CrossMark

Semidefinite programming relaxation methods for global optimization problems with sparse polynomials and unbounded semialgebraic feasible sets

V. Jeyakumar[1] · S. Kim[2] · G. M. Lee[3] · G. Li[1]

Abstract We propose a hierarchy of semidefinite programming (SDP) relaxations for polynomial optimization with sparse patterns over unbounded feasible sets. The convergence of the proposed SDP hierarchy is established for a class of polynomial optimization problems. This is done by employing known sums-of-squares sparsity techniques of Kojima and Muramatsu Comput Optim Appl 42(1):31–41, (2009) and Lasserre SIAM J Optim 17:822–843, (2006) together with a representation theorem for polynomials over unbounded sets obtained recently in Jeyakumar et al. J Optim Theory Appl 163(3):707–718, (2014). We demonstrate that the proposed sparse SDP hierarchy can solve some classes of large scale polynomial optimization problems with unbounded feasible sets using the polynomial optimization solver SparsePOP developed by Waki et al. ACM Trans Math Softw 35:15 (2008).

Keywords Global continuous optimization · Sparse polynomial optimization · Structured sparsity · Sums of squares polynomials · Semidefinite programming relaxations

✉ V. Jeyakumar
v.jeyakumar@unsw.edu.au

S. Kim
skim@ewha.ac.kr

G. M. Lee
gmlee@pknu.ac.kr

G. Li
g.li@unsw.edu.au

[1] Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia

[2] Department of Mathematics, Ewha W. University, 52 Ewhayeodae-gil, Sudaemoon-gu, Seoul 120-750, Korea

[3] Department of Applied Mathematics, Pukyong National University, Busan 48513, Korea

1 Introduction

The optimal value of a polynomial optimization over a compact semialgebraic set can be approximated as closely as desired by solving a hierarchy of semidefinite programming (SDP) relaxations and the convergence is finite generically under a mild assumption called Archimedean condition which requires the compactness of the feasible region (see [19,20, 24,27,28,30,31]). It is known that the size of the SDP relaxations of the hierarchy, known now as the Lasserre hierarchy, rapidly grows as the number of variables and the relaxation order increase, preventing applications of the hierarchy to large scale polynomial optimization problems as the size of the SDP relaxations are too large to solve. A great deal of attention has recently been focused on reducing the size of these SDP relaxations. This has led to a sparse variant of the Lasserre hierarchy that allowed applications to various large scale polynomial optimization problems over compact semialgebraic feasible sets [5,8,16,18,36] such as the sensor network localization problems [17,25], hybrid system identification problems [7], optimal power flow problems [9], robust optimization problems [1], portfolio optimization problems [15] and filter design problems [22].

More recently, the standard Lasserre hierarchy of SDP relaxations has been shown to extend to polynomial optimization over unbounded semialgebraic feasible sets via suitable modifications [11]. The purpose of this paper is to present a convergent *sparse SDP hierarchy* for solving polynomial optimization problems with sparse patterns and unbounded semialgebraic feasible sets, extending the unbounded version of the Lasserre hierarchy [11]. More specifically, we make the following contributions to global polynomial optimization.

(1) We present a sparse SDP hierarchy for solving polynomial optimization problems with unbounded feasible sets incorporating the objective function in the construction of quadratic modules that generate the sequence of SDP relaxations. This approach extends the relaxation scheme, developed for convex polynomial optimization over noncompact sets [12]. This is done by employing known sums-of-squares sparsity techniques [18,21] together with a representation theorem for polynomials over unbounded sets, established recently in [11].

(2) By solving some numerical test problems, we illustrate that our sparse SDP hierarchy can easily be adapted with the current large scale polynomial optimization solver Sparse-POP [37] to solve some classes of large scale polynomial optimization problems with unbounded feasible sets. We also apply our SDP hierarchy to solve a class of sparse polynomial optimization problems with unbounded feasible sets and hidden coercivity.

The organization of the paper is as follows. In Sect. 2, we fix the notation and recall some basic facts on polynomial optimization. In Sect. 3, we present our sparse SDP hierarchy for polynomial optimization with unbounded feasible sets and establish its convergence. In Sect. 4, we illustrate how our proposed scheme works by solving various large scale numerical test problems.

2 Preliminaries

Throughout this paper, $\mathbb{R}^n$ denotes the Euclidean space with dimension n. The inner product in $\mathbb{R}^n$ is defined by $\langle x, y \rangle := x^T y$ for all $x, y \in \mathbb{R}^n$. The non-negative orthant of $\mathbb{R}^n$ is denoted by $\mathbb{R}^n_+$ and is defined by $\mathbb{R}^n_+ := \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_i \geq 0\}$. The closed ball with center x and radius r is denoted by $\overline{\mathbb{B}}(x, r)$. We use $\mathbf{e}_n$ to denotes the vector in $\mathbb{R}^n$ whose elements are all one. For a multi-index $\alpha = (\alpha_1, \ldots, \alpha_n)$, the monomial x^α is defined by $x^\alpha = x_1^{\alpha_1}, \ldots, x_n^{\alpha_n}$.

Denote by $\mathbb{R}[x]$ the ring of polynomials in $x := (x_1, x_2, \ldots, x_n)$ with real coefficients. The degree of a real polynomial f is denoted by $\deg f$. We say that a real polynomial $f \in \mathbb{R}[x]$ is a *sum of squares* (SOS) polynomial if there exist real polynomials f_j, $j = 1, \ldots, r$, such that $f = \sum_{j=1}^{r} f_j^2$. The set of all sum of squares real polynomials with variable x is denoted by $\Sigma^2[x]$. The set of all sum of squares real polynomial with variable x and degree at most d is denoted by $\Sigma_d^2[x]$. An important property of SOS polynomials is that checking a given polynomial with degree d is a sum of squares polynomial with degree d or not is equivalent to solving a semidefinite linear programming problem (see [10,13,14,20,24,28,32]).

The *quadratic module* $\mathbf{M}(g; h) = \mathbf{M}(-g_1, \ldots -g_m; h_1, \ldots, h_q) \subset \mathbb{R}[x]$ associated with the semi-algebraic set K, defined by

$$K := \{x : g_j(x) \leq 0, \quad j = 1, \ldots, m; \quad h_s(x) = 0, \ s = 1, \ldots, q\}, \qquad (2.1)$$

is given by

$$\mathbf{M}(g; h) := \left\{ \sigma_0 - \sum_{j=1}^{m} \sigma_j(x) g_j(x) + \sum_{s=1}^{q} \phi(x) h_s(x) : \right.$$
$$\left. \sigma_j \in \Sigma^2[x], \ j = 0, 1, \ldots, m; \quad \phi_s \in \mathbb{R}[x], \ s = 1, \ldots, q \right\}.$$

We note that any member of the quadratic module $\mathbf{M}(g; h)$ is a polynomial which is non-negative over the set K. The quadratic module $\mathbf{M}(g; h)$ is called *Archimedean* [20,24] if there exists a polynomial $\psi \in \mathbf{M}(g; h)$ such that $\{x : \psi(x) \geq 0\}$ is compact. A necessary condition for the quadratic module $\mathbf{M}(g; h)$ to be *Archimedean* is that the set K is a compact set.

Lemma 2.1 (Putinar positivstellensatz) [29] *Let $f, g_j, h_s, j = 1, \ldots, m, s = 1, \ldots, q$, be real polynomials with $K := \{x : g_j(x) \leq 0, j = 1, \ldots, m, h_s(x) = 0, s = 1, \ldots, q\} \neq \emptyset$. Let $\mathbf{M}(g; h)$ be Archimedean. If $f(x) > 0$ for all $x \in K$, then $f \in M(g; h)$.*

We now introduce a sparse version of Putinar positivstellensatz which was derived by Lasserre [21] and improved later on by Kojima and Muramatsu [18]. To do this, we need the definitions of the support of a polynomial and the running intersection property. Recall that, for a polynomial $f(x) = \sum_{\alpha} f_{\alpha} x^{\alpha}$ on $\mathbb{R}^n$ with degree d, the support of f is denoted by $\operatorname{supp} f$ and is defined by

$$\operatorname{supp} f = \left\{ \alpha \in (\mathbb{N} \cup \{0\})^d : f_{\alpha} \neq 0 \right\}.$$

Definition 2.1 (Running intersection property) Let I_l, $l = 1, \ldots, p$ be index sets such that $\bigcup_{l=1}^{p} I_l = \{1, \ldots, n\}$. We say that the *running intersection property* holds for I_l, $l = 1, \ldots, p$, if for each $l = 1, \ldots, p - 1$, there exists $s \leq l$ such that

$$I_{l+1} \cap \left(\bigcup_{j=1}^{l} I_j \right) \subseteq I_s.$$

We note that the running intersection property has its roots in graph theory, and can be satisfied for many polynomial optimization problems with structured sparsity patterns. In particular, It is known that maximal cliques of a chordal graph (after possible reordering) satisfy the running intersection property [35, Theorem 3.5]. For a detailed discussion, see [18].

Lemma 2.2 (Sparse version of Putinar positivstellensatz [18,21]) *Let* $f(x) = \sum_{l=1}^{p} f^l(x)$, $g_j(x) = \sum_{\alpha} (g_j)_{\alpha} x^{\alpha}$, $j = 1, \ldots, m$, *and* $h_s(x) = \sum_{\alpha} (h_s)_{\alpha} x^{\alpha}$, $s = 1, \ldots, q$, *be polynomials on* $\mathbb{R}^n$ *with degree* d. *Let* I_l *be a set of indexes such that* $\mathrm{supp} f^l \subseteq I_l \subseteq \{1, \ldots, n\}$, $l = 1, \ldots, p$ *and* $\bigcup_{l=1}^{p} I_l = \{1, \ldots, n\}$. *Suppose that for each* $j = 1, \ldots, m$ *and* $s = 1, \ldots, q$, $\mathrm{supp} g_j \cup \mathrm{supp} h_s \subseteq I_l$ *for some* $l \in \{1, \ldots, p\}$, *and the running intersection property holds for* I_l, $l = 1, \ldots, p$. *Let* $K := \{x : g_j(x) \le 0, j = 1, \ldots, m, h_s(x) = 0, s = 1, \ldots, q\} \ne \emptyset$ *and let* $\mathbf{M}(g; h)$ *be Archimedean. If* $f(x) > 0$ *on* K, *then*

$$f = \sum_{l=1}^{p} \left(\sigma_{0l} - \sum_{j=1}^{m} \sigma_{jl} g_j + \sum_{s=1}^{q} \phi_{sl} h_s \right)$$

where σ_{jl}, $j = 0, 1, \ldots, m$, *are SOS polynomials with variables* $\{x_i : i \in I_l\}$, *and* ϕ_{sl}, $s = 1, \ldots, q$, *are real polynomials with variables* $\{x_i : i \in I_l\}$.

We note that, the assumption "for each $j = 1, \ldots, m$ and $s = 1, \ldots, q$, $\mathrm{supp} g_j \cup \mathrm{supp} h_s \subseteq I_l$ for some $l \in \{1, \ldots, p\}$" and the running intersection property are automatically satisfied in the special case where $p = 1$ and $I_1 = \{1, \ldots, n\}$. So, in this case, the sparse version of Putinar positivstellensatz reduces to Putinar positivstellensatz.

2.1 Coercivity of polynomials

We now recall the definitions for coercive polynomials. This definition plays an important role in our analysis for positivity over noncompact sets later on.

Definition 2.2 (Coercivity and strong coercivity) Let $f(x) = \sum_{\alpha} f_{\alpha} x^{\alpha}$ be a polynomial on $\mathbb{R}^n$ with degree d. Let $f(x) = \sum_{i=0}^{d} f_i(x)$ where each f_i is a homogeneous polynomial of degree i, $i = 0, 1, \ldots, d$. We say the polynomial f is

- **coercive** if $f(x) \to +\infty$ whenever $\|x\| \to +\infty$;
- **s-strongly coercive** for some $s \in \{1, \ldots, d\}$ if $f_s(x) > 0$ for all $x \ne 0$ and $f_i(x) \ge 0$ for all $s + 1 \le i \le d$ and $x \in \mathbb{R}^n$;
- **strongly coercive** if f is d-strongly coercive.

It follows from the definition that an s-strongly coercive polynomial, $s = 1, \ldots, d$, must be coercive. On the other hand, the converse is not true. As an example, the 2-dimensional Rosenbrock function

$$f(x_1, x_2) = 1 + \left(x_2 - x_1^2 \right)^2 + (1 - x_2)^2$$

is a coercive polynomial which is not s-strongly coercive for $s = 1, 2, 3, 4$. We also note that it was shown in [11] that the strong coercivity can be numerically checked by solving semidefinite programming problems. Furthermore, any polynomial of the form $\sum_{i=1}^{n} a_i x_i^d + \sum_{|\alpha| \le d-1} h_{\alpha} x^{\alpha}$ where $|\alpha| = \sum_{i=1}^{n} \alpha_i$ with $\alpha = (\alpha_1, \ldots, \alpha_n) \in (\mathbb{N} \cup \{0\})^n$, $d \in 2\mathbb{N}$ and $a_i > 0$, $i = 1, \ldots, n$, is a strongly coercive polynomial.

The following proposition shows that a coercive polynomial is always level-bounded. Moreover, the corresponding bound for the lower level set can be computed using the coefficients of the underlying polynomial if the polynomial is assumed to be s-strongly coercive for some $s = 1, \ldots, d$.

Proposition 2.1 (Boundness of the lower level set via coercivity) *Let* $f(x) = \sum_{\alpha} f_{\alpha} x^{\alpha}$ *be a coercive polynomial on* $\mathbb{R}^n$ *with degree* d. *Let* $f(x) = \sum_{i=0}^{d} f_i(x)$ *where each* f_i *is a homogeneous polynomial with degree* i, $i = 0, 1, \ldots, d$. *Then, for each* $c \in \mathbb{R}$, *the lower*

level set $\{x : f(x) \leq c\}$ is a compact set. Furthermore, if f is s-strongly coercive for some $s = 1, \ldots, d$, then $\{x : f(x) \leq c\} \subseteq \bar{\mathbb{B}}(0, r)$ where

$$r = \max\left\{1, \frac{c + \sum_{0 \leq |\alpha| \leq s-1} |f_\alpha|}{\rho_s}\right\} \quad \text{and} \quad \rho_s = \min\{f_s(x) : \|x\| = 1\}.$$

Proof Fix any $c \in \mathbb{R}$. Then, the lower level set $\{x : f(x) \leq c\}$ is a compact set. To see this, we suppose on the contrary that there exists $\{x_n\}_{n=1}^\infty$, with $f(x_n) \leq c$ and $\{x_n\}$ is unbounded. By passing to subsequence if necessary, we may assume that $\|x_n\| \to +\infty$. As f is a coercive polynomial, we must have $f(x_n) \to +\infty$. This contradicts the fact that $f(x_n) \leq c$ for all $n \in \mathbb{N}$.

To see the second assertion, we assume that f is s-strongly coercive for some $s = 1, \ldots, d$. Let $a \in \mathbb{R}^n$ be any point such that $f(a) \leq c$. Then,

$$\rho_s \|a\|^s \leq f_s(a) \leq \sum_{j=s}^d f_j(a) = f(a) - \sum_{j=0}^{s-1} f_j(a) \leq c - \sum_{j=0}^{s-1} f_j(a) \leq c + \sum_{0 \leq |\alpha| \leq s-1} |f_\alpha| \, |a^\alpha|.$$

This gives us that either $\|a\| \leq 1$ or

$$\rho_s \|a\|^s \leq c + \sum_{0 \leq |\alpha| \leq s-1} |f_\alpha| \, |a^\alpha| \leq c + \sum_{0 \leq |\alpha| \leq s-1} |f_\alpha| \|a\|^{|\alpha|}$$

$$\leq \left(c + \sum_{0 \leq |\alpha| \leq s-1} |f_\alpha| \right) \|a\|^{s-1},$$

where the second inequality follows from the fact that $|x^\alpha| \leq \|x\|^{|\alpha|}$ for all $\|x\| \geq 1$ and the last inequality is from the fact that $\|x\|^{|\alpha|} \leq \|x\|^{s-1}$ for all $|\alpha| \leq s - 1$ and $\|x\| \geq 1$. So, we have

$$\|a\| \leq \max\left\{1, \frac{c + \sum_{0 \leq |\alpha| \leq s-1} |f_\alpha|}{\rho_s}\right\},$$

and hence, the conclusion follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Corollary 2.1 Let $f(x) = \sum_\alpha f_\alpha x^\alpha$ and $g_j(x) = \sum_\alpha (g_j)_\alpha x^\alpha$, $j = 1, \ldots, m$, be polynomials on $\mathbb{R}^n$ with degree d.

(i) If there exist $\mu_j \geq 0$, $j = 0, 1, \ldots, m$, such that $\mu_0 f + \sum_{j=1}^m \mu_j g_j$ is coercive, then, for each $c \in \mathbb{R}$, the set $\{x : g_j(x) \leq 0, j = 1, \ldots, m, f(x) \leq c\}$ is a compact set.

(ii) If $\mu_0 f + \sum_{j=1}^m \mu_j g_j$ is s-strongly coercive for some $s \in \{1, \ldots, d\}$, then

$$\{x : g_j(x) \leq 0, \quad j = 1, \ldots, m, f(x) \leq c\} \subseteq \bar{\mathbb{B}}((0, r),$$

where

$$r = \max\left\{1, \frac{\mu_0 c + \sum_{0 \leq |\alpha| \leq s-1} |\mu_0 f_\alpha + \sum_{j=1}^m \mu_j (g_j)_\alpha|}{\rho_s}\right\}$$

and

$$\rho_s = \min\left\{\left(\mu_0 f + \sum_{j=1}^m \mu_j g_j\right)_s (v) : \|v\| = 1\right\}.$$

Proof Note that $\{x : g_j(x) \leq 0, j = 1, \ldots, m, f(x) \leq c\} \subseteq \{x : \mu_0 f(x) + \sum_{j=1}^m \mu_j g_i(x) \leq \mu_0 c\}$. The conclusion follows by applying the Proposition 2.1 with f replaced by $\mu_0 f + \sum_{j=1}^m \mu_j g_i$.   □

3 A sparse hierarchy for optimization over unbounded sets

Consider the polynomial optimization problem

$$(P) \min \sum_{l=1}^p f^l(x)$$
$$\text{s.t. } g_j(x) \leq 0, \quad j = 1, \ldots, m,$$
$$h_s(x) = 0, \quad s = 1, \ldots, q,$$

where f^l, g_j, h_s are (nonconvex) polynomials on $\mathbb{R}^n$, $l = 1, \ldots, p$, $j = 1, \ldots, m$ and $s = 1, \ldots, q$. Let the feasible set be denoted by K, that is, $K = \{x : g_j(x) \leq 0, j = 1, \ldots, m, h_s(x) = 0, s = 1, \ldots, q\}$.

Let $f(x) = \sum_{l=1}^p f^l(x)$, $g_j(x) = \sum_\alpha (g_j)_\alpha x^\alpha$, $j = 1, \ldots, m$ and $h_s(x) = \sum_\alpha (h_s)_\alpha x^\alpha$, $s = 1, \ldots, q$, be polynomials on $\mathbb{R}^n$ with degree d. For simplicity, here we assume that the objective polynomial functions and the constraint polynomial functions have the same degree because one can let d be the maximum of the degree of f, g_j, $j = 1, \ldots, m$ and h_s, $s = 1, \ldots, q$, and regard f, g_j and h_s as polynomials with degree d.

Let I_l be a set of indices such that $\text{supp} f^l \subseteq I_l \subseteq \{1, \ldots, n\}$, $l = 1, \ldots, p$ and $\bigcup_{l=1}^p I_l = \{1, \ldots, n\}$. Let $x^0 \in \{x : g_j(x) \leq 0, j = 1, \ldots, m\}$ and let c be a number such that $c > \max_{1 \leq l \leq p} \{f^l(x^0)\}$. For each integer k, we define the truncated sparse version of the quadratic module $\bar{\mathbf{M}}_k$ by

$$\bar{\mathbf{M}}_k := \left\{ \sum_{l=1}^p \left(\sigma_{0l} - \sum_{j=1}^m \sigma_{jl} g_j + \bar{\sigma}_l (c - f^l) + \sum_{s=1}^q \phi_{sl} h_s \right) \mid \sigma_{0l}, \sigma_{jl}, \bar{\sigma}_l \in \Sigma^2[\underline{x_l}], \phi_s \in \mathbb{R}[\underline{x_l}] \right.$$

$$\left. \deg \sigma_{0l} \leq 2k, \deg \sigma_{jl} g_j \leq 2k, \deg \bar{\sigma}_l(c - f^l) \leq 2k, \deg \phi_{sl} h_s \leq 2k \right\},$$

where for each $l = 1, \ldots, p$, $\Sigma^2[\underline{x_l}]$ denotes the set of all SOS polynomials with variable $\{x_i : i \in I_l\}$ and $\mathbb{R}[\underline{x_l}]$, denotes the set of all real polynomials with variable $\{x_i : i \in I_l\}$.

Consider the following relaxation problem

$$\bar{f}_k^* := \sup \{ \mu \in \mathbb{R} \mid f - \mu \in \bar{\mathbf{M}}_k \}. \tag{3.1}$$

By construction, $\bar{f}_k^* \leq \bar{f}_{k+1}^* \leq \cdots \leq \min(P)$. Note that, if we set $\bar{\sigma}_l \equiv 0$, $l = 1, \ldots, p$, then the hierarchy (3.1) reduces to the known sparse SDP hierarchy proposed in [17,18].

Theorem 3.1 (Convergence of sparse SDP hierarchy) *Let* $f(x) = \sum_{l=1}^p f^l(x)$, $g_j(x) = \sum_\alpha (g_j)_\alpha x^\alpha$, $j = 1, \ldots, m$ *and* $h_s(x) = \sum_\alpha (h_s)_\alpha x^\alpha$, $s = 1, \ldots, q$, *be polynomials on* $\mathbb{R}^n$ *with degree* d. *Let* I_l *be a set of indexes such that* $\text{supp} f^l \subseteq I_l \subseteq \{1, \ldots, n\}$, $l = 1, \ldots, p$ *and* $\bigcup_{l=1}^p I_l = \{1, \ldots, n\}$. *Consider the SDP hierarchy (3.1) and denote its optimal value by* $\bar{f}_k^*$. *Let* $K = \{x : g_j(x) \leq 0, j = 1, \ldots, m, h_s(x) = 0, s = 1, \ldots, q\}$, $x^0 \in K$ *and* $c > \max_{1 \leq l \leq p} \{f^l(x^0)\}$.

Suppose that for each $j = 1, \ldots, m$ *and* $s = 1, \ldots, q$, $\text{supp} g_j \cup \text{supp} h_s \subseteq I_l$ *for some* $l \in \{1, \ldots, p\}$, *and the running intersection property holds for* I_l, $l = 1, \ldots, p$. *If, for each* $l =$

$1, \ldots, p$, there exist $\mu_{0l} \geq 0, l = 1, \ldots, p, \mu_j \geq 0, j = 1, \ldots, m$ and $\tau_s \in \mathbb{R}, s = 1, \ldots, q$, such that $\sum_{l=1}^{p} \mu_{0l} f^l + \sum_{j=1}^{m} \mu_j g_j + \sum_{s=1}^{q} \tau_s h_s$ is coercive, then $\lim_{k \to \infty} \bar{f}_k^* = \min(P)$.

Proof Let $\epsilon > 0$. Then, we have $f - \min(P) + \epsilon > 0$ over the feasible set K. As there exist $\mu_{0l} \geq 0, l = 1, \ldots, p, \mu_j \geq 0, j = 1, \ldots, m$ and $\tau_s \in \mathbb{R}$, $s = 1, \ldots, q$, such that $\sum_{l=1}^{p} \mu_{0l} f^l + \sum_{j=1}^{m} \mu_j g_j + \sum_{s=1}^{q} \tau_s h_s$ is coercive, we see that $\sum_{l=1}^{p} \mu_{0l}(f^l - c) + \sum_{j=1}^{m} \mu_j g_j + \sum_{s=1}^{q} \tau_s h_s$ is also coercive. It follows from Proposition 2.1 that $\{x : \sum_{l=1}^{p} \mu_{0l}(c - f^l(x)) + \sum_{j=1}^{m} \mu_j(-g_j)(x) + \sum_{s=1}^{q}(-\tau_s)h_s \geq 0\}$ is compact. So, by definition, $\mathbf{M}(-g_1, \ldots, -g_m, c - f^1, \ldots, c - f^p; h_1, \ldots, h_q)$ is Archimedean. Note that $f - \min(P) + \epsilon > 0$ over $\hat{K}$ where

$$\hat{K} = \{x : g_j(x) \leq 0, j = 1, \ldots, m, f^l(x) - c \leq 0, l = 1, \ldots, p, h_s(x) = 0, s = 1, \ldots, q\}.$$

Hence, by Lemma 2.2, we obtain that

$$f - \min(P) + \epsilon = \sum_{l=1}^{p} \left(\sigma_{0l} - \sum_{j=1}^{m} \sigma_{jl} g_j + \bar{\sigma}_l(c - f^l) + \sum_{s=1}^{q} \phi_{sl} h_s \right),$$

for some sum-of squares polynomials $\sigma_{0l}, \ldots, \sigma_{ml}, \bar{\sigma}_l$ with variables $\{x_i : i \in I_l\}$ and real polynomials ϕ_{sl} with variables $\{x_i : i \in I_l\}, l = 1, \ldots, p$. Thus, for each $\epsilon > 0$, there exists $k_0 \in \mathbb{N}$ such that $\bar{f}_{k_0}^* \geq \min(P) - \epsilon$. On the other hand, from the construction of the hierarchy, we see that $\bar{f}_k^* \leq \bar{f}_{k+1}^* \leq \cdots \leq \min(P)$. Therefore, the conclusion follows. □

Remark 3.1 It is worth noting that, if a polynomial f on $\mathbb{R}^n$ is convex in the sense that $\nabla^2 f(x)$ is positive semi-definite for all $x \in \mathbb{R}^n$ and there exists $x^* \in \mathbb{R}^n$ such that $\nabla^2 f(x^*)$ is positive definite, then it is coercive (for example see [12]). Therefore, our coercivity assumption of Theorem 3.1 that, "there exist $\mu_{0l} \geq 0, l = 1, \ldots, p \mu_j \geq 0, j = 1, \ldots, m$ and $\tau_s \in \mathbb{R}$, $s = 1, \ldots, q$ such that $\sum_{l=1}^{p} \mu_{0l} f^l + \sum_{j=1}^{m} \mu_j g_j + \sum_{s=1}^{q} \tau_q h_q$ is coercive" is satisfied, if one of the polynomials $f^l, g_j, h_s, l = 1, \ldots, p, j = 1, \ldots, m, s = 1, \ldots, q$, is convex and has a positive definite Hessian at some point $x^* \in \mathbb{R}^n$.

In the special case where $p = 1, I_1 = \{1, \ldots, n\}$ and each $h_s \equiv 0, s = 1, \ldots, q$, the optimization problem (P) reduces to

$$\min_{x \in \mathbb{R}^n} f(x)$$
$$\text{s.t.} \quad g_j(x) \leq 0, \quad j = 1, \ldots, m.$$

In this case, the sparse truncated quadratic module $\bar{\mathbf{M}}_k$ reduces to the truncated quadratic module $\mathbf{M}_k$ generated by the polynomials $c - f$ and $-g_1, \ldots, -g_m$ given by

$$\mathbf{M}_k := \left\{ \sigma_0 + \sum_{j=1}^{m} \sigma_j g_j + \bar{\sigma}(c - f) \mid \sigma_0, \sigma, \bar{\sigma} \in \Sigma^2[\underline{x}] \subset \mathbb{R}[\underline{x}], \right.$$

$$\left. \deg \sigma_0 \leq 2k, \ \deg \sigma_j g_j \leq 2k, \ \text{and} \ \deg \bar{\sigma}(c - f) \leq 2k \right\},$$

and the corresponding relaxation problem collapses to

$$f_k^* := \sup \{ \mu \in \mathbb{R} \mid f - \mu \in \mathbf{M}_k \}. \tag{3.2}$$

So, Theorem 3.1 collapses to the convergence result of the dense SDP hierarchy for polynomial optimization problem with noncompact sets proposed in [11].

Remark 3.2 (**Locating a global minimizer**) It is worth noting that, in addition to the assumptions of Theorem 3.1, if we further assume that problem (P) has a unique solution $\bar{x}$, then we can also find the global minimizer $\bar{x}$ by solving a sequence of semidefinite programming problems which are the dual problems of the equivalent SDP reformulation of (3.2). The details are omitted here and they can be found in [20,21].

4 Numerical experiments

In this Section, we show the effectiveness of the proposed sparse SDP hierarchy in Sect. 4 by solving some numerical test problems with unbounded feasible sets. All the numerical tests were conducted on a computer, with a 2.8 GHz Intel Core i7 and 8 GB RAM, equipped with Matlab 7.14 (R2012a).

The purpose of the numerical experiments is to illustrate how our proposed sparse SDP hierarchy works for solving polynomial optimization problems with unbounded feasible sets. Therefore, we first select some known test problems which are nonconvex coercive polynomials and test them by minimizing them over unbounded feasible sets. The selected test problems are:

(1) a nonconvex quadratic programming problem with an unbounded feasible set, in which Lasserre's hierarchy is known to fail;
(2) the Rosenbrock function over the nonnegative orthant;
(3) the Chain-wood function over the nonnegative orthant;

For the numerical tests, we first form the proposed sparse SDP hierarchy (3.1) for the above test problems. Then, we use a Matlab software SparsePOP [37] to solve the corresponding sparse SDP problems. Let $f(x) = \sum_{l=1}^{q} f^l(x)$. We note that the proposed sparse SDP hierarchy (3.1) can be obtained from the existing sparse SDP hierarchy used in [18,36,37] designed for problems with compact feasible sets by adding additional constraints $f^l(x) \le c$, $l = 1, \ldots, q$, to the test problems with unbounded feasible sets, where c is appropriately chosen. SparsePOP can solve polynomial optimization problems exploiting the sparsity described in Sect. 4 by setting the parameter *sparseSW = 1*, and can also implement Lasserre's hierarchy by setting the parameter *sparseSW = 0*. In addition, a parameter, called the relaxation order, can be chosen in SparsePOP, depending on the degree of the polynomial optimization problem. The larger value for the relaxation order is used, the better approximation to the optimal value of the polynomial optimization problem can be expected.

In SparsePOP, the accuracy of an obtained objective value is computed by

$$\text{Relative objective error (Rel.Er)} = \frac{\text{POP.objValue} - \text{SDP.objValue}}{\max\{1, |\text{POP.objValue}|\}},$$

where POP.objValue means the value of the objective function of a polynomial optimization problem computed using a candidate for the optimal solution, and SDP.objValue the value of the objective function of the SDP relaxation of the polynomial optimization problem. Moreover, CPU time reported in the subsequent discussion is measured in seconds. For details, we refer to [37].

Table 1 Numerical test on problem (QP2)

	RelaxOrder	c	Optimal val.	Optimal sol.	Rel.Er	CPU
DenseHierarchy (3.2)	2	40	1.1729	–	3.9e+0	0.27
DenseHierarchy (3.2)	3	40	27.9626	(5.1924,1.000)	5.6e−5	0.31

4.1 A 2-dimensional QP with an unbounded feasible set

Consider the following nonconvex quadratic optimization problem:

$$
\begin{aligned}
(QP) \min_{(x_1,x_2)\in\mathbb{R}^2} \quad & x_1^2 + x_2^2 \\
\text{s.t.} \quad & x_2^2 - 1 \geq 0 \\
& x_1^2 - Mx_1x_2 - 1 \geq 0 \\
& x_1^2 + Mx_1x_2 - 1 \geq 0.
\end{aligned}
\tag{4.1}
$$

It was shown in [23,26] that the global minimizers are $(\pm\frac{M+\sqrt{M^2+4}}{2}, \pm 1)$ and its global minimum is $1 + \frac{\sqrt{2M^2+4+2M\sqrt{M^2+4}}}{4}$. For instance, if $M = 5$, then, the global minimizers are $(\pm 5.1926, \pm 1)$ and its global minimum is 27.9629.

Let $M = 5$. It was shown in [26] that Lasserre's hierarchy only provides a lower bound 2, no matter how large the relaxation order k is chosen. As explained in [26], the main reason why the Lasserre's hierarchy fails to achieve the global minimum of this quadratic problem is that the feasible set of this nonconvex quadratic problem is unbounded.

Notice that the objective $f(x_1, x_2) := x_1^2 + x_2^2$ is coercive. Thus, we can apply the proposed sparse hierarchy. Since the problem has only two variables without sparsity, we test the problem using our dense SDP hierarchy (3.2) with $c = 40$. Note that $c > f(6, 1)$ and the point $(6, 1)$ is feasible for this quadratic optimization problem.

As shown in Table 1, by running the sparsePOP with sparseSW = 0 and the relaxation order 4, the relaxation problem in the dense hierarchy (3.2) solves the original problem and returns a good approximation to the true global minimizer (5.1926, 1.000). Table 1 summarizes the optimal values for the dense SDP hierarchy (3.1) with relaxation order $k = 2, 3, 4$, which illustrates the effectiveness of our approach.

4.2 The Rosenbrock function over nonnegative orthant

The Rosenbrock function is described as

$$
f_R(x_1, \ldots, x_n) = 1 + \sum_{i=2}^{n} \left((x_i - x_{i-1}^2)^2 + (1 - x_i)^2 \right), \quad n \geq 2.
$$

Clearly, f_R is a SOS nonconvex polynomial, and is coercive. We add constraints to the Rosenbrock function to have a polynomial optimization problem with an unbounded region as follows:

$$
\begin{aligned}
(EP_R) \min_{x\in\mathbb{R}^n} \quad & f_R(x) \\
\text{s.t.} \quad & x_i \geq 0, \ i = 1, \ldots, n.
\end{aligned}
$$

It can be easily verified that this problem has a unique global minimizer $\mathbf{e}_n := \underbrace{(1, \ldots, 1)}_{n}$.

Table 2 Numerical tests on the problem (EP_R)

	n	RelaxOrder	c	Optimal sol.	Rel.Er	CPU
SparseHierarchy (3.1)	500	2	2	$\mathbf{e}_{500}$	4.1e−5	81.8
SparseHierarchy (3.1)	5000	2	2	$\mathbf{e}_{5000}$	1.0e−3	148.5
SparseHierarchy (3.1)	10,000	2	2	$\mathbf{e}_{10,000}$	2.1e−3	308.5
SparseHierarchy (3.1)	20,000	2	2	$\mathbf{e}_{20,000}$	4.3e−3	772.2

Let $I_l = \{l, l+1\}, l = 1, \ldots, n-1$ and $g_j(x) = x_j, j = 1, \ldots, n$. Then, the assumptions in Theorem 3.1 are satisfied as

$$\mathrm{supp} g_j = \{j\} \subseteq \begin{cases} I_j, & \text{if } j = 1, \ldots, n-1, \\ I_{n-1}, & \text{if } j = n, \end{cases}$$

and for each $l = 1, \ldots, n-1$,

$$I_{l+1} \cap \left(\bigcup_{j=1}^{l} I_j \right) = \{l\} \subseteq I_l.$$

Therefore, according to Theorem 3.1, the optimal value of the proposed sparse SDP hierarchy (3.1) converges to the optimal value of the global minimum of (EP_R).

We now test our proposed sparse SDP hierarchy (3.1) using the global optimization problem (EP_R) for different dimension n with $c = 2$. From Table 2, we see that for $n = 10,000$ or $n = 20,000$, the sparsePOP returns an accurate solution for the relaxation order 2 in the proposed sparse hierarchy Sparse SDP hierarchy (3.1).

When the dimension n of the problem is large, directly applying the dense SDP hierarchy proposed in [11] without exploiting the sparsity leads to very large SDP problems which cannot be handled by a SDP solver such as SeDuMi [34]. Indeed, we confirm in our numerical computation that the dense SDP hierarchy proposed in [11] can only be used to solve (EP_R) up to dimension 20. The larger problems than dimension 20 resulted in out-of-memory error.

4.3 Chained-wood function over nonnegative orthant

Let $n \in 4\mathbb{N}$ where $\mathbb{N}$ denotes the set of integers. Consider the Chained-wood function given by

$$f_C(x_1, \ldots, x_n) = 1 + \sum_{i \in J} \left((x_{i+1} - x_i^2)^2 + (1 - x_i)^2 + 90(x_{i+3}^2 - x_{i+2})^2 + (x_{i+2} - 1)^2 \right.$$
$$\left. + 10(x_{i+1} + x_{i+3} - 2)^2 + \frac{1}{10}(x_{i+1} - x_{i+3})^2 \right),$$

where $J = \{1, 3, 5, \ldots, n-3\}$. The function f_C is a SOS nonconvex polynomial, and is coercive. Adding nonnegative constraints for the variables $x_i, i = 1, \ldots, n$ results in the following polynomial optimization problem:

$$(EP_C) \min_{x \in \mathbb{R}^n} f_C(x)$$
$$\text{s.t.} \quad x_i \geq 0, \ i = 1, \ldots, n.$$

Table 3 Numerical tests on the problem (EP_C)

	n	RelaxOrder	c	Optimal sol.	Rel.Er	CPU
SparseHierarchy (3.1)	1000	2	1001	$(\mathbf{e}_{1000})$	2.4e−4	38.9
SparseHierarchy (3.1)	5000	2	5001	$(\mathbf{e}_{5000})$	1.2e−3	175.2
SparseHierarchy (3.1)	10,000	2	10,001	$(\mathbf{e}_{10,000})$	2.4e−3	392.3
SparseHierarchy (3.1)	20,000	2	20,001	$(\mathbf{e}_{20,000})$	4.8e−3	1049.8

Clearly, the feasible set of this polynomial optimization is unbounded, and this problem has a unique global minimizer $\mathbf{e}_n := \underbrace{(1, \ldots, 1)}_{n}$.

Note that f_C can be equivalently rewritten as

$$f_C(x_1, \ldots, x_n) = 1 + \sum_{l=1}^{\frac{n}{2}-1} \left((x_{2l} - x_{2l-1}^2)^2 + (1 - x_{2l-1})^2 + 90(x_{2l+2}^2 - x_{2l+1})^2 \right.$$

$$\left. + (x_{2l+1} - 1)^2 + 10(x_{2l} + x_{2l+2} - 2)^2 + \frac{1}{10}(x_{2l} - x_{2l+2})^2 \right).$$

For each $l = 1, \ldots, \frac{n}{2} - 1$, let $I_l = \{2l - 1, 2l, 2l + 1, 2l + 2\}$ and $g_j(x) = x_j$, $j = 1, \ldots, n$. Then, the assumptions in Theorem 3.1 are satisfied as $\text{supp} g_j = \{j\} \subseteq I_{[\frac{j}{2}]}$, where $[\frac{j}{2}]$ is the largest integer that is smaller than $\frac{j}{2}$, and for each $l = 1, \ldots, n - 1$,

$$I_{l+1} \cap \left(\bigcup_{j=1}^{l} I_j \right) = \{2l + 1, 2l + 2\} \subseteq I_l.$$

Therefore, Theorem 3.1 implies that the optimal value of the proposed sparse SDP hierarchy (3.1) converges to the optimal value of the global minimum of (EP_R).

We now test our proposed sparse SDP hierarchy (3.1) on (EP_C) for different values of n with $c = n + 1$. From Table 3, we see that for dimension $n = 10,000$ or $n = 20,000$, the sparsePOP with the relaxation order =2 returns an accurate solution in the proposed sparse hierarchy Sparse SDP hierarchy (3.1).

The numerical experiment using the dense SDP hierarchy proposed in [11] could solve (EP_C) up to only $n = 10$. We observe again that much larger problems can be solved by the sparse SDP hierarchy.

4.4 Further examples

In this part, for further illustration, we consider the following polynomial optimization problem:

$$(P_0) \quad \min_{(x^l, w^l) \in \mathbb{R}^{n_l} \times \mathbb{R}^{n_l}, l=1, \ldots, q} \sum_{l=1}^{q} \sum_{i \in I_l} (w_i^l)^2$$

$$\text{s.t.} \quad g_j^l(x^l) \leq 0, \quad j = 1, \ldots, m, \quad l = 1, \ldots, q$$

$$(1 - w_i^l) x_i^l = 0, \quad i \in I_l, \quad l = 1, \ldots, q,$$

$$\|x^l\|_2^2 \leq M^l, \quad l = 1, \ldots, q,$$

where $I_l \subseteq \{1, \ldots, n\}$, $l = 1, \ldots, q$, $\bigcup_{l=1}^q I_l = \{1, \ldots, n\}$, $n_l = |I_l|$, that is, n_l is the cardinality of $I_l, l = 1, \ldots, q$, $I_l = \{i_1, \ldots, i_{n_l}\}$ and g_j^l are polynomials, $j = 1, \ldots, m$, $l = 1, \ldots, q$.

Note that problem (P_0) is a polynomial optimization reformulation of the sparse optimization problem with polynomial inequality constraints. Problems of this kind arise in signal processing and statistics [3,4] and have received a great deal of attention in the literature. For a recent comprehensive survey, see [2]. Other related references can be found in [33]. The relationship between problem (P_0) with the sparse optimization problem has been briefly outlined in the Appendix.

We also note that, there are other efficient approaches for solving the problem (P_0) (or the corresponding equivalent sparse optimization problem) such as the branch and bound approach and the techniques of mathematical programming with equilibrium constraints (for example see [6]). Due to the limitation of the SDP solvers, our approach cannot compete with these ad-hoc but numerically tractable approaches. The purpose of this section is to illustrate that our proposed sparse SDP hierarchy can be used to solve problem (P_0) with large size where the usual dense SDP hierarchy may not be directly applicable.

We now introduce a SDP hierarchy for problem (P_0) using the results in Sect. 3. It is worth noting that, problem (P_0) is a minimization problem with variables $(x, w) \in \mathbb{R}^n \times \mathbb{R}^n$. Thus, the feasible set of (P_0) can be unbounded in general (for example, simply take $g_j \equiv 0$. Then, $(w, 0)$ is feasible for problem (P_0) for any $w \in \mathbb{R}^n$).

Let $f(x, w) = \sum_{l=1}^q f_l(x^l, w^l)$, where $f_l(x^l, w^l) = \sum_{i \in I_l} (w_i^l)^2$ and $x = (x^1, \ldots, x^q) \in \mathbb{R}^{\sum_{l=1}^q n_l}$ and $w = (w^1, \ldots, w^q) \in \mathbb{R}^{\sum_{l=1}^q n_l}$. For each $l = 1, \ldots, q$, define $\bar{g}_{il}(x^l, w^l) = (1 - w_i^l)x_i^l$, $i \in I_l$ and $G_l(x^l, w^l) := \|x^l\|_2^2 - M^l$. Let c be a number such that $c > n^l$, $l = 1, \ldots, q$. For each integer k, we define the following sparse truncated quadratic module $\hat{\mathbf{M}}_k$ generated by the polynomials $c - f^l, -g_j^l, j = 1, \ldots, m, \bar{g}_{il}, i \in I_l$ and $G_l, l = 1, \ldots, q$, as follows:

$$\hat{\mathbf{M}}_k := \left\{ \sum_{l=1}^q \left(\sigma_{0l} - \sum_{j=1}^m \sigma_{jl} g_j^l + \sum_{i=1}^{I_l} h_{il} \bar{g}_{il} + \sigma_l(-G_l) + \bar{\sigma}_l \left(c - f^l \right) \right) \right.$$

$$\left. | \sigma_{0l}, \sigma_{il}, \sigma_l, \bar{\sigma}_l \in \Sigma^2 \left[\underline{x^l}, \underline{w^l} \right], \right.$$

$$h_{il} \in \mathbb{R} \left[\underline{x^l}, \underline{w^l} \right], \ \deg \sigma_{0l} \le 2k, \ \deg \sigma_{jl} g_{jl} \le 2k \ \deg h_{il} \bar{g}_{il} \le 2k, \text{ and}$$

$$\left. \deg \sigma_l G_l \le 2k, \deg \bar{\sigma} \left(c - f^l \right) \le 2k \right\}.$$

Consider the following relaxation problem

$$f_k^* := \sup \left\{ \mu \in \mathbb{R} \mid f - \mu \in \hat{\mathbf{M}}_k \right\}. \tag{4.2}$$

Then, one can show that $\lim_{k \to \infty} f_k^* = \min(P_0)$ if the running intersection property holds.

Proposition 4.1 (Convergence of the sparse SDP hierarchy value for (P_0)) *For problem (P_0) and the SDP hierarchy (4.2), assume that the running intersection property holds for $I_l, l = 1, \ldots, p$. Then, $\lim_{k \to \infty} f_k^* = \min(P_0)$.*

Table 4 Numerical tests on the polynomial optimization problem with hidden coercivity

Problem	q, n	c	M	Optimal val.	Rel.Er	CPU
SparseHierarchy (4.2)	200, 1600	5	100	200	3.3e−9	173.9
SparseHierarchy (4.2)	1000, 8000	5	100	1000	1.1e−5	470.2

n is the number of variables. The relaxation order 2 was used

Proof Note that the problem (P_0) can be equivalently rewritten as

$$\min_{(x^l, w^l) \in \mathbb{R}^{|I_l|} \times \mathbb{R}^{|I_l|}} \sum_{l=1}^{q} f^l(x^l, w^l) = \sum_{l=1}^{q} \sum_{i \in I_l} (w_i^l)^2$$

$$\text{s.t.} \quad g_j^l(x^l) \leq 0, \quad j = 1, \ldots, m, l = 1, \ldots, q$$

$$\bar{h}_{il}(x^l, w^l) := (1 - w_i^l) x_i^l = 0, \quad i \in I_l, l = 1, \ldots, q$$

$$G_l(x^l, w^l) := \|x^l\|_2^2 - M^l \leq 0.$$

As $\bigcup_{l=1}^{p} I_l = \{1, \ldots, n\}$, $\sum_{l=1}^{p} f^l(x^l, w^l) + \sum_{l=1}^{p} G_l(x^l, w^l)$ is strongly coercive. So, the assumptions in Theorem 3.1 are satisfied. Therefore, Theorem 3.1 implies that $\lim_{k \to \infty} f_k^* = \min(P_0)$. $\qquad \square$

We now illustrate the SDP hierarchy using a numerical example. For the numerical test on the effectiveness of the proposed SDP hierarchy, we let $q \in \mathbb{N}$. For each $l = 1, \ldots, q$, we generate a 3-by-4 matrix A_l containing random values drawn from the uniform distribution on the interval $[0, 1]$ and define $b_l = 4A_l(:, 1)$, where $A_l(:, 1)$ denotes the first column of A_l. Let $I_l = \{4l - 3, 4l - 2, 4l - 1, 4l\}$, $l = 1, \ldots, q$. Denote $x^l = (x_i^l)_{i \in I_l} \in \mathbb{R}^4$ and $w^l = (w_i^l)_{i \in I_l} \in \mathbb{R}^4$, $l = 1, \ldots, q$. Consider the problem:

$$(EP) \quad \min_{\substack{x = (x^1, \ldots, x^q) \in \mathbb{R}^{4q} \\ w = (w^1, \ldots, w^q) \in \mathbb{R}^{4q}}} \sum_{l=1}^{q} \sum_{i \in I_l} (w_i^l)^2$$

$$\text{s.t.} \quad A_l x^l = b_l, \quad l = 1, \ldots, q$$

$$(1 - w_i^l) x_i^l = 0, \quad i \in I_l, l = 1, \ldots, q,$$

$$\|x^l\|_2^2 \leq 100, \quad l = 1, \ldots, q.$$

It is not hard to see that the optimal solution of (EP) is (x^*, w^*) with optimal value q, where $x^* := [x^{1*}, \ldots, x^{q*}]$ with $x^{l*} = [4, 0, 0, 0]$ and $w^* := [w^{1*}, \ldots, w^{q*}]$ with $w^{l*} = [1, 0, 0, 0]$, $l = 1, \ldots, q$. Moreover, it can be verified that the running intersection property holds. So, the optimal value of the proposed sparse SDP hierarchy (4.2) converges to the optimal value of (EP).

As shown in Table 4, for $q = 200$ and $q = 1000$, in our numerical experiment, sparsePOP with the relaxation order 2, $c = 5$ and $M = 100$, returns an accurate optimal value and an accurate optimal solution x^* in the proposed Sparse SDP hierarchy (4.2).

We also notice that the numerical experiment using the dense SDP hierarchy proposed in [11] could solve the above problem with number of blocks q up to only 5. We observe again that much larger problems can be solved by the sparse SDP hierarchy.

Table 5 Numerical tests on the polynomial optimization problem with hidden coercivity

Problem	q, n	c	M	Optimal val.	Rel.Er	CPU
SparseHierarchy (4.2)	200, 1600	5	100	200	5.9e−6	107.6
SparseHierarchy (4.2)	1000, 8000	5	100	1000	1.8e−6	493.5

n is the number of variables. The relaxation order 2 was used

Finally, we note that we can obtain optimal solution for (EP) using the solver sparsePOP. For example, consider problem (EP) where $q = 200$,

$$A_l = A := \begin{pmatrix} 2 & -1 & 30 & 3 \\ 3 & 3 & 44 & 2 \\ -2 & 7 & -40 & -6 \end{pmatrix},$$

$$b_l = b := 4A(:, 1) = \begin{pmatrix} 8 \\ 12 \\ -8 \end{pmatrix} \quad \text{for all } l = 1, \ldots, 200.$$

Then, the true unique solution of this problem is (x^*, w^*) where $x^* := [x^{1*}, \ldots, x^{q*}]$ with $x^{l*} = [4, 0, 0, 0]$, and $w^* := [w^{1*}, \ldots, w^{q*}]$ with $w^{l*} = [1, 0, 0, 0]$ $l = 1, \ldots, 200$. In our numerical experiment, sparsePOP with the relaxation order 2 and $c = 5$, returns an accurate optimal value 200 and an accurate optimal solution (x^*, w^*), in the proposed Sparse SDP hierarchy (4.2), as shown in Table 5.

Appendix

The polynomial optimization problem (P_0) has a close relationship with the problem of finding the solution with the least number of nonzero components which satisfies a system of polynomial inequalities and simple bounds. Mathematically, the problem of finding the solution with the least number of nonzero components which satisfies a system of polynomial inequalities and simple bounds, can be formulated as

$$(P_0') \quad \min_{(x^1, \ldots, x^q) \in \mathbb{R}^{n_1} \times \ldots \times \mathbb{R}^{n_q}} \sum_{l=1}^{q} \|x^l\|_0$$

$$\text{s.t.} \quad g_j^l(x^l) \leq 0, \quad j = 1, \ldots, m, \, l = 1, \ldots, q,$$

$$\|x^l\|_2^2 \leq M^l, \quad l = 1, \ldots, q,$$

where $n_l \in \mathbb{N}$, $l = 1, \ldots, q$, and $\|x\|_0$ denotes the l_0-seminorm of the vector $x \in \mathbb{R}^n$, which gives the number of nonzero components of the vector x.

In the case where $q = 1$, $I_1 = \{1, \ldots, n\}$, $g_j(x) = a_j^T x - b_j$, $j = 1, \ldots, m$, $g_j(x) = -(a_j^T x - b_j)$, $j = m + 1, \ldots, 2m$, the problem (P_0') collapses to the sparse optimization problem which finds the solution with the least number of nonzero components satisfying simple bounds as well as linear equations $Ax = b$ with more unknowns than equalities:

$$(P_1) \quad \min_{x \in \mathbb{R}^n} \|x\|_0$$

$$\text{s.t.} \quad Ax = b,$$

$$\|x\|_2^2 \leq M,$$

where $A = (a_1, \ldots, a_m)^T \in \mathbb{R}^{m \times n}$ $(m \leq n)$, $b = (b_1, \ldots, b_m)^T \in \mathbb{R}^m$. We note that the standard sparse optimization problem which is given by

$$(P_2) \quad \min_{x \in \mathbb{R}^n} \ \|x\|_0$$
$$\text{s.t.} \quad Ax = b,$$

arises in signal processing and was examined, for example, [2–4,33]. Moreover, problems (P_1) and (P_2) have the same optimal value if $M > \|x^*\|_2^2$ for some solution x^* of problem (P_2).

In fact, the problem (P_0) and problem (P_0') are equivalent in the sense that $\min(P_0) = \min(P_0')$ and $(x^{1*}, \ldots, x^{q*}) \in \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_q}$ is a solution of problem (P_0) if and only if $(x^{l*}, w^{l*}) \in \mathbb{R}^{n_l} \times \mathbb{R}^{n_l}$, $l = 1, \ldots, q$, is a solution of problem (P_0') where $w^{l*} = (w_{i_1}^{l*}, \ldots, w_{i_{n_l}}^{l*}) \in \mathbb{R}^{n_l}$ is defined by

$$w_i^{l*} = \begin{cases} 1 & \text{if } x_i^{l*} \neq 0, \\ 0 & \text{if } x_i^{l*} = 0. \end{cases} \qquad i \in I_l, \ l = 1, \ldots, q. \tag{4.3}$$

References

1. Andersen, M., Vandenberghe, L., Dahl, J.: Linear matrix inequalities with chordal sparsity patterns and applications to robust quadratic optimization. In: Proceedings of IEEE International Symposium on Computer-Aided Control System Design (CACSD) (2010)
2. Bruckstein, A.M., Donoho, D.L., Elad, M.: From sparse solutions of systems of equations to sparse modeling of signals and images. SIAM Rev. **51**(1), 34–81 (2009)
3. Candés, E.J., Tao, T.: Decoding by linear programming. IEEE Trans. Inf. Theory **51**, 4203–4215 (2005)
4. Candés, E.J., Wakin, M.B., Boyd, S.: Enhancing sparsity by re-weighted l_1 minimization. J. Fourier Anal. Appl. **14**(5), 877–905 (2008)
5. Demmel, J., Nie, J.W.: Sparse SOS relaxations for minimizing functions that are summations of small polynomials. SIAM J. Optim. **19**(4), 1534–1558 (2008)
6. Feng, M., Mitchell, J.E., Pang, J.S., Shen, X., Wachter, A.: Complementarity formulations of l_0-norm optimization problems. http://www.optimization-online.org/DB_HTML/2013/09/4053.html
7. Feng, C., Lagoa, C., Sznaier, M.: Hybrid system identification via sparse polynomial optimization. In: American Control Conference (ACC), pp. 160–165 (2010)
8. Fukuda, M., Kojima, M., Murota, K., Nakata, K.: Exploiting sparsity in semidefinite programming via matrix completion I: general framework. SIAM J. Optim. **11**, 647–674 (2000)
9. Ghaddar, B., Marecek, J., Mevissen, M.: Optimal power flow as a polynomial optimization problem. IEEE Trans. Power Syst. **99**, 1–8 (2015)
10. Hà, H.V., Phạm, T.S.: Representations of positive polynomials and optimization on noncompact semi-algebraic sets. SIAM J. Optim. **20**, 3082–3103 (2010)
11. Jeyakumar, V., Lasserre, J.B., Li, G.: On polynomial optimization over non-compact semi-algebraic sets. J. Optim. Theory Appl. **163**(3), 707–718 (2014)
12. Jeyakumar, V., Li, G., Pham, T.S.: Convergence of the Lasserre hierarchy of SDP relaxations for convex polynomial programs without compactness. Oper. Res. Lett. **42**, 34–40 (2014)
13. Jeyakumar, V., Li, G.: A new class of alternative theorems for SOS-convex inequalities and robust optimization. Appl. Anal. **94**, 56–74 (2015)
14. Jeyakumar, V., Li, G.: Exact SDP relaxations for classes of nonlinear semidefinite programming problems. Oper. Res. Lett. **40**, 529–536 (2012)
15. Kleniati, P., Parpas, P., Rustem, B.: Partitioning procedure for polynomial optimization. J. Glob. Optim. **48**, 549–567 (2010)
16. Kim, S., Kojima, M., Mevissen, M., Yamashita, M.: Exploiting sparsity in linear and nonlinear matrix inequalities via positive semidefinite matrix completion. Math. Program. **129**, 33–68 (2011)
17. Kim, S., Kojima, M., Waki, H., Yamashita, M.: Algorithm 920: SFSDP: a sparse version of full semidefinite programming relaxation for sensor network localization problems. ACM Trans. Math. Softw. **38**, 1–19 (2012)

18. Kojima, M., Muramatsu, M.: A note on sparse SOS and SDP relaxations for polynomial optimization problems over symmetric cones. Comput. Optim. Appl. **42**(1), 31–41 (2009)
19. Lasserre, J.B.: Global optimization with polynomials and the problem of moments. SIAM J. Optim. **11**, 796–817 (2001)
20. Lasserre, J.B.: Moments, Positive Polynomials and Their Applications. Imperial College Press, London (2009)
21. Lasserre, J.B.: Convergent SDP-relaxations in polynomial optimization with sparsity. SIAM J. Optim. **17**, 822–843 (2006)
22. Lu, W.-S., Hinamoto, T.: Design of FIR filters with discrete coefficients via polynomial programming: towards the global solution. In: IEEE International Symposium on Circuit and Systems, pp. 2048–2051 (2007)
23. Luo, Z.-Q., Sidiropoulos, N.D., Tseng, P., Zhang, S.: Approximation bounds for quadratic optimization with homogeneous quadratic constraints. SIAM J. Optim. **18**, 1–28 (2007)
24. Marshall, M.: Positive Polynomials and Sums of Squares, Mathematical Surveys and Monographs, vol. 146. American Mathematical Society, Providence (2008)
25. Nie, J.W.: Sum of squares method for sensor network localization. Comput. Optim. Appl. **43**(2), 151–179 (2009)
26. Nie, J.W.: An exact Jacobian SDP relaxation for polynomial optimization. Math. Program. **137**(1–2), 225–255 (2013)
27. Nie, J.: Optimality conditions and finite convergence of Lasserre's hierarchy. Math. Program. **146**, 97–121 (2014)
28. Parrilo, P.A.: Structured semidefinite programs and semi-algebraic geometry methods in robustness and optimization, Ph.D. thesis, California Institute of Technology (2000)
29. Putinar, M.: Positive polynomials on compact semi-algebraic sets. Indiana Univ. Math. J. **41**, 49–95 (1993)
30. Scheiderer, C.: Sums of squares on real algebraic curves. Math. Z. **245**, 725–760 (2003)
31. Schmüdgen, K.: The K-moment problem for compact semi-algebraic sets. Math. Ann. **289**, 203–206 (1991)
32. Schweighofer, M.: Global optimization of polynomials using gradient tentacles and sums of squares. SIAM J. Optim. **17**, 920–942 (2006)
33. Simon, F., Holger, R.: A Mathematical Introduction to Compressive Sensing, Applied and Numerical Harmonic Analysis, Chapter 15. Birkhauser/Springer, New York (2013)
34. Strum, J.F.: SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. Optim. Methods Softw. **11—-12**, 625–653 (1999)
35. Vandenberghe, L., Andersen, M.: Chordal graphs and semidefinite optimization. Found. Trends Optim. **1**, 241–433 (2014)
36. Waki, H., Kim, S., Kojima, M., Muramatsu, M.: Sums of squares and semidefinite programming relaxations for polynomial optimization problems with structured sparsity. SIAM J. Optim. **17**, 218–242 (2006)
37. Waki, H., Kim, S., Kojima, M., Muramtasu, M., Sugimoto, H.: Algorithm 883: SparsePOP: a sparse semidefinite programming relaxation of polynomial optimization problems. ACM Trans. Math. Softw. **35**, 15 (2008)