J Glob Optim (2010) 48:549–567
DOI 10.1007/s10898-010-9529-5

Partitioning procedure for polynomial optimization

Polyxeni-Margarita Kleniati · Panos Parpas · Berç Rustem

Received: 15 January 2010 / Accepted: 26 January 2010 / Published online: 13 February 2010

Abstract We consider the problem of finding the minimum of a real-valued multivariate polynomial function constrained in a compact set defined by polynomial inequalities and equalities. This problem, called polynomial optimization problem (POP), is generally nonconvex and has been of growing interest to many researchers in recent years. Our goal is to tackle POPs using decomposition, based on a partitioning procedure. The problem manipulations are in line with the pattern used in the generalized Benders decomposition, namely projection followed by relaxation. Stengle's and Putinar's Positivstellensätze are employed to derive the feasibility and optimality constraints, respectively. We test the performance of the proposed partitioning procedure on a collection of benchmark problems and present the numerical results.

Keywords Polynomial optimization · Positivstellensatz · Sum of squares · Benders decomposition

1 Introduction

Global optimization of polynomials and semidefinite programming have attracted considerable attention in the last decade. Both areas of research have numerous applications of interest in fields such as finance, statistics, control theory, and combinatorial optimization.

The goal of global optimization is the computation of global optimal solutions of nonconvex functions constrained in a specified domain [6]. Semidefinite programming (SDP) is linear programming over positive semidefinite matrices [4].

P.-M. Kleniati (✉) · B. Rustem
Department of Computing, Imperial College London, 180 Queen's Gate, SW7 2AZ London, UK
e-mail: pk2003@doc.ic.ac.uk

P. Parpas
Energy Initiative Engineering Systems Division, Massachusetts Institute of Technology, 400 Main Street, Cambridge, MA 02139, UK

Consider the polynomial optimization problem (POP) below:

$$p^* = \min_{x \in \mathscr{K}} p(x),$$

where $p(x)$ is a real-valued multivariate polynomial $p(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ and $\mathscr{K}$ is a compact semialgebraic set.[1] POPs are generally characterized by nonconvexities, which make them to be global optimization problems. As shown by Lasserre [14] and Parrilo [20], one is able to convexify a POP by employing the moment problem and its interaction with positive polynomials and semidefinite programming. In particular, one can approximate p^* by solving a sequence of (convex) semidefinite relaxations of increasing size. The relaxations can be solved efficiently by interior-point methods in polynomial time [2]. The sequence of solutions of the relaxations provide monotonically increasing lower bounds to the global optimal value p^* of the POP. These bounds converge asymptotically to p^* [14]. In the rest of the paper, this procedure is referred to as the SDP relaxation technique.

In this work, we aim at tackling POPs using decomposition, based on a partitioning procedure. The problem manipulations carried out are in line with the pattern used in the generalized Benders decomposition, namely projection followed by relaxation [10]. In particular, we partition the polynomial variables into two disjoint subsets and assign them to a *subproblem* and a *master problem*, respectively. By solving a series of subproblems and relaxed master problems we obtain upper and lower bounds, respectively, of the global optimal objective value of the POP. Namely, at each iteration, the subproblem first is checked for feasibility. If infeasible, we apply Stengle's Positivstellensatz to the semialgebraic set of the subproblem and derive a feasibility constraint. If feasible, we apply the sum of square decomposition for multivariate polynomials based on Putinar's Positivstellensatz and derive an optimality constraint. The use of these theorems was motivated by the SDP relaxation technique for POPs [14,20]. The optimal objective value of the subproblem gives an upper bound of the optimal value of the POP. Then, the relaxed master problem, which is augmented by either the feasibility or the optimality constraint, gives a lower bound of the optimal objective value of the POP.

Our partitioning procedure was inspired by the generalized Benders decomposition, which is applicable to convex nonlinear optimization problems [10]. The proposed algorithm was also motivated by Wolsey [32] and Floudas et al. [8,29], who have developed decomposition-based algorithms for nonconvex nonlinear optimization problems. On the one hand, Wolsey [32] employs general duality theory to formulate the optimality constraints and yields functional multipliers in place of the constant multipliers found in the convex analog.[2] As shown in Theorems 5 and 6, we too derive polynomial multipliers, which are non constant but functional multipliers. On the other hand, Floudas et al. [8,29] consider nonconvex nonlinear problems that satisfy some convexity conditions when a subset of variables is fixed. In this case convex duality theory is employed and the resulting multipliers are in line with the convex analog. The latter decomposition-based algorithm is known as the *Global OPtimization algorithm* or briefly *GOP*. The same authors also applied GOP to univariate POPs [30].

Contribution Our algorithm is an extension of the generalized Benders decomposition for convex optimization [10] to the global optimization of POPs. Accompanying theoretical results are also stated regarding the formulation of feasibility and optimality constraints. In addition to these results, we prove in Theorem 7 that our procedure terminates without cycling and attains ε-global optimality. Moreover, asymptotic ε-convergence is shown

[1] A set is *compact* if it is both closed and bounded; a set is *semialgebraic* if it is a Boolean combination of polynomial inequalities and equalities [21, Sect. 2.1].

[2] "Convex analog" refers to the generalized Benders decomposition [10].

in Theorem 8. The asymptoticity comes from the underlying SDP relaxation technique. Nevertheless, practice demonstrated that the algorithm generally terminates in a finite number of iterations.

This paper is organized as follows. In Sect. 2, we discuss the SDP relaxation technique for global optimization of polynomials as this is essential to the theoretical development of our procedure. In Sect. 3, we introduce the partitioning procedure for POPs. Theoretical derivation of the master problem, as well as convergence of our procedure are some of the main topics of this section. In Sect. 4, we apply our method to a collection of well-known test problems and we present the results. Sect. 5 summarizes.

Notation By $\mathbb{R}[x] = \mathbb{R}[x, \ldots, x_n]$ we denote the *polynomial ring* over $\mathbb{R}$ in n variables. In addition, we use $\Sigma^2 \subseteq \mathbb{R}[x]$ to denote the set of squares of polynomials in this polynomial ring. The same symbols appended by a subscript, e.g $\mathbb{R}_d[x]$ and Σ^2_d, denote the set of polynomials up to degree d and the set of squares of polynomials up to degree $2d$, respectively.

2 Optimization over POPs: relevant theory

In this section, we discuss the semidefinite relaxation technique for solving polynomials optimization problems. For a thorough investigation, the interested reader is referred to [14,20] and the references therein. In addition, [16] consists of a detailed survey on the topic.

Let us consider the POP below,

$$\begin{aligned} p^* &= \min_x p(x) \\ \text{s.t. } g_i(x) &\geq 0, \quad i = 1, \ldots, m, \\ h_j(x) &= 0, \quad j = 1, \ldots, r. \end{aligned} \tag{1}$$

The objective function $p(x) \in \mathbb{R}[x]$ is a polynomial of degree d_0 and the feasible set is a basic closed semialgebraic set,

$$\mathcal{K} = \{x \in \mathbb{R}^n \mid g_1(x) \geq 0, \ldots, g_m(x) \geq 0, \quad h_1(x) = 0, \ldots, h_r(x) = 0\}, \tag{2}$$

where $g_1 \ldots, g_m, h_1, \ldots, h_r \in \mathbb{R}[x]$ are polynomials of degrees $d_1, \ldots, d_{m+r}$, respectively. In general, when one deals with an optimization problem, two issues are of major importance. The first is to determine whether or not the feasible set is empty and the second is to compute the optimal solution vector and the optimal objective value over this set, if it is nonempty. For problem (1), one can check whether or not the set $\mathcal{K}$ is empty by employing the Stengle's Positivstellensatz. This is stated in the following theorem.

Theorem 1 (Stengle's Positivstellensatz [25]) *The set $\mathcal{K}$ is empty if and only if there exist sum-of-squares polynomial multipliers $\sigma_I(x) \in \Sigma^2$, $I \subseteq \{1, \ldots, m\}$, and polynomial multipliers $t_j \in \mathbb{R}[x]$, $j = 1, \ldots, r$, such that:*

$$-1 = \sum_{I \subseteq \{1,\ldots,m\}} \sigma_I(x) g_I(x) + \sum_{j=1}^{r} t_j(x) h_j(x), \quad \forall x \in \mathbb{R}^n. \tag{3}$$

The first expression in the right-hand side of (3) corresponds to an element of the *preordering* $\mathscr{P}^{\mathrm{PO}}(g_1,\ldots,g_m) \subseteq \mathscr{P}(\mathscr{K})$, a cone generated by $g_1,\ldots,g_m$ [21]:

$$\mathscr{P}^{\mathrm{PO}}(g_1,\ldots,g_m) = \left\{ \sum_{I\subseteq\{1,\ldots,m\}} \sigma_I(x) g_I(x) \ \mid \ \sigma_I \in \Sigma^2 \right\}, \tag{4}$$

where $\mathscr{P}(\mathscr{K})$ is the convex cone of positive polynomials on $\mathscr{K}$ [16]. The second expression in the right-hand side of (3) is an element of the *ideal* $\mathscr{I}$ generated by $h_1,\ldots,h_r$ [5]:

$$\mathscr{I}(h_1,\ldots,h_r) = \left\{ \sum_{j=1}^{r} t_j(x) h_j(x) \ \mid \ t_j \in \mathbb{R}[x] \right\}. \tag{5}$$

The preordering is also used in Schmüdgen's Positivstellensatz to represent a positive polynomial on a compact semialgebraic set [23]. As far as computing for the optimal solution and objective value is concerned, when the feasible set $\mathscr{K}$ is nonempty, computing for the global optimal solution of the POP (1) can be a hard problem. This is due to the fact that the set $\mathscr{K}$ and the function $p(x) : \mathscr{K} \to \mathbb{R}$ are usually nonconvex. According to the SDP relaxation technique, to convexify the problem one can write:

$$\begin{aligned} p^* = {} & \sup\ \gamma \\ & \text{s.t. } \ p(x) - \gamma \in \mathscr{P}(\mathscr{K}). \end{aligned} \tag{6}$$

However, the cone $\mathscr{P}(\mathscr{K})$ is difficult to describe and yields intractable problems. For this reason, one can consider instead the *quadratic module* $\mathscr{P}^{\mathrm{QM}}(g_1,\ldots,g_m) \subseteq \mathscr{P}(\mathscr{K})$, a cone generated by $g_1,\ldots,g_m$:

$$\mathscr{P}^{\mathrm{QM}}(g_1,\ldots,g_m) = \left\{ \sigma_0 + \sum_{i=1}^{m} \sigma_i(x) g_i(x) \ \mid \ \sigma_0,\sigma_1,\ldots,\sigma_m \in \Sigma^2 \right\}. \tag{7}$$

As a result, one obtains the following approximate problem:

$$\begin{aligned} \gamma^* = {} & \sup\ \gamma \\ & \text{s.t. } \ p(x) - \gamma \in \mathscr{P}^{\mathrm{QM}}(g_1,\ldots,g_m) + \mathscr{I}(h_1,\ldots,h_r). \end{aligned} \tag{8}$$

The quadratic module used in the formulation above is due to Putinar's Positivstellensatz, which is a refinement of Schmüdgen's Positivstellensatz, using less sum-of-squares polynomial multipliers to represent a positive polynomial on a semialgebraic set. To achieve this, it is based on a stronger assumption than just compactness of $\mathscr{K}$ which we state below.

Assumption 1 ([14]) *The set $\mathscr{K}$ is compact and there exists a polynomial $u(x) : \mathbb{R}^n \to \mathbb{R}$ such that $\{u(x) \geq 0\}$ is compact and*

$$u(x) = u_0(x) + \sum_{i=1}^{m} u_i(x) g_i(x) + \sum_{j=1}^{r} t_j(x) h_j(x), \tag{9}$$

for all $x \in \mathbb{R}^n$, where the polynomials $u_i(x)$, $i = 0,\ldots,m$, are sum of squares.

Assumption 1 may be stronger than compactness but it is not a restrictive one. In particular, it is satisfied if there exists a polynomial $g_i(x)$ such that the set $\{g_i(x) \geq 0\}$ is compact. One can also add an extra inequality to ensure the satisfiability of the assumption, such as the constraint $M^2 - \|x\|^2 \geq 0$ for M sufficiently large [14,24].

Theorem 2 (Putinar's Positivstellensatz [22,18]) *If Assumption 1 holds, every real polynomial f, positive on $\mathscr{K}$, possesses a representation:*

$$f(x) = \sigma_0(x) + \sum_{i=1}^{m} \sigma_i(x) g_i(x) + \sum_{j=1}^{r} t_j(x) h_j(x), \tag{10}$$

for all $x \in \mathbb{R}^n$, where $\sigma_0, \ldots, \sigma_m$ are sum of squares.

Such a representation is a certificate for the nonnegativity of f on $\mathscr{K}$. Since $\mathscr{P}^{\mathrm{QM}}(\mathscr{K}) \subseteq \mathscr{P}(\mathscr{K})$, the objective value of problem (8) is a lower bound on the optimal objective value p^*, i.e. $\gamma^* \leq p^*$ [16]. In view of this and if Assumption 1 is satisfied, the global optimal value of problem (1) is approximated by the hierarchy of bounds:

$$\begin{aligned} \gamma_\omega = & \sup_{\substack{\sigma_i \in \Sigma^2_{\omega - \lceil \frac{d_i}{2} \rceil} \\ t_j \in \mathbb{R}_{2\omega - d_j}[x]}} \gamma \\ & \text{s.t. } p(x) - \gamma = \sigma_0(x) + \sum_{i=1}^{m} \sigma(x) g_i(x) + \sum_{j=1}^{r} t_j(x) h_j(h). \end{aligned} \tag{11}$$

The constraint in (11) is known as sum-of-squares (SOS) constraint. Observe that all summands in the SOS constraint have bounded degrees:

$$\deg(\sigma_0), \deg(\sigma_1 g_1), \ldots, \deg(\sigma_m g_m), \deg(t_1 h_1), \ldots, \deg(t_r h_r) \leq 2\omega,$$

for any $\omega \geq \max\{\lceil \frac{d_0}{2} \rceil, \lceil \frac{d_1}{2} \rceil, \ldots, \lceil \frac{d_{m+r}}{2} \rceil\}$, where ω is the *relaxation order*. When ω is fixed, γ_ω is efficiently computed via SDP. Notice that $\sup_\omega \gamma_\omega = \gamma^*$. Moreover, by increasing ω and solving the corresponding convex approximate problems (11), one attains asymptotic convergence of γ^* to p^*, as stated in the following theorem.

Theorem 3 ([14,16]) *If the semialgebraic set $\mathscr{K}$ satisfies Assumption 1, then problems (1) and (8) have the same optimal values, i.e.*

$$\lim_{\omega \to \infty} \gamma_\omega = \gamma^* = p^*.$$

3 Partitioning procedure for polynomial optimization

The essence of decomposition schemes, such as the generalized Benders decomposition for convex programs [10], is to initially derive the so-called *master problem* such that it is equivalent to the original problem, and secondly employ a series of subproblems in order to solve the master problem. We consider the following POP:

$$\begin{aligned} p^* = & \min_{x,y} p(x, y) \\ & \text{s.t. } g_i(x, y) \geq 0, \quad i = 1, \ldots, m, \\ & h_j(x, y) = 0, \quad j = 1, \ldots, r, \\ & x \in X, \; y \in Y, \end{aligned} \tag{12}$$

where $p, g_1, \ldots, g_m, h_1, \ldots, h_r \in \mathbb{R}[x]$. Also, $x = (x, y) \in \mathbb{R}^n$ and the sets $X \subseteq \mathbb{R}^{n_1}$ and $Y \subseteq \mathbb{R}^{n_2}$, where $n = n_1 + n_2$, are assumed to be compact. The feasible region of our problem is a basic closed semialgebraic set,

$$\mathcal{K} = \{(x, y) \in X \times Y \subseteq \mathbb{R}^n \mid g_i(x, y) \geq 0,\ i = 1, \ldots, m,\ h_j(x, y) = 0,\ j = 1, \ldots, r\}. \tag{13}$$

We assume that the set $\mathcal{K}$ is non-empty and compact.

3.1 Derivation of the master problem

If we apply the concept of projection [9], often referred to as partitioning, we can express problem (12) as a problem in y-space as follows.

$$\begin{aligned} p^* = \ & \min_y \quad v(y) \\ & \text{s.t.} \quad y \in Y \cap V, \end{aligned} \tag{14}$$

where

$$\begin{aligned} v(y) &= \inf_{x \in X} p(x, y) \\ \text{s.t. } g_i(x, y) &\geq 0,\ i = 1, \ldots, m, \\ h_j(x, y) &= 0,\ j = 1, \ldots, r, \end{aligned} \tag{15}$$

and

$$V = \{y \mid g_i(x, y) \geq 0,\ i = 1, \ldots, m,\ h_j(x, y) = 0,\ j = 1, \ldots, r,\ \text{for some } x \in X\}. \tag{16}$$

Observe that $v(y)$ is the optimal value of (12) for a given fixed y. Hence, $v(y)$ is an upper bound on p^*. To obtain $v(y)$ for a fixed y, we first have to solve the POP below:

$$\begin{aligned} \min_{x \in X} \quad & p(x, y) \\ \text{s.t.} \quad & g_i(x, y) \geq 0,\ i = 1, \ldots, m, \\ & h_j(x, y) = 0,\ j = 1, \ldots, r. \end{aligned} \tag{17}$$

The set V introduced in (16) consists of those values of y for which (17) is feasible and $Y \cap V$ is the projection of the feasible region of (12) onto y-space. The projection of a semialgebraic set is also semialgebraic [19]. The feasible region of subproblem (17), with respect to a fixed y, is the following semialgebraic set:

$$\mathcal{K}(y) = \{x \in X \subseteq \mathbb{R}^{n_1} \mid g_i(x, y) \geq 0,\ i = 1, \ldots, m,\ h_j(x, y) = 0,\ j = 1, \ldots, r\}. \tag{18}$$

The assumption that $\mathcal{K}$ is compact implies that the sets $\mathcal{K}(y)$, $y \in Y \cap V$, are also compact. Moreover, for each set $\mathcal{K}(y)$, with respect to a fixed y, the inequality constraints induce a polynomial cone, that is either expressed by the preordering in (4) or by the quadratic module in (7), and the equality constraints induce a polynomial ideal expressed in (5). For a fixed y, the preordering, quadratic module and polynomial ideal, respectively, follow:

$$\mathscr{P}^{\text{PO}}_{\mathscr{K}(y)} = \left\{ \sum_{I \subseteq \{1,\ldots,m\}} \sigma_I^{(y)}(x) g_I(x,y) \mid \sigma_I^{(y)} \in \Sigma^2,\ I \subseteq \{1,\ldots,m\} \right\},$$

$$\mathscr{P}^{\text{QM}}_{\mathscr{K}(y)} = \left\{ \sigma_0^{(y)}(x) + \sum_{i=1}^{m} \sigma_i^{(y)}(x) g_i(x,y) \mid \sigma_i^{(y)} \in \Sigma^2,\ i = 0,\ldots,m \right\},$$

$$\mathscr{I}_{\mathscr{K}(y)} = \left\{ \sum_{j=1}^{r} t_j^{(y)}(x) h_j(x,y) \mid t_j^{(y)} \in \mathbb{R}[x],\ j = 1,\ldots,r \right\}.$$

Following the idea underlying the generalized Benders decomposition, the three following manipulations are essential to derive the master problem: (i) projection; (ii) dual representation of V; (iii) dual representation of $v(y)$. Earlier we expressed problem (12) as a problem onto y-space. In other words, by projection we managed to represent (12) in terms of (14) and the first problem manipulation has been completed. Problem (14) is equivalent to (12) and therefore, solving (14) is akin to solving (12).

Theorem 4 (Projection [9]) *Problem* (12) *is infeasible or has unbounded value if and only if the same is true of problem* (14). *If* (x^*, y^*) *is optimal in* (12) *then* y^* *must be optimal in* (14). *If* y^* *is optimal in* (14) *and* x^* *achieves the infimum in* (15) *for* $y = y^*$, *then* (x^*, y^*) *is optimal in* (12). *If* y^* *is* ε_1*-optimal in* (14) *and* x^* *is* ε_2*-optimal in* (17), *then* (x^*, y^*) *is* $(\varepsilon_1 + \varepsilon_2)$*-optimal in* (12).

Next, we would like to invoke a dual representation of the set V in (16). To achieve this, we introduce the next theorem that aims at expressing formally the set of feasibility constraints.

Theorem 5 (Feasibility Constraints) *Assume that* X *is a nonempty set. A point* $\hat{y} \in Y$ *is also in the set* V *if and only if* $\hat{y}$ *satisfies the (infinite) system:*

$$\inf_{x \in X} \left\{ -\sum_{I \subseteq \{1,\ldots,m\}} \sigma_I(x) g_I(x,y) - \sum_{j=1}^{r} t_j(x) h_j(x,y) \right\} \le 0, \tag{19}$$

for all polynomial multipliers $\sigma_I(x) \in \Sigma^2$, $I \subseteq \{1,\ldots,m\}$,[3] *and* $t_j \in \mathbb{R}[x]$, $j = 1,\ldots,r$.

Proof Let $\hat{y} \in Y$. If $\hat{y} \in V$ then $\hat{y}$ satisfies (19). To prove the converse let us assume that $\hat{y}$ satisfies conditions (19) and that $\hat{y} \notin V$. Since $\hat{y} \notin V$ the set $\mathscr{K}(\hat{y}) = \{x \in X \mid g_i(x,\hat{y}) \ge 0,\ i = 1,\ldots,m,\ h_j(x,\hat{y}) = 0,\ j = 1,\ldots,r\}$ is empty. Then according to Theorem 1, there exist polynomial multipliers $\sigma_I \in \Sigma^2$, $I \subseteq \{1,\ldots,m\}$, and $t_j \in \mathbb{R}[x]$, $j = 1,\ldots,r$, such that:

$$\sum_{I \subseteq \{1,\ldots,m\}} \sigma_I(x) g_I(x,\hat{y}) + \sum_{j=1}^{r} t_j(x) h_j(x,\hat{y}) = -1, \quad \forall x \in X. \tag{20}$$

But this contradicts our assumption, namely that $\hat{y}$ satisfies conditions (19); hence, $\hat{y} \in V$. □

Success of proving infeasibility of the subproblem gives the necessary (polynomial) multipliers and the feasibility constraint which is added to the master problem. Failure to prove infeasibility provides points in the solution set. This is tackled by the following and last manipulation, namely by invoking a dual representation of $v(y)$. Such a dual representation is the sum-of-squares formulation based on Putinar's Positivstellensatz [14].

[3] Notice that $I \subseteq \{1,\ldots,m\}$, hence the index I starts with zero value corresponding to the empty set. The first term of the sum in (19) is then $\sigma_0(x) g_0(x,y)$, where $g_0(x,y) = 1$.

Assumption 2 ([14]) *Let $y \in Y \cap V$. The set $\mathcal{K}(y)$ is compact*[4] *and we assume there exists a polynomial $s^{(y)}(x) : \mathbb{R}^{n_1} \to \mathbb{R}$ such that $\{s^{(y)}(x) \geq 0\}$ is compact and*

$$s^{(y)}(x) = s_0^{(y)}(x) + \sum_{i=1}^{m} s_i^{(y)}(x) g_i(x, y) + \sum_{j=1}^{r} t_j^{(y)}(x) h_j(x, y), \tag{21}$$

for all $x \in X$, where the polynomials $s_i^{(y)}(x)$, $i = 0, \ldots, m$, are sum of squares.

Theorem 6 (Optimality Constraints) *If Assumption 2 holds for every $y \in Y \cap V$, then the optimal value of (17) equals that of its dual on $Y \cap V$:*

$$v(y) = \sup_{\substack{\sigma_i \in \Sigma^2 \\ t_j \in \mathbb{R}[x]}} \left\{ \inf_{x \in X} \{p(x, y) - \sum_{i=0}^{m} \sigma_i(x) g_i(x, y) - \sum_{j=1}^{r} t_j(x) h_j(x, y)\} \right\}, \tag{22}$$

where $g_0(x, y) = 1$.

Proof Let γ^* be the optimal objective value of v(y). If we fix the degree of the summands involved to be less than 2ω we get the corresponding relaxation of order ω which has optimal objective value $\gamma_\omega \leq \gamma^*$. Since it is assumed that the feasible set of the subproblem (17) satisfies Assumption 2, and this is directly analogous to a semialgebraic set $\mathcal{K}$, as in (2), satisfying Assumption 1 in Theorem 3, we can conclude that there is indeed asymptotic convergence to the global optimal value: $\lim_{\omega \to \infty} \gamma_\omega = \gamma^* = v(y)$. □

Using Theorem 6 we rewrite problem (14) as follows:

$$\begin{aligned} &\min_{y \in Y \cap V} \; v(y) \\ &\text{s.t.} \quad v(y) = \sup_{\substack{\sigma_i \in \Sigma^2 \\ t_j \in \mathbb{R}[x]}} \left\{ \inf_{x \in X} \{p(x, y) - \sum_{i=0}^{m} \sigma_i(x) g_i(x, y) - \sum_{j=1}^{r} t_j(x) h_j(x, y)\} \right\}. \end{aligned} \tag{23}$$

Next by introducing a scalar variable z, using the definition of supremum, and replacing the constraint $y \in Y \cap V$ by conditions (19) we derive the following formulation:

$$\begin{aligned} &\min_{y,z} \; z \\ &\text{s.t.} \quad z \geq \inf_{x \in X} \left\{ p(x, y) - \sum_{i=0}^{m} \sigma_i(x) g_i(x, y) - \sum_{j=1}^{r} t_j(x) h_j(x, y) \right\}, \quad \forall \sigma_i \in \Sigma^2, \; t \in \mathbb{R}[x], \\ &\qquad 0 \geq \inf_{x \in X} \left\{ - \sum_{I \subseteq \{1,\ldots,m\}} \sigma_I(x) g_I(x, y) - \sum_{j=1}^{r} t_j(x) h_j(x, y) \right\}, \quad \forall \sigma_I \in \Sigma^2, \; t \in \mathbb{R}[x], \end{aligned} \tag{24}$$

which is equivalent to (12) and is our master problem. The first and the second sets of constraints in (24) represent the sets of optimality and feasibility constraints, respectively. The set of constraints of the master problem therefore, consists of a complete set of optimality constraints and a separate complete set of feasibility constraints. Observe that the feasibility constraints include all the 2^m square-free products of the inequality constraints $g_i(x, y)$,

[4] The sets $\mathcal{K}(y)$ are compact for all $y \in Y \cap V$ by construction.

$i = 1, \ldots, m$, due to the Stengle's Positivstellensatz, plus r polynomial constraints due to the equality constraints. Thus, for the generation of each feasibility constraint we need $2^m + r$ polynomial multipliers. On the other hand, the optimality constraints include significantly fewer polynomial multipliers, namely $m + 1 + r$, since these are generated based on the Putinar's Positivstellensatz.

The master problem (24) has an infinite number of constraints. For this reason relaxation is adopted as the solution strategy [9]. In other words, we begin by solving a relaxed version of (24), the so-called *relaxed master problem*, ignoring all but few constraints and if the resulting solution does not satisfy all of the ignored constraints we generate and add to the relaxed master problem one violated constraint, either to the set of feasibility constraints or to the set of optimality constraints. We continue this way until a termination criterion is satisfied which signals that the obtained solution is optimal within an acceptable accuracy. The equivalence of the master problem to the original POP implies that every time we solve a relaxed version of the master problem we get a lower bound on the optimal value of (12). Hence, solving a series of relaxed master problems yields a sequence of monotonically increasing lower bounds on the global optimal value p^*.

3.2 Algorithm

The assumptions made to develop our partitioning procedure are: (i) the feasible set $\mathscr{K}$ is nonempty and compact; (ii) the sets X, Y are compact. Finally, we consider that an initial point $\hat{y} \in Y \cap V$ is known. For convenience of the reader, we introduce the following definitions:

$$f^{\text{opt}}(\sigma_0, \ldots, \sigma_m, t_1, \ldots, t_r, y) := \inf_{x \in X} \left\{ p(x, y) - \sum_{i=0}^{m} \sigma_i(x) g_i(x, y) - \sum_{j=1}^{r} t_j(x) h_j(x, y) \right\},$$

$$f^{\text{feas}}(\sigma_0, \ldots, \sigma_{2^m-1}, t_1, \ldots, t_r, y) := \inf_{x \in X} \left\{ - \sum_{I \subseteq \{1, \ldots, m\}} \sigma_I(x) g_i(x, y) - \sum_{j=1}^{r} t_j(x) h_j(x, y) \right\},$$

where f^{opt} (f^{feas}) represents the optimality (feasibility) function, $y \in Y$, σ_i, $\sigma_I \in \Sigma^2$, $i = 0, \ldots, m$, $I = 0, \ldots, 2^m - 1$, and $t_j \in \mathbb{R}[x]$, $j = 1, \ldots, r$. Then, the partitioning procedure for polynomial optimization is stated below.

Algorithm 1 Partitioning Procedure for POPs

Step 1 Initialize y to $\hat{y}^1$, where $\hat{y}^1 \in Y \cap V$. Initialize the iteration counter, i.e. $k = 1$, and set the lower (LB) and upper (UB) bounds to minus infinity ($-\infty$) and plus infinity (∞), respectively. Set $n_{\text{opt}} = 1$ and $n_{\text{feas}} = 0$, where n_{opt} is the counter for the optimality constraints and n_{feas} is the counter for infeasibility constraints. Determine the convergence tolerance parameter $\varepsilon > 0$.

Step 2 Solve the subproblem (17) with $y = \hat{y}^1$ to compute the optimal solution vector $\bar{x}^1$ and the optimal polynomial multipliers $(\sigma_i^1(x), t_j^1(x))$, $i = 0, \ldots, m$, $j = 1, \ldots, r$.

Observe that the problem is feasible since we choose $\hat{y}^1 \in Y \cap V$. Then, generate the optimality constraint,[5]

$$z \geq \inf_{x \in X} \left\{ p(x, y) - \sum_{i=0}^{m} \sigma_i^1(x) g_i(x, y) - \sum_{j=1}^{r} t_j^1(x) h_j(x, y) \right\}. \tag{25}$$

Update the upper bound: $\mathrm{UB} = v(\hat{y}^1)$.

Step 3 Solve the relaxed master problem,

$$\begin{aligned} \min_{y,z} \ & z \\ \text{s.t.} \ & z \geq f^{\text{opt}}(\sigma_0^k, \ldots, \sigma_m^k, t_1^k, \ldots, t_r^k, y), \qquad k = 1, \ldots, n_{\text{opt}}, \\ & 0 \geq f^{\text{feas}}(\sigma_0^k, \ldots, \sigma_{2^m-1}^k, t_1^k, \ldots, t_r^k, y), \ k = 1, \ldots, n_{\text{feas}}. \end{aligned} \tag{26}$$

Let $(\hat{y}^{k+1}, \hat{z}^{k+1})$ be the optimal solution of problem (26). If $\hat{z}^{k+1} \geq \mathrm{UB} - \varepsilon$, stop. Else update the lower bound: $\mathrm{LB} = \hat{z}^{k+1}$ and increase the iteration counter $k = k+1$. Go to Step 4.

Step 4 Solve the subproblem (17) with $y = \hat{y}^k$.

Step 4.1 If feasible, compute the optimal solution vector $\bar{x}^k$, the optimal polynomial multipliers $(\sigma_i^k(x), t_j^k(x))$, $i = 0, \ldots, m$, $j = 1, \ldots, r$ and the objective value $v(\hat{y}^k)$. If $\mathrm{LB} \geq v(\hat{y}^k) - \varepsilon$, stop. Else generate the optimality constraint,

$$z \geq \inf_{x \in X} \left\{ p(x, y) - \sum_{i=0}^{m} \sigma_i^k(x) g_i(x, y) - \sum_{j=1}^{r} t_j^k(x) h_j(x, y) \right\}. \tag{27}$$

Increase the optimality counter $n_{\text{opt}} = n_{\text{opt}} + 1$. Update the upper bound $\mathrm{UB} = v(\hat{y}^k)$ only if necessary, i.e. if $v(\hat{y}^k)$ is less than the last stored upper bound value. Go to Step 3.

Step 4.2 If infeasible, compute polynomial multipliers $(\sigma_I^k(x), t_j^k(x))$, $I \subseteq \{1, \ldots, m\}$, $j = 1, \ldots, r$, that satisfy the following:

$$-1 = \sum_{I \subseteq \{1,\ldots,m\}} \sigma_I(x) g_I(x, \hat{y}^k) + \sum_{j=1}^{r} t_j(x) h_j(x, \hat{y}^k), \ \forall x \in X, \tag{28}$$

[5] In practice, the optimality constraint is generated as follows:

$$z \geq p(\bar{x}^1, y) - \sum_{i=0}^{m} \sigma_i^1(\bar{x}^1) g_i(\bar{x}^1, y) - \sum_{j=1}^{r} t_j^1(\bar{x}^1) h_j(\bar{x}^1, y),$$

where $\bar{x}^1$ is the optimal solution vector of subproblem (17) with $y = \hat{y}^1$. Similar remark applies to Step 4.1.

and generate the feasibility constraint,[6]

$$0 \geq \inf_{x \in X} \left\{ - \sum_{I \subseteq \{1,\ldots,m\}} \sigma_I^k(x) g_i(x, y) - \sum_{j=1}^{r} t_j^k(x) h_j(x, y) \right\}. \quad (29)$$

Increase the infeasibility counter $n_{\text{feas}} = n_{\text{feas}} + 1$. Go to Step 3.

Remark 1 At Step 4 we deal with a POP with a size smaller than the original one. In order to handle this problem and to compute the appropriate polynomial multipliers we employ the SDP relaxation technique described in Sect. 2. As this technique may not guarantee global optimality and successful extraction of minimizers at some specific relaxation of the SDP hierarchy, we may need to increase the relaxation order and try as many SDP relaxations as necessary to detect global optimality. The same applies to Step 3.[7]

3.3 Theoretical convergence

Theorem 7 *The partitioning procedure for polynomial optimization terminates either at Step 3 or at Step 4.1 without cycling. Termination implies that a global optimum* (x^*, y^*, z^*) *has been reached.*

Proof (Motivated by Theorem 3.4 in [32].) The solution (z^k, y^k) of the relaxed master problem (26) at iteration k is not going to be repeated at the next iteration:

1. As $z^k < \inf_{x \in X} \left\{ p(x, y^k) - \sum_{i=0}^{m} \sigma_i^k(x) g_i(x, y^k) - \sum_{j=1}^{r} t_j^k(x) h_j(x, y^k) \right\}$, the new optimality constraint cuts off (z^k, y^k).
2. As $0 < \inf_{x \in X} \left\{ - \sum_{I \subseteq \{1,\ldots,m\}} \sigma_I^k(x) g_I(x, y^k) - \sum_{j=1}^{r} t_j^k(x) h_j(x, y^k) \right\}$, the new feasibility constraint cuts off (z^k, y^k).

To prove optimality, we should note that at each iteration the relaxed master problem is a relaxation of (14), therefore, z^k is always a lower bound on the optimal value of (14). In view of Theorem 4, termination at Step 4.1, i.e. $z^* \geq v(y^*) - \varepsilon$, implies that y^* is ε-optimal in (14) since z^* is a lower bound on the optimal value of (14), i.e. $z^* \leq p^*$. Thus, any optimal solution x^* of (17) for $y = y^*$ yields an ε-optimal solution (x^*, y^*) of (12). Termination at Step 3 is similar, except that UB plays the role of $v(y^*)$. The UB is set at the least optimal value computed after one execution of Step 2 and successive executions of Step 4.1. In other words, the UB is the best known upper bound on the optimal value of (14), i.e. $UB \geq p^*$. When $z^* \geq UB - \varepsilon$, the subproblem (17) corresponding to the UB and the current relaxed master problem yield an ε-optimal solution (x^*, y^*) of (12). □

[6] In practice, the feasibility constraint is generated as follows:

$$0 \geq - \sum_{I \subseteq \{1,\ldots,m\}} \sigma_I^k(\bar{x}^{\text{UB}}) g_i(\bar{x}^{\text{UB}}, y) - \sum_{j=1}^{r} t_j^k(\bar{x}^{\text{UB}}) h_j(\bar{x}^{\text{UB}}, y),$$

where $\bar{x}^{\text{UB}}$ is the optimal solution vector corresponding to the best upper bound that has been computed up to iteration k.

[7] The resulting orders of relaxation, namely those at which global optimality is detected, are denoted by ω^{out} and ω^{in}, corresponding to Steps 3 and 4, respectively.

Theorem 8 (Asymptotic ε-Convergence) *Assume that Y is a nonempty compact subset of V, X is a nonempty compact set, and the set $M(y)$ of optimal polynomial multipliers for (17) is nonempty for all $y \in Y$ and is uniformly bounded in some neighborhood of each such point (let this be labelled as the "uniform boundedness assumption"). Then, for any given ε, the partitioning procedure for polynomial optimization converges.*

Proof (This is based on the finite ε-convergence proof in [10].) We fix ε arbitrarily and suppose that no termination is achieved. Let $\langle z^k, y^k \rangle$ be the sequence of optimal solutions to (26) at successive iterations. This sequence, or a subsequence, converges to a point (z^*, y^*), since $\langle z^k \rangle$ is a nondecreasing sequence bounded above and the sequence $\langle y^k \rangle$ belongs to the compact set Y. Next, as the number of iterations increases, the accumulation of constraints in (26) results in the following:

$$z^{k+1} \geq f^{\text{opt}}(\sigma_0^k, \ldots, \sigma_m^k, t_1^k, \ldots, t_r^k, y^{k+1}). \tag{30}$$

Introducing the set $M(y)$ of tuples holding the polynomial multipliers $\sigma_i, t_j, i = 0, \ldots, m$, $j = 1, \ldots, r$, as follows:

$$M(y) = \{(\sigma_i, t_j), i = 0, \ldots, m, j = 1, \ldots, r \mid f^{\text{opt}}(\sigma_0, \ldots, \sigma_m, t_1, \ldots, t_r, y) = v(y)\}, \tag{31}$$

as $\sigma_i(x) \in \Sigma^2, i = 0, \ldots, m$, and $t_j(x) \in \mathbb{R}[x], j = 1, \ldots, r$, correspond to the optimal polynomial multipliers of (22). By the uniform boundedness assumption of the set $M(y)$ we can conclude that the sequence $\langle(\sigma_i, t_j)\rangle$ produced by successive executions of Step 4.1 converges to a single tuple of polynomial multipliers (σ_i^*, t_j^*), $i = 0, \ldots, m$, $j = 1, \ldots, r$. Thus, by the continuity of the optimality function f^{opt}, as $k \to \infty$, Eq. 30 yields

$$z^* \geq f^{\text{opt}}(\sigma_0^*, \ldots, \sigma_m^*, t_1^*, \ldots, t_r^*, y^*). \tag{32}$$

To complete the proof, we first need to show that $(\sigma_0^*, \ldots, \sigma_m^*, t_1^*, \ldots, t_r^*) \in M(y^*)$ and then that v is upper semicontinuous at y^*. The set $M(y)$ is an upper-semicontinuous mapping at y^* by Theorem 1.5 in [17]. Because of this upper semicontinuity of $M(y)$, we can conclude that with increasing k, $y^k \to y^*$ and $(\sigma_i^k, t_j^k) \to (\sigma_i^*, t_j^*)$ with $(\sigma_i^k, t_j^k) \in M(y^k)$ for each k, imply that $(\sigma_i^*, t_j^*) \in M(y^*)$. Now that $(\sigma_0^*, \ldots, \sigma_m^*, t_1^*, \ldots, t_r^*) \in M(y^*)$ holds, representation (31) and Theorem 4.2 in [14] imply the following:

$$f^{\text{opt}}(\sigma_0^*, \ldots, \sigma_m^*, t_1^*, \ldots, t_r^*, y^*) = v(y^*). \tag{33}$$

As a result, from (32) and (33), we obtain: $z^* \geq v(y^*)$.

Then, using Lemma 1.3 from [17], it can be proved that v is (numerically) upper semicontinuous at y^*. Given the upper semicontinuity of v at y^*, we can conclude that $z^k \geq v(y^k) - \varepsilon$ for all k sufficiently large, which contradicts the assumption that the termination criterion is not met. Hence, termination of the partitioning procedure is achieved after a large number of iterations. Therefore, the proof is complete. □

As we employ the sum of squares formulation (22), only asymptotic convergence is guaranteed. On the other hand, in the original procedure, when convex programs are involved, finite convergence is ensured [3,10]. Finally, regarding the uniform boundedness assumption in the statement of Theorem 8, it does not appear to be limiting as we show below.

Bounded polynomial multipliers As the asymptotic convergence in Theorem 8 hinges on the uniform boundedness assumption of the set $M(y)$ of optimal polynomial multipliers of (17), we intend to introduce the well-known perturbation function of problem (17) and its tight lower bound expression to reach the conclusion that the polynomial multipliers at successive iterations are bounded. We start with the general case and consider the POP:

$$\begin{aligned} p^* &= \inf_x p(x) \\ \text{s.t. } G(x) &\geq b, \\ x &\in X, \end{aligned} \tag{34}$$

where $X \subseteq \mathbb{R}^n, b \in \mathbb{R}^m, G : \mathbb{R}^n \to \mathbb{R}^m$ and $f : \mathbb{R}^n \to \mathbb{R}$. Then, we introduce the perturbation function of (34) defined as follows:

Definition 1 The function $\phi : \mathbb{R}^m \to \mathbb{R}$ given by

$$\begin{aligned} \phi(d) &= \inf_x p(x) \\ \text{s.t. } G(x) &\geq d, \\ x &\in X, \end{aligned} \tag{35}$$

is the *perturbation* function of problem (34).

We then define function $F : \mathbb{R}^m \to \mathbb{R}$:

$$F(d) = \sigma_0(x) + \sigma(x)^{\mathrm{T}}(d - b) + \gamma, \tag{36}$$

where $\sigma(x) = [\sigma_1(x), \ldots, \sigma_m(x)]^{\mathrm{T}}$, $\sigma_i(x) \in \Sigma^2$, $i = 0, \ldots, m$ and $\gamma \in \mathbb{R}$. By applying Lemma 2.8 from [27] we get $F(d) \leq \phi(d)$ for all $d \in \mathbb{R}^m$. Next, by using the definition of *conjugate* function from [28], namely $\phi^*(F) = \sup_{d\in\mathbb{R}^m}\{F(d) - \phi(d)\}$; and applying statement (b) of Proposition 6.7 in [27], i.e. $\phi^*(F) = F(b) - \phi(b)$, we derive the following inequality:

$$F(d) - F(b) \leq \phi(d) - \phi(b), \quad \forall\, d \in \mathbb{R}^m, \tag{37}$$

which is a *subdifferential* inequality. Finally, by substituting $F(d)$ as defined in (36), we get $\phi(d) \geq \phi(b) + \sigma(x)^{\mathrm{T}}(d - b)$ for all $d \in \mathbb{R}^m$; or for $b = 0$ we get:

$$\phi(d) \geq \phi(0) + \sigma(x)^{\mathrm{T}} d, \quad \forall\, d \in \mathbb{R}^m, \tag{38}$$

which is the tight lower bound expression of (35). Finally, we return to our problem (17) and define its perturbation function:[8]

$$\begin{aligned} \psi(y, d) &= \inf_{x\in X} p(x, y) \\ \text{s.t. } g_i(x, y) &\geq d_i, \quad i = 1, \ldots, m. \end{aligned} \tag{39}$$

Using the derivations above in conjunction with Lemma 2.1 from [10] we obtain the following corollary, which shows that the polynomial multipliers at successive iterations of our partitioning procedure are bounded.

Corollary 1 *Assume that X is a nonempty compact set. If there exists a point $\bar{x} \in X$ such that $g_i(\bar{x}, y^*) > 0$, $i = 1, \ldots, m$, then the set $M(y)$ of optimal polynomial multipliers for (17) is uniformly bounded in some open neighborhood of y^*.*

8 Without loss of generality, we eliminated equality constraints for simplicity.

Proof The set $\Omega(y,d) = \{x \in X \mid g_i(x,y) \geq d_i,\ i = 1,\dots,m\}$ is nonempty for all (y,d) in some open neighborhood U of $(y^*,0)$. This claim arises from our assumption that $g_i(\bar{x},y^*) > 0$, $i = 1,\dots,m$, for a point $\bar{x} \in X$. By applying Theorem 1.4 of [17], or Theorem 7 of [12], we conclude that $\psi(y,d)$ in (39) is continuous on U. Next, let U_1 be an open neighborhood of y^* and $U_2 = \{d \in \mathbb{R}^m \mid 0 \leq d_i \leq \delta,\ i = 1,\dots,m\}$ such that $\delta > 0$ and $\bar{U}_1 \times U_2 \subset U$, where $\bar{U}_1$ is the closure of U_1, and define the following lower and upper bounds of $\psi(y,d)$:

$$\psi_1^* = \min_{y,d} \psi(y,d) \text{ s.t. } y \in \bar{U}_1, \quad d \in U_2, \tag{40}$$

$$\psi_2^* = \max_{y,d} \psi(y,d) \text{ s.t. } y \in \bar{U}_1, \quad d \in U_2. \tag{41}$$

Since, $\psi(y,d)$ is continuous on the closed set $\bar{U}_1 \times U_2$ we have that $-\infty \leq \psi_1^* \leq \psi_2^* \leq \infty$. Now let us take any point $y^k \in U_1$ and $\sigma^k(x) = [\sigma_1^k(x),\dots,\sigma_m^k(x)]^{\mathrm{T}}$ be the respective optimal polynomial multipliers of problem (17) with $y = y^k$. By inequality (38) we get:

$$\psi(y,d) \geq \psi(y,0) + \sigma^k(x)^{\mathrm{T}} d, \quad \forall\, d \in \mathbb{R}^m.$$

If we take $d = \delta e_i$, $i = 1,\dots,m$, where e_i denotes the ith unit vector in $\mathbb{R}^m$, we conclude that:

$$0 \leq \sigma_i^k(x) \leq \frac{\psi(y,\delta e_i) - \psi(y,0)}{\delta} \leq \frac{\psi_2^* - \psi_1^*}{\delta},$$

where $\frac{\psi_2^* - \psi_1^*}{\delta}$ is clearly a constant which proves that the polynomial multipliers $\sigma_1^k,\dots,\sigma_m^k \in M(y^k)$ are indeed uniformly bounded for any $y^k \in U_1$. □

4 Computational experience

Our procedure was implemented in `Matlab`. We integrated the solver `GloptiPoly` [11] into our procedure so as to handle the polynomial optimization subprograms that occur at Step 3 and Step 4. In particular, each time `GloptiPoly` is called from our procedure it either builds the ω^{out}th (if called at Step 3) or the ω^{in}th (if called at Step 4) semidefinite relaxation of the input polynomial subproblem. `GloptiPoly` then solves the corresponding relaxation with the help of the semidefinite programming solver `SeDuMi` [26]. The parameters ω^{out} and ω^{in} are fixed externally, i.e. by the user, to the same value. However, if the global optimality criterion is not met either at an execution of Step 3, or at an execution of Step 4 of the algorithm, the corresponding relaxation order, ω^{out} or ω^{in}, respectively, is increased automatically until a global optimal solution is computed. What is more, at Step 4 the solution of the primal problem solved by `SeDuMi` yields the coefficients of the polynomial multipliers needed for the generation of the optimality (feasibility) constraints. Let us consider an example from [7].

Example 1 (Test problem name: Bex2_1_2)

$$\begin{array}{ll} \min & -0.5(x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2) - \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ \text{s.t.} & 6x_1 + 3x_2 + 3x_3 + 2x_4 + x_5 \leq 6.5, \\ & 10x_1 + 10x_3 + x_6 \leq 20, \\ & 0 \leq x_i \leq 1, \quad i = 1,\dots,5, \\ & 0 \leq x_6 \leq 20, \end{array} \tag{42}$$

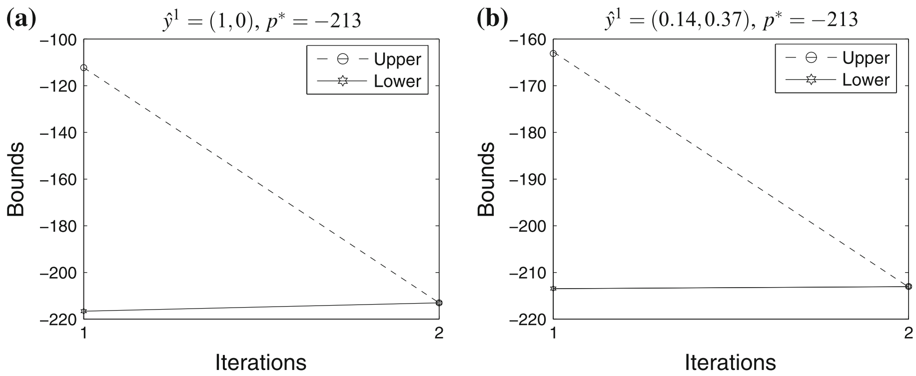

Fig. 1 Convergent bounds for Example 1 (Bex2_1_2)

where $\mathbf{c} = [10.5, 7.5, 3.5, 2.5, 1.5, 10]^{\mathrm{T}}$. The known global optimal solution of the above problem is attained at point $x_1 = x_3 = 0, x_2 = x_4 = x_5 = 1,\ x_6 = 20$, and the global optimal objective value is $p^* = -213$. Initially, we reformulated the problem in order to incorporate the bounds on the variables into the set of inequality constraints. Next, we scaled the problem so that all variables belonged to the closed interval [0, 1]. Our method partitioned the set of variables into the sets $\mathbf{x} = (x_2, x_4, x_5, x_6)$ and $\mathbf{y} = (x_1, x_3)$. Finally, the accuracy was specified to $\varepsilon = 10^{-6}$ and the starting point fixed to $\hat{\mathbf{y}}^{(1)} = (x_1^{(1)}, x_3^{(1)}) = (1, 0)$. We computed the optimal solution of problem (42) with $\mathbf{y} = \hat{\mathbf{y}}^{(1)}$ and it was $\bar{\mathbf{x}}^{(1)} = (x_2^{(1)}, x_4^{(1)}, x_5^{(1)}, x_6^{(1)}) = (0.17, 0, 0, 10)$, as well as its optimal objective value: $v(\hat{\mathbf{y}}^{(1)}) = -112.26$. Observe that $v(\hat{\mathbf{y}}^{(1)})$ is an upper bound to p^*. Based on the computed information, the relaxed master problem after we added the first optimality constraint, was:

$$\begin{array}{ll} \min & z \\ \text{s.t.} & z - 0.5242x_1 - 0.5208x_3 + 0.0025x_1^2 + 0.0025x_3^2 + 1.0830 \geq 0, \\ & 0 \leq x \leq 1, \\ & 0 \leq x_3 \leq 1. \end{array} \tag{43}$$

The solution[9] of the relaxed master problem (43) was $(\hat{z}^{(2)}, \hat{\mathbf{y}}^{(2)}) = (-216.6, 0, 0)$. As can be seen, $\hat{z}^{(2)}$ is a lower bound on the optimal objective value p^*. Since the termination criterion had not yet been met, the iteration counter was increased and the problem (42) was solved with $\mathbf{y} = \hat{\mathbf{y}}^{(2)}$. The second subproblem yielded an optimal objective value $v(\hat{\mathbf{y}}^{(2)}) = -213$ and optimal solution vector $\bar{\mathbf{x}}^{(2)} = (x_2^{(2)}, x_4^{(2)}, x_5^{(2)}, x_6^{(2)}) = (1, 1, 1, 20)$. The subsequent relaxed master problem was augmented with a new optimality constraint:

$$\begin{array}{ll} \min & z \\ \text{s.t.} & z - 0.5242x_1 - 0.5208x_3 + 0.0025x_1^2 + 0.0025x_3^2 + 1.0830 \geq 0, \\ & z - 0.22x_1 - 0.26x_3 + 0.0025x_1^2 + 0.0025x_3^2 + 1.065 \geq 0, \\ & 0 \leq x \leq 1, \\ & 0 \leq x_3 \leq 1, \end{array} \tag{44}$$

which gave a lower bound $\hat{z}^{(3)}$ equal to -213 and optimal solution vector $\hat{\mathbf{y}}^{(3)} = (x_1^{(3)}, x_3^{(3)}) = (0, 0)$. At this point of progress our termination criterion was met and the procedure stopped

9 Notice that although the constraints in problems (43) and (44) appear scaled, due to the internal scaling performed, the reported optimal values are unscaled.

Table 1 Partitioning procedure for POPs and `GloptiPoly` on test problems

Problem	n	m	p^*	ω	`GloptiPoly`	Partitioning procedure for POPs					
					$p^*_{\omega,bmrk}$	p^*_ω	$(\omega^{\text{in}}, \omega^{\text{out}})$	**iters**	**cputime/iter**	ε_{p^*}	ε_{x^*}
Wol3_5	4	3	−7	3	−7	$\{-9, \ldots, -7\}$	(2,1)	2	0.67	$3.32e-09$	$5.37e-09$
Wol4_2	4	6	−13	2	−13	$\{-13, \ldots, -13\}$	(1,1)	$1\frac{1}{2}$	1.12	$8.37e-10$	$2.65e-09$
Bex2_1_1	5	1	−17	3	−17	$\{-454.955, \ldots, -17\}$	(3,4)	3	4.71	$3.76e-09$	$7.49e-09$
Bex2_1_2	6	2	−213	2	−213	$\{-217, \ldots, -213\}$	(2,2)	2	2.89	$7.29e-10$	$3.46e-08$
Bex2_1_4	6	5	−11	2	−11	$\{-11, \ldots, -11\}$	(2,1)	1	3.68	$7.13e-08$	$8.02e-10$
Bex2_1_5	10	11	−268.0146	2	−268.015	$\{-284.236, \ldots, -268.015\}$	(2,2)	3	10.70	$3.26e-07$	$5.86e-06$
Bex3_1_2	5	6	−30665.54	2	−30665.5	$\{-39492.7, \ldots, -30665.5\}$	(2,1)	4	2.54	$8.90e-10$	$3.25e-09$
Bex9_1_1	13	12	−13	2	−13	$\{-45868.1, \ldots, -13\}$	(1,1)	3	2.83	$7.62e-10$	$5.93e-03$
Bex9_1_2	10	9	−16	3	–	$\{-16.5556, \ldots, -16\}$	(1,1)	2	0.88	–	–
Bex9_1_4	10	9	−37	3	–	$\{-106.968, \ldots, -37\}$	(1,2)	2	1.38	–	–
Bex9_2_4	8	7	0.5	2	0.499889	$\{0.24,\ldots,0.49\}$	(2,1)	6	21.04	$9.31e-03$	$4.02e-02$
Bex9_2_5	8	7	5.0	2	5.00004	$\{-46.1465, \ldots, 5\}$	(1,2)	6	2.31	$7.62e-06$	$3.55e-04$
Bex9_2_8	6	5	1.5	2	1.5	$\{1.185,\ldots,1.5\}$	(2,1)	3	2.23	$4.95e-09$	$3.72e-09$
meanvar	8	2	5.2434	2	5.2434	$\{5.2434,\ldots,5.2434\}$	(1,1)	1	0.00	$1.50e-09$	$1.05e-05$
Bst_bpaf1a	10	10	–	2	−45.3797	$\{-45.3797, \ldots, -45.3797\}$	(1,1)	$1\frac{1}{2}$	2.36	$3.47e-09$	$7.38e-08$
Bst_bpaf1b	10	10	–	2	−42.9626	$\{-42.9626, \ldots, -42.9626\}$	(1,1)	$1\frac{1}{2}$	2.38	$8.77e-09$	$6.60e-08$
Bst_e05	5	3	–	3	7049.25	$\{2739.29,\ldots,7049.23\}$	(1,1)	$24\frac{1}{2}$	1.89	$3.34e-06$	$9.25e-04$
Bst_e07	10	7	–	2	−1809.18	$\{-2283.98, \ldots, -1809.2\}$	(1,3)	2	1.81	$9.34e-06$	$8.77e-03$
Bst_jcbpaf2	10	13	–	2	−794.856	$\{-1259.28, \ldots, -794.944\}$	(1,1)	20	3.35	$1.11e-04$	$1.86e-02$
st_e21	6	6	–	2	−14.1	$\{-16.7, \ldots, -14.1\}$	(1,1)	$2\frac{1}{2}$	0.37	$3.80e-10$	$5.28e-10$
st_glmp_kk90	5	7	–	2	3	$\{-146.315, \ldots, 3\}$	(2,1)	2	1.66	$7.23e-11$	$1.95e-09$

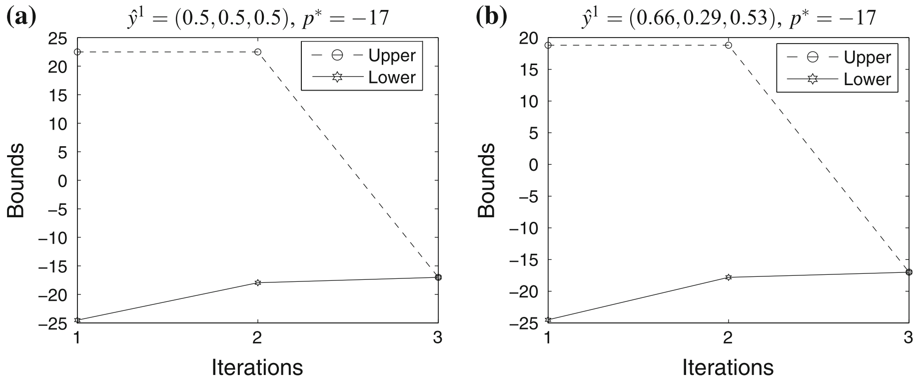

Fig. 2 Convergent bounds for test problem Bex2_1_1

with global optimal solution vector (0, 1, 0, 1, 1, 20) and global optimal objective value equal to $p^* = -213$. The progression of upper and lower bounds computed by our method for two different starting points, i.e. $\hat{y}^1 = (1, 0)$ and $\hat{y}^1 = (0.14, 0.37)$, is depicted in Fig. 1. Another graphical example of convergent bounds for the test problem named Bex2_1_1 is given in Fig. 2. To test our code, we applied it to a collection of test problems obtained from [1] and two bilinear problems from [32]. The results are summarized in Table 1 and are compared with the results computed from direct application of GloptiPoly (Version 2.3.0). The metrics we used to compare our results with the results from GloptiPoly were:

$$\varepsilon_{p^*} = \frac{|\ p^*_{\omega,\text{bmrk}} - p^*_\omega\ |}{\max\{1, |\ p^*_{\omega,\text{bmrk}}\ |\}}, \quad \varepsilon_{x^*} = \max\left\{\frac{|\ x^*_{i,\text{bmrk}} - x^*_i\ |}{\max\{1, |\ x^*_{i,\text{bmrk}}\ |\}}\right\}, \tag{45}$$

where ‘bmrk’ stands for benchmark and x^*_i ($x^*_{i,\text{bmrk}}$) corresponds to the ith element of the solution vector x^* (x^*_{bmrk}). In Table 1 the first column holds the name of the test problem. The second, third and fourth columns hold the number of the variables, the number of constraints, the best global optimal objective value known so far, and the relaxation order at which GloptiPoly computes the global optimal solution. The next column holds the actual objective value computed by GloptiPoly. The following six columns include the results computed by our method. In particular, p^*_ω corresponds to the optimal objective value but in this column the sequence of computed lower bounds is also reported. The upper bound is close to the lower bound within accuracy $\varepsilon = 10^{-6}$, hence its value is implied. The subsequent column reports the maximun relaxation order needed in the subproblems and the relaxed master problems, i.e. ω^{in} and ω^{out}, respectively, until global optimality is reached. The number of iterations, the average cpu time per iteration, and the values of the metrics (45) are stated in the last four columns. Observe that when a rational iteration number is reported, it means that our procedure terminated at Step 4.1, while an integer iteration number means that our procedure terminated at Step 3. Note also that the times reported aim only at giving an indicative idea of the average time spent per iteration by our procedure. Due to the testing status of our code, several redundant file input/output operations increase the actual time spent.

In all cases our procedure and GloptiPoly gave equally satisfactory results. In two cases, i.e. Problems Bex9_1_2 and Bex9_1_4, our procedure outperformed GloptiPoly. However, the POPs we handled using our procedure are still small to medium size. Further improvements are needed in order to tackle larger problems. For instance, we can customize the partitioning of variables according to the sparse structure of the problem, if sparsity is present [13,15,31]. The sparsity pattern of the POP would help us partition the set of variables, not randomly, but based on its sparsity structure, in more than one subset. This controlled partitioning of polynomial variables would produce several even smaller POP subproblems at Step 4, which makes these subproblems easier to solve individually.

5 Conclusions and future plans

In this paper it is intended to show that the generalized Benders decomposition can be extended to POPs using Stengle's Positivstellensatz in place of the Farkas lemma and Putinar's Positivstellensatz in place of the convex duality theorem. The lack of convexity assumptions in our problems and the polynomial functions involved necessitates the use of the sum-of-squares representation. This representation yields polynomial functions instead of constant multipliers for the generation of feasibility and optimality constraints. The theoretical results in this work are in line with those presented in [32], where nonconvex problems are also tackled and the use of duality theory for general programs produces functions in place of constant multipliers. However, the sum-of-squares representation only ensures asymptotic convergence of our procedure. Finite convergence remains to be investigated in the future. From the computational perspective, the numerical results presented appear to be satisfactory. To improve the efficiency of our procedure for sparse POPs, we intend to replace the optimality and feasibility constraints introduced here by their sparse analog based on the results introduced in [15,31]. Such an amendment is expected to simplify the optimality/feasibility constraints formulation. Hence, this would result in even smaller and easier to tackle master problems. In addition, sparsity of the POP should enable us to derive smaller-size subproblems, which could be solved more efficiently, instead of partitioning the set of variables randomly.

Acknowledgements The authors would like to thank two anonymous referees. Their thoughtful comments and stimulating suggestions helped improve the paper. The authors are also thankful to Fook Kong for his remarks that helped enhance the text.

References

1. GLOBAL Library. http://www.gamsworld.org/global/globallib/globalstat.htm (2008)
2. Ben-Tal, A., Nemirovski, A.: Lectures on modern convex optimization. MPS/SIAM Series on Optimization, SIAM, Philadelphia, PA (2001)
3. Benders, J.F.: Partitioning procedures for solving mixed-variables programming problems. Comput. Manag. Sci. **2**(1):3–19 (2005), reprinted from Numer. Math. **4**(1962), pp. 238–252
4. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
5. Cox, D., Little, J., O'Shea, D.: Ideals, Varieties, and Algorithms. Undergraduate Texts in Mathematics. Springer, New York (1992)
6. Floudas, C.A.: Deterministic global optimization, Nonconvex Optimization and its Applications, vol 37. Kluwer, Dordrecht, Theory, methods and applications (2000)
7. Floudas, C.A., Pardalos, P.M.: A Collection of Test Problems for Constrained Global Optimization Algorithms, Lecture Notes in Computer Science, vol 455. Springer, Berlin (1990)
8. Floudas, C.A., Visweswaran, V.: A global optimization algorithm (GOP) for certain classes of nonconvex NLPs: I. Theory Comput. Chem. Eng. **14**(12), 1397–1417 (1990)
9. Geoffrion, A.M.: Elements of large-scale mathematical programming. I. Concepts. Manag. Sci. **16**, 652–675 (1970)
10. Geoffrion, A.M.: Generalized benders decomposition. J. Optim. Theory Appl. **10**, 237–260 (1972)
11. Henrion, D., Lasserre, J.B.: GloptiPoly: global optimization over polynomials with Matlab and SeDuMi. ACM Trans. Math. Softw. **29**(2), 165–194 (2003)

12. Hogan, W.W.: Point-to-set maps in mathematical programming. SIAM Rev. **15**, 591–603 (1973)
13. Kleniati, P.M., Parpas, P., Rustem, B.: Decomposition-based method for sparse semidefinite relaxations of polynomial optimization problems. J. Optim. Theory Appl. doi:10.1007/s10957-009-9624-2 (2009)
14. Lasserre, J.B.: Global optimization with polynomials and the problem of moments. SIAM J. Optim. **11**(3), 796–817 (2000/2001)
15. Lasserre, J.B.: Convergent SDP-relaxations in polynomial optimization with sparsity. SIAM J. Optim. **17**(3), 822–843 (2006)
16. Laurent, M.: Sums of squares, moment matrices and optimization over polynomials. In: Putinar, M., Sullivant, S. (eds.) Emerging applications of algebraic geometry, IMA Vol. in Math. and its Appl., vol 149, pp. 157–270. Springer, New York (2009)
17. Meyer, R.: The validity of a family of optimization methods. SIAM J. Control **8**, 41–54 (1970)
18. Nie, J., Schweighofer, M.: On the complexity of Putinar's Positivstellensatz. J. Complex. **23**(1), 135–150 (2007)
19. Parrilo, P.A.: Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization. Ph.D. thesis, California Institute of Technology (2000)
20. Parrilo, P.A.: Semidefinite programming relaxations for semialgebraic problems. Math. Program. **96**(2, Ser. B), 293–320 (2003)
21. Prestel, A., Delzell, C.N.: Positive Polynomials. Springer Monographs in Mathematics. Springer, Berlin (2001)
22. Putinar, M.: Positive polynomials on compact semi-algebraic sets. Indiana Univ. Math. J. **42**(3), 969–984 (1993)
23. Schmüdgen, K.: The K-moment problem for compact semi-algebraic sets. Math. Ann. **289**(2), 203–206 (1991)
24. Schweighofer, M.: Optimization of polynomials on compact semialgebraic sets. SIAM J. Optim. **15**(3), 805–825 (2005)
25. Stengle, G.: A nullstellensatz and a positivstellensatz in semialgebraic geometry. Math. Ann. **207**, 87–97 (1974)
26. Sturm, J.F.: Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. Optim. Methods Softw. **11/12**(1–4), 625–653 (1999)
27. Tind, J., Wolsey, L.A.: An elementary survey of general duality theory in mathematical programming. Math. Program. **21**(3), 241–261 (1981)
28. Tuy, H.: Convex Analysis and Global Optimization, Nonconvex Optimization and its Applications, vol 22. Kluwer, Dordrecht (1998)
29. Visweswaran, V., Floudas, C.A.: A global optimization algorithm (GOP) for certain classes of nonconvex NLPs: II. Application of theory and test problems. Comput. Chem. Eng. **14**(12), 1419–1434 (1990)
30. Visweswaran, V., Floudas, C.A.: Unconstrained and constrained global optimization of polynomial functions in one variable. J. Global Optim. **2**(1), 73–99 (1992)
31. Waki, H., Kim, S., Kojima, M., Muramatsu, M.: Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity. SIAM J. Optim. **17**(1), 218–242 (2006)
32. Wolsey, L.A.: A resource decomposition algorithm for general mathematical programs. Math. Program. Stud. **14**, 244–257 (1981)