$, we have $\mathbf{x}/x_0 \in K$ and

$$
\tilde{f}(\tilde{\mathbf{x}}) - \gamma (x_0)^d = (x_0)^d (f(\mathbf{x}/x_0) - \gamma) > 0. \tag{4.5}
$$

Therefore, $\tilde{f}(\tilde{\mathbf{x}}) - \gamma x_0^d > 0$ on $\widetilde{K}_s^2$ for any $\gamma < f_{\min}$. Moreover, 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 clearly Archimedean in $\mathbb{R}[\mathbf{u}(\ell)]$. It follows from Theorem 2.1 that for any $\gamma < f_{\min}$,

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

As a result, we obtain $\bar{f}_k \to f_{\min}$ as $k \to \infty$. ◻

Remark 4.3. If the flat truncation condition (3.17) is satisfied for the kth moment relaxation (4.4), then $\bar{f}_k = f_{\min}$ and we can extract minimizers of (1.1) via a similar procedure as described in Section 3.3.

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

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

THEOREM 4.4. *Assume that $f > 0$ 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*

$$\tilde{f} = \sum_{\ell=1}^{q} \left(\sum_{u_i\in\mathbf{u}(\ell)} \sigma_{\ell,i}(1-u_i^2) + \sum_{j\in I_\ell} \tau_{\ell,j} h_j' \right). \tag{4.6}$$

Proof. Let

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

We show that $\tilde{f} > 0$ on S. Take any $(\tilde{\mathbf{x}}, \mathbf{w}) \in S$. If $x_0 \neq 0$, then we have $\tilde{f}(\tilde{\mathbf{x}}) = x_0^d f(\mathbf{x}/x_0) > 0$; If $x_0 = 0$, then $\mathbf{x} \neq \mathbf{0}$ and $\tilde{f}(\tilde{\mathbf{x}}) = f^{(\infty)}(\mathbf{x}) > 0$ since $f^{(\infty)}$ is positive definite. 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 2.1. ◻

The following theorem addresses the constrained case. As the proof is quite similar to that of Theorem 4.4, we omit it for cleanliness.

THEOREM 4.5. *Notations follow Section 4.1. Suppose that K is closed at infinity. If $f > 0$ on K and f is positive definite at infinity on K, then there exists $\sigma_\ell \in$ 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*

$$\tilde{f} = \sum_{\ell=1}^{q} \left(\sigma_\ell + \sum_{j\in I_\ell} \tau_{\ell,j} h_j' \right). \tag{4.7}$$

4.3. An alternative sparse homogenized Moment-SOS hierarchy without perturbations. In the description of $\widetilde{K}_s^2$, there is a spherical constraint $\|\mathbf{w}\|^2 = p-1$ involving all auxiliary variables. When the csp of (1.1) contains a lot of variable cliques, this constraint would lead to a variable clique of big size in the csp of $\widetilde{K}_s^2$, which could significantly increase the computational complexity of the hierarchy (4.3)–(4.4). To address this issue, in the following we propose an alternative sparse homogenized Moment-SOS reformulation without perturbations.

Suppose $(\mathbf{x}(1),\ldots,\mathbf{x}(p))$ is a csp of (1.1). Recall that for $i\in[n]$, p_i denotes the frequency of the variable x_i occurring in $\mathbf{x}(1),\ldots,\mathbf{x}(p)$. Define

$$\widetilde{K}_s^3 \coloneqq \left\{ (\tilde{\mathbf{x}},\mathbf{w})\in\mathbb{R}^{n+p} \;\middle|\; \begin{array}{l} x_0\ge 0,\quad \tilde{g}_j(\tilde{\mathbf{x}})\ge 0,\quad j\in[m],\\ \sum\limits_{x_i\in\mathbf{x}(1)} \frac{1}{p_i}x_i^2+\frac{1}{p}x_0^2 = w_1^2,\\ \sum\limits_{x_i\in\mathbf{x}(2)} \frac{1}{p_i}x_i^2+\frac{1}{p}x_0^2 + w_1^2 = w_2^2,\\ \qquad\vdots\\ \sum\limits_{x_i\in\mathbf{x}(p)} \frac{1}{p_i}x_i^2+\frac{1}{p}x_0^2 + w_{p-1}^2 = 1,\\ \mathbf{1}-\tilde{\mathbf{x}}^2\ge\mathbf{0},\quad \mathbf{1}-\mathbf{w}^2\ge\mathbf{0}, \end{array} \right\} \tag{4.8}$$

where $\mathbf{w}\coloneqq(w_1,\ldots,w_{p-1})$. Consider the following sparse homogenized reformulation for (1.1):

$$\begin{cases} \sup & \gamma\\ \text{s.t.} & \tilde{f}(\tilde{\mathbf{x}})-\gamma x_0^d\ge 0,\quad \forall(\tilde{\mathbf{x}},\mathbf{w})\in\widetilde{K}_s^3. \end{cases} \tag{4.9}$$

Similarly, one can verify that (4.9) has the same optimal value with (1.1). Furthermore, the asymptotic convergence of the corresponding sparse homogenized Moment-SOS hierarchy for (4.9) as well as related sparse Positivstellensätze can be established using similar arguments as in the previous subsections.

So far, we have discussed how to exploit correlative sparsity for homogenized polynomial optimization but do not touch term sparsity. Actually, correlative sparsity and term sparsity can be exploited simultaneously to gain further reductions on the size of SDP relaxations arising from Moment-SOS hierarchies. We refer the reader to [25, 48] for details.

5. Numerical examples. In this section, we present numerical results on solving POPs with three sparse homogenized Moment-SOS hierarchies. All numerical experiments are performed on a desktop computer with Intel(R) Core(TM) i9-10900 CPU@2.80GHz and 64G RAM. To model the homogenized hierarchies, we use the Julia package `TSSOS`[1] [24], relying on `Mosek` 10.0 [1] as an SDP backend with default settings. Unless otherwise specified, we set $\epsilon=10^{-4}$ for relaxations (3.3)–(3.4). We do not implement and compare with the approach proposed in [27] since it is limited to problems of modest size. Notations are listed in Table 1.

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 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 [39]. This example is a variation of Example 1 in [27]. As shown in [27],* f *is nonnegative and* $f\notin\Sigma[\mathbf{x}(1)]+\Sigma[\mathbf{x}(2)]$. *By solving* (3.3) *with* $\epsilon=0$ *and* $k=5$, *we obtain* $f_5\approx-1.6\times10^{-7}$, *which confirms* $f_5=f_{\min}=0$ *(up to numerical round-off errors).*

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

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 in Section 4.3
bold font	global optimality being certified
*	indicating unknown termination status
**	infeasible SDP
-	returning an out of memory error

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 csp $(\mathbf{x}(1), \mathbf{x}(2), \mathbf{x}(3))$*, where* $f = f_1 + f_2 + f_3$ $(x_0 \coloneqq 1)$

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

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

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	0.02	0.5497	0.03
3	0.5497	0.21	**0.6927**	13.3	**0.6927**	0.37	**0.6927**	0.15	**0.6927**	0.20
4	0.5864*	0.73	**0.6927**	683	**0.6927**	3.27	**0.6927**	1.38	**0.6927**	1.77

EXAMPLE 5.3. *Let*

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

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 can draw the following conclusions: (1) Without homogenization, the sparse hierarchy converges slowly; (2) By exploiting sparsity, we gain a significant speed-up and reach relaxations of higher orders; (3) HSSOS1 achieves the optimum at* $k = 4$ *and HSSOS3 achieves the optimum at* $k = 3$*, whereas HSSOS2 gives wrong answers due to numerical issues when* $k \geq 3$*.*

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	0.09	1.1804	0.11
3	1.1804	0.07	-	-	1.1895	0.34	1.1969*	1.18	**1.1900**	0.96
4	1.1809	0.40	-	-	**1.1900**	1.48	1.4871*	17.4	1.1901	5.94

5.2. Constrained polynomial optimziation.

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 csp $(\mathbf{x}(1), \mathbf{x}(2), \mathbf{x}(3))$*:*

$$\left\{ \begin{array}{ll} \inf & x_1^2 + 3x_2^2 - 2x_2 x_3^2 + x_3^4 - x_2(x_4^2 + x_5^2) \\ \text{s.t.} & x_1^2 - 2x_1 x_2 - 1 \geq 0,\, x_1^2 + 2x_1 x_2 - 1 \geq 0, \\ & x_2^2 - 1 \geq 0,\, x_2 - x_6^2 - x_7^2 \geq 0. \end{array} \right. \tag{5.1}$$

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 can draw the following conclusions: (1) Without homogenization, the sparse hierarchy converges slowly; (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 csp $(\mathbf{x}(1), \mathbf{x}(2))$*:*

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

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	0.05	2.0000	0.01	2.0048	0.01	2.0000	0.01	2.0000	0.01
3	2.0343*	0.05	5.1310*	0.06	2.8286	0.05	4.9449*	0.09	4.9678	0.09
4	2.1950*	0.07	**6.8284**	0.19	4.1178	0.19	**6.8284**	0.15	**6.8284**	0.18

where

$$\begin{aligned} f_1 &= x_1^4x_2^2 + x_2^4x_3^2 + x_3^4x_1^2 - 3(x_1x_2x_3)^2 + x_2^2 + x_7^2(x_1^2 + x_2^2 + x_3^2), \\ f_2 &= x_4^2x_5^2(10 - x_6^2) + x_7^2(x_4^2 + 2x_5^2 + 3x_6^2). \end{aligned}$$

For this problem, the optimal value is 0 *[38]. 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 looser bounds; (2) By exploiting sparsity, we gain some speed-up; (3) HSOS achieves the optimum at* $k = 4$*, and both HSSOS2 and HSSOS3 achieves the optimum at* $k = 5$*, while HSSOS1 converges to a near neighbourhood of* $f_{\min}$ *at* $k = 4$*.*

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	-1065*	0.24	-1106*	0.20
4	**	0.19	**-1.6e-8**	2.71	0.0001	0.82	-0.0002	1.37	-0.0002	1.77
5	-4.0e5	0.89	**-9.8e-9**	33.4	0.0001	5.33	**1.1e-7**	6.98	**1.4e-7**	6.38

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

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

For $i \in [p]$*, let*

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

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

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

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 looser bounds;*

(2) For $p = 2, 3$, HSOS achieves the optimum while for $p \geq 4$, HSOS runs out of memory; (3) By exploiting sparsity, we improve the scalability of the homogenization approach and still obtain good bounds.

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**	15.5	**6.1488**	12.8
3	-18999*	4.14	**9.2232**	2763	9.1475	20.1	9.2227	20.6	9.2228	31.3
4	-26984*	5.14	-	-	12.196	29.4	12.294	30.4	12.295	55.1
5	-31198*	6.54	-	-	15.246	39.0	15.365	39.4	15.364	69.4
10	-80847*	12.8	-	-	30.491	101	30.543	122	30.504	170

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 $f_i = \|A_i \mathbf{x}(i)^2\|^2 + b_i^\intercal \mathbf{x}(i)^2$. Consider the POP with csp $(\mathbf{x}(1), \mathbf{x}(2), \ldots, \mathbf{x}(p))$:

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

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 a significant speed-up; (3) Both HSOS and HSOSS2 do not scale well with the problem size (HSOS runs out of memory when $n \geq 40$ and HSOSS2 runs out of memory when $n \geq 100$); (4) Both HSSOS1 and HSOSS3 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**	1.21	**8.8328**	1.52
40	11.813	0.35	-	-	17.481	1.51	18.059	29.7	17.856	3.59
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

6. Applications to trajectory optimization. Trajectory optimization plays an essential role in the fields of robotics and control [3]. [41] demonstrates applying the sparse Moment-SOS hierarchy to specific trajectory optimization problems with compact feasible sets tends to yield tight solutions. However, assuming pre-defined 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 [2]. In such contexts, the ability to relax the compactness assumption while still achieving tight solutions is desirable. In this section, we explore two trajectory optimization problems with unbounded feasible sets: (1) block-moving with minimum work using direct collocation; (2) optimal control of Van der Pol oscillator with direct multiple shooting. We compare the performance of SSOS, HSSOS1, and HSSOS3 (noting that HSOS and HSSOS2 do not scale with large clique numbers).

6.1. Block-moving with minimum work. The continuous time version of block-moving with minimum work is shown as follows [14]:

$$\left\{\begin{array}{cl} \min\limits_{u(\tau),x_1(\tau),x_2(\tau)} & \int_{\tau=0}^{1} |u(\tau)x_2(\tau)|\,\mathrm{d}\tau \\ \text{s.t.} & \frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = f(x,u) = \begin{bmatrix} x_2 \\ u \end{bmatrix}, \\ & x_1(0)=0, x_2(0)=0, x_1(1)=1, x_2(1)=0, \end{array}\right. \tag{6.1}$$

where x_1 and x_2 are the block's position and velocity respectively. Starting from the origin in state space $\mathbf{x}(0) = [0,0]^\intercal$, 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:

$$\left\{\begin{array}{cl} \min\limits_{\substack{u_k,\ k=0,\ldots,N \\ x_{k,1},x_{k,2},\ k=0,\ldots,N \\ s_{k,1},s_{k,2},\ k=0,\ldots,N}} & \sum_{k=0}^{N} (s_{k,1}+s_{k,2})\cdot h \\ \text{s.t.} & s_{k,1} \geq 0, s_{k,2} \geq 0, \quad k=0,\ldots,N, \\ & s_{k,1} - s_{k,2} = u_k \cdot x_{k,2}, \quad k=0,\ldots,N, \\ & \dot{x}_{k,c} = f(x_{k,c},u_{k,c}), \quad k=0,\ldots,N-1, \\ & x_{0,1}=0, x_{0,2}=0, x_{N,1}=1, x_{N,2}=0, \end{array}\right. \tag{6.2}$$

where N is the total time steps and h is the time step. Since the terminal time is fixed as 1, $N \cdot h = 1$ should hold. Here $\dot{x}_{k,c}$, $x_{k,c}$, and $u_{k,c}$ in (6.2) stems from the collocation constraints.

It should be noted that (6.2) is a non-convex problem due to the inclusion of quadratic equality constraints. (6.2) exhibits a chain-like csp. Specifically, if we consider the kth clique as $(u_{k-1}, x_{k-1}, s_{k-1}, u_k, x_k, s_k), k = 1,\ldots,N$, the RIP is satisfied due to the Markov property. Setting relaxation order k to 2, we test the three algorithms' 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 sub-optimality gap η between SDP's solution and the solution refined by nonlinear programming solvers (here we use MATLAB's fmincon). Denote SDP's optimal value as f_{lower} and fmincon's 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 HSSOS3 achieves tight solutions in all parameter settings, with the sub-optimiality 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.

TABLE 8
Results of the block-moving example

N	$u_{\max}$	SSOS			HSSOS1			HSSOS3		
		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

6.2. Optimal control of Van der Pol. Now we consider the optimal control problem for a Van der Pol oscillator [6], a highly nonlinear and potentially unstable system. Its continuous time dynamics is

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

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

$$\begin{cases} \min\limits_{\substack{u_k,\ k=0,\ldots,N-1 \\ x_k,\ k=0,\ldots,N}} & \sum_{k=0}^{N-1}(u_k^2 + \|x_k\|^2)\cdot h + \|x_N\|^2 \cdot dt \\ \text{s.t.} & x_{k+1} = x_k + f(x_k, u_k)\cdot h, \quad k = 0,\ldots,N-1, \\ & u_{\max}^2 - u_k^2 \geq 0, \quad k = 0,\ldots,N-1, \\ & x_0 = x_{\text{init}}, \end{cases} \tag{6.5}$$

where N is 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 $\{(x_{k-1}, u_{k-1}, x_k)\}_{k=1}^N$.

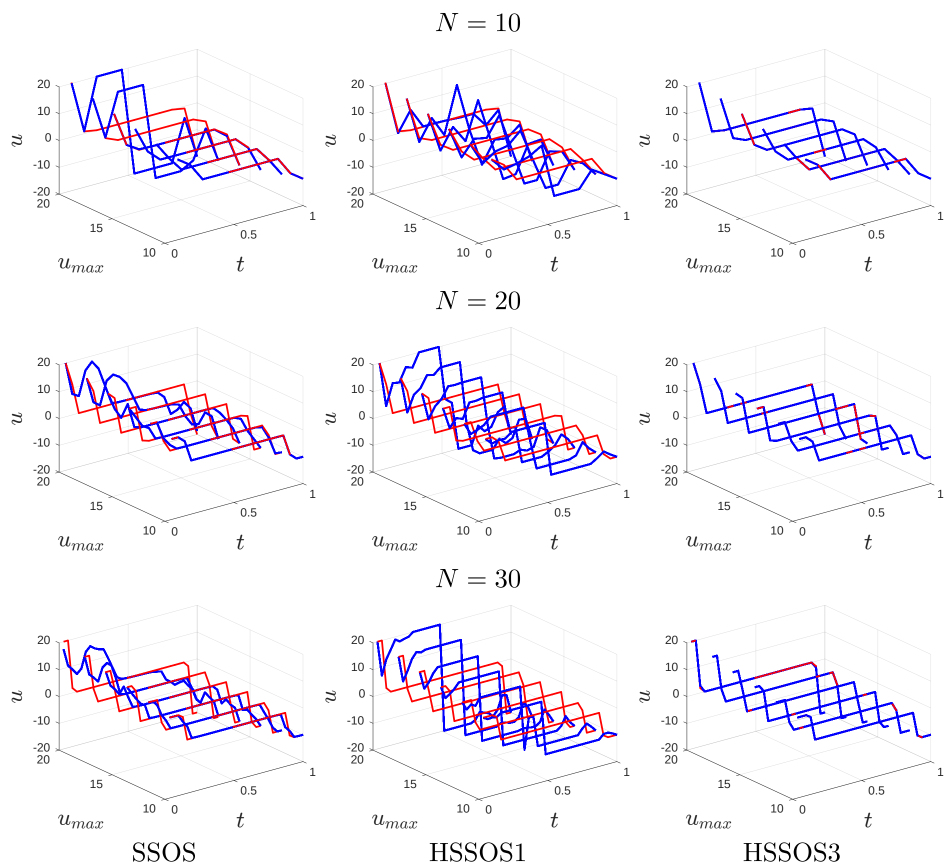

FIG. 1. *Comparison between SDP's solutions (blue lines) and solutions refined by fmincon (red lines) in the block-moving example. In HSSOS3, red lines and blue lines are nearly indistinguishable, indicating the attainment of tight solutions.*

However, (6.5) does not fulfill the Archimedeanness assumption since the variables $\{x_k\}_{k=1}^N$ are not subject to any bound. With the relaxation order $k = 2$, we incrementally vary N from 10 to 100 in steps of 10. At each N, the performance of three algorithms 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 HSSOS3 are better than those of SSOS and HSSOS1 in terms of achieving one or two order of magnitudes lower sub-optimality 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 algorithms, they all converge to identical refined solutions.

7. Conclusions and discussions. In this paper, we propose the sparse homogenized Moment-SOS hierarchies to solve sparse polynomial optimization with unbonunded sets. We have shown the asymptotic convergence under the RIP and extensive numerical experiments demonstrate the power of our approach in solving problems with up to thousands of variables. Furthermore, 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

Table 9
Results of the Van der Pol example

N	SSOS			HSSOS1			HSSOS3		
	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

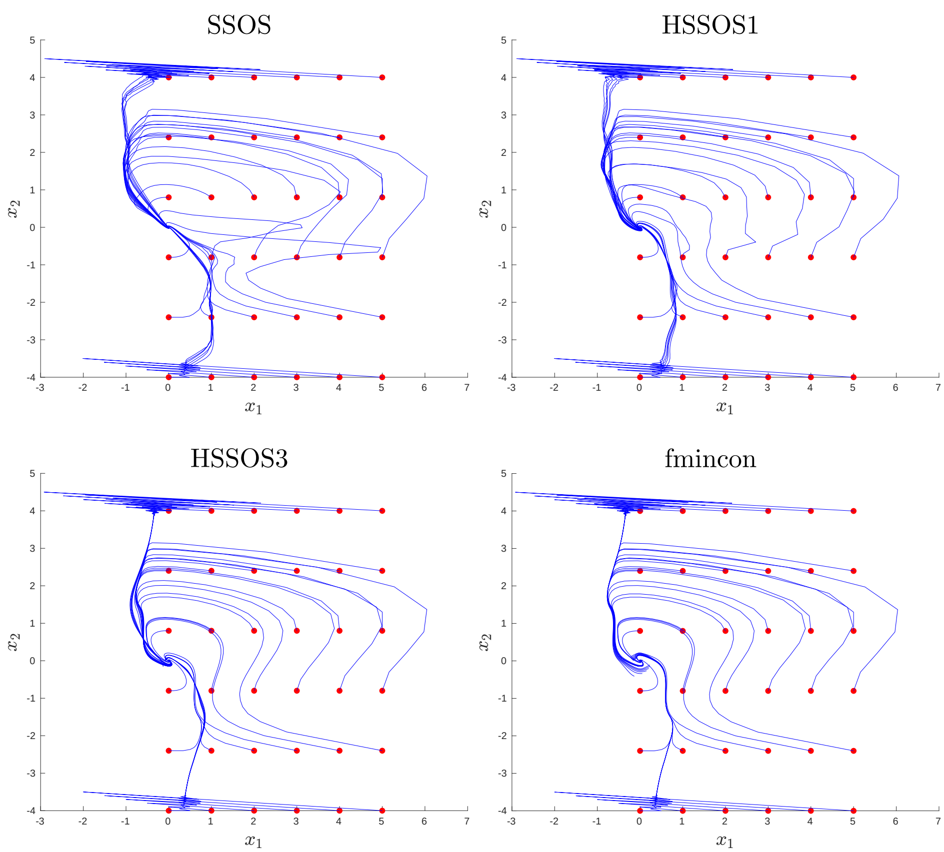

Fig. 2. *Comparison between SDP's solutions and solutions refined by fmincon in the Van der Pol example. Notably, all three algorithms' initial guesses lead to the same refined trajectories. Among these initial guesses, the one offered by HSSOS3 is of the best quality.*

hierarchy with correlative sparsity (2.7)–(2.8) are obtained in [17]. It is promising to get similar convergence rates for our sparse homogenized hierarchies with additional considerations on the behaviour of f at infinity of K. In Section 4, we propose two sparse homogenized Moment-SOS hierarchies without perturbations at the price of possibly increasing the maximal clique size. When tackling application problems, an interesting question is to explore how to construct correlative sparsity patterns with a small maximal clique size as the computational cost of sparse relaxations largely depends on this quantity.

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, H. Frenk, K. Roos, T. Terlaky, and S. Zhang, eds., vol. 33 of Applied Optimization, Springer US, 2000, pp. 197–232.

[2] A. Aydinoglu, P. Sieg, V. M. Preciado, and M. Posa, *Stabilization of complementarity systems via contact-aware controllers*, IEEE Transactions on Robotics, 38 (2021), pp. 1735–1754.

[3] J. T. Betts, *Survey of numerical methods for trajectory optimization*, Journal of guidance, control, and dynamics, 21 (1998), pp. 193–207.

[4] R. Curto and L. Fialkow, *Truncated K-moment problems in several variables*, Journal of Operator Theory, (2005), pp. 189–226.

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

[6] J. Guckenheimer, K. Hoffman, and W. Weckesser, *The forced van der pol equation i: The slow flow and its bifurcations*, SIAM Journal on Applied Dynamical Systems, 2 (2003), pp. 1–35.

[7] H. V. Ha and T. S. Pham, *Solving polynomial optimization problems via the truncated tangency variety and sums of squares*, Journal of Pure and Applied Algebra, 213 (2009), pp. 2167–2176.

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

[9] L. Huang, J. Nie, and Y.-X. Yuan, *Finite convergence of Moment-SOS relaxations with non-real radical ideals*, arXiv preprint arXiv:2309.15398, (2023).

[10] L. Huang, J. Nie, and Y.-X. Yuan, *Generalized truncated moment problems with unbounded sets*, Journal of Scientific Computing, 95 (2023), p. 15.

[11] L. Huang, J. Nie, and Y.-X. Yuan, *Homogenization for polynomial optimization with unbounded sets*, Mathematical Programming, 200 (2023), pp. 105–145.

[12] V. Jeyakumar, J. Lasserre, and G. Li, *On polynomial optimization over non-compact semialgebraic sets*, Journal of Optimization Theory and Applications, 163 (2014), pp. 707–718.

[13] C. Josz and D. K. Molzahn, *Lasserre hierarchy for large scale polynomial optimization in real and complex variables*, SIAM Journal on Optimization, 28 (2018), pp. 1017–1048.

[14] M. Kelly, *An introduction to trajectory optimization: How to do your own direct collocation*, SIAM Rev., 59 (2017), pp. 849–904, https://api.semanticscholar.org/CorpusID:33541404.

[15] I. Klep, V. Magron, and J. Povh, *Sparse Noncommutative Polynomial Optimization*, Mathematical Programming, (2021), pp. 1–41.

[16] M. Kojima and M. Muramatsu, *A note on sparse SOS and SDP relaxations for polynomial optimization problems over symmetric cones*, Computational Optimization and Applications, 42 (2009), pp. 31–41.

[17] M. Korda, V. Magron, and R. Zertuche, *Convergence rates for sums-of-squares hierarchies with correlative sparsity*, Mathematical Programming, (2024).

[18] J. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM Journal on Optimization, 11 (2001), pp. 796–817.

[19] J. Lasserre, *Convergent SDP-relaxations in polynomial optimization with sparsity*, SIAM Journal on optimization, 17 (2006), pp. 822–843.

[20] J. Lasserre, *An Introduction to Polynomial and Semi-algebraic Optimization*, vol. 52, Cambridge University Press, 2015.

[21] M. Laurent, *Semidefinite representations for finite varieties*, Mathematical programming, 109 (2007), pp. 1–26.

[22] M. Laurent, *Sums of squares, moment matrices and optimization over polynomials*, Emerging

applications of algebraic geometry, (2009), pp. 157–270.
[23] V. Magron, *Interval enclosures of upper bounds of roundoff errors using semidefinite programming*, ACM Transactions on Mathematical Software (TOMS), 44 (2018), pp. 1–18.
[24] V. Magron and J. Wang, *TSSOS: a Julia library to exploit sparsity for large-scale polynomial optimization*, 2021, https://arxiv.org/abs/2103.00915.
[25] V. Magron and J. Wang, *Sparse polynomial optimization: theory and practice*, World Scientific, 2023.
[26] N. Mai and V. Magron, *On the complexity of Putinar-Vasilescu's Positivstellensatz*, Journal of Complexity, (2022). Accepted for publication.
[27] N. Mai, V. Magron, and J. Lasserre, *A sparse version of Reznick's Positivstellensatz*, Mathematics of Operations Research, 48 (2023), pp. 812–833.
[28] N. H. A. Mai, J. Lasserre, and V. Magron, *Positivity certificates and polynomial optimization on non-compact semialgebraic sets*, Mathematical Programming, (2021), pp. 1–43.
[29] M. Marshall, *Positive Polynomials and Sums of Squares*, no. 146, American Mathematical Soc., 2008.
[30] M. Newton and A. Papachristodoulou, *Sparse polynomial optimisation for neural network verification*, Automatica, 157 (2023), p. 111233.
[31] J. Nie, *Optimality conditions and finite convergence of Lasserre's hierarchy*, Mathematical programming, 146 (2014), pp. 97–121.
[32] J. Nie, *Tight relaxations for polynomial optimization and Lagrange multiplier expressions*, Mathematical Programming, 178 (2019), pp. 1–37.
[33] J. Nie, *Moment and Polynomial Optimization*, 2023.
[34] T. Pham, *Tangencies and polynomial optimization*, Mathematical Programming, 199 (2023), pp. 1239–1272.
[35] M. Putinar, *Positive polynomials on compact semi-algebraic sets*, Indiana University Mathematics Journal, 42 (1993), pp. 969–984.
[36] M. Putinar and F. Vasilescu, *Positive polynomials on semi-algebraic sets*, Comptes Rendus de l'Académie des Sciences-Series I-Mathematics, 328 (1999), pp. 585–589.
[37] M. Putinar and F. Vasilescu, *Solving moment problems by dimensional extension*, Comptes Rendus de l'Académie des Sciences-Series I-Mathematics, 328 (1999), pp. 495–499.
[38] Z. Qu and X. Tang, *A correlative sparse lagrange multiplier expression relaxation for polynomial optimization*, arXiv preprint arXiv:2208.03979, (2022).
[39] B. Reznick, *Some concrete aspects of Hilbert's 17th problem*, Contemporary Mathematics, 253 (2000), pp. 251–272.
[40] M. Schweighofer, *Global optimization of polynomials using gradient tentacles and sums of squares*, SIAM Journal on Optimization, 17 (2006), pp. 920–942.
[41] S. Teng, A. Jasour, R. Vasudevan, and M. Ghaffari, *Convex geometric motion planning on lie groups via moment relaxation*, 2023, https://arxiv.org/abs/2305.13565.
[42] K. Toh, *Some numerical issues in the development of SDP algorithms*, Informs OS Today, 8 (2018), pp. 7–20.
[43] H. H. Vui and P. T. Son, *Global optimization of polynomials using the truncated tangency variety and sums of squares*, SIAM Journal on Optimization, 19 (2008), pp. 941–951.
[44] H. Waki, S. Kim, M. Kojima, and M. Muramatsu, *Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity*, SIAM Journal on Optimization, 17 (2006), pp. 218–242.
[45] J. Wang and V. Magron, *Exploiting term sparsity in noncommutative polynomial optimization*, Computational Optimization and Applications, 80 (2021), pp. 483–521.
[46] J. Wang, V. Magron, and J. Lasserre, *Chordal-TSSOS: a moment-SOS hierarchy that exploits term sparsity with chordal extension*, SIAM Journal on Optimization, 31 (2021), pp. 114–141.
[47] J. Wang, V. Magron, and J. Lasserre, *TSSOS: a moment-SOS hierarchy that exploits term sparsity*, SIAM Journal on optimization, 31 (2021), pp. 30–58.
[48] J. Wang, V. Magron, J. Lasserre, and N. Mai, *CS-TSSOS: Correlative and term sparsity for large-scale polynomial optimization*, ACM Trans. Math. Softw., 48 (2022), pp. 1–26.
[49] J. Wang, V. Magron, and J. B. Lasserre, *Certifying global optimality of AC-OPF solutions via sparse polynomial optimization*, Electric Power Systems Research, 213 (2022), p. 108683.
[50] J. Wang, C. Schlosser, M. Korda, and V. Magron, *Exploiting term sparsity in moment-sos hierarchy for dynamical systems*, IEEE Transactions on Automatic Control, (2023).