A hierarchy of spectral relaxations for polynomial optimization

Ngoc Hoang Anh Mai¹ · Jean-Bernard Lasserre¹,² · Victor Magron¹,²

Received: 17 July 2020 / Accepted: 21 April 2023 / Published online: 22 May 2023
© Springer-Verlag GmbH Germany, part of Springer Nature and Mathematical Optimization Society 2023

Abstract
We show that (1) any constrained polynomial optimization problem (POP) has an equivalent formulation on a variety contained in an Euclidean sphere and (2) the resulting semidefinite relaxations in the moment-SOS hierarchy have the constant trace property (CTP) for the involved matrices. We then exploit the CTP to avoid solving the semidefinite relaxations via interior-point methods and rather use ad-hoc spectral methods for minimizing the largest eigenvalue of a matrix pencil. Convergence to the optimal value of the semidefinite relaxation is guaranteed. As a result we obtain a hierarchy of nonsmooth “spectral relaxations” of the initial POP. Efficiency and robustness of this spectral hierarchy is tested against several equality constrained POPs on a sphere as well as on a sample of randomly generated quadratically constrained quadratic problems.

Keywords Polynomial optimization · Moment-SOS hierarchy · Maximal eigenvalue minimization · Limited-memory bundle method · Nonsmooth optimization · Semidefinite programming

Mathematics Subject Classification 90C23 · 90C22 · 90C25 · 14Q30

Ngoc Hoang Anh Mai
ngoc-hoang.mai@laas.fr

Jean-Bernard Lasserre
jean-bernard.lasserre@laas.fr

Victor Magron
victor.magron@laas.fr

¹ CNRS, LAAS, 7 avenue du Colonel Roche, 31400 Toulouse, France
² Université de Toulouse, LAAS, 31400 Toulouse, France
1 Introduction

The moment-sums of squares (moment-SOS) hierarchy for solving polynomial optimization problems (POP) consists of solving a sequence of semidefinite programming (SDP) relaxations of increasing size. Thanks to powerful positivity certificates from real algebraic geometry, its associated monotone sequence of optimal values converges to the global optimum [25]. Even though this procedure is efficient, with generically finite convergence [40], it suffers from two main drawbacks:

1. In view of the current status of SDP solvers, it is limited to problems of modest size unless some sparsity and/or symmetry can be exploited.
2. When solving the SDP relaxations of the hierarchy by interior-point methods (as do most current SDP solvers) the computational cost is quite high.

Recent efforts have tried to overcome these drawbacks:

(a) By designing computationally cheaper hierarchies of convex relaxations based on alternative positivity certificates such as the bounded degree SOS hierarchy [31], nonnegative circuits relying on geometric programming [13] or second-order cone programming [51], and arithmetic–geometric-exponentials [6] relying on relative entropy programming.
(b) By exploiting certain sparsity patterns in the POP formulation, based on correlative sparsity [26, 50] or term sparsity [52–54], possibly combined with (a).
(c) By exploiting a Constant Trace Property (CTP) of semidefinite relaxations associated with POPs coming from combinatorial optimization [17, 56]. This permits to solve the semidefinite relaxation with ad-hoc method, like, e.g., limited-memory bundle methods, instead of the costly interior-point methods.

The present paper is part of the latter type-(c) efforts.

1.1 Background on SDP with CTP

One way to exploit the CTP of matrices in SDPs is to consider the dual which reduces to minimize the maximum eigenvalue of a symmetric matrix pencil [17]. For problems of moderate size one may solve the latter problem with interior-point methods [2]. However for larger-scale instances, running a single iteration becomes computationally too demanding and therefore one has to use an alternative method, and in particular first-order methods.

To solve large-scale instances of this maximal eigenvalue minimization problem, two types of first-order methods can be used: subgradient descent or variants of the mirror-prox algorithm [39], and spectral bundle methods [17]. In other methods of interest based on non-convex formulations [4, 21], the problem is directly solved over the set of low rank matrices. These latter approaches are particularly efficient for problems where the solution is low rank, e.g., for matrix completion or combinatorial relaxations.

Despite their empirical efficiency, the computational complexity of spectral bundle and low rank methods is still not completely understood. This is in contrast with
methods based on stochastic smoothing results for which explicit computational complexity estimates are available. For instance in [10] smooth stochastic approximations of the maximum eigenvalue function are obtained via rank-one Gaussian perturbations. In [45] Newton’s method is used, assuming that the multiplicity of the maximal eigenvalue is known in advance.

By combining quasi-Newton methods (e.g., Broyden–Fletcher–Goldfarb–Shanno (BFGS) method or its so-called “Limited-memory” version (L-BFGS) [42]) with adaptive gradient sampling [5, 24], convergence guarantees are obtained for certain non smooth problems while keeping good empirical performance [7, 34].

Another hybrid method is the Limited-Memory Bundle Method (LMBM) which combines L-BFGS with bundle methods [14, 15]: Briefly, L-BFGS is used in the line search procedure to determine the step sizes in the bundle method. LMBM enjoys global convergence for locally Lipschitz continuous functions which are not necessarily differentiable.

Finally the more recent SketchyCGAL algorithm [56] also uses limited memory and arithmetic. It combines a primal-dual optimization scheme together with a randomized sketch for low-rank matrix approximation. Assuming that zero duality gap holds, it provides a near-optimal low-rank approximation. A variant of SketchyCGAL can handle SDPs with bounded (instead of constant) trace property.

Concerning SDPs coming from relaxations in polynomial optimization, Malick and Henrion [19, Sect. 3.2.3] have used the CTP to provide an efficient algorithm for unconstrained polynomial optimization problems. At last but not least, the CTP trivially holds for Shor’s relaxation [48] of combinatorial optimization problems formulated as linear-quadratic POPs on the discrete hypercube $\{-1, 1\}^n$. This fact has been exploited in Helmberg and Rendl [17] to avoid solving the associated SDP via interior-point methods.

1.2 Contribution

A novelty with respect to previous (c)-efforts is to show that every POP on a compact basic semialgebraic set has an equivalent equality constrained POP formulation on an Euclidean sphere (possibly after adding some artificial variables) such that each of its semidefinite relaxations in the moment-SOS hierarchy has the CTP. We call CTP-POP such a formulation of POPs. Therefore to solve each semidefinite relaxation of a CTP-POP one may avoid the computationally costly interior-point methods in some cases. Indeed as the dual reduces to minimize the largest eigenvalue of a matrix pencil, one may rather use efficient ad-hoc non smooth methods as those invoked above.

Main results

(I) In Sect. 3.1, we prove that each semidefinite moment relaxation indexed by $k \in \mathbb{N}$:
\[-\tau_k = \sup_{X \in S_k} \langle C_k, X \rangle > A_k X = b_k, \quad X \succeq 0,\]

of the moment-SOS hierarchy associated with an equality constrained POP on an Euclidean sphere of \(\mathbb{R}^n\) has CTP (see Lemma 5), i.e.,

\[\forall X \in S_k, \quad A_k X = b_k \Rightarrow \text{trace}(X) = a_k,\]

where \(A_k : S_k \to \mathbb{R}^{m_k}\) is a linear operator with \(S_k\) being the set of real symmetric matrices of size \((n+k)n\), \(C_k \in S_k\) and \(b_k \in \mathbb{R}^{m_k}\) with \(m_k = O((n+k)^2)\). Following the framework by Helmberg and Rendl [17], SDP (1) boils down to minimizing the largest eigenvalue of a matrix pencil:

\[-\tau_k = \inf_{z \in \mathbb{R}^{m_k}} a_k \lambda_1 \left( C_k - A_k^\top z \right) + b_k^\top z,\]

where \(\lambda_1(A)\) stands for the largest eigenvalue of \(A\) and \(A_k^\top\) denotes the adjoint operator of \(A_k\).

Hence (2) form what we call a hierarchy of (non smooth, convex) spectral relaxations of the equality constrained POP on a sphere. Convergence of \((\tau_k)_{k \in \mathbb{N}}\) to the optimal value \(f^*\) of the initial POP is guaranteed with rate at least \(O(k^{-1/c})\) (see Theorem 2), where \(c\) depends only on the polynomials describing the cost and constraints of the POP.

In addition, existence of an optimal solution of the spectral relaxation (2) is guaranteed for sufficiently large \(k\) under certain conditions on the POP (see Proposition 5). Finally, when the set of global minimizers of the equality constrained POP on the sphere is finite, we also describe how to obtain an optimal solution \(x^*\) via an optimal solution \(\bar{z}\) of (2).

(II) In Sect. 3 we prove that any POP on a compact basic semialgebraic set (including a ball constraint \(R - \|x\|^2 \geq 0\)) has an equivalent equality constrained POP (called CTP-POP) on a sphere of \(\mathbb{R}^{n+l_g+1}\), where \(l_g\) is the number of inequality constraints of the initial POP. This CTP-POP can be solved by using spectral relaxations (2).

(III) We describe Algorithm 3 to handle a given equality constrained POP on the sphere. It consists of handling each semidefinite relaxation (1) by solving the spectral formulation (2), with a nonsmooth optimization procedure chosen in advance by the user in our software library, called SpectralSOS. This library supports the three optimization subroutines LMBM [14, 15], proximal bundle (PB) [17], and SketchyCGAL [56]. Our default method in Algorithm 3 is LMBM.

(IV) Finally, efficiency and robustness of SpectralPOP are illustrated in Sect. 4 on extensive benchmarks. We solve several (randomly generated) dense equality constrained QCQPs on the unit sphere by running Algorithm 3 and compare results with those obtained with the standard moment-SOS hierarchy. Suprisingly SpectralPOP can provide the optimal value as well as an optimal solution with high accuracy, and up to twenty five times faster than the semidefinite hierarchy. For instance, SpectralPOP can solve the first relaxation of minimization problem of dense quadratic polynomials on the unit sphere with up to \(n = 500\) variables in about 35 s and up to 1500 variables.
in about 7000 s on a standard laptop computer. We emphasize that for some problems not randomly generated and scaled so as to fit our optimization framework on the unit sphere, we could observe a lack of high precision after transferring results (of the scaled formulation) back to the unscaled initial formulation.

We also provide extended applications of spectral relaxations to the following three decision problems: deciding nonnegativity of even degree forms, deciding convexity of even degree forms and deciding copositivity of real symmetric matrices, with very satisfactory results.

In [17], Helmberg and Rendl propose a spectral bundle method (based on Kiwiel’s proximal bundle method [23]) to solve an SDP relying on the maximal eigenvalue minimization problem of the form (2). This method works better than interior-point algorithms for very large-scale SDPs, when the number of trace equality constraints is not larger than the size of the positive semidefinite matrix (e.g., Shor’s relaxation of MAXCUT problems). However this method is not always more efficient than interior-point solvers (e.g., SDPT3) for instance when the SDPs involve a number of trace equality constraints which is larger than the size of the positive semidefinite matrix, as reported in [16, Table 1–6]. Unfortunately this latter type of SDP is the generic form of moment-SOS relaxations for POPs and thus is not suitable to be solved by Helmberg–Rendl’s spectral bundle method. By contrast with previous works, our numerical results show that the combination between Helmberg–Rendl’s spectral formulation and LMBM is cheaper and faster than Mosek (the currently fastest SDP solver based on interior-point method) while maintaining the same accuracy when solving moment relaxations of equality constrained POPs on a sphere.

2 Background and preliminary results

With \( x = (x_1, \ldots, x_n) \), let \( \mathbb{R}[x] \) stand for the ring of real polynomials and let \( \Sigma[x] \subset \mathbb{R}[x] \) be its subset of SOS polynomials. We denote by \( \mathbb{R}[x]_t \) and \( \Sigma[x]_t \) their respective restrictions to polynomials of degree at most \( t \) and \( 2t \). Given \( \alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n \), let \( |\alpha| := \alpha_1 + \cdots + \alpha_n \). We denote by \( \mathbb{N}_d^n \) the set of all vectors in \( \mathbb{N}^n \) whose coordinates sum up to a value of at most \( d \). Let \( (x^\alpha)_{\alpha \in \mathbb{N}^n} \) be the canonical basis of monomials for \( \mathbb{R}[x] \) (ordered according to the graded lexicographic order) and \( v_t(x) \) be the vector of monomials up to degree \( t \), with length \( \{ n \} \) := \( (n+1)! \). A polynomial \( p \in \mathbb{R}[x] \), is written as \( p(x) = \sum_{\alpha \in \mathbb{N}^n} p_\alpha x^\alpha = p^T v_t(x) \), where \( p = (p_\alpha) \in \mathbb{R}\{n\} \) is its vector of coefficients in the canonical basis. The \( l_1 \)-norm of a polynomial \( p \) is given by the \( l_1 \)-norm of its vector of coefficients \( p \), that is \( |p|_1 := \sum_\alpha |p_\alpha| \). Given \( a \in \mathbb{R}^n \), the \( l_2 \)-norm of \( a \) is \( |a|_2 := (a_1^2 + \cdots + a_n^2)^{1/2} \). We denote by \( \mathbb{N}^* \) the set \( \{ t \in \mathbb{N} : t \neq r \} \) with \( * \in \{ >, <, =, \geq, \leq \} \). For every \( l \in \mathbb{N}^*>0 \), let \( [l] := \{ 1, \ldots, l \} \) and \( [0] := \emptyset \).

2.1 Riesz linear functional

Given a real-valued sequence \( y = (y_\alpha)_{\alpha \in \mathbb{N}^n} \), define the Riesz linear functional \( L_y : \mathbb{R}[x] \rightarrow \mathbb{R}, f \mapsto L_y(f) := \sum_\alpha f_\alpha y_\alpha \). A real infinite (resp. finite) sequence \( (y_\alpha)_{\alpha \in \mathbb{N}^n} \) (resp. \( (y_\alpha)_{\alpha \in \mathbb{N}^*} \)) has a representing measure if there exists a finite Borel measure \( \mu \).
such that $y_\alpha = \int_{\mathbb{R}^n} x^\alpha d\mu(x)$ is satisfied for every $\alpha \in \mathbb{N}^n$ (resp. $\alpha \in \mathbb{N}_0^n$). In this case, $(y_\alpha)_{\alpha \in \mathbb{N}^n}$ is called the moment sequence of $\mu$.

2.2 Moment matrices

The moment matrix of degree $d$ associated with a real-valued sequence $y = (y_\alpha)_{\alpha \in \mathbb{N}^n}$ and $d \in \mathbb{N}^{>0}$, is the real symmetric matrix $M_d(y)$ of size $\{n\}^d$, with entries $(y_\alpha + \beta)_{\alpha, \beta \in \mathbb{N}^n}$.

2.3 Localizing matrices

The localizing matrix of degree $d$ associated with $y = (y_\alpha)_{\alpha \in \mathbb{N}^n}$ and $p = \sum_{\gamma} p_\gamma x^\gamma \in \mathbb{R}[x]$, is the real symmetric matrix $M_d(p y)$ of size $\{n\}^d$ with entries $(\sum_{\gamma} p_\gamma y_\gamma + \alpha + \beta)_{\alpha, \beta \in \mathbb{N}^n}$.

2.4 General POPs on basic compact semialgebraic sets

A polynomial optimization problem is of the form

$$f^* := \inf \{ f(x) | x \in S(g, h) \},$$

where $S(g, h)$ is a basic semialgebraic set defined as follows:

$$S(g, h) := \{ x \in \mathbb{R}^n : g_i(x) \geq 0, i \in [l_g]; h_j(x) = 0, j \in [l_h] \}$$

for some polynomials $f, g_i, h_j \in \mathbb{R}[x]$. Let $g := \{ g_i \}_{i \in [l_g]}$ and $h := \{ h_j \}_{j \in [l_h]}$. For $p \in \mathbb{R}[x]$, let $\lceil p \rceil := \lceil \deg(p)/2 \rceil$.

If $S(g, h)$ is nonempty and compact, then $f^* < \infty$ and POP (3) has at least one global minimizer. Next, as we are concerned with POPs on compact feasible sets, we assume that $S(g, h) \subset B^*_R$, where $B^*_R := \{ x \in \mathbb{R}^n : R - \|x\|_2^2 \geq 0 \}$. In addition, if $l_g \neq 0$ then we may and will assume that $g_1 := R - \|x\|_2^2$.

2.4.1 Second-order sufficient conditions

Given $(\lambda_i)_{i \in [l_g]}$ and $(\gamma_j)_{j \in [l_h]}$, let:

$$x \mapsto \mathcal{L}(x, \lambda, \gamma) := f(x) - \sum_{i \in [l_g]} \lambda_i g_i(x) - \sum_{j \in [l_h]} \gamma_j h_j(x), \quad x \in \mathbb{R}^n.$$

Given $x \in S(g, h)$, let $J(x) := \{ i \in [l_g] : g_i(x) = 0 \}$.

**Definition 1** (see [43, Chapter 2]) We say that the second-order sufficient conditions hold at $x^* \in S(g, h)$ if:

- **KKT-Lagrange multipliers** There exist $\lambda^*_i \geq 0, i \in [l_g]$, and $\gamma^*_j \in \mathbb{R}, j \in [l_h]$, such that $\nabla_x \mathcal{L}(x^*, \lambda^*, \gamma^*) = 0$ and $\lambda^*_i g_i(x^*) = 0$ for all $i \in [l_g]$. 
– **Linear independence constraint qualification**: The family
\[
\{ \nabla g_i(x^*), \nabla h_j(x^*) \}_{i \in J(x^*), j \in [l_h]}
\]
is linearly independent.

– **Strict complementarity** \( \lambda^*_i + g_i(x^*) > 0 \), for all \( i \in [l_g] \).

– **Positive definiteness of Hessian matrices** \( u^T \nabla^2 L(x^*, \lambda^*, \gamma^*) u > 0 \) for all \( u \neq 0 \) such that \( u^T \nabla g_i(x^*) = 0 \) and \( u^T \nabla h_j(x^*) = 0 \), \( i \in J(x^*), j = 1, \ldots, l_h \).

### 2.4.2 The moment-SOS hierarchy

Given \( k \in \mathbb{N} \), the set
\[
Q(g, h) := \left\{ \sigma_0 + \sum_{i=1}^{l_g} \sigma_i g_i + \sum_{j=1}^{l_h} \psi_j h_j \mid \sigma_0 \in \Sigma[x], \sigma_i \in \Sigma[x], \psi_j \in \mathbb{R}[x] \right\}.
\]
is the *quadratic module* associated with the semialgebraic set \( S(g, h) \), while the set
\[
Q_k(g, h) := \left\{ \sigma_0 + \sum_{i=1}^{l_g} \sigma_i g_i + \sum_{j=1}^{l_h} \psi_j h_j \mid \sigma_0 \in \Sigma[x]_k, \sigma_i \in \Sigma[x]_{k-\lceil \log_2 l_i \rceil}, \psi_j \in \mathbb{R}[x]_{2(k-\lceil \log_2 l_j \rceil)} \right\},
\]
is its truncated version at order \( k \). Notice that \( g_1 (= R - \| x \|^2) \in Q(g, h) \) and therefore \( Q(g, h) \) is Archimedean [27].

Let \( c_\alpha := \frac{|\alpha|!}{\alpha_1! \cdots \alpha_n!} \) for each \( \alpha \in \mathbb{N}^n \). Let \( \| p \| := \max_{\alpha} \frac{|p|_{\alpha}}{c_\alpha} \), for a given \( p \in \mathbb{R}[x] \). As a consequence of Nie-Schweighofer’s main result in [41, Theorem 8], one obtains the following result:

**Lemma 1** Let \( f^* \) be as in (3) with \( S(g, h) \neq \emptyset \) as in (4). There exists \( c > 0 \) depending on \( g \) and \( h \) such that for \( k \in \mathbb{N} \) with \( k \geq c \exp((2d^2n^2)^c) \), one has
\[
(f - f^*) + 6d^3 n^2 \| f \| \log(k/c)^{-1/c} \in Q_k(g, h).
\]

Next, consider the hierarchy of semidefinite programs (SDP) indexed by \( k \in \mathbb{N} \):
\[
\rho_k := \sup \{ \xi \in \mathbb{R} \mid f - \xi \in Q_k(g, h) \}.
\]

By invoking Lemma 1, one obtains the convergence behavior of the sequence \((\rho_k)_{k \in \mathbb{N}}\) in the following result.

**Theorem 1** Let \( f^* \) be as in (3) with \( S(g, h) \neq \emptyset \) as in (4). Then:

1. For all \( k, l \in \mathbb{N} \), \( \rho_k \leq \rho_{k+l} \leq f^* \).
2. The sequence \((\rho_k)_{k \in \mathbb{N}}\) converges to \( f^* \) with rate at least \( O(\log(k/c)^{-1/c}) \).
Set
\[ k_{\text{min}} := \max_{i,j} \{ \lceil g_i \rceil, \lceil h_j \rceil \}. \]  
(6)

For every \( k \geq k_{\text{min}} \) the dual of (5) reads
\[
\tau_k := \inf_{y \in \mathbb{R}^n} L_y(f) : \quad \text{s.t. } \begin{align*}
M_k(y) &\succeq 0; \quad y_0 = 1 \\
M_{k-[g_i]}(g_i y) &\succeq 0, \quad i \in [l_g], \\
M_{k-[h_j]}(h_j y) &= 0, \quad j \in [l_h].
\end{align*}
\]  
(7)

here we slightly abuse terminology and say that there is a “zero duality gap” between (5) and (7) if \( \tau_k = \rho_k \) and \( \tau_k \in \mathbb{R} \) (the abuse of terminology is due to the fact that zero duality gap can occur with both values being infinite).

Slater’s condition on either (5) or (7) is a well-known sufficient condition to ensure zero duality gap. However, in case of equality constraints in the description (4) of \( S(g, h) \), Slater’s condition does not hold for (7).

**Proposition 1** (Josz–Henrion [20]) Let \( f^* \) be as in (3) with \( S(g, h) \neq \emptyset \) as in (4). Zero duality gap between the primal (5) and dual (7) holds for sufficiently large \( k \in \mathbb{N} \), i.e., \( \rho_k = \tau_k \) and \( \tau_k \in \mathbb{R} \). Moreover, SDP (7) has an optimal solution.

In [20] the authors prove that the set of optimal solutions of (7) is compact and therefore (7) has an optimal solution. Although there exist situations where SDP (5) has no optimal solution (see for instance the end of [40, Sect. 3]), the following proposition ensures the existence of an optimal solution under mild assumptions:

**Proposition 2** (Lasserre [25, Theorem 3.4a]) If \( S(g, h) \) has nonempty interior, then Slater’s condition on the dual (7) holds for \( k \geq k_{\text{min}} \), where \( k_{\text{min}} \) is defined as in (6).

In this case, \( \rho_k = \tau_k \), \( \tau_k \in \mathbb{R} \) and the primal (5) has an optimal solution.

Let \( \delta_a \) stand for the Dirac measure at point \( a \in \mathbb{R}^n \). The following result is a consequence of Curto–Fialkow’s Flat Extension Theorem [8, 32].

**Proposition 3** Let \( y^* \) be an optimal solution of SDP (7) at some order \( k \in \mathbb{N} \), and assume that the flat extension condition holds, i.e., \( \text{rank}(M_{k-w}(y^*)) = \text{rank}(M_k(y^*)) =: r \), with \( w := \max_{i,j} \{ \lceil g_i \rceil, \lceil h_j \rceil \} \). Then \( y^* \) has a representing \( r \)-atomic measure \( \mu = \sum_{t=1}^r \lambda_t \delta_{a(t)} \), where \( (\lambda_1, \ldots, \lambda_r) \) belong to standard \( (r-1) \)-simplex and \( \{ a^{(1)}, \ldots, a^{(r)} \} \subset S(g, h) \). Moreover, \( \tau_k = f^* \) and \( a^{(1)}, \ldots, a^{(r)} \) are all global minimizers of POP (3).

Henrion and Lasserre [18] provide a numerical algorithm to extract the \( r \) minimizer \( a^{(1)}, \ldots, a^{(r)} \) from \( M_k(y^*) \) when the assumptions of Proposition 3 hold.

The following proposition provides a sufficient condition to ensure finite convergence of the sequence \( (\tau_k)_{k \in \mathbb{N}} \).

**Proposition 4** The following statements are true:
1. (Nie [40]) The equality $\tau_k = f^*$ occurs generically for some $k \in \mathbb{N}$.

2. (Lasserre [28, Theorem 7.5]) If (i) $Q(g, h)$ is Archimedean, (ii) the ideal $\langle h \rangle$ is real radical, and (iii) the second-order sufficient conditions (see Definition 1) hold at every global minimizer of POP (3), then $\tau_k = \rho_k = f^*$ for some $k \in \mathbb{N}$ and both primal-dual (5)–(7) have optimal solutions.

3. (Lasserre et al. [30, Proposition 1.1] and [28, Theorem 6.13]) If $V(h)$ defined as in (9) is finite, $\tau_k = \rho_k = f^*$ for some $k \in \mathbb{N}$ and both primal-dual (5)–(7) have optimal solutions. In this case, the flatness condition holds at order $k$.

The first statement of Proposition 4 means that with fixed $f \in \mathbb{R}[x]$, the equality $\tau_k = f^*$ holds for $k \in \mathbb{N}$ sufficiently large on a Zariski open set (the complement of the zeros of a polynomial) in the space of the coefficients of $g_i, h_j$ with given degrees.

Note that the real radical property is not generic and so the condition “$\langle h \rangle$ is real radical” must be checked case by case. On the other hand, if $V(h)$ is the real zero set of a squared system of polynomial equations, i.e., $l_h = n$, then generically $V(h)$ has a finite number of points.

### 2.5 POPs on a variety contained in a sphere

We consider a special form of POP (3) which is of the form

$$f^* := \inf \{ f(x) \mid x \in V(h) \},$$

where $V(h)$ is the real variety defined by:

$$V(h) := \{ x \in \mathbb{R}^n : h_j(x) = 0; \ j = 1, \ldots, l_h \},$$

for some set of polynomials $h := \{h_j\} \subset \mathbb{R}[x]$. We assume that $h_1 := \frac{\bar{R} - \|x\|^2}{2}$ for some $\bar{R} > 0$, so that $V(h) \subset \partial B_{\bar{R}}^n$, where $\partial B_{\bar{R}}^n := \{ x \in \mathbb{R}^n : \bar{R} - \|x\|^2 = 0 \}$. By assuming that $V(h) \neq \emptyset$, $f^* < \infty$ and POP (8) has at least one global minimizer.

Given $k \in \mathbb{N}$, define the truncated preorder of order $k$ associated with the variety $V(h)$ in (9) as follows:

$$P_k(h) := \left\{ \sigma_0 + \sum_{j=1}^{l_h} \psi_j h_j \sigma_0 \in \Sigma[x]_k, \ \psi_j \in \mathbb{R}[x]_{2(l_h - \lfloor h_j \rfloor)}, \ j \in [l_h] \right\}.$$

**Remark 1** For every $k \in \mathbb{N}$, $P_k(h)$ is also the truncated quadratic module $Q_k(h)$ associated with the semialgebraic set $V(h) = S(\emptyset, h)$.

As a consequence of Schweighofer’s main result in [47, Theorem 4], one obtains the following result:

**Lemma 2** Let $f^*$ be as in (8) with $V(h)$ as in (9). There exists $c > 0$ depending on $h$ such that for $k \in \mathbb{N}$ with $k \geq cd^* n^{cd}$, one has

$$(f - f^*) + cd^* n^{2d} \| f \| k^{-1/c} \in P_k(h).$$
Note that in the case of polynomial optimization on the sphere (i.e., $h = \{ R - \| x \|_2^2 \}$ for some $R > 0$), one can take $c = 1$ in Lemma 2, as a consequence of the convergence result from [12].

Next, consider the hierarchy of semidefinite programs (SDP) indexed by $k \in \mathbb{N}$:

$$
\rho_k := \sup \{ \xi \in \mathbb{R} : f - \xi \in P_k(h) \}. \quad (10)
$$

For every $k \in \mathbb{N}$, the dual of $(10)$ reads

$$
\tau_k := \inf_{y \in \mathbb{R}^{[l_h]}} L_y(f) \quad \text{s.t. } M_k(y) \succeq 0; \ y_0 = 1 \quad \text{and} \quad M_{k-[h_j]}(h_j y) = 0, \ j \in [l_h]. \quad (11)
$$

By invoking Lemma 2 and Proposition 4, one obtains the convergence behavior of the sequence $(\rho_k)_{k \in \mathbb{N}}$ in the following result.

**Theorem 2** Let $f^*$ be as in (8) with $V(h) \neq \emptyset$ as in (9). Then:

1. For all $k \in \mathbb{N}$, $\rho_k \leq \rho_{k+1} \leq f^*$.
2. The sequence $(\rho_k)_{k \in \mathbb{N}}$ converges to $f^*$ with rate at least $O(k^{-1/c})$.
3. If the ideal $\langle h \rangle$ is real radical and the second-order sufficiency conditions (Definition 1) hold at every global minimizer of POP (8) then $\tau_k = \rho_k = f^*$ for some $k$ and (10) has an optimal solution, i.e., $f - f^* \in P_k(h)$.
4. If $V(h)$ defined as in (9) is finite, $\tau_k = \rho_k = f^*$ for some $k \in \mathbb{N}$ and both primal-dual (5)–(7) have optimal solutions. In this case, the flatness condition holds at order $k$.

with $V(h)$ in lieu of $S(g, h)$, zero duality gap as well as analogues of Proposition 1 and 3, also hold.

### 2.6 Spectral minimizations of SDP

Let $s, l, s^j \in \mathbb{N}^>0$, $j \in [l]$, be fixed such that $s = \sum_{j=1}^l s^j$. Let $S$ be the set of real symmetric matrices of size $s$ in a block diagonal form:

$$
X = \text{diag}(X_1, \ldots, X_l), \quad (12)
$$

such that $X_j$ is of size $s^j$, $j \in [l]$. Let $S^+$ be the set of all $X \in S$ such that $X \succeq 0$, i.e., $X$ has only nonnegative eigenvalues. Then $S$ is a Hilbert space with scalar product $\langle A, B \rangle = \text{trace}(B^T A)$ and $S^+$ is a self-dual cone.

Let us consider the following SDP:

$$
-\tau = \sup_{X \in S} \{ \langle C, X \rangle : A X = b, \ X \succeq 0 \}, \quad (13)
$$

where $A : S \to \mathbb{R}^m$ is a linear operator of the form

$$
A X = \{ \langle A_1, X \rangle, \ldots, \langle A_m, X \rangle \},
$$

$\Box$ Springer
with \( A_i \in \mathcal{S}, i \in [m], \) \( C \in \mathcal{S} \) is the cost matrix and \( b \in \mathbb{R}^m \) is the right-hand-side vector.

The dual of SDP (13) reads:

\[-\rho = \inf_z \left\{ b^T z : A^T z - C \succeq 0 \right\}, \tag{14}\]

where \( A^T : \mathbb{R}^m \to \mathcal{S} \) is the adjoint operator of \( A \), i.e., \( A^T z = \sum_{i=1}^m z_i A_i \).

The following assumption will be used in the next two sections:

**Assumption 1** Consider the following conditions:

1. Zero duality gap of primal-dual (13)–(14) holds, i.e., \( \tau = \rho \) and \( \tau \in \mathbb{R} \).
2. Primal attainability: SDP (13) has an optimal solution.
3. Dual attainability: SDP (14) has an optimal solution.
4. Constant trace property (CTP): There exists \( a > 0 \) such that

\[ \forall X \in \mathcal{S}, \quad A X = b \Rightarrow \text{trace}(X) = a. \tag{15} \]

5. Bounded trace property (BTP): There exists \( a > 0 \) such that

\[ \forall X \in \mathcal{S}, \quad A X = b \Rightarrow \text{trace}(X) \leq a. \tag{16} \]

In Assumption 1, conditions 1 and 5 (or conditions 1 and 4) imply condition 2. Indeed, if condition 5 holds, the feasible set of (13) is compact and if condition 1 holds, the feasible set of (13) is nonempty. Moreover, conditions 2 and 5 (or conditions 2 and 4) imply condition 1. Indeed, if conditions 2 and 5 hold, the set of optimal solutions of (13) is nonempty and bounded. Then Trnovska’s result [49, Corollary 1] yields condition 1.

**Example 1** We provide below two SDPs, showing first that condition 4 of Assumption 1 does not always imply the existence of a Slater point for primal problem (13), and next that condition 4 does not always imply conditions 1 and 2.

1. Let \( X = \text{diag}(x_1, x_2, x_3) \) be the matrix variable and let \( A X = b \) be the affine constraints encoding \( x_1 + x_2 = 1 \) and \( x_3 = 0 \). It implies that \( A X = b \Rightarrow \text{trace}(X) = 1 \), thus condition 4 holds. However, there is no Slater point for primal problem (13) in this case since every feasible solution is of the form \( \text{diag}(x_1, x_2, 0) \) which cannot be positive definite.

2. Let \( X = \text{diag}(x_1, x_2) \) be the matrix variable and let \( A X = b \) be the affine constraints encoding \( x_1 + x_2 = 1 \) and \( x_1 - x_2 = 3 \). (This case possibly happens for SOS relaxations of low order.) Although we get \( A X = b \Rightarrow \text{trace}(X) = 1 \) (i.e., condition 4 holds), the feasible set

\[ \{X = \text{diag}(x_1, x_2) : X \succeq 0, \ A X = b\} \tag{17} \]

is empty. It is due to the fact that the system: \( x_1 + x_2 = 1 \) and \( x_1 - x_2 = 3 \) implies \( \text{diag}(x_1, x_2) = \text{diag}(2, -1) \), which has trace 1 but is not positive semidefinite. Thus conditions 1 and 2 do not hold.
Remark 2 Condition 4 of Assumption 1 implies that the dual problem (14) has a Slater point. (However, Example 1 is showing that under this condition, the primal (13) has no Slater’s point and is not feasible in general.) So strong duality does always hold under condition 4, even though the optimal value could be both positive infinities. Note that conditions 4 and 5 do not assume the feasibility of the primal (13). And condition 1 is not just a zero duality gap, but also, both values are finite. Here a zero duality gap could mean both values are positive infinities.

Remark 3 If condition 5 of Assumption 1 holds, by adding a slack variable $y$ and noting $Y = \text{diag}(X, y)$, we obtain an equivalent SDP of (13) as follows:

$$-\tau = \sup_{Y \in \hat{S}} \left\{ \langle \hat{C}, Y \rangle : \langle \hat{A}_i, Y \rangle = b_i, \ Y \succeq 0, \ \text{trace}(Y) = a \right\},$$  

(18)

where $\hat{S} = \{\text{diag}(X, y) : X \in S, \ y \in \mathbb{R}\}$, $\hat{C} = \text{diag}(C, 0)$ and $\hat{A}_i = \text{diag}(A_i, 0)$. Obviously, SDP (18) has CTP.

2.6.1 SDP with constant trace property (CTP)

Recall that $\lambda_1(A)$ stands for the largest eigenvalue of a real symmetric matrix $A$.

Lemma 3 Let conditions 1 and 4 of Assumption 1 hold and let $\varphi : \mathbb{R}^m \to \mathbb{R}$ be the function:

$$z \mapsto \varphi(z) := a\lambda_1(C - A^\top z) + b^\top z.$$  

(19)

Then:

$$-\tau = \inf_{z} \{ \varphi(z) \ z \in \mathbb{R}^m \}.$$  

(20)

Moreover if condition 3 of Assumption 1 holds, i.e., SDP (14) has an optimal solution then problem (20) has an optimal solution.

The proof of Lemma 3 is postponed to Appendix A.1.1.

Next, we describe Algorithm 1 to solve SDP (13), which is based on nonsmooth first-order optimization methods (e.g., LMBM [15, Algorithm 1]). As shown later on in Sect. 4, this algorithm works well in almost all cases and with significantly lower computational cost when compared to the (currently fastest) SDP solver Mosek 9.1.

Algorithm 1 SDP-CTP

Input: SDP (13) with unknown optimal value and optimal solution;
method (T) for solving convex nonsmooth unconstrained optimization problems (NSOP).
Output: optimal value $-\tau$ and optimal solution $X^*$ of SDP (13).

1: Compute the optimal value $-\tau$ and an optimal solution $\bar{z}$ of the NSOP (20) by using method (T);
2: Compute a normalized eigenvector $u$ corresponding to $\lambda_1(C - A^\top \bar{z})$ and set $X^* = a \bar{u}u^\top$. 
For $X \in S$, the Frobenius norm of $X$ is defined by $\|X\|_F := \sqrt{\langle X, X \rangle}$. We denote by $\|A\|$ the operator norm of $A$, i.e., $\|A\| := \max_{X \in S} \|AX\|_2 / \|X\|_F$.

**Remark 4** Before running Algorithm 1, we scale the problem’s input as follows: $\|C\|_F = \|A\| = a = 1$ and $\|A_1\|_F = \cdots = \|A_m\|_F$.

The fact that Algorithm 1 is well-defined under certain conditions is a corollary of Lemmas 3, 8 and 9.

**Corollary 1** Let conditions 1 and 4 of Assumption 1 hold. Assume that the method $(T)$ is globally convergent for NSOP (20) (e.g., $(T)$ is LMBM). Then output $-\tau$ of Algorithm 1 is well-defined. Moreover, if condition 3 of Assumption 1 holds, the vector $\bar{z}$ mentioned at Step 1 of Algorithm 1 exists. In this case, if $\lambda_1(C - A^T \bar{z})$ has multiplicity 1, the output $X^*$ of Algorithm 1 is well-defined.

When $\lambda_1(C - A^T \bar{z})$ has multiplicity larger than 1, one can obtain a dual matrix $G$ corresponding to an optimal solution $X^*$ of SDP (13) as in the following corollary:

**Corollary 2** Let conditions 1 and 4 of Assumption 1 hold. Let $\bar{z}$ be an optimal solution of the NSOP (20). Define

$$U := C - A^T \bar{z},$$
$$G := \lambda_1(U)I - U,$$

where $I$ is the identity matrix. Then $G$ is positive semidefinite and satisfies

$$G = A^T z^* - C,$$

for some optimal solution $z^*$ of (14).

**Proof** It is not hard to prove that $G \succeq 0$. Let us prove the other statement. By using Farkas’ lemma, there exists $y \in \mathbb{R}^m$ such that $A^T y = I$ and $y^T b = a$ (see [16, Sect. 2]). Then $G = \lambda_1(U)A^T y - C + A^T \bar{z} = A^T (\lambda_1(U)y + \bar{z}) - C = A^T z^* - C$, where $z^* := \lambda_1(U)y + \bar{z}$. Since $b^T z^* = \lambda_1(U)b^T y + b^T \bar{z} = \varphi(\bar{z}) = -\tau$, $z^*$ is an optimal solution of (14).

2.6.2 Largest eigenvalue computation

Step 1 of Algorithm 1 requires the largest eigenvalue and corresponding eigenvectors of $C - A^T \bar{z}$ to evaluate the function $\varphi$ (resp. $\psi$) and a subgradient of the subdifferential $\partial \varphi$ (resp. $\partial \psi$) given in Proposition 8 (resp. Proposition 9) at $z$. Fortunately, solving the eigenvalue problem for $C - A^T \bar{z} \in S$ can be done on every block of $C - A^T \bar{z}$. Indeed, with $X \in S$ as in (12),

$$\lambda(X) = \lambda(X_1) \cup \cdots \cup \lambda(X_t),$$

where $\lambda(A)$ is the set of all eigenvalues $\lambda_1(A) \geq \cdots \geq \lambda_t(A)$ for every real symmetric matrix $A$ of size $t$. In particular,

$$\lambda_1(X) = \max\{\lambda_1(X_1), \ldots, \lambda_1(X_t)\}.$$
If \( u \in \mathbb{R}^{s(j)} \) is an eigenvector of \( X_j \) corresponding to the eigenvalue \( \lambda_i(X_j) \) for some \( i \in [s(j)] \) and \( j \in [l] \), by adding zero entries in \( u \),

\[
\tilde{u} = \left( 0_{\mathbb{R}^{s(1)+\ldots+s(j-1)}}, u, 0_{\mathbb{R}^{s(j+1)+\ldots+s(l)}} \right)
\]

is an eigenvector of \( X = \text{diag}(X_1, \ldots, X_l) \) corresponding to \( \lambda_i(X_j) \).

The interested reader can refer to [33, 46] to solve largest eigenvalue problems of symmetric matrices of large sizes.

**Remark 5** Let conditions 1, 2 and 5 of Assumption 1 hold. We keep all notation from Remark 3. By applying Lemma 3 for SDP (18) with CTP, one has

\[
-\tau = \inf_{z} \left\{ a\lambda_1 \left( \hat{C} - \hat{A}^\top z \right) + b^\top z \mid z \in \mathbb{R}^m \right\},
\]

where \( \hat{A}^\top z = \sum_{i=1}^m z_i \hat{A}_i \). Note that \( \hat{C} - \hat{A}^\top z = \text{diag}(C - A^\top z, 0) \). It implies that \( \lambda_1(\hat{C} - \hat{A}^\top z) = \max\{\lambda_1(C - A^\top z), 0\} \). Thus, (23) can be rewritten as

\[
-\tau = \inf_{z} \left\{ a \max\{\lambda_1(C - A^\top z), 0\} + b^\top z \mid z \in \mathbb{R}^m \right\}.
\]

In the next section, we consider the spectral formulation (24) introduced by Ding et al. in [11, Sect. 6].

### 2.6.3 SDP with bounded trace property (BTP)

In the last subsection, we have seen that SDPs with CTP can be solved efficiently with first-order methods. Similar results can be obtained for the larger class of SDPs with the weaker bounded trace property (BTP). In particular the semidefinite relaxations of the Moment-SOS hierarchy associated with a POP on a compact semialgebraic set have the BTP. So in principle there is no need to add auxiliary “slack” variables to obtain an equivalent CTP-POP, as shown in Remark 3. However, numerical experiments of Sect. 4 suggest that the CTP is a highly desirable property that justifies addition of auxiliary variables.

The analogue of Lemma 3 for BTP reads:

**Lemma 4** Let conditions 1, 2 and 5 of Assumption 1 hold, and let \( \psi : \mathbb{R}^m \to \mathbb{R} \) be the function:

\[
\psi(z) := a \max\left\{ \lambda_1 \left( C - A^\top z \right), 0 \right\} + b^\top z.
\]

Then

\[
-\tau = \inf_{z} \left\{ \psi(z) \mid z \in \mathbb{R}^m \right\}.
\]

Moreover if condition 3 of Assumption 1 holds, then problem (26) has an optimal solution.
The proof of Lemma 4 is postponed to Appendix A.1.2.

We next describe Algorithm 2 to solve SDP (13). As Algorithm 1, it is also based on nonsmooth optimization methods such as LMBM.

Algorithm 2 SDP-BTP

**Input:** SDP (13) with unknown optimal value and optimal solution; method (T) for solving convex NSOP.

**Output:** optimal value $-\tau$ and optimal solution $X^\star$ of SDP (13).

1: Compute the optimal value $-\tau$ and an optimal solution $\bar{z}$ of NSOP (20) by using method (T);
2: Compute $\lambda_1(C - A^\top \bar{z})$ and a corresponding normalized eigenvector $u$;
3: Let $\bar{\xi} > 0$ such that

$$
\bar{\xi} = \begin{cases} 
0 & \text{if } \lambda_1(C - A^\top \bar{z}) < 0, \\
\zeta & \text{if } \lambda_1(C - A^\top \bar{z}) = 0, \\
a & \text{otherwise}
\end{cases}
$$

for some $\zeta \in [0, 1]$ such that $X^\star = \bar{\xi} uu^\top$ satisfies $AX^\star = b$.

The next result is a consequence of Lemma 4, 10 and 11.

**Corollary 3** Let conditions 1, 2 and 5 of Assumption 1 hold. Assume that method (T) is globally convergent for NSOP (26) (e.g., (T) is LMBM). Then the output $-\tau$ of Algorithm 2 is well-defined. Moreover, if condition 3 of Assumption 1 holds, the vector $\bar{z}$ from Step 1 of Algorithm 2 exists. In this case, if $\lambda_1(C - A^\top \bar{z})$ has multiplicity 1, the output $X^\star$ of Algorithm 2 is well-defined.

3 Application to polynomial optimization

We consider the following POP:

$$
f^\star := \inf \{ f(x) : x \in S(g, h) \},
$$

where $S(g, h)$ is defined as in (4) with $l_g$ (resp. $l_h$) being the number of inequality (resp. equality) constraints. Assume that $S(g, h) \subset B_n^R$.

**Remark 6** By setting $X := \text{diag}(M_k(y), M_{k-[g_1]}(g_1 y), \ldots, M_{k-[g_l]}(g_l y))$ and using the upper bound $\text{trace}(X) \leq \tilde{a}_k$ with

$$
\tilde{a}_k := R^k \left( \binom{n+k}{n} + \sum_{i=1}^{l_g} \|g_i\|_1 \binom{n+k-[g_i]}{n} \right),
$$

SDP (7) can be converted to an equivalent SDP with BTP, thanks to the absolute upper bound for each moment variable $|y_\alpha| \leq R^{|\alpha|/2}$, $\alpha \in \mathbb{N}^n$. In principle, we can solve this SDP by applying directly Algorithm 2. However, in our experiments presented in Sect. 4 this method is not only inefficient but also provides output with low accuracy.
In order to overcome the accuracy issue mentioned in Remark 6, we convert every POP to a CTP-POP (i.e., a new POP formulation with CTP) by adding slack variables associated with inequality constraints. In the sequel, we consider three particular cases: equality constrained POPs on a sphere in Sect. 3.1, constrained POPs with single inequality (ball) constraint in Sect. 3.2, and constrained POPs on a ball in Sect. 3.3.

### 3.1 Equality constrained POPs on a sphere

Assume that \( l_g = 0 \) and \( h_1 = \bar{R} - \|x\|_2^2 \). Note that \( \|x\|_2^2 = x_1^2 + \cdots + x_n^2 \) is a quadratic polynomial. In this case, we consider equality constrained POPs on a sphere, presented in Sect. 2.5. We propose to reduce SDP (11) to an NSOP. For each \( k \in \mathbb{N} \), let \( (\theta_{k,\alpha})_{\alpha \in \mathbb{N}_k^n} \) be a sequence of positive real numbers such that

\[
\left(1 + \|x\|_2^2\right)^k = \sum_{\alpha \in \mathbb{N}_k^n} \theta_{k,\alpha} x_\alpha^2, \tag{30}
\]

and define the diagonal matrix

\[
P_{n,k} := \text{diag} \left( \left( \theta_{k,\alpha}^{1/2} \right)_{\alpha \in \mathbb{N}_k^n} \right). \tag{31}
\]

For every \( k \in \mathbb{N} \), since \( P_{n,k} > 0 \), SDP (11) is equivalent to SDP:

\[
\tau_k = \inf_{y \in \mathbb{R}^{\left\lfloor \frac{n}{2} \right\rfloor}} L_y(f) \quad \text{s.t.} \quad y_0 = 1; \quad P_{n,k} M_k(y) P_{n,k} \succeq 0, \quad M_{k-[h_j]} (h_j y) = 0, \quad j \in [l_h]. \tag{32}
\]

For every \( k \in \mathbb{N} \), let \( a_k := (\bar{R} + 1)^k \). We will use the following lemma:

**Lemma 5** For all \( k \in \mathbb{N} \),

\[
\text{trace}(P_{n,k} M_k(y) P_{n,k}) = a_k.
\]

**Proof** Let \( k \in \mathbb{N} \) be fixed. From \( M_{k-1}((\bar{R} - \|x\|_2^2) y) = 0 \), \( L_y(\rho(\bar{R} - \|x\|_2^2)) = 0 \), for every \( \rho \in \mathbb{R}[x]_{2(k-1)} \). For every \( r \in \mathbb{N} \leq k-1 \), by choosing \( \rho = \|x\|_2^{2r} \),

\[
L_y \left( \|x\|_2^{2(r+1)} \right) = -L_y \left( \|x\|_2^{2r} \left( \bar{R} - \|x\|_2^2 \right) \right) + \bar{R} L_y \left( \|x\|_2^{2r} \right) = \bar{R} L_y \left( \|x\|_2^{2r} \right).
\]
By induction, \( L_y(\|x\|_2^{2r}) = \bar{R} L_y(\|x\|_2^{2(r-1)}) = \cdots = \bar{R}^r L_y(\|x\|_2^{2 \times 0}) = \bar{R}^r y_0 = \bar{R}^r \), for every \( r \in \mathbb{N} \leq k \). Thus,

\[
\begin{align*}
\text{trace}(P_{n,k} M_k(y) P_{n,k}) &= \sum_{\alpha \in \mathbb{N}_k^n} \theta_{k,\alpha}^{1/2} \gamma_{2,\alpha}^{1/2} = L_y((1 + \|x\|_2^2)^k) = L_y \left( \sum_{r=0}^{k} \binom{k}{r} \|x\|_2^{2r} \right) \\
&= \sum_{r=0}^{k} \binom{k}{r} \bar{R}^r (\|x\|_2^2)^r = \sum_{r=0}^{k} \binom{k}{r} \bar{R}^r = (\bar{R} + 1)^k.
\end{align*}
\]

\( \square \)

For each \( k \in \mathbb{N} \), we denote by \( S_k \) the set of symmetric matrices of size \( \omega_k = \{ n \} \) and let \( (A, B) = \text{trace}(B^T A) \) be the usual scalar product on \( S_k \). For every \( k \in \mathbb{N} \), letting \( X = P_{n,k} M_k(y) P_{n,k} \), \( \text{(32)} \) can be written in the form:

\[
-\tau_k = \sup_{X \in S_k} \{ \langle C_k, X \rangle : A_k X = b_k, \ X \succeq 0 \}, \quad \text{(33)}
\]

where \( A_k : S_k \to \mathbb{R}^{m_k} \) is a linear operator of the form

\[
A_k X = \left[ \langle A_{k,1}, X \rangle, \ldots, \langle A_{k,m_k}, X \rangle \right],
\]

with \( A_{k,i} \in S_k, i \in [m_k] \), \( C_k \in S_k \) is the cost matrix and \( b_k \in \mathbb{R}^{m_k} \) is the right-hand-side vector. Appendix A.2 describes how to reduce SDP (31) to the form (33).

For every \( k \in \mathbb{N} \), the dual of SDP (33) reads:

\[
-\rho_k = \inf_z \left\{ b_k^T z : A_k^T z - C_k \succeq 0, \right\} \quad \text{(34)}
\]

where \( A_k^T : \mathbb{R}^{m_k} \to S_k \) is the adjoint operator of \( A_k \), i.e., \( A_k^T z = \sum_{i=1}^{m_k} z_i A_{k,i} \).

From Lemma 5 and since \( h_1 = \bar{R} - \|x\|_2^2 \), it implies that for every \( k \in \mathbb{N} \),

\[
\forall X \in S_k, \ A_k X = b_k \Rightarrow \text{trace}(X) = a_k. \quad \text{(35)}
\]

We guarantee zero duality gap, primal attainability, and dual attainability for primal-dual (33)–(34) in the following proposition:

**Proposition 5** \( \) Let \( f^* \) be as in (8). Then:

1. Zero duality gap holds for primal-dual (33)–(34) for large enough \( k \in \mathbb{N} \).
2. SDP (33) has an optimal solution for large enough \( k \in \mathbb{N} \).
3. Assume that one of the following two conditions holds:

(a) $\langle h \rangle$ is real radical and the second-order sufficiency conditions (Definition 1) hold at every global minimizer of (8);
(b) $V(h)$ is finite.

Then SDP (34) has an optimal solution for large enough $k \in \mathbb{N}$. In this case, $\tau_k = \rho_k = f^*$.

Proof Since (11) (resp. (10)) and (33) (resp. (34)) are equivalent, the first and second statements follow from Proposition 1. The third statement is due to Theorem 2. \hfill $\square$

By replacing $(A_k, A_{k,i}, b_k, C_k, S_k, \omega_k, m_k, \tau_k, \rho_k, a_k)$ by $(A, A_i, b, C, S, s, m, \tau, \rho, a)$, primal-dual (33)–(34) becomes primal-dual (13)–(14), we then go back to Sect. 2.6 with $l = 1$.

We illustrate the conversion from SDP (11) to SDP (33) in the following example.

Example 2 Consider a simple example of POP (8) with $n = 1$:

$$-1 = \inf\{x : 1 - x^2 = 0\}.$$  

Then the second order moment relaxation ($k = 2$) has the form:

$$\tau_2 = \inf_y y_1$$

s.t. $$\begin{bmatrix} y_0 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{bmatrix} \succeq 0, \begin{bmatrix} y_0 - y_2 & y_1 - y_3 \\ y_1 - y_3 & y_2 - y_4 \end{bmatrix} = 0, \ y_0 = 1.$$  

It can be rewritten as

$$\tau_2 = \inf_y y_1$$

s.t. $$\begin{bmatrix} 1 & y_1 & 1 \\ y_1 & 1 & y_1 \\ 1 & y_1 & 1 \end{bmatrix} \succeq 0,$$

by removing equality constraints. Obviously, the positive semidefinite matrix of this form has trace 3.

In a different way, according to Appendix A.2, let us note

$$X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_0 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

to obtain

$$-\tau_2 = \sup_{X \in S_2} \{\langle C, X \rangle : \langle A_i, X \rangle = b_i, i \in [5], X \succeq 0\},$$
A hierarchy of spectral relaxations for polynomial...

where \( b_1 = \cdots = b_4 = 0, b_5 = 1 \) and

\[
C = -\sqrt{2} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},
A_1 = \sqrt{2} \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix},
A_2 = \frac{1}{2} \begin{bmatrix} 2 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix},
A_3 = \sqrt{2} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix},
A_4 = \frac{1}{2} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -2 \end{bmatrix},
A_5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.
\]

Remark that for any \( X \in S_2 \),

\[
(\langle A_i, X \rangle = b_i, \ i \in [5]) \Rightarrow \text{trace}(X) = 4.
\]

Next, we present an alternative iterative method, stated in Algorithm 3, to solve (8), based on nonsmooth optimization methods, e.g., LMBM. It performs well in practice for most cases and with significantly lower computational cost when compared to the (currently fastest) SDP solver Mosek 9.1.

**Algorithm 3 SpectralPOP-CTP**

**Input:** POP (8) with unknown optimal value \( f^* \) and optimal solutions; method (D) for solving SDP with CTP.

**Output:** increasing real sequence \( (\tau_k)_{k \in \mathbb{N}} \) and \( x^* \in \mathbb{R}^n \).

1: for \( k \in \mathbb{N} \) do
2: Compute the optimal value \( -\tau_k \) and an optimal solution \( X^* \) of SDP (33) by using method (D);
3: Set \( M_k(y^*) := P_{-1}^{n,k} X^* P_{-1}^{n,k} \) (relying on (32)) and extract an atom \( x^* \) by using Henrion-Lasserre’s algorithm in [18] from \( M_k(y^*) \);
4: If \( x^* \) exists, set \( \tau_{k+j} = \tau_k, \ j \in \mathbb{N}^+, \) and terminate.

Note that one can choose method (D) in Algorithm 3 as Algorithm 1 with LMBM solver or SketchyCGAL.

**Remark 7** In practice, to verify that an atom \( x^* \) extracted in Step 3 of Algorithm 3 is an approximate optimal solution of POP (8), with given \( \varepsilon \in (0, 1) \), we check the following inequalities:

\[
|f(x^*) - \tau_k| \leq \varepsilon \|f\|_{\max} \text{ and } |h_j(x^*)| \leq \varepsilon \|h_j\|_{\max}, \ j \in [I_g],
\]

where \( \|p\|_{\max} := \max_{\alpha} |p_\alpha| \) for any \( p \in \mathbb{R}[x] \). We take \( \varepsilon = 0.01 \) for the experiments in Sect. 4.

Following Proposition 3, Corollary 1 and Proposition 5, we obtain the following corollary:

**Corollary 4** (i) Sequence \( (\tau_k)_{k \in \mathbb{N}} \) of Algorithm 3 is well defined and \( \tau_k \uparrow f^* \) as \( k \to \infty \).

(ii) Assume that condition (a) or (b) of Proposition 5.3 holds. If there exists an optimal solution \( y^* \) of SDP (11) for some order \( k \in \mathbb{N} \) such that the flat extension condition
holds, $x^*$ exists at the $k$-th iteration of Algorithm 3. In this case, if $X^*$ in the first step of Algorithm 3 is well-defined, Algorithm 3 terminates at the $k$-th iteration, $x^*$ is an optimal solution of POP (8) and $f^* = \tau_k$.

In Corollary 4, the flat extension condition implies that the SOS problem (10) has an optimal solution (due to [25, Theorem 3.4 (b)] and $\tau_k = \rho_k$), so that SDP (34) has an optimal solution. In this case, $X^*$ exists, which in turn implies the existence of $x^*$.

In Step 4 of Algorithm 3, if the atom $x^*$ exists, then we do not need to increase the relaxation order $k$. It is due to the fact that $f^* = \tau_k \leq \tau_{k+1} \leq \cdots \leq f^*$.

Remark 8 When Algorithm 1 with LMBM solver is used for method (D) in Algorithm 3, we have the following cases:

1. If the SDP relaxation (11) is exact then the value is $f^*$ and one indeed may expect that generically the moment matrix is rank-one, which will yield $X^* = uu^T$ for some $u$. Thus, $X^*$ in the first step of Algorithm 3 is well-defined.

2. If the SDP relaxation (11) is not exact then we only use the relaxation value as a (supposedly accurate) lower bound on the global minimum $f^*$.

Remark 9 In practice, we use the following heuristic extraction algorithm when method (D) in Algorithm 3 is Algorithm 1 with LMBM solver:

1. Obtain a dual matrix $\tilde{G}$ corresponding to an optimal solution $X^*$ of SDP (33) based on Corollary 2;

2. Set $\tilde{G} := P_{n,k}GP_{n,k}$;

3. Obtain an atom $x^*$ by using the extraction algorithm of Henrion and Lasserre in [18], where the matrix $V$ in [18, (6)] is taken such that the columns of $V$ form a basis of the null space $\{u \in \mathbb{R}^{\omega_k} : \tilde{G}u = 0\}$;

4. Verify that $x^*$ is an approximate optimal solution of POP (8) as in Remark 7.

This heuristic extraction algorithm performs practically well when the moment matrices are not rank-one. Note that $\tilde{G}$ obtained in Step 2 is a Gram matrix corresponding to some moment matrix $M_k(y^*)$. Step 3 is well-defined when the complementary slackness, i.e., $\tilde{G}M_k(y^*) = 0$, and the strict complementarity, i.e., rank $\tilde{G} + \text{rank } M_k(y^*) = \omega_k$, hold (see [11, Sect. 1.3]). In this case, the range of $M_k(y^*)$, which is the linear span of the columns of $V$ in [18, (6)], is identical with the null space of $\tilde{G}$.

In the two following subsections, we consider POPs on general compact sets as stated in Sect. 2.4.

### 3.2 Constrained POPs with single inequality (ball) constraint

Assume that $l_g = 1$ and $g_1 = R - \|x\|_2^2$. In this case, $g = \{R - \|x\|_2^2\}$. Let us show that POP (3) can be reduced to an equality constrained POP on a sphere. By adding one slack variable $x_{n+1}$, the inequality constraint $R - \|x\|_2^2 \geq 0$ can be rewritten as an equality constraint $R - \|x\|_2^2 - x_{n+1}^2 = 0$ and so

$$f^* := \inf \{ f(x) \ (x, x_{n+1}) \in V(\tilde{h}) \},$$

(36)
where \( \bar{h} := h \cup \{ R - \| x \|_2^2 - x_{n+1}^2 \} \subset \mathbb{R}[x, x_{n+1}] \).

Notice that:

– If \( \tilde{x}^* = (x^*, x_{n+1}^*) \) is an optimal solution of POP (36), \( x^* \) is an optimal solution of POP (3).

– Conversely, if \( x^* \) is an optimal solution of POP (3), then

\[
\tilde{x}^* := \left( x^*, \sqrt{R - \| x^* \|_2^2} \right)
\]

is an optimal solution of POP (36).

Let us define \( \bar{n} := n + 1 \) and \( \tilde{x} := (x, x_{n+1}) \) to ease notation. For every \( k \in \mathbb{N} \), consider the order \( k \) moment relaxation of (36):

\[
\tilde{\tau}_k = \inf_{y \in \mathbb{R}[\bar{x}]} L_y(f) \quad \text{s.t.} \quad y_0 = 1, \; M_k(y) \succeq 0, \; M_{k-1}(R - \| \bar{x} \|_2^2) y = 0, \; M_{k-\lceil h_j \rceil}(h_j y) = 0, \; j \in \lbrack l_h \rbrack.
\]  

(37)

The corresponding dual SOS problem indexed by \( k \in \mathbb{N} \) reads:

\[
\tilde{\rho}_k := \sup \{ \xi \in \mathbb{R} \mid f - \xi \in P_k(\bar{h}) \},
\]

(38)

where \( P_k(\bar{h}) \) is the truncated preorder of all polynomials of the form

\[
\sigma_0 + \psi_0 \left( R - \| \bar{x} \|_2^2 \right) + \sum_{j=1}^{l_h} \psi_j h_j,
\]

with \( \sigma_0 \in \Sigma[\bar{x}]_k, \psi_0 \in \mathbb{R}[\bar{x}]_{2(k-1)}, \) and \( \psi_j \in \mathbb{R}[\bar{x}]_{2(k-\lceil h_j \rceil)}, \; j \in \lbrack l_h \rbrack. \)

The following lemma will be used later on:

**Lemma 6** If \( f - f^* \in Q_k(g, h) \) for some \( k \in \mathbb{N} \) then \( f - f^* \in P_k(\bar{h}) \).

**Proof** By assumption, there exist polynomials \( \sigma_0 \in \Sigma[x]_k, \sigma_1 \in \Sigma[x]_{k-1}, \) and \( \psi_j \in \mathbb{R}[x]_{2(k-\lceil h_j \rceil)}, \; j \in \lbrack l_h \rbrack \) such that

\[
f - f^* = \sigma_0 + \sigma_1 \left( R - \| x \|_2^2 \right) + \sum_{j=1}^{l_h} \psi_j h_j = \sigma_0 + \sigma_1 x_{n+1}^2
\]

\[
+ \sigma_1 \left( R - \| \bar{x} \|_2^2 \right) + \sum_{j=1}^{l_h} \psi_j h_j,
\]

yielding the result. \( \Box \)
Zero duality gap, primal attainability, and dual attainability for primal-dual (37)–(38) are guaranteed in the following proposition:

**Proposition 6** Let $f^*$ be as in (3) with $g = \{ R - \| x \|^2 \}$. Then:

1. Zero duality gap holds for primal-dual (37)–(38) for large enough $k \in \mathbb{N}$.
2. SDP (37) has an optimal solution for large enough $k \in \mathbb{N}$.
3. Assume that one of the following two conditions holds:
   
   (a) $Q(g, h)$ is Archimedean, the ideal $(h)$ is real radical, and the second-order sufficiency conditions (Definition 1) hold at every global minimizer of POP (3);
   
   (b) $V(h)$ is finite.

Then SDP (38) has an optimal solution for large enough $k \in \mathbb{N}$. In this case, $\bar{\tau}_k = \bar{\rho}_k = f^*$.

**Proof** The first and second statement follow from Proposition 1, after replacing $S(g, h)$ by $V(\bar{h})$. The third statement is due to Proposition 4 and Lemma 6. \qed

For every $k \in \mathbb{N}$, according to Lemma 5, if $M_{k-1}((R - \| \bar{x} \|^2) y) = 0$ and $y_0 = 1$, then one has

$$\text{trace}(P_{\bar{n}, k} M_k(y) P_{\bar{n}, k}) = (R + 1)^k,$$

where $P_{\bar{n}, k}$ is defined as in (30) after replacing $n$ by $\bar{n}$. Thus SDP (37) has the CTP. We now do a similar process as in Sect. 3.1.

Next, we present an iterative method, stated in Algorithm 4, to solve (3) with $g = \{ R - \| x \|^2 \}$, based on a nonsmooth optimization method such as LMBM.

**Algorithm 4 SpectralPOP-CTP-WithSingleBallConstraint**

**Input:** POP (3) with $g = \{ R - \| x \|^2 \}$, unknown optimal value $f^*$ and optimal solutions; method (D) for solving SDP with CTP.

**Output:** increasing real sequence $(\bar{\tau}_k)_{k \in \mathbb{N}}$ and $x^* \in \mathbb{R}^n$.

1: for $k \in \mathbb{N}$ do
2: Compute the optimal value $-\bar{\tau}_k$ and an optimal solution $y^*$ of SDP (37) with CTP (39) by using method (D);
3: Extract an atom $\bar{x}^* = (x^*, x_{n+1}^*)$ by using Henrion-Lasserre’s algorithm in [18] from $M_k(y^*)$;
4: If $\bar{x}^*$ exists, set $\bar{\tau}_{k+j} = \bar{\tau}_k$, $j \in \mathbb{N}^>0$, and terminate.

Note that one can choose method (D) in Algorithm 4 as Algorithm 1 with LMBM solver or SketchyCGAL.

Following Proposition 3, Corollary 1 and Proposition 6, we obtain the following corollary:

**Corollary 5** (i) Sequence $(\bar{\tau}_k)_{k \in \mathbb{N}}$ of Algorithm 4 is well defined and $\bar{\tau}_k \uparrow f^*$ as $k \to \infty$.

(ii) Assume that condition (a) or (b) of Proposition 6.3 holds. If there exists an optimal solution $y^*$ of SDP (37) for some order $k \in \mathbb{N}$ such that the flat extension condition holds, $x^*$ exists at the $k$-th iteration of Algorithm 4. In this case, if Algorithm 4 terminates at the $k$-th iteration, $x^*$ is an optimal solution of POP (3) and $f^* = \bar{\tau}_k$. 

\[ \hfill \]
3.3 Constrained POPs on a ball

Assume that $l_g > 1$ and $g_1 = R - \|x\|_2^2$. Let us show that POP (3) can be reduced to an equality constrained POP on a sphere. After adding $l_g$ slack variables $x_{n+i}$, $i \in [l_g]$, every inequality constraint $g_i(x) \geq 0$ can be rewritten as an equality constraint $g_i(x) = x_{i+n}^2$ and so

$$f^* := \inf \left\{ f(x) : (x, x_{n+1}, \ldots, x_{n+l_g}) \in V(\hat{h}) \right\},$$

where $\hat{h} := h \cup \{ g_i - x_{i+n}^2 : i \in [l_g] \} \subset \mathbb{R}[x, x_{n+1}, \ldots, x_{n+l_g}]$.

Let us take upper bounds $b_i := \sup\{g_i(x) : x \in S([g_1], h), i \in [l_g] \}$. For every $i \in [l_g]$, the bound $b_i$ can be computed by solving the order $k$ moment relaxation:

$$-b_i = \inf_{y \in \mathbb{R}^{n+1}} L_y(-g_i)$$

s.t.

$$y_0 = 1, \quad M_k(y) \geq 0,$$

$$M_{k-1}((R - \|x, x_{n+1}\|_2^2) y) = 0, \quad M_{k-[h_j]}(h y) = 0, \quad j \in [l_h],$$

based on the spectral minimization method presented in the previous section.

For every $(x, x_{n+1}, \ldots, x_{n+l_g}) \in V(\hat{h})$, $x \in S(g, h)$ and $x_{n+i}^2 = g_i(x) \leq b_i$, $i \in [l_g]$, since $S(g, h) \subset S([g_1], h)$. Therefore

$$\|x\|_2^2 + \sum_{i=1}^{l_g} x_{i+n}^2 \leq \bar{R} \quad \text{with} \quad \bar{R} := R + \sum_{i=1}^{l_g} b_i.$$  \hspace{1cm} (41)

Equivalently $V(\hat{h}) \subset \mathcal{B}_{\bar{R}}^{n+l_g}$ and after adding one more slack variable $x_{n+l_g+1}$:

$$f^* := \inf\{ f(x) : \bar{x} \in V(\hat{h}) \},$$  \hspace{1cm} (42)

where $\bar{x} := (x, x_{n+1}, \ldots, x_{n+l_g+1})$ and

$$\hat{h} := h \cup \{ \bar{R} - \|\bar{x}\|_2^2 \} = h \cup \{ g_i - x_{i+n}^2 : i \in [l_g] \} \cup \{ \bar{R} - \|\bar{x}\|_2^2 \} \subset \mathbb{R}[\bar{x}].$$

Notice that:

- If $\bar{x}^* = (x^*, x_{n+1}^*, \ldots, x_{n+l_g+1}^*)$ is an optimal solution of POP (42), $x^*$ is an optimal solution of POP (3).
- Conversely, if $x^*$ is an optimal solution of POP (3), then

$$\bar{x}^* := \left( x^*, \sqrt{g_1(x^*)}, \ldots, \sqrt{g_{l_g}(x^*)}, \sqrt{\bar{R} - \sum_{i=1}^{l_g} g_i(x^*) - \|x^*\|_2^2} \right)$$
is an optimal solution of POP (42).

Note $\bar{n} := n + l_g + 1$ for simplicity. For every $k \in \mathbb{N}$, consider the order $k$ moment relaxation of (42):

$$\bar{\tau}_k = \inf_{y \in \mathbb{R}^{[\bar{n}]}} L_y(f)$$

s.t. $y_0 = 1$, $M_k(y) \succeq 0,$ $M_{k-[g_i]} \left( \left( g_i - x_{n+i}^2 \right) y \right) = 0$, $i \in [l_g],$

$$M_{k-\left( \{ h_j \} \right)} \left( h_j y \right) = 0,$$ $j \in [l_h].$

The corresponding dual SOS problem indexed by $k \in \mathbb{N}$ reads:

$$\bar{\rho}_k := \sup \{ \xi \in \mathbb{R} \mid f - \xi \in P_k(\tilde{h}) \},$$

where $P_k(\tilde{h})$ is the truncated preordering of all polynomials of the form

$$\sigma_0 + \sum_{i=1}^{l_g} \psi_i \left( g_i - x_{n+i}^2 \right) + \sum_{j=1}^{l_h} \psi_{l_k+1+j} h_j$$

with $\sigma_0 \in \Sigma[\tilde{x}], \psi_i \in \mathbb{R}[\tilde{x}]_{2(k-[g_i])}, \psi_{l_k+1} \in \mathbb{R}[\tilde{x}]_{2(k-1)}$, and $\psi_{l_k+1+j} \in \mathbb{R}[\tilde{x}]_{2(k-[h_j])}, j \in [l_h].$

We will use the following lemma later on:

**Lemma 7** If $f - f^* \in Q_k(g, h)$ for some $k \in \mathbb{N}$ then $f - f^* \in P_k(\tilde{h})$.

**Proof** By assumption, there exist polynomials $\sigma_0 \in \Sigma[x], \sigma_i \in \Sigma[x]_{k-[g_i]}, i \in [l_g]$, and $\psi_j \in \mathbb{R}[x]_{2(k-[h_j])}, j \in [l_h]$ such that

$$f - f^* = \sigma_0 + \sum_{i=1}^{l_g} \sigma_i g_i + \sum_{j=1}^{l_h} \psi_j h_j.$$

It implies that

$$f - f^* = \sigma_0 + \sum_{i=1}^{l_g} \sigma_i x_{i+n}^2 + \sum_{i=1}^{l_g} \sigma_i (g_i - x_{i+n}^2) + 0 \times \left( \tilde{R} - \| \tilde{x} \|_2^2 \right) + \sum_{j=1}^{l_h} \psi_j h_j,$$

yielding the result. \qed

Zero duality gap, primal attainability, and dual attainability for primal-dual (43)–(44) are guaranteed in the following proposition:

**Proposition 7** Let $f^*$ be as in (3). Then:

1. Zero duality gap holds for primal-dual (43)–(44) for large enough $k \in \mathbb{N}$. 
2. SDP (43) has an optimal solution for large enough \( k \in \mathbb{N} \).
3. Assume one of the following two conditions holds:
   
   (a) \( Q(g, h) \) is Archimedean, the ideal \( (h) \) is real radical, and the second-order sufficiency conditions (Definition 1) hold at every global minimizer of POP (3);
   
   (b) \( V(h) \) is finite.

   Then SDP (44) has an optimal solution for large enough \( k \in \mathbb{N} \). In this case, \( \bar{\tau}_k = \bar{\rho}_k = f^* \).

   **Proof**  The first and second statement follow from Proposition 1 after replacing \( S(g, h) \) by \( V(\bar{h}) \). The third statement is due to Proposition 4 and Lemma 7. \( \square \)

   For every \( k \in \mathbb{N} \), according to Lemma 5, if \( M_{k-1}((\bar{R} - \| \bar{x} \|^2_2) y) = 0 \) and \( y_0 = 1 \),

   \[
   \text{trace}(P_{\bar{n},k} M_k(y)P_{\bar{n},k}) = (\bar{R} + 1)^k, \tag{45}
   \]

   where \( P_{\bar{n},k} \) is defined as in (30) with \( n \) replaced by \( \bar{n} \). Thus SDP (43) has the CTP. It remains to follow a process which is similar to the one from Sect. 3.1.

   Next, we present an iterative method, stated in Algorithm 5, to solve POP (3) with \( g = R - \| x \|^2_2 \), based on nonsmooth optimization methods such as LMBM.

   **Algorithm 5** SpectralPOP-CTP-WithBallConstraint

   **Input:** POP (3) with \( g_1 = R - \| x \|^2_2 \), unknown optimal value \( f^* \) and optimal solutions; method (D) for solving SDP with CTP.

   **Output:** increasing real sequence \( (\bar{\tau}_k)_{k \in \mathbb{N}} \) and \( x^* \in \mathbb{R}^n \).

   1: for \( k \in \mathbb{N} \) do
   2: Compute the optimal value \( b_i \) of SDP (40) with CTP, \( i \in [l_g] \), by using method (D) and set \( \bar{R} := R + \sum_{i=1}^{l_g} b_i \);
   3: Compute the optimal value \(-\bar{\tau}_k\) and an optimal solution \( y^* \) of SDP (43) with CTP (45) by using method (D);
   4: Extract an atom \( \tilde{x}^* = (x^*, x^*_{n+1}, \ldots, x^*_{n+l_g+1}) \) by using Henrion-Lasserre’s algorithm in [18] from \( M_k(y^*) \);
   5: If \( \tilde{x}^* \) exists, set \( \bar{\tau}_{k+j} = \bar{\tau}_k, j \in \mathbb{N}^+ \), and terminate.

As in the single (ball) constraint case, one can choose method (D) in Algorithm 5 as Algorithm 1 with LMBM solver or SketchyCGAL.

   Following Proposition 3, Corollary 1 and Proposition 7, we obtain the following corollary:

   **Corollary 6**  (i) The sequence \( (\bar{\tau}_k)_{k \in \mathbb{N}} \) of Algorithm 5 is well defined and \( \bar{\tau}_k \uparrow f^* \) as \( k \to \infty \).

   (ii) Assume that either condition (a) or condition (b) of Proposition 7.3 holds. If there exists an optimal solution \( y^* \) of SDP (43) at order \( k \in \mathbb{N} \) such that the flat extension condition holds, then \( x^* \) exists at the \( k \)-th iteration of Algorithm 5. In this case, Algorithm 5 terminates at the \( k \)-th iteration, \( x^* \) is an optimal solution of POP (3) and \( f^* = \bar{\tau}_k \).
4 Numerical experiments

Let us report numerical results obtained while relying on algorithms from Sect. 3 to solve equality constrained QCQPs on a sphere, quartic minimization problems on the unit sphere as well as further applications to three well-known NP-hard optimization problems on the unit sphere: deciding nonnegativity/convexity of even degree forms and copositivity of real symmetric matrices.

The experiments are performed in Julia 1.3.1 with the following packages:

- SumOfSquare.jl [55] is a modeling library to write and solve SDP relaxations of POPs, based on JuMP.jl and the SDP solver Mosek 9.1.
- LMBM.jl solves unconstrained NSOPs with the limited-memory bundle method of Haarala et al. [14, 15]. LMBM.jl calls Karmitsa’s Fortran implementation of LMBM algorithm [22].
- SketchyCGAL is a MATLAB package to handle SDP problems with CTP or BTP, implemented by Yurtsever et al. [56]. We have implemented a Julia version (SketchyCGAL.jl) of SketchyCGAL to ensure fair comparison with LMBM.jl and SumOfSquare.jl. In this section, SketchyCGAL is used as a solver for SDP (33) in Algorithm 3 instead of Algorithm 1 or 2.

We also use the package Arpack.jl, which is based on the implicitly restarted Lanczos’s algorithm, to compute the largest eigenvalues and the corresponding eigenvectors of real symmetric matrices of (potentially) large size.

When POPs have equality constraints, SumOfSquare.jl uses reduced forms with Groebner basis instead of creating SOS multipliers, in order to reduce solving time.

The implementation of algorithms described in Sect. 3 can be downloaded from the link: https://github.com/maihoanganh/SpectralSOS.

We use a desktop computer with an Intel(R) Core(TM) i7-8665U CPU @ 1.9GHz × 8 and 31.2 GB of RAM. The notation for our numerical results are given in Table 1.

4.1 Random dense equality constrained QCQPs on the unit sphere

4.1.1 Test problems

We construct several instances of POP (8) as follows:

1. Take \( h_1 = 1 - \|x\|^2 \) and choose \( f \), \( h_j \), \( j \in [l_h]\backslash\{1\} \) with degrees at most 2;
2. Each coefficient of the objective function \( f \) is taken randomly in \((-1, 1)\) with respect to the uniform distribution;
3. Select a random point \( a \in \mathbb{R}^n \) in the unit sphere;
4. For every \( j \in [l_h]\backslash\{1\} \), all non-constant coefficients of \( h_j \) are taken randomly in \((-1, 1)\) with respect to the uniform distribution, and the constant coefficient of \( h_j \) is chosen such that \( h_j(a) = 0 \).

Note that by construction, \( a \) is a feasible solution. We use the method presented in Sect. 3.1 (actually the \( k \)-th iteration of Algorithm 3) to solve these problems. Numerical results are displayed in Table 2 for the case \( l_h = 1 \) and Tables 3, 4 for the case \( l_h = \lceil n/4 \rceil \). For these results, we use the Julia version of SketchyCGAL, which runs...
A hierarchy of spectral relaxations for polynomial...

Table 1 Notation

| Symbol | Description |
|--------|-------------|
| \( n \) | The number of variables of the POP |
| \( l_g \) | The number of inequality constraints of the POP |
| \( l_h \) | The number of equality constraints of the POP |
| \( k \) | The order of the moment-SOS relaxation or the iteration of Algorithm 3 |
| \( s \) | The size of the positive semidefinite matrix involved in the SDP relaxation |
| \( m \) | The number of trace equality constraints of the SDP relaxation |
| SumOfSquares | SDP relaxation modeled by SumOfSquares.jl and solved by Mosek 9.1 |
| CTP | The method described either in Sect. 3.1, Sect. 3.2 or Sect. 3.3 |
| BTP | The method described in Remark 6 |
| LMBM | SDP relaxation solved by spectral minimization, described in Sect. 2.6 with the LMBM solver |
| SketchyCGAL | SDP relaxation solved by SketchyCGAL |
| SpectralPOP | SDP relaxation handled by CTP or BTP method, with LMBM or SketchyCGAL solver |
| Val | The optimal value of the SDP relaxation |
| Gap | The relative optimality gap w.r.t. SumOfSquares, defined by \( \text{gap} = \frac{|\text{val} - \text{val(SumOfSquares)}|}{|\text{val(SumOfSquares)}|} \) |
| * | There exists at least one optimal solution of the POP, which can extracted by Henrion-Lasserre’s algorithm in [18] |
| Time | The total computation time of the SDP relaxation in seconds |
| - | The calculation did not finish in 3000 s or ran out of memory |
Table 2  Numerical results for random dense equality constrained QCQPs on the unit sphere, described in Sect. 4.1, with \((l_g, l_h) = (0, 1)\) and \(k = 1\)

| POP size \(n\) | SumOfSquares (Mosek) | SpectralPOP (CTP) | LMBM | SketchyCGAL |
|-----------------|-----------------------|-------------------|------|-------------|
|                 | \(n\) Val | Time | Val | Val | Time | Val | Time |
| 50              | 6.24844* | 0.3  | -6.24844* | 0.7  | -.11351 | 0.7 |
| 75              | 7.25326* | 2    | -7.25326* | 0.7  | -6.95325 | 0.8 |
| 100             | 7.00957* | 8    | -7.00957* | 0.9  | -6.75991 | 1   |
| 125             | 9.76963* | 23   | -9.76963* | 1    | -9.39907 | 1   |
| 150             | 8.49449* | 64   | -8.49449* | 1    | -8.15382 | 2   |
| 175             | 10.7286* | 140  | -10.72866* | 1    | -10.1323 | 2   |
| 200             | 11.3521* | 300  | -11.3521* | 2    | -10.4724 | 3   |
| 250             | 13.8881* | 1152 | -13.8881* | 4    | -13.5571 | 5   |
| 300             | 13.9957* | 3708 | -13.9958* | 6    | -13.8327 | 12  |
| 400             | -      | -    | -15.7584* | 15   | -15.5036 | 28  |
| 500             | -      | -    | -17.5838* | 35   | -17.2513 | 65  |
| 700             | -      | -    | -22.3584* | 218  | -22.0710 | 355 |
| 900             | -      | -    | -25.6117* | 621  | -25.2435 | 947 |
| 1200            | -      | -    | -28.3170* | 1401 | -27.8270 | 2074|
| 1500            | -      | -    | -30.8475* | 7120 | -30.2347 | 9020|

Fig. 1  Efficiency and accuracy comparison for Table 2

much faster than the MATLAB version without compromising accuracy (Figs. 1, 2, 3).

4.1.2 Efficiency comparison

In Table 2, we minimize quadratic polynomials on the unit sphere. The SDP relaxation for a POP in \(n\) variables involves a matrix of size \(n + 1\) and 2 trace equality constraints. In this table, LMBM is the fastest SDP solver while Mosek (the SDP solver used by SumOfSquares) is the slowest. It is due to the fact that Mosek relies on interior-
Table 3 Numerical results for random dense equality constrained QCQPs on the unit sphere, described in Sect. 4.1, with \((l_g, l_h) = (0, [n/4])\) and \(k = 1\)

| POP size | SumOfSquares (Mosek) | SpectralPOP (CTP) | SketchyCGAL |
|----------|-----------------------|-------------------|-------------|
| n        | \(l_h\)   | Val    | Time | Val    | Time | Val    | Time |
| 50       | 14        | -4.26516 | 0.4   | -4.26516 | 1     | -4.24511 | 1     |
| 60       | 16        | -6.42900* | 1     | -6.42929* | 1     | -6.36177 | 2     |
| 70       | 19        | -5.08320  | 3     | -5.08322  | 2     | -5.01911 | 3     |
| 80       | 21        | -5.35178  | 5     | -5.35178  | 2     | -5.29900 | 4     |
| 100      | 26        | -7.50097  | 15    | -7.50097  | 10    | -7.42432 | 11    |
| 120      | 31        | -5.89903  | 33    | -5.89903  | 12    | -5.81244 | 18    |
| 150      | 39        | -7.44920  | 127   | -7.44921  | 26    | -7.32154 | 36    |
| 200      | 51        | -8.93976  | 363   | -8.93976  | 51    | -8.79487 | 71    |
| 300      | 76        | -12.4295  | 3753  | -12.4295  | 530   | -12.2180 | 480   |
| 400      | 101       | –       | –     | -14.7190  | 2553  | -14.4830 | 2318  |

Fig. 2 Efficiency and accuracy comparison for Table 3

Fig. 3 Efficiency and accuracy comparison for Table 4
Table 4  Numerical results for random dense equality constrained QCQPs on the unit sphere, described in Sect. 4.1, with \((l_g, l_h) = (0, \lceil n/4 \rceil)\) and \(k = 2\)

| POP size | SDP size | SumOfSquares (Mosek) | SpectralPOP (CTP) | SketchyCGAL |
|----------|----------|----------------------|-------------------|-------------|
|          |          | Val      | Time  | Val      | Time  | Val      | Time  |
| \(n\)   | \(l_h\)  | \(s\)    | \(m\)  |          |      |          |      |
| 5        | 3        | 21       | 169    | - 2.43,822* | 0.04 | - 2.43,823* | 1     | - 2.43,982 | 1 |
| 10       | 4        | 66       | 1475   | - 1.4206*  | 0.4  | - 1.42,013* | 1     | - 1.41,268 | 1 |
| 15       | 5        | 136      | 6121   | - 2.87,129* | 7    | - 2.87,142* | 2     | - 2.86,744 | 6 |
| 20       | 6        | 231      | 17,557 | - 3.28,734* | 73   | - 3.28,736* | 5     | - 3.27,733 | 26 |
| 25       | 8        | 351      | 40,834 | - 3.32,902* | 592  | - 3.32,918* | 13    | - 3.31,634 | 65 |
| 30       | 8        | 496      | 81,345 | - 4.34,398* | 4678 | - 4.34,407* | 60    | - 4.32,974 | 294 |
| 35       | 10       | 666      | 146,521| - 4.77,580* |      | - 4.77,580* | 275   | - 4.75,946 | 450 |
| 40       | 11       | 861      | 244,812| -         |      | - 2.95,099  | 390   | - 2.91,856 | 1225 |
| 45       | 13       | 1081     | 386,999| -         |      | - 3.95,743  | 1588  | - 3.88,533 | 2905 |
| 50       | 14       | 1326     | 582,115| -         |      | - 4.01846   | 6126  | -         | -  |
point methods based on second order conditions to solve SDP while LMBM and SketchyCGAL only rely on algorithms based on first order conditions. Note that we use the same modeling technique to generate the SDP-CTP relaxation solved with either SketchyCGAL or LMBM, so both related modeling times are the same. The solving time of SketchyCGAL is a bit larger than the one of LMBM in this case.

In Tables 3, 4, we consider random equality constrained QCQPs and solve their first \((k = 1)\) and second \((k = 2)\) order moment relaxation, respectively. In Table 3, the size of the positive semidefinite matrix (resp. the number of trace equality constraints) involved in the SDP relaxation is equal to \(n + 1\) (resp. \(l_h + 1\)). In Table 4, the matrices involved in the SDP relaxation have size \(\binom{n}{4}\) and the number of trace equality constraints is \(O(\binom{n}{4}^2)\), due to (60). Thus, the number of trace equality constraints for these SDP relaxations is more than 200 times larger than the matrix size, for almost all instance of Table 4. LMBM and SketchyCGAL still happen to be faster than SumOfSquares in both Tables 3, 4, but LMBM is not much more efficient than SketchyCGAL. The most expensive step performed by Mosek (used by SumOfSquares) is to solve a system of linear equations coming from certain complementarity conditions (see page 13 in [9] for more details). The linear system becomes harder to solve when the number of trace equality constraints is larger. This is in contrast with LMBM, which does not need to solve any such large size linear system of equations. By comparison with LMBM, SketchyCGAL may perform a larger number of operations [56, Algorithm 6.1], as emphasized later on.

### 4.1.3 Accuracy comparison

When \(n \leq 300\) in Table 2, \(n \leq 300\) in Table 3 or \(n \leq 30\) in Table 4, LMBM converges to the exact optimal value of POPs with high accuracy, similarly to SumOfSquares. Both LMBM and SumOfSquares can extract at least one approximate optimal solution by Henrion-Lasserre’s algorithm [18], when \(n \leq 300\) in Table 2 or \(n \leq 35\) in Table 4. Moreover, LMBM can provide an approximate optimal solution even for large-scale problems with \(n = 1500\) in Table 2 (resp. \(n = 40\) in Table 4) and in a case in Table 3. Unfortunately SketchyCGAL cannot do the extraction procedure successfully, because of its inaccurate output.

### 4.1.4 Storage and evaluation comparisons

In Tables 5 and 6, we display some additional information related to Mosek, LMBM and SketchyCGAL, for the rows \(n = 5, 10, 15, 20, 25\) of Table 4:

- **Storage**;
- \(#A\): the number of evaluations of the linear operator \(A\) in SDP (13);
- \(#A^\top\): the number of evaluations of the adjoint operator \(A^\top\);
- \(s_{\text{max}}\): the largest size of symmetric matrices of which eigenvalues and eigenvectors are computed;
- \(N_{\text{eig}}\): the number of symmetric matrices of which eigenvalues and eigenvectors are computed.
Table 5 indicates that SumOfSquares requires a bit lower storage than LMBM only for the cases \( n = 5, 10 \). However, SketchyCGAL requires a bit smaller storage than LMBM. Note that SketchyCGAL performs a large number of evaluations of both \( A \) and \( A^\top \) while relying on three specific primitive computations (see [56, Sect. 2.3]). Compared to SketchyCGAL, LMBM performs a smaller number of evaluations of both \( A \) and \( A^\top \). For instance, the number of evaluations of LMBM is ten times smaller than the one of SketchyCGAL for the row \( n = 25 \) of Table 6. Because of the large number \( m \) of trace equality constraints, the evaluations of \( A \) and \( A^\top \) in SDP relaxations of POPs is more expensive than the simple one related to the first order SDP relaxation of MAXCUT, which is solved very efficiently by SketchyCGAL (see [56, Sect. 2.5]).

These specific behaviors mainly come from the subroutines used by LMBM and SketchyCGAL to compute eigenvalues and eigenvectors. While LMBM computes directly the largest eigenvector (and corresponding eigenvalue) of the matrix \( C - A^\top z \) involved in the nonsmooth function from (19), SketchyCGAL computes indirectly the smallest eigenvalue of the matrix \( C + A^\top (y + \beta (z - b)) \) in Step 8 of [56, Algorithm 6.1] while relying on the so-called “ApproxMinEvec” subroutine. When the ApproxMinEvec subroutine is implemented via [56, Algorithm 4.2], SketchyCGAL provides approximations of the smallest eigenvalue and eigenvector of each matrix \( C + A^\top (y + \beta (z - b)) \) by using the randomized Lanczos method. It only requires to compute the smallest eigenvalue and eigenvector of a tridiagonal matrix of small size (e.g. \( s_{\text{max}} = 42 \) when \( n = 25 \) in Table 6 while the value \( s_{\text{max}} \) of LMBM is 351).
Besides, SketchyCGAL computes $v_i^T (C + A^T (y + \beta (z - b))) v_i$ while relying on three primitive computations (see [56, (2.4)] for more details), which yields a large number of evaluations of $A^T$. For instance, $\# A^T = 6134$ for SketchyCGAL when $n = 25$ while the value $\# A^T$ is 337 for LMBM. Because of its slow convergence, SketchyCGAL also performs a larger number of iterations in Step 6 of [56, Algorithm 6.1].

Based on the above comparison, we emphasize that LMBM is cheaper and faster than Mosek or SketchyCGAL while LMBM ensures the same accuracy as Mosek when solving SDP relaxations of equality constrained QCQPs on the unit sphere.

### 4.2 Random dense QCQPs on the unit ball

#### 4.2.1 Test problems

We construct several samples of POP (3) as follows:

1. Take $g_1 = 1 - \|x\|_2^2$ and choose $f, g_i, i \in [l_g] \backslash \{1\}$, and $h_j, j \in [l_h]$ with degrees at most 2;
2. Each coefficient of the objective function $f$ is taken randomly in $(-1, 1)$ with respect to the uniform distribution;
3. Select a random point $a \in \mathbb{R}^n$ in the unit ball, with respect to the uniform distribution;
4. For each $i \in [l_g] \backslash \{1\}$, all non-constant coefficients of $g_i$ are taken randomly in $(-1, 1)$ with respect to the uniform distribution, and the constant coefficient of $g_i$ is chosen such that $g_i(a) > 0$;
5. For $j \in [l_h]$, all non-constant coefficients of $h_j$ are taken randomly in $(-1, 1)$ with respect to the uniform distribution, and the constant coefficient of $h_j$ is chosen such that $h_j(a) = 0$.

Numerical results are displayed in Table 7 for the case $(l_g, l_h) = (1, \lceil n/4 \rceil)$ and Table 8 for the case $(l_g, l_h) = (\lceil n/8 \rceil, \lceil n/8 \rceil)$. We recall the following notation:

- CTP (LMBM): the SDP relaxation is solved via the method described in Sect. 3.2 (the $k$-th iteration of Algorithm 4) or Sect. 3.3 (the $k$-th iteration of Algorithm 5) with the LMBM solver.
- BTP (LMBM): the SDP relaxation is solved via the method described in Remark 6 with the LMBM solver (Algorithm 2).

In Tables 7 and 8, SumOfSquares and BTP solve relaxations involving matrices with the same size, corresponding exactly to the size of the moment relaxation (7) (Figs. 4, 5).

#### 4.2.2 Efficiency and accuracy comparisons

In Table 7, we consider POPs which involve a single inequality (ball) constraint. In this case, CTP (LMBM) is the most efficient and accurate solver. Numerical results

---

1. The vector $v_i$ is updated in Step 6 of [56, Algorithm 4.2] within the loop from Step 5 of [56, Algorithm 4.2]
Table 7  Numerical results of random dense QCQPs on the unit ball, described in Sect. 4.2, with $(l_g, l_h) = (1, \lceil n/4 \rceil)$, and $k = 2$

| POP size | SDP size (CTP) | SDP size (Mosek) | SumOfSquares (Mosek) | SpectralPOP CTP (LMBM) | SpectralPOP BTP (LMBM) |
|----------|----------------|-----------------|----------------------|------------------------|------------------------|
| n | l_h | s | m | Val | Time | Val | Time | Val | Time |
| 5 | 2 | 28 | 281 | $-0.33,125^*$ | 0.07 | $-0.33,126^*$ | 1 | $-0.98,254$ | 0.7 |
| 10 | 3 | 78 | 2029 | $-2.30,410^*$ | 0.5 | $-2.30,411^*$ | 1 | $-3.64,371$ | 5 |
| 15 | 4 | 153 | 7702 | $-2.26,182^*$ | 10 | $-2.26,195^*$ | 2 | $-4.21,202$ | 134 |
| 20 | 5 | 253 | 21,000 | $-2.24,031^*$ | 112 | $-2.24,033^*$ | 4 | $-5.77,860$ | 1722 |
| 25 | 7 | 378 | 47,251 | $-2.88,952^*$ | 1484 | $-2.88,770^*$ | 15 | $-6.81,243$ | 16,185 |
| 30 | 7 | 528 | 92,049 | $-4.15,791^*$ | 4694 | $-4.15,798^*$ | 49 | $-4.15,798$ | 49 |
| 35 | 9 | 703 | 163,097 | $-4.10,015$ | 150 | $-4.10,015$ | 150 | $-4.10,015$ | 150 |
| 40 | 10 | 903 | 269,095 | $-4.47,927$ | 694 | $-4.47,927$ | 694 | $-4.47,927$ | 694 |
| 45 | 12 | 1128 | 421,121 | $-5.50,988$ | 849 | $-5.50,988$ | 849 | $-5.50,988$ | 849 |
| 50 | 13 | 1378 | 628,369 | $-5.52,884$ | 2086 | $-5.52,884$ | 2086 | $-5.52,884$ | 2086 |
Fig. 4 Efficiency and accuracy comparison for Table 7

Table 8 Numerical results of random dense QCQPs on the unit ball, described in Sect. 4.2, with \((l_g, l_h) = ([n/8], [n/8]), \) and \(k = 2\)

| POP size | SDP size (CTP) | SumOfSquares (Mosek) | SpectralPOP (CTP (LMBM)) | BTP (LMBM) |
|----------|----------------|-----------------------|---------------------------|------------|
|          | \(n\) | \(l_g\) | \(l_h\) | \(s\) | \(m\) | Val | Time | Val | Time | Val | Time |
|          | 10   | 2     | 2     | 105  | 3711 | −3.17792 | 0.2 | −3.18434 | 31  | −4.71914 | 6   |
|          | 15   | 2     | 2     | 190  | 11,781 | −2.14424 | 6   | −2.16250 | 69  | −4.45435 | 649 |
|          | 20   | 3     | 3     | 190  | 11,781 | −2.92513 | 190 | −3.09124 | 469 | −81.3719 | 2573 |

emphasize that SumOfSquares and CTP (LMBM) behave in a similar way as in Table 4. This indicates that converting a POP with a single inequality (ball) constraint to a CTP-POP by adding one slack variable, and solving the resulting SDP-CTP relaxation by means of spectral methods allows one to reduce the computing time while ensuring the same accuracy as the one obtained with SumOfSquares (Mosek). Note that when we use the method described in Sect. 3.2, the constant trace in (39) is always equal to \(2^k\), which is independent of \(n\).
Table 9 Subgradient norms computed during the last 10 iterations of CTP (LMBM) for the experiments from Tables 7 and 8 with $n = 10$

|          | Table 7 | Table 8 |
|----------|---------|---------|
|          | 0.035   | 0.871   |
|          | 0.033   | 0.959   |
|          | 0.028   | 0.947   |
|          | 0.022   | 0.792   |
|          | 0.024   | 1.684   |
|          | 0.019   | 0.794   |
|          | 0.015   | 0.579   |
|          | 0.009   | 0.559   |
|          | 0.007   | 0.230   |
|          | 0.004   | 0.916   |

In Table 8, CTP (LMBM) provides inaccurate output as it only yields lower bounds, while SumOfSquares still preserves accuracy. Moreover, CTP (LMBM) is less efficient than SumOfSquares. We also emphasize that when one relies on the method stated in Sect. 3.3, we obtain a value of $\bar{R}$ in (41), for the sphere constraint of CTP-POP, which becomes larger when $n$ increases. It implies that the constant trace factor $(\bar{R} + 1)^k$ in (45) has a polynomial growth rate in $\bar{R}$. Thus we minimize a nonsmooth function of the form (19) with a large constant trace factor $a$. The norm of the subgradient of this function at a point near its minimizers is rather large, which prevents LMBM to perform properly its minimization, by contrast with Table 7. This difference of magnitude is shown in Table 9, where we compute the subgradient norms during the last 10 iterations of CTP (LMBM) for the experiments from Tables 7 and 8 with $n = 10$.

In both Tables 7 and 8, BTP (LMBM) has the worst performance in terms on efficiency and accuracy. The trace bound (29) obtained in Remark 6 is usually much larger than the “exact” trace of the optimal solution of the SDP relaxation. The same issue occurs for the subgradient norm of the nonsmooth function at a point near its minimizers.

According to our experience, LMBM is suitable for spectral minimization of SDP problems with trace bounds which are small enough and close to the exact trace value of the optimal solution.

4.3 Random dense quartics on the unit sphere

4.3.1 Test problems

We construct several instances of POP (8) as follows:

1. Take $l_h = 1$ and $h_1 = 1 - \|x\|^2_2$ and choose $f$ with degree at most 4;
2. Each coefficient of the objective function $f$ is taken randomly in $(-1, 1)$ with respect to the uniform distribution.

We use the method presented in Sect. 3.1 to solve these problems. The corresponding numerical results are displayed in Table 10 (Fig. 6).

4.3.2 Efficiency and accuracy comparisons

Table 10 indicates that LMBM is about twice faster than SumOfSquares when $n \geq 20$ as well as SketchyCGAL. While SketchyCGAL can be rather inaccurate, LMBM has an accuracy which is similar to SumOfSquares (Mosek), yielding the ability to extract optimal solutions of POPs.
Table 10  Numerical results for random dense quartics on the unit sphere, described in Sect. 4.3, with \((l_g, l_h) = (0, 1)\) and \(k = 2\)

| POP size | SDP size | SumOfSquares (Mosek) | SpectralPOP (CTP) | LMBM | SketchyCGAL |
|----------|----------|-----------------------|-------------------|------|-------------|
|          |          |                       | Val               | Val  | Val         | Time  | Time  |
|          |          |                       |                   |      |             |       |       |
| 5        | 21       | 127                   | - 2.92,483*       | - 2.92,485* | - 2.84,710 | 1     |       |
| 10       | 66       | 1277                  | - 3.59,964*       | - 3.59,964* | - 3.48,501 | 2     |       |
| 15       | 136      | 5577                  | - 4.18,773*       | - 4.18,778* | - 4.03,882 | 18    |       |
| 20       | 231      | 16,402                | - 3.92,438*       | - 3.92,440* | - 3.67,857 | 87    |       |
| 25       | 351      | 38,377                | - 6.36,891*       | - 6.36,894* | - 5.93,774 | 251   |       |
For comparisons in Sects. 4.2 and 4.3, the coefficients of $f$ have been randomly generated in $(-1, 1)$. However, for some non random problems that were scaled so as to fit the framework of optimization on the unit sphere, we could observe a lack of precision after transferring results (of the scaled formulation) back to results in the unscaled initial formulation.

In the next three subsections, we consider further applications of the minimization of forms on the unit sphere listed in [29].

### 4.4 Deciding the nonnegativity of even degree forms

Given $p \in \mathbb{R}[x]$, we recall that $p$ is a form of degree $d$ if $p = \sum_{|\alpha|=d} p_\alpha x^\alpha$ for some $d \in \mathbb{N}$ and $p_\alpha \in \mathbb{R}$. Given a form $p \in \mathbb{R}[x]$, $p$ is nonnegative on $\mathbb{R}^n$ iff $p$ is nonnegative on the unit sphere. Moreover, given a polynomial $f \in \mathbb{R}[x]_{2d}$, $f$ is nonnegative on $\mathbb{R}^n$ iff its homogenization $\tilde{f} := x_{n+1}^{2d} f(x_{n+1})$ is nonnegative on $\mathbb{R}^{n+1}$.

Note that $\tilde{f}$ is a form. Thus, in order to verify the nonnegativity of the polynomial $f$, we only verify the nonnegativity of its homogenization $\tilde{f}$ on the unit sphere in $\mathbb{R}^{n+1}$. Namely, given a form $f \in \mathbb{R}[x]$ of degree $2d$, we consider the following POP:

$$f^* := \min_{x \in \mathbb{R}^n} \{ f(x) : \|x\|_2^2 = 1 \}. \quad (46)$$

Note that if $d = 1$, problem (46) boils down to computing the minimal eigenvalue of the Gram matrix associated to $f$. Thus, we consider the case where $d \geq 2$.

#### 4.4.1 Test problems

We construct several instances of the form $f$ of degree $2d$ as follows:

1. Take $f_\alpha$ randomly in $(-1, 1)$ with respect to the uniform distribution, for each $\alpha \in \mathbb{N}^n$ with $|\alpha| = 2d$.
2. Set $f := \sum_{|\alpha|=2d} f_\alpha x^\alpha$.

We use the method presented in Sect. 3.1 to solve problem (46). The corresponding numerical results are displayed in Table 11.
Table 11 Numerical results for deciding the nonnegativity of even degree forms, described in Sect. 4.4, with $k = d$

| POP size | SDP size | SumOfSquares | SpectralPOP (CTP) |
|----------|----------|--------------|-------------------|
| $n$      | $d$      | $s$          | $m$               | (Mosek)         | (LMBM)         | Val     | Time | Val     | Time |
| 5        | 2        | 21           | 127               | $-2.62,353^*$    | 0.01           | $-2.62,353^*$ | 1     |
| 10       | 2        | 66           | 1277              | $-6.31,389^*$    | 0.4            | $-6.31,390^*$ | 1     |
| 15       | 2        | 126          | 5577              | $-12.1405^*$     | 6              | $-12.1405^*$ | 2     |
| 20       | 2        | 231          | 16402             | $-19.9981^*$     | 76             | $-19.9981^*$ | 2     |
| 25       | 2        | 351          | 38377             | $-29.4812^*$     | 633            | $-29.4812^*$ | 4     |
| 30       | 2        | 496          | 77377             | $-40.6934^*$     | 3471           | $-40.6934^*$ | 9     |
| 35       | 2        | 666          | 140527            | $-54.2561^*$     | –              | $-54.2561^*$ | 24    |
| 40       | 2        | 861          | 236202            | $-69.4700^*$     | –              | $-69.4700^*$ | 57    |
| 45       | 2        | 1081         | 374027            | $-86.4113^*$     | –              | $-86.4113^*$ | 127   |
| 50       | 2        | 1326         | 564877            | $-105.532^*$     | –              | $-105.532^*$ | 250   |
| 5        | 3        | 56           | 1261              | $-1.56,744$      | 0.1            | $-1.60,032$ | 23    |
| 5        | 4        | 126          | 7177              | $-1.35,267$      | 1              | $-1.35,315$ | 235   |

4.4.2 Efficiency and accuracy comparisons

Table 11 shows that LMBM is much faster than Mosek when $n \geq 20$ and $d = 2$. In these cases, we were able to extract the solution of the resulting POPs. One can then conclude that $f$ is not globally nonnegative since it has negative value at its atoms. For higher values of $d = 3, 4$, LMBM becomes less efficient and accurate than Mosek.

4.5 Deciding the convexity of even degree forms

Recall that a polynomial $p \in \mathbb{R}[x]$ is convex if $p(tx + (1-t)y) \leq tp(x) + (1-t)p(y)$ for all $x, y \in \mathbb{R}^n$ and $t \in [0, 1]$. Moreover, $p \in \mathbb{R}[x]$ is convex iff the polynomial $f(x, y) := p(y) - p(x) - \nabla p(x)^T (y - x)$ is globally nonnegative, where $\nabla p$ stands for the gradient of $p$. If $p$ is a form of degree $d$, $f$ is also a form of degree $d$. In this case, the nonnegativity of $f$ can be verified on the unit sphere in $\mathbb{R}^{2n}$.

Given a form $p \in \mathbb{R}[x]$ of degree $2d$, consider the following POP:

$$ f^* := \min_{x, y \in \mathbb{R}^n} p(y) - p(x) - \nabla p(x)^T (y - x) $$

s.t. $\|x, y\|_2^2 = 1$.  \hspace{1cm} (47)

From the previous discussion, $p$ is convex iff $f^* \geq 0$.

4.5.1 Test problems

We construct several instances of the form $p$ of degree $2d$ as follows:
Table 12 Numerical results for deciding the convexity of even degree forms, described in Sect. 4.5, with $k = d$

| POP size | SDP size | SumOfSquares (Mosek) | SpectralPOP (CTP) (LMBM) |
|----------|----------|----------------------|----------------------------|
| $n$ | $d$ | $s$ | $m$ | Val | Time | Val | Time |
| 5 | 2 | 66 | 1277 | $-3.87,350^*$ | 0.2 | $-3.87,350^*$ | 1 |
| 7 | 2 | 120 | 4321 | $-4.44,260^*$ | 3 | $-4.44,280^*$ | 11 |
| 10 | 2 | 231 | 16,402 | $-4.98,855^*$ | 63 | $-4.98,883^*$ | 17 |
| 12 | 2 | 325 | 32,826 | $-4.47,495^*$ | 336 | $-4.49,239^*$ | 91 |
| 5 | 3 | 286 | 34,035 | $-3.89,581^*$ | 53 | $-3.89,586^*$ | 70 |

1. Take $p_{\alpha}$ randomly in $(-1, 1)$ with respect to the uniform distribution, for each $\alpha \in \mathbb{N}^n$ with $|\alpha| = 2d$.

2. Set $p := \sum_{|\alpha|=2d} p_{\alpha} x^\alpha$.

We use the method presented in Sect. 3.1 to solve problem (47). The corresponding numerical results are displayed in Table 12.

4.5.2 Efficiency and accuracy comparisons

Table 12 shows that LMBM is about 3 times faster than Mosek when $n \geq 10$ and $d = 2$. In these cases, one concludes as above that $p$ is nonconvex since $f^*$ is negative. For $d = 3$ and $n = 5$, LMBM still returns a value with reasonably high accuracy ($10^{-5}$) even though it is slower than Mosek. In this case, one can also certify that $p$ is nonconvex.

4.6 Deciding the copositivity of real symmetric matrices

Given a symmetric matrix $A \in \mathbb{R}^{n \times n}$, we say that $A$ is copositive if $u^T A u \geq 0$ for all $u \in \mathbb{R}_{+}^n$. Consider the following POP:

$$f^* := \min_{x \in \mathbb{R}^n} \{ (x^2)^T A x^2 : \|x\|_2 = 1 \},$$

(48)

where $x^2 := (x_1^2, \ldots, x_n^2)$. The matrix $A$ is copositive iff $f^* \geq 0$.

4.6.1 Test problems

We construct several instances of the matrix $A$ as follows:

1. Take $B_{ij}$ randomly in $(-1, 1)$ with respect to the uniform distribution, for all $i, j \in \{1, \ldots, n\}$.

2. Set $B := (B_{ij})_{1 \leq i, j \leq n}$ and $A := \frac{1}{2}(B + B^T)$.

We use the method presented in Sect. 3.1 to solve problem (48). The corresponding numerical results are displayed in Table 13.
Table 13: Numerical results for deciding the copositivity of real symmetric matrices, described in Sect. 4.6, with \( k = 2 \)

| POP size | SDP size | SumOfSquares (Mosek) | SpectralPOP (CTP) (LMBM) |
|----------|----------|----------------------|--------------------------|
| \( n \)  | \( s \)  | \( m \)  | Val | Time | Val | Time |
| 10       | 66       | 1277                | −0.89102^* | 0.2 | −0.89102^* | 1 |
| 15       | 136      | 5577                | −0.91701^* | 6 | −0.91701^* | 9 |
| 20       | 231      | 16,402              | −0.98474^* | 57 | −0.98474^* | 19 |
| 25       | 351      | 38,377              | −0.97873^* | 558 | −0.97873^* | 28 |

### 4.6.2 Efficiency and accuracy comparisons

Table 13 indicates that LMBM is twice faster than Mosek when \( n \geq 20 \). In all cases, we can extract the solutions of the resulting POP and certify that \( A \) is not copositive since \( f^\star \) is negative.

### 5 Conclusion

We have presented a nonsmooth hierarchy of SDP relaxations for optimizing polynomials on varieties contained in a Euclidean sphere. The advantage of this hierarchy is to circumvent the hard constraints involved in the standard SDP hierarchy (11) by minimizing the maximal eigenvalue of a matrix pencil. This in turn boils down to solving an unconstrained convex nonsmooth optimization problem by LMBM and to computing largest eigenvalues by means of the modified Lanczos’s algorithm. Our numerical experiments indicate that solving this nonsmooth hierarchy is more efficient and more robust than solving the classical semidefinite hierarchy by interior-point methods, at least for a class of interesting POPs, including equality constrained QCQPs on the sphere, QCQPs with a single inequality (ball) constraint, and minimization of quartics on the sphere. Our CTP framework can be further applied for an interesting class of noncommutative polynomial optimization problems [35], in particular for eigenvalue maximization problems arising from quantum information theory, where the variables are unitary operators [38]. A topic of future investigation is to better handle the case of POPs involving several inequalities. Our current method transforms such a POP into a CTP-POP by adding a slack variable for each inequality. One promising workaround would be to exploit the inherent sparse structure of this CTP-POP. Another similar investigation track would be to exploit the CTP of SDP relaxations resulting from polynomial optimization problems with sparse input data. These research lines are pursued in our follow-up paper [36].

Eventually, we have tried to use spectral methods to solve SDP relaxations of QCQPs involving inequalities only, systems of polynomial equations, MAXCUT problems, 0/1 linear constrained quadratic problems and computation of stability numbers of graphs. However, our preliminary experiments for these problems have not been convincing in terms of efficiency and accuracy. In order to improve upon these results,
one possible remedy would be to index the moment matrices by alternative Legendre/Chebychev bases, rather than with the standard monomial basis.

**Acknowledgements** The authors would like to thank the associate editor and anonymous reviewers, whose insightful comments and careful proof-checks helped to improve the paper.

**Funding** The first author was supported by the MESRI funding from EDMITT. The second author was supported by the FMJH Program PGMO (EPICS project) and EDF, Thales, Orange et Criteo, as well as from the Tremplin ERC Stg Grant ANR-18-ERC2-0004-01 (T-COPS project). This work has benefited from the European Union’s Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie Actions, grant agreement 813211 (POEMA) as well as from the AI Interdisciplinary Institute ANITI funding, through the French “Investing for the Future PIA3” program under the Grant agreement n° ANR-19-PI3A-0004. The third author was supported by the European Research Council (ERC) under the European’s Union Horizon 2020 research and innovation program (Grant Agreement 666981 TAMING).

**Data availability** All data analyzed during this study are publicly available.

**Declarations**

**Conflict of interest** The authors declare that they have no conflict of interest.

**Code availability** The full code was made available for review. We remark that a set of packages were used in this study, that were either open source or available for academic use. Specific references are included in this published article.

**A Appendix**

**A.1 Spectral minimizations of SDP**

In this section, we provide the proofs of lemmas stated in Sects. 2.6.1 and 2.6.3. First we recall the following useful properties of $S$ and $S^+$:

- If $X = \text{diag}(X_1, \ldots, X_l) \in S$,

$$X \succeq 0 \iff X_j \succeq 0, \quad j \in [l] \quad \text{and} \quad \text{trace}(X) = \sum_{j=1}^{l} \text{trace}(X_j). \quad (49)$$

- If $A = \text{diag}(A_1, \ldots, A_l) \in S$ and $B = \text{diag}(B_1, \ldots, B_l) \in S$,

$$\langle A, B \rangle = \sum_{j=1}^{l} \langle A_j, B_j \rangle. \quad (50)$$

**A.1.1 SDP with constant trace property**

**Proof of Lemma 3** The proof of (20) is similar in spirit to the one of Helmberg and Rendl in [17, Sect. 2]. Here, we extend this proof for SDP (13), which involves a
block-diagonal positive semidefinite matrix. From (13),

\[-\tau = \sup_{X \in S} \{ \langle C, X \rangle : AX = b, \ \text{trace}(X) = a, \ X \succeq 0 \}.\]

The dual of this SDP reads:

\[-\rho = \inf_{(z, \xi)} \{ b^\top z + a\xi : A^\top z + \xi I - C \succeq 0 \},\]

where I is the identity matrix of size s. From this,

\[-\rho = \inf_{(z, \xi)} \{ b^\top z + a\xi : \xi \geq \lambda_1(C - A^\top z) \}
= \inf \{ a\lambda_1(C - A^\top z) + b^\top z : z \in \mathbb{R}^m \}.
\]

Since \(\rho = \tau\), (20) follows. For the second statement, let \(z^*\) be an optimal solution of SDP (14). Then \(b^\top z^* = -\rho = -\tau\). In addition, \(C - A^\top z^* \preceq 0\) implies that

\[\lambda_1(C - A^\top z^*) \leq 0,\]

so that \(\varphi(z^*) \leq -\tau\). Note that (20) indicates that \(\varphi(z^*) \geq -\tau\). Thus, \(\varphi(z^*) = -\tau\), yielding the second statement.

The following proposition recalls the differentiability properties of \(\varphi\).

**Proposition 8** The function \(\varphi\) in (19) has the following properties:

1. \(\varphi\) is convex and continuous but not differentiable.
2. The subdifferential of \(\varphi\) at \(z\) reads:

\[\partial \varphi(z) = \{b - aAW : W \in \text{conv}(\Gamma(C - A^\top z))\},\]

where for each \(A \in S\),

\[\Gamma(A) := \{uu^\top \ A = \lambda_1(A)u \ | \ u \parallel_2 = 1 \}.\]

**Proof** Properties 1–2 are from Helmberg–Rendl [17, Sect. 2] (see also [44, (4)]).

The following result is useful to recover an optimal solution of SDP (13) from an optimal solution of NSOP (20).

**Lemma 8** Let \(\tilde{z}\) be an optimal solution of NSOP (20). Then:

1. There exists \(X^* \in \text{conv}(\Gamma(C - A^\top \tilde{z}))\) such that \(AX^* = b\).
2. Let \(u\) be a normalized eigenvector corresponding to \(\lambda_1(C - A^\top \tilde{z})\). If \(\lambda_1(C - A^\top \tilde{z})\) has multiplicity 1 then \(X^* = auu^\top\), thus \(X^*\) is an optimal solution of SDP (13).
Proof By [1, Theorem 4.2], $0 \in \partial \varphi(\bar{z})$. Combining this with Proposition 8.2, the first statement follows.

We next prove the second statement. Assume that $\lambda_1(C - A^T \bar{z})$ has multiplicity 1. Then $\Gamma(C - A^T \bar{z}) = \{uu^T\}$. The first statement implies that $X^* = auu^T$. From this, $X^* \succeq 0$ and $AX^* = b$, so that $X^*$ is a feasible solution of SDP (13). Moreover,

$$
\langle C, X^* \rangle = \langle C - A^T \bar{z}, X^* \rangle + \langle A^T \bar{z}, X^* \rangle = a\langle C - A^T \bar{z}, uu^T \rangle + \bar{z}^T (AX^*) = au^T (C - A^T \bar{z})u + \bar{z}^T b = a\lambda_1(C - A^T \bar{z}) + \bar{z}^T b = \varphi(\bar{z}) = -\tau.
$$

Thus, $\langle C, X^* \rangle = -\tau$, yielding the second statement.

To obtain a convergence guarantee when solving NSOP (20) by LMBM [15, Algorithm 1], we need the following technical lemma:

Lemma 9 When applied to problem NSOP (20), the LMBM algorithm is globally convergent.

Proof The convexity of $\varphi$ yields that $\varphi$ is weakly upper semismooth on $\mathbb{R}^m$ according to [37, Proposition 5]. From this, $\varphi$ is upper semidifferentiable on $\mathbb{R}^m$ by using [3, Theorem 3.1]. Combining this with the fact that $\varphi$ is bounded from below on $\mathbb{R}^m$, the result follows thanks to [3, Sect. 5] (see also the final statement of [1, Sect. 14.2]).

A.1.2 SDP with bounded trace property

Proof of Lemma 4 Let $X^*$ be an optimal solution of SDP (13) and set $a := \text{trace}(X^*)$. By Condition 4 of Assumption 1, one has

$$
a \geq a > 0. \quad (53)
$$

Similarly to the proof of Lemma 3, one obtains:

$$
-\tau = \inf\{a\lambda_1(C - A^T z) + b^T z : z \in \mathbb{R}^m\}. \quad (54)
$$

Let us prove that

$$
\psi(z) \geq -\tau, \forall z \in \mathbb{R}^m. \quad (55)
$$

Let $z \in \mathbb{R}^m$ be fixed and consider the following two cases:

- Case I $\lambda_1(C - A^T z) > 0$. By (53) and (54),

$$
\psi(z) = a\lambda_1(C - A^T z) + b^T z \geq a\lambda_1(C - A^T z) + b^T z \geq -\tau.
$$

Thus, $\psi(z) \geq -\tau$. 

\(\blacksquare\) Springer
– Case 2 $\lambda_1(C - A^\top z) \leq 0$. Then $A^\top z - C \succeq 0$ and $\psi(z) = b^\top z \geq -\rho = -\tau$ by (14).

Let $(z^{(j)})_{j \in \mathbb{N}}$ be a minimizing sequence of SDP (14). Then $\lambda_1(C - A^\top z^{(j)}) \leq 0$, $j \in \mathbb{N}$, since $A^\top z^{(j)} - C \succeq 0$ and $b^\top z^{(j)} \to -\tau$ as $j \to \infty$ since $\tau = \rho$. It implies that $\psi(z^{(j)}) = b^\top z^{(j)} \to -\tau$ as $j \to \infty$. From this and by (55), the first statement follows.

For the second statement, let $z^\star$ be an optimal solution of SDP (14). Since $A^\top z - C \succeq 0$, $\lambda_1(C - A^\top z) \leq 0$ and thus $\psi(z^\star) = b^\top z^\star = -\rho = -\tau$. Thus, $z^\star$ is an optimal solution of (26), yielding the second statement. \qed

We consider the differentiability properties of $\psi$ in the following proposition:

**Proposition 9** The function $\psi$ has the following properties:

1. $\psi$ is convex and continuous but not differentiable.
2. The subdifferential of $\psi$ at $z$ reads:

$$
\partial \psi(z) = \begin{cases} 
\{b\} & \text{if } \lambda_1(C - A^\top z) < 0, \\
b - aAW & W \in \text{conv}(\Gamma(C - A^\top z)) & \text{if } \lambda_1(C - A^\top z) = 0, \\
b - \zeta aAW & \zeta \in [0, 1], \ W \in \text{conv}(\Gamma(C - A^\top z)) \text{ otherwise,} 
\end{cases}
$$

where $\Gamma(.)$ is defined as in (52).

**Proof** Note that $\psi$ is the maximum of two convex functions, i.e.,

$$
\psi(z) = \max\{\varphi_1(z), \varphi_2(z)\},
$$

with $\varphi_1(z) = a\lambda_1(C - A^\top z) + b^\top z$ and $\varphi_2(z) = b^\top z$. Thus, $\psi$ is convex and

$$
\partial \psi(z) = \begin{cases} 
\partial \varphi_1(z) & \text{if } \varphi_1(z) > \varphi_2(z), \\
\text{conv}(\partial \varphi_1(z) \cup \partial \varphi_2(z)) & \text{if } \varphi_1(z) = \varphi_2(z), \\
\partial \varphi_2(z) & \text{otherwise.}
\end{cases}
$$

Note that $\partial \varphi_2(z) = \{b\}$ and $\partial \varphi_1(z)$ is computed as in formula (51). Thus, the result follows. \qed

The following theorem is useful to recover an optimal solution of SDP (13) from an optimal solution of NSOP (26).

**Lemma 10** Assume that $\bar{z}$ is an optimal solution of NSOP (26). The following statements are true:

1. There exists

$$
X^\star \begin{cases} 
= 0 & \text{if } \lambda_1(C - A^\top \bar{z}) < 0, \\
\in \zeta a\text{conv}(\Gamma(C - A^\top \bar{z})) & \text{if } \lambda_1(C - A^\top \bar{z}) = 0, \\
\in a\text{conv}(\Gamma(C - A^\top \bar{z})) & \text{otherwise,}
\end{cases}
$$

for some $\zeta \in [0, 1]$ such that $AX^\star = b$. \qed

Springer
2. Let $u$ be a normalized eigenvector corresponding to $\lambda_1(C - A^T\bar{z})$. If $\lambda_1(C - A^T\bar{z})$ has multiplicity 1, then $X^* = \bar{\xi}u^T$ with $\bar{\xi}$ defined as in (27) and $X^*$ is an optimal solution of SDP (13).

Proof Due to [1, Theorem 4.2], $0 \in \partial \psi(\bar{z})$. From this and by Proposition 9.2, the first statement follows. Let us prove the second statement. Assume that $\lambda_1(C - A^T\bar{z})$ has multiplicity 1. Then one has $\Gamma(C - A^T\bar{z}) = \{ uu^T \}$, yielding $X^* = \bar{\xi}u^T$ with $\bar{\xi}$ defined as in (27), so that $X^* \succeq 0$. From this and since $AX^* = b$, $X^*$ is a feasible solution of SDP (13). Moreover,  

$$
\langle C, X^* \rangle = \langle C - A^T\bar{z}, X^* \rangle + \langle A^T\bar{z}, X^* \rangle
= \bar{\xi} \langle C - A^T\bar{z}, uu^T \rangle + \bar{z}^T(AX^*)
= \bar{\xi}u^T(C - A^T\bar{z})u + \bar{z}^Tb
= \lambda_1(C - A^T\bar{z})\bar{\xi} + \bar{z}^Tb
= a \max\{\lambda_1(C - A^T\bar{z}), 0\} + \bar{z}^Tb = \psi(\bar{z}) = -\tau.
$$

Thus, $\langle C, X^* \rangle = -\tau$, yielding the second statement. 

The next result proves that when applied to NSOP (26), the LMBM algorithm [15, Algorithm 1] converges.

Lemma 11 LMBM applied to NSOP (26) is globally convergent.

The proof of Lemma 11 is similar to Lemma 9.

A.2 Converting moment relaxations to standard SDP

We will present a way to transform SDP (31) to the form (33) recalled as follows:

$$
-\tau_k = \sup_{X \in S_k} \{ \langle C, X \rangle : \langle A_j, X \rangle = b_j, j \in [m], X \succeq 0 \}.
$$

Let $k \in \mathbb{N}$ be fixed. We will prove that there exists $A_j \in S_k$, $j \in [r]$, such that $X = P_kM_k(y)P_k$ for some $y \in \mathbb{R}^{\binom{n}{2}}$ if and only if $\langle A_j, X \rangle = 0$, $j \in [r]$. Let $V = \{ P_kM_k(z)P_k : z \in \mathbb{R}^{\binom{n}{2}} \}$. Then $V$ is a linear subspace of $S_k$ and $\dim(V) = \{ \binom{n}{2} \}$. We take a basis $A_1, \ldots, A_r$ of the orthogonal complement $V^\perp$ of $V$. Notice that

$$
r = \dim(V^\perp) = \dim(S_k) - \{ \binom{n}{2k} \} = \frac{1}{2} \left\{ \binom{n}{k} \right\} + 1 - \{ \binom{n}{2k} \}.
$$

with $X \in S_k$, it implies that $X \in V$ if and only if $\langle A_j, X \rangle = 0$, $j \in [r]$.

Let us find such a basis $A_1, \ldots, A_r$. Let $A = (A_{\alpha,\beta})_{\alpha,\beta \in \mathbb{N}_k^2} \in V^\perp$. Then for all $X = (X_{\alpha,\beta})_{\alpha,\beta \in \mathbb{N}_k^2} \in V$, $\langle A, X \rangle = 0$. Note that if $X = P_kM_k(y)P_k$, then one has

$$
X_{\alpha,\beta} = w_{\alpha,\beta}y_{\alpha + \beta}, \forall \alpha, \beta \in \mathbb{N}_k^2,
$$

where the elements $w_{\alpha,\beta}$ are defined as in (27) and

$$
\sum_{\alpha,\beta \in \mathbb{N}_k^2} w_{\alpha,\beta} = 0.
$$
with \( w_{\alpha, \beta} := \theta_k^{1/2} \theta_{k, \beta}^{1/2} \), for all \( \alpha, \beta \in \mathbb{N}_k^n \). It implies that

\[
0 = \sum_{\alpha, \beta \in \mathbb{N}_k^n} w_{\alpha, \beta} A_{\alpha, \beta} y_{\alpha+\beta}, \quad \forall y \in \mathbb{R}^{\binom{n}{2}}.
\]

Let \( \gamma \in \mathbb{N}_{2k}^n \) be fixed and let \( y \in \mathbb{R}^{\binom{n}{2}} \) be such that for \( \xi \in \mathbb{N}_{2k}^n \),

\[
y_\xi = \begin{cases} 
0 & \text{if } \xi \neq \gamma, \\
1 & \text{otherwise}.
\end{cases}
\]

Then

\[
0 = \sum_{\alpha, \beta \in \mathbb{N}_k^n} w_{\alpha, \beta} A_{\alpha, \beta} = \sum_{\alpha, \beta \in \mathbb{N}_k^n, \alpha+\beta=\gamma} w_{\alpha, \beta} A_{\alpha, \beta} + 2 \sum_{\alpha, \beta \in \mathbb{N}_k^n, \alpha=\beta} w_{\alpha, \beta} A_{\alpha, \beta}.
\]

If \( \gamma \notin 2\mathbb{N}^n \), we do not have the first term in the latter equality. Let us define

\[
\Lambda_{\gamma} := \{A_{\alpha, \beta} : \alpha, \beta \in \mathbb{N}_k^n, \alpha + \beta = \gamma, \alpha \leq \beta\}.
\]

Note that \( \Lambda_{\gamma} \) consists of values of \( A \) indexed by pairs of vectors \((\alpha, \beta)\) satisfying the lexicographic order relation \( \alpha \leq \beta \). Moreover, it can be rewritten as \( \Lambda_{\gamma} = \{A_{\alpha_j, \beta_j}, j \in [t]\} \) where \((\alpha_1, \beta_1) < \cdots < (\alpha_t, \beta_t)\) and \( t = |\Lambda_{\gamma}| \). Thus, if \( t \geq 2 \), we can choose \( A \) such that for all \( \alpha, \beta \in \mathbb{N}_k^n \),

\[
A_{\alpha, \beta} = \begin{cases} 
w_{\alpha_\mu, \beta_\mu} & \text{if } \alpha_1 = \beta_1 \text{ and } (\alpha, \beta) = (\alpha_1, \beta_1), \\
\frac{1}{2} w_{\alpha_\mu, \beta_\mu} & \text{if } \alpha_1 < \beta_1 \text{ and } (\alpha, \beta) \in \{(\alpha_1, \beta_1), (\beta_1, \alpha_1)\}, \\
-w_{\alpha_1, \beta_1} & \text{if } \alpha_\mu = \beta_\mu \text{ and } (\alpha, \beta) = (\alpha_\mu, \beta_\mu), \\
-\frac{1}{2} w_{\alpha_1, \beta_1} & \text{if } \alpha_\mu < \beta_\mu \text{ and } (\alpha, \beta) \in \{(\alpha_\mu, \beta_\mu), (\beta_\mu, \alpha_\mu)\}, \\
0 & \text{otherwise},
\end{cases}
\]

for some \( \mu \in [t]\setminus\{1\} \). We denote by \( B_{\gamma} \) the set of all such above \( A_{\alpha, \beta} \) satisfying \( t = |\Lambda_{\gamma}| \geq 2 \), otherwise let \( B_{\gamma} = \emptyset \). Then \( |B_{\gamma}| = |\Lambda_{\gamma}| - 1 \). From this and since \( (B_{\gamma})_{\gamma \in \mathbb{N}_{2k}^n} \) is a sequence of pairwise disjoint subsets of \( \mathcal{S}_k \),

\[
\left| \bigcup_{\gamma \in \mathbb{N}_{2k}^n} B_{\gamma} \right| = \sum_{\gamma \in \mathbb{N}_{2k}^n} |B_{\gamma}| = \sum_{\gamma \in \mathbb{N}_{2k}^n} |\{(\alpha, \beta) \in (\mathbb{N}_k^n)^2 : \alpha + \beta = \gamma, \alpha \leq \beta\}| - \binom{n}{2k}.
\]

It must be equal to \( r \) as in (56). We just proved that \( \bigcup_{\gamma \in \mathbb{N}_{2k}^n} B_{\gamma} \) is a basis of \( \mathcal{V}^\perp \). Now we assume that \( \bigcup_{\gamma \in \mathbb{N}_{2k}^n} B_{\gamma} = \{A_1, \ldots, A_r\} \).
Let us rewrite the constraints

$$M_k^{-[h_j]}(h_j \ y) = 0, \ j \in [l_h]. \quad (57)$$

as $\langle A_j, X \rangle = 0$, $j = r + 1, \ldots, m - 1$ with $X = P_kM_k(y)P_k$. From (57),

$$\sum_{\gamma \in \mathbb{N}^n_2(h_j)} h_{j,\gamma} y_{\alpha+\gamma} = 0, \ \alpha \in \mathbb{N}^n_{2(k-\lceil h_j \rceil)}, \ j \in [l_h]. \quad (58)$$

Let $j \in [l_h]$ and $\alpha \in \mathbb{N}^n_{2(k-\lceil h_j \rceil)}$ be fixed. We define $\tilde{A} = (\tilde{A}_{\mu,v})_{\mu,v \in \mathbb{N}^n_k}$ as follows:

$$\tilde{A}_{\mu,v} = \begin{cases} h_{j,\gamma} & \text{if } \mu = v, \mu + v = \alpha + \gamma, \\ \frac{1}{2} h_{j,\gamma} & \text{if } \mu \neq v, \mu + v = \alpha + \gamma, \\ 0 & \text{otherwise.} \end{cases}$$

(59)

Then (58) implies that $\langle \tilde{A}, M_k(y) \rangle = 0$. Since $M_k(y) = P_k^{-1}XP_k^{-1}$,

$$0 = \langle \tilde{A}, P_k^{-1}XP_k^{-1} \rangle = \langle P_k^{-1}\tilde{A}P_k^{-1}, X \rangle = \langle A, X \rangle,$$

where $A := P_k^{-1}\tilde{A}P_k^{-1}$, yielding the statement. Thus, we obtain the constraints $\langle A_j, X \rangle = 0$, $j \in [m - 1]$.

The final constraint $y_0 = 1$ can be rewritten as $\langle A_m, X \rangle = 1$ with $A_m \in S_k$ having zero entries except the top left one $[A_m]_{0,0} = 1$. Thus, we select $b$ such that all entries of $b$ are zeros except $b_m = 1$.

The number $m$ (or $m_k$ when plugging the relaxation order $k$) of equality trace constraints $\langle A_j, X \rangle = b_j$ is:

$$m = \frac{1}{2} \left\lceil \frac{n}{k} \right\rceil \left( \left\lceil \frac{n}{k} \right\rceil + 1 \right) - \left\lceil \frac{n}{2k} \right\rceil + 1 + \sum_{j=1}^{l_h} \left\lceil \frac{n}{2(k-\lceil h_j \rceil)} \right\rceil. \quad (60)$$

The function $-L_y(f) = -\sum_{\gamma} f_{\gamma} y_{\gamma}$ is equal to $\langle C, X \rangle$ with $C := P_k^{-1}\tilde{C}P_k^{-1}$, where $\tilde{C} = (\tilde{C}_{\mu,v})_{\mu,v \in \mathbb{N}^n_k}$ is defined by:

$$\tilde{C}_{\mu,v} = \begin{cases} -f_{\gamma} & \text{if } \mu = v, \mu + v = \gamma, \\ -\frac{1}{2} f_{\gamma} & \text{if } \mu \neq v, \mu + v = \gamma, \\ 0 & \text{otherwise.} \end{cases}$$

(61)
References

1. Bagirov, A., Karmitsa, N., Mäkelä, M.M.: Introduction to Nonsmooth Optimization: Theory, Practice and Software. Springer, New York (2014)
2. Ben-Tal, A., Nemirovski, A.: Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications, vol. 2. Siam, Philadelphia (2001)
3. Bhaint, A.: Optimization of upper semidifferential functions. J. Optim. Theory Appl. 44(4), 545–568 (1984)
4. Burer, S., Monteiro, R.D.: A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. Math. Program. 95(2), 329–357 (2003)
5. Burke, J.V., Lewis, A.S., Overton, M.L.: A robust gradient sampling algorithm for nonsmooth, nonconvex optimization. SIAM J. Optim. 15(3), 751–779 (2005)
6. Chandrasekaran, V., Shah, P.: Relative entropy relaxations for signomial optimization. SIAM J. Optim. 26(2), 1147–1173 (2016)
7. Curtis, F.E., Que, X.: A quasi-Newton algorithm for nonconvex, nonsmooth optimization with global convergence guarantees. Math. Program. Comput. 7(4), 399–428 (2015)
8. Curto, R.E., Fialkow, L.A.: Truncated K-moment problems in several variables. J. Oper. Theory 189–226 (2005)
9. Dahl, J.: Semidefinite Optimization using MOSEK. ISMP, Berlin (2012)
10. d’Aspremont, A., El Karoui, N.: A stochastic smoothing algorithm for semidefinite programming. SIAM J. Optim. 24(3), 1138–1177 (2014)
11. Ding, L., Yurtsever, A., Cevher, V., Tropp, J.A., Udell, M.: An optimal-storage approach to semidefinite programming using approximate complementarity. SIAM J. Optim. 31(4), 2695–2725 (2021)
12. Doherty, A.C., Wehner, S.: Convergence of SDP hierarchies for polynomial optimization on the hypersphere. arXiv preprint arXiv:1210.5048, (2012)
13. Dressler, M., Ilum, S., de Wolff, T.: A positivstellensatz for sums of nonnegative circuit polynomials. SIAM J. Appl. Algebra Geom. 1(1), 536–555 (2017)
14. Haarala, M., Miettinen, K., Mäkelä, M.M.: New limited memory bundle method for large-scale nonsmooth optimization. Optim. Methods Softw. 19(6), 673–692 (2004)
15. Haarala, N., Miettinen, K., Mäkelä, M.M.: Globally convergent limited memory bundle method for large-scale nonsmooth optimization. Math. Program. 109(1), 181–205 (2007)
16. Helmberg, C., Overton, M.L., Rendl, F.: The spectral bundle method with second-order information. Opt. Methods Softw. 29(4), 855–876 (2014)
17. Helmberg, C., Rendl, F.: A spectral bundle method for semidefinite programming. SIAM J. Optim. 10(3), 673–696 (2000)
18. Henrion, D., Lasserre, J.-B.: Detecting global optimality and extracting solutions in gloptipoly. Posit. Polynomials Control 312, 293–310 (2005)
19. Henrion, D., Malick, J.: Projection methods in conic optimization. Handb. Semidefinite Conic Polynomial Optim. 565–600 (2012)
20. Josz, C., Henrion, D.: Strong duality in Lasserre’s hierarchy for polynomial optimization. Optim. Lett. 10(1), 3–10 (2016)
21. Journée, M., Bach, F., Absil, P.-A., Sepulchre, R.: Low-rank optimization for semidefinite convex problems. arXiv preprint arXiv:0807.4423, (2008)
22. Karmitsa, N.: LMBM-FORTRAN subroutines for Large-Scale nonsmooth minimization: user’s manual. TUCS Techn. Rep. 77(856) (2007)
23. Kiwiel, K.C.: Proximity control in bundle methods for convex nondifferentiable minimization. Math. Program. 46(1–3), 105–122 (1990)
24. Kiwiel, K.C.: Convergence of the gradient sampling algorithm for nonsmooth nonconvex optimization. SIAM J. Optim. 18(2), 379–388 (2007)
25. Lasserre, J.B.: Global optimization with polynomials and the problem of moments. SIAM J. Optim. 11(3), 796–817 (2001)
26. Lasserre, J.-B.: Convergent SDP-relaxations in polynomial optimization with sparsity. SIAM J. Optim. 17(3), 822–843 (2006)
27. Lasserre, J.-B.: Moments, Positive Polynomials and Their Applications, vol. 1. World Scientific, Singapore (2010)
28. Lasserre, J.B.: An Introduction to Polynomial and Semi-algebraic Optimization, vol. 52. Cambridge University Press, Cambridge (2015)
29. Lasserre, J.B.: Homogeneous polynomials and spurious local minima on the unit sphere. Optim. Lett. 1–14 (2021)
30. Lasserre, J.B., Laurent, M., Rostalski, P.: Semidefinite characterization and computation of zero-dimensional real radical ideals. Found. Comput. Math. 8(5), 607–647 (2008)
31. Lasserre, J.B., Toh, K.-C., Yang, S.: A bounded degree SOS hierarchy for polynomial optimization. EURO J. Comput. Optim. 5(1–2), 87–117 (2017)
32. Laurent, M.: Revisiting two theorems of Curto and Fialkow on moment matrices. Proc. Am. Math. Soc. 133(10), 2965–2976 (2005)
33. Lehoucq, R.B., Sorensen, D.C., Yang, C.: ARPACK Users’ Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods. SIAM, Philadelphia (1998)
34. Lewis, A.S., Overton, M.L.: Nonsmooth optimization via quasi-Newton methods. Math. Program. 141(1–2), 135–163 (2013)
35. Mai, N.H.A., Bhardwaj, A., Magron, V.: The constant trace property in noncommutative optimization. In: Proceedings of the 2021 on International Symposium on Symbolic and Algebraic Computation, ISSAC ’21, pp. 297–304, New York, NY, USA. Association for Computing Machinery (2021)
36. Mai, N.H.A., Lasserre, J.-B., Magron, V., Wang, J.: Exploiting constant trace property in large-scale polynomial optimization. ACM Trans. Math. Softw. 48(4), 1–39 (2022)
37. Mifflin, R.: An algorithm for constrained optimization with semismooth functions. Math. Oper. Res. 2(2), 191–207 (1977)
38. Navascués, M., Pironio, S., Acín, A.: A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. New J. Phys. 10(7), 073013 (2008)
39. Nemirovsky, A., Yudin, D.: Problem complexity and method efficiency in optimization. Nauka (1983)
40. Nie, J.: Optimality conditions and finite convergence of Lasserre’s hierarchy. Math. Program. 146(1–2), 97–121 (2014)
41. Nie, J., Schweighofer, M.: On the complexity of Putinar’s Positivstellensatz. J. Complex. 23(1), 135–150 (2007)
42. Nocedal, J.: Updating quasi-Newton matrices with limited storage. Math. Comput. 35(151), 773–782 (1980)
43. Nocedal, J., Wright, S.: Numerical Optimization. Springer, New York (2006)
44. Overton, M.L.: Large-scale optimization of eigenvalues. SIAM J. Optim. 2(1), 88–120 (1992)
45. Overton, M.L., Womersley, R.S.: Second derivatives for optimizing eigenvalues of symmetric matrices. SIAM J. Matrix Anal. Appl. 16(3), 697–718 (1995)
46. Saad, Y.: Numerical Methods for Large Eigenvalue Problems, revised SIAM, Philadelphia (2011)
47. Schweighofer, M.: On the complexity of Schmüdgen’s positivstellensatz. J. Complex. 20(4), 529–543 (2004)
48. Shor, N.Z.: Quadratic optimization problems. Sov. J. Comput. Syst. Sci. 25, 1–11 (1987)
49. Tmovska, M.: Strong duality conditions in semidefinite programming. J. Electr. Eng. 56(12), 1–5 (2005)
50. Waki, H., Kim, S., Kojima, M., Muramatsu, M.: Sums of squares and semidefinite programming relaxations for polynomial optimization problems with structured sparsity. SIAM J. Optim. 17(1), 218–242 (2006)
51. Wang, J., Magron, V.: A second order cone characterization for sums of nonnegative circuits. In: Proceedings of the 45th International Symposium on Symbolic and Algebraic Computation, pp. 450–457 (2020)
52. Wang, J., Magron, V., Lasserre, J.-B.: Chordal-TSSOS: a moment-SOS hierarchy that exploits term sparsity with chordal extension. SIAM J. Optim. 31(1), 114–141 (2021)
53. Wang, J., Magron, V., Lasserre, J.-B.: TSSOS: A Moment-SOS hierarchy that exploits term sparsity. SIAM J. Optim. 31(1), 30–58 (2021)
54. Wang, J., Magron, V., Lasserre, J.B., Mai, N.H.A.: CS-TSSOS: correlative and term sparsity for large-scale polynomial optimization. ACM Trans. Math. Softw. 48(4), 1–26 (2022)
55. Weisser, T., Legat, B., Coey, C., Kapelevich, L., Vielma, J.P.: Polynomial and moment optimization in Julia and JuMP. In: JuliaCon (2019)
56. Yurtsever, A., Tropp, J.A., Fercq, O., Udell, M., Cevher, V.: Scalable semidefinite programming. SIAM J. Math. Data Sci. 3(1), 171–200 (2021)
