Journal of Optimization Theory and Applications (2025) 205:22
https://doi.org/10.1007/s10957-025-02631-6

Granularity for Mixed-Integer Polynomial Optimization Problems

Carl Eggen[1] · **Oliver Stein**[2] · **Stefan Volkwein**[1]

Received: 1 August 2024 / Accepted: 29 January 2025
© The Author(s) 2025

Abstract
Finding good feasible points is crucial in mixed-integer programming. For this purpose we combine a sufficient condition for consistency, called granularity, with the moment-/sum-of-squares-hierarchy from polynomial optimization. If the mixed-integer problem is granular, we obtain feasible points by solving continuous polynomial problems and rounding their optimal points. The moment-/sum-of-squares-hierarchy is hereby used to solve those continuous polynomial problems, which generalizes known methods from the literature. Numerical examples from the MINLPLib illustrate our approach.

Keywords Mixed-integer nonlinear programming · Granularity · Rounding · Polynomial optimization · Semidefinite programming

Mathematics Subject Classification 90C10 · 90C11 · 90C22 · 90C23 · 90C31

1 Introduction

Mixed-integer nonlinear minimization problems are a fundamental class of optimization problems that involve both continuous and discrete variables as well as nonlinear constraints. Due to their inherent complexity, solving mixed integer problems to global

Communicated by Arvind Raghunathan.

✉ Carl Eggen
carl.eggen@uni-konstanz.de

Oliver Stein
stein@kit.edu

Stefan Volkwein
stefan.volkwein@uni-konstanz.de

1 Department of Mathematics and Statistics, University of Konstanz, Universitätsstraße 10, 78464 Konstanz, Germany

2 Institute for Operations Research, Karlsruhe Institute of Technology (KIT), Blücherstraße 17, 76185 Karlsruhe, Germany

Published online: 06 March 2025

optimality can be computationally very challenging [7, 8]. Usually, these problems are solved by a branch-and-bound framework, whose performance particularly relies on good upper bounds for the objective function [2]. These upper bounds are usually given by the evaluation of the objective function at feasible points.

In this article, we improve a method to find feasible points for mixed-integer problems with polynomial objective function and polynomial constraints, so called mixed-integer polynomial optimization problems (MIPOP). For this purpose a condition of consistency called *granularity* is used and generalized by using the moment-/sos-hierarchy from polynomial optimization.

The notion of granularity for mixed-integer optimization was introduced in [18] and expanded in [19] and [17]. Such an optimization problem and its feasible set are considered granular when a certain nonempty inner parallel set exists within the continuously relaxed feasible set. In this case a sufficient condition for the existence of feasible points is obtained, the difficulties imposed by the integrality constraints can be relaxed and a feasible point may be extracted.

For approximating the inner parallel set a global optimization method for polynomial optimization problems (POP) can be used, the so called moment-/sos-hierarchy introduced by Lasserre in [9]. It relies on nonnegativity certificates for polynomials and the theory of moments. For a self-contained introduction and for an overview about the topic, we refer to [10, 11, 13, 20]. Theoretically, if some mild conditions hold, the moment-/sos-hierarchy can be used to approximate a POP arbitrarily well. In practice instead, solving this hierarchy of approximations may be rather computationally expensive, since each step in the hierarchy is formulated as a semidefinite program (SDP) and the sizes of the involved semidefinite matrices grow rapidly in size for higher steps. In the case that some sparsity structure appears in POP, the issue of growing complexity can be partially tackled and even instances with several thousand variables and constraints can be solved [15].

Our contribution is to combine these two concepts. We calculate an inner approximation of the inner parallel set for MIPOPs with a modification of the original moment-/sos-hierarchy introduced in [12]. This inner approximation is described by polynomials. If this inner approximation is nonempty, this directly implies the nonemptiness of the inner parallel set and hence the granularity of the original MIPOP. Now, a feasible point may be extracted by using a local NLP-solver. This approach generalizes the approximation techniques for calculating the inner parallel set from [19].

The outline of our article is as follows: Firstly, the definitions and notations for MIPOPs and for the moment-/sos-hierarchy are introduced in Sect. 2. Then the inner parallel set and the notion of granularity are defined (cf. Section 3). Further, an exact description of the inner parallel set in the convex case is given (cf. Theorem 3.1). In Sect. 4 we explain, how polynomials, which describe an inner approximation of the inner parallel set, can be calculated by a modification of the moment-/sos-hierarchy. In the linear case (cf. Section 5) this approximation is exact and coincides with the description in the literature [18]. For the nonlinear case we show that our approach generalizes the method for approximating the inner parallel set in [19] and an example is given for illustration (cf. Example 6.1). With all this, we can set up our algorithm (cf. Section 8) and propose some enhancements, especially for the binary case (cf.

Sect. 7). Finally, the performance of our algorithm is compared to the solver SCIP [3] on examples from the MINLPLib [16] (cf. Section 9).

2 Definitions and Notation

We consider polynomial mixed-integer nonlinear problems of the form

$$\begin{aligned} \min_{(x,y)\in\mathbb{R}^n\times\mathbb{Z}^m} \ & f(x,y) \\ \text{s.t. } & g_j(x,y)\le 0,\ \ j=1,\dots,k \\ & x^l \le x \le x^u, \\ & y^l \le y \le y^u, \end{aligned} \tag{MIPOP}$$

where $f, g_1, \dots, g_k \in \mathbb{R}[x, y]$ are real polynomials in n continuous variables $x = (x_1, \dots, x_n)$ and m integer variables $y = (y_1, \dots, y_m)$. All variables shall be bounded by known lower bounds $x^l \in \mathbb{R}^n$, $y^l \in \mathbb{Z}^m$ and upper bounds $x^u \in \mathbb{R}^n$, $y^u \in \mathbb{Z}^m$. Then let the feasible set of (MIPOP) be denoted by M, the NLP relaxation of (MIPOP) by the minimization of f over

$$\widehat{M} := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : \forall j : g_j(x, y) \le 0, x^l \le x \le x^u, y^l \le y \le y^u\},$$

and the box of the variable domains by

$$B := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : x^l \le x \le x^u, y^l \le y \le y^u\}.$$

To formulate the moment-/sos-hierarchy, we first introduce some algebraic definitions. For simplicity, we present them for the ring of real polynomials $\mathbb{R}[x]$ in the variables $x = (x_1, \dots, x_n)$, though they will later be used for polynomial rings with additional variables.

Let $r \in \mathbb{N}_0$. Then $\mathbb{R}[x]_r$ defines the vector space of real polynomials up to degree r in $x = (x_1, \dots, x_n)$ variables, $\Sigma[x]$ the convex cone of sums of squares of real polynomials and $\Sigma_r[x]$ the subcone of sums of squares (sos) of real polynomials up to degree $2r$. These are defined as follows:

$$\Sigma[x] = \left\{ \sum_{i=1}^{k} u_i^2 : k \in \mathbb{N}, u_i \in \mathbb{R}[x] \right\},$$

$$\Sigma_r[x] = \left\{ \sum_{i=1}^{k} u_i^2 : k \in \mathbb{N}, u_i \in \mathbb{R}[x]_r \right\}.$$

The quadratic module generated by polynomials $p_1, \dots, p_k \in \mathbb{R}[x]$ is defined by

$$\mathcal{M}(p_1, \dots, p_k) := \left\{ u_0 + \sum_{j=1}^{k} u_j p_j : u_0, u_j \in \Sigma[x] \right\} \subseteq \mathbb{R}[x]$$

and the corresponding r-truncated quadratic module by

$$\mathcal{M}_r(p_1, \dots, p_k) := \left\{ u_0 + \sum_{j=1}^{k} u_j p_j : u_0, u_j \in \Sigma_{r-\nu_j}[x] \right\} \subseteq \mathbb{R}[x]_{2r}$$

for $r \in \mathbb{N}$, $\nu_0 = 0$ and $\nu_j = \lceil \deg(p_j)/2 \rceil$. Further, let $\zeta := (\zeta_\alpha) \subseteq \mathbb{R}$ be a sequence indexed by the monomial basis (x^α) and $L_\zeta : \mathbb{R}[x] \to \mathbb{R}$ the corresponding Riesz functional defined by

$$f = \sum_\alpha f_\alpha x^\alpha \mapsto L_\zeta(f) = \sum_\alpha f_\alpha \zeta_\alpha. \tag{1}$$

Then the r-truncated moment matrix $\mathcal{M}_r(\zeta)$ is defined by

$$\mathcal{M}_r(\zeta)_{\alpha,\beta} := L_\zeta(x^{\alpha+\beta}) = \zeta_{\alpha+\beta}, \quad \alpha, \beta \in \mathbb{N}_r^n$$

with columns and rows indexed by the monomial basis $(x^\alpha)_{|\alpha| \le r}$ and $\mathbb{N}_r^n := \{\alpha \in \mathbb{N}^n : |\alpha| \le r\}$ ($|\alpha| = \sum_i \alpha_i$). For $g = \sum_\gamma g_\gamma x^\gamma \in \mathbb{R}[x]$ we also define the localizing matrix

$$\mathcal{M}_r(g\zeta)_{\alpha,\beta} = L_\zeta(g(x) x^{\alpha+\beta}) = \sum_\gamma g_\gamma \zeta_{\alpha+\beta+\gamma}, \quad \alpha, \beta \in \mathbb{N}_r^n.$$

3 The Inner Parallel Set and Granularity

The inner parallel set and the notion of granularity were introduced by Neumann et al. for feasible rounding in mixed-integer optimization [18] and in the following we stick to the notation therein and in [19].

The inner parallel set of (MIPOP) is defined as

$$\widehat{M}^- := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : (x, y) + K \subseteq \widehat{M}\}$$

with $K := \{0\} \times B_\infty(\frac{1}{2})$ and

$$B_\infty(\tfrac{1}{2}) := B_\infty(0, \tfrac{1}{2}) = \left\{ z \in \mathbb{R}^m : \forall i : |z_i| \le \tfrac{1}{2} \right\} \subseteq \mathbb{R}^m.$$

One can easily see that any rounding of a point (x, y) in $\widehat{M}^-$ lies in the original feasible set M. More precisely, $(\check{x}, \check{y})$ is a rounding for a point $(x, y) \in \widehat{M}^-$ if and only if

$$\check{x} = x \quad \text{and} \quad \check{y} \in \mathbb{Z}^m, |\check{y}_i - y_i| \le \tfrac{1}{2},\ i = 1, \ldots, m$$

hold. Note that roundings might not be unique. Now, it is easy to see that

$$(\check{x}, \check{y}) \in ((x, y) + K) \cap (\mathbb{R}^n \times \mathbb{Z}^m) \subseteq \widehat{M} \cap (\mathbb{R}^n \times \mathbb{Z}^m) = M$$

holds. This yields the following results.

Lemma 3.1 *[18, Lemma 2.1] For any point $(x, y) \in \widehat{M}^-$, any of its roundings $(\check{x}, \check{y})$ lies in M.*

Proposition 3.1 *[19, Proposition 2.1] If the inner parallel set $\widehat{M}^-$ is nonempty, then also M is nonempty.*

Proposition 3.1 motivates the definition of granularity.

Definition 3.1 [19, Definition 2.1] We call the set M *granular*, if the inner parallel set $\widehat{M}^-$ of M is nonempty. Moreover, we call a problem MIPOP *granular*, if its feasible set M is granular.

Further, for each constraint g_j analogous definitions can be established

$$\begin{aligned} G_j &:= \{(x, y) \in B : g_j(x, y) \le 0\}, \\ G_j^- &:= \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : (x, y) + K \subseteq G_j\} \end{aligned}$$

for all $1 \le j \le k$. Note that the definition of G includes the box B and thus slightly differs from the one in [19]. Then

$$\widehat{M}^- = \Big(\bigcap_{j=1}^{k} G_j\Big)^- = \bigcap_{j=1}^{k} G_j^- \tag{2}$$

holds. The first equality is only the definition of $\widehat{M}^-$, whereas the second one easily follows from the definition of the inner parallel set.

Thus, for calculating the inner parallel set $\widehat{M}^-$ it is sufficient to concentrate on the single sets G_j^-. Using again the fact that taking intersections and inner parallel sets can be interchanged, we obtain

$$\begin{aligned} G_j^- =& (B \cap \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : g_j(x, y) \le 0\})^- \\ =& B^- \cap \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : g_j(x, y) \le 0\}^- \end{aligned}$$

with the inner parallel set

$$B^- = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : \forall z \in B_\infty(\tfrac{1}{2}) : x^l \le x \le x^u, y^l \le y + z \le y^u\}$$

$$=\{(x,y)\in\mathbb{R}^n\times\mathbb{R}^m : x^l\le x\le x^u,\ y^l+\tfrac{1}{2}e\le y\le y^u-\tfrac{1}{2}e\}$$

of B, where $e\in\mathbb{R}^m$ denotes the vector of ones. We therefore obtain the description

$$G_j^- =\{(x,y)\in B^- : \forall z\in B_\infty(\tfrac{1}{2}) : g_j(x,y+z)\le 0\}$$

of G_j^- by a semi-infinite constraint.

For nonlinear polynomials g these semi-infinite constraints are usually difficult to handle. Hence, one would like to generate tight inner approximations of the sets G_j^- with easy descriptions. This can be accomplished by finding suitable polynomials $h_j\in\mathbb{R}[x,y]$ such that

$$T_j^- = \{(x,y)\in B^- : h_j(x,y)\le 0\}\subseteq G_j^-. \tag{3}$$

By (2) and (3) it holds

$$T^- := \bigcap_{j=1}^{k} T_j^- \subseteq \widehat{M}^-.$$

If the intersection T^- is nonempty, then a feasible point for (MIPOP) can simply be found by rounding any point in T^-, as seen above. Therefore, solving the continuous problem

$$\begin{aligned}\min_{(x,y)\in B^-}\ & f(x,y)\\ \text{s.t. } & h_j(x,y)\le 0,\ \ j=1,\dots,k\end{aligned} \tag{NLP}$$

and evaluating f at any rounding of any of its optimal points yield a possibly tight upper bound for the minimal value of the original mixed-integer problem (MIPOP).

Remark 3.1 The best candidates for the polynomials h_j are the functions

$$J_j : \mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R},\ (x,y)\mapsto \max_{z\in B_\infty(\frac{1}{2})} g_j(x,y+z).$$

These functions are well defined, since the maximum is attained on the compact box $B_\infty(\frac{1}{2})$, and yield an exact description of the inner parallel set, but they are only continuous but not differentiable in general. This is a well known fact from parametric optimization (e.g., [1, Theorem 4.2.1]), but since the proof and the counterexample are short, we will give them here. Note that the continuity of J_j implies the closedness of $G_j^- = \{(x,y)\in B^- : J_j(x,y)\le 0\}$.

Lemma 3.2 *Let $g\in\mathbb{R}[x,y]$. Then the function*

$$J : \mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R},\ (x,y)\mapsto \max_{z\in B_\infty(\frac{1}{2})} g(x,y+z)$$

is continuous.

Proof We show lower and upper semicontinuity separately. Let $(x_k, y_k)_k$ be a sequence in $\mathbb{R}^n \times \mathbb{R}^m$ converging to a point $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$. Further, choose $z \in B_\infty(\frac{1}{2})$, resp. $z_k \in B_\infty(\frac{1}{2})$, such that $J(x, y) = g(x, y + z)$, resp. $J(x_k, y_k) = g(x_k, y_k + z_k)$.

Lower semicontinuity: Choose a subsequence $(x_{k_l}, y_{k_l})_l$ of $(x_k, y_k)_k$ such that $\liminf_{k\to\infty} J(x_k, y_k) = \lim_{l\to\infty} J(x_{k_l}, y_{k_l})$. Then

$$\begin{aligned}\liminf_{k\to\infty} J(x_k, y_k) &= \lim_{l\to\infty} J(x_{k_l}, y_{k_l}) = \lim_{l\to\infty} g(x_{k_l}, y_{k_l} + z_{k_l}) \\ &\geq \lim_{l\to\infty} g(x_{k_l}, y_{k_l} + z) = g(x, y + z) = J(x, y).\end{aligned}$$

Upper semicontinuity [12, Lemma 1]: Choose a subsequence $(x_{k_j}, y_{k_j})_j$ of $(x_k, y_k)_k$ such that $\limsup_{k\to\infty} J(x_k, y_k) = \lim_{j\to\infty} J(x_{k_j}, y_{k_j})$. Using the Bolzano-Weierstrass theorem, we can assume $(x_{k_j}, y_{k_j}, z_{k_j})_j \to (x, y, z^*)$ for $j \to \infty$ for some $z^* \in B_\infty(\frac{1}{2})$. Then

$$\begin{aligned}\limsup_{k\to\infty} J(x_k, y_k) &= \lim_{j\to\infty} J(x_{k_j}, y_{k_j}) = \lim_{l\to\infty} g(x_{k_j}, y_{k_j} + z_{k_j}) \\ &= g(x, y + z^*) \leq g(x, y + z) = J(x, y).\end{aligned}$$

The continuity of J follows. □

Example 3.1 Consider the polynomial $g(y) = y^2 - 1$. Then

$$J(y) = \max_{z\in B_\infty(\frac{1}{2})} g(y + z) = \max\{g(y - \tfrac{1}{2}), g(y + \tfrac{1}{2})\}$$

is not differentiable at 0 and, in particular, not a polynomial.

As touched on in Example 3.1, one can give a polynomial description of the inner parallel set in the case that all g_j's are convex. But the number of polynomials needed for this description grows exponentially in the number of integer variables. Hence, this result is less suited for practical computation in the case of a high number of integer variables.

Theorem 3.1 *Write $B = B_x \times B_y \subseteq \mathbb{R}^n \times \mathbb{R}^m$, where B_x, resp. B_y, is the box corresponding to the bounds on the variables x, resp. y. Let $g \in \mathbb{R}[x, y]$ be convex on B_y for each $x \in B_x$. Then*

$$G^- = \{(x, y) \in B^- : \forall i \in \{1, \ldots, 2^m\} : g(x, y + v^i) \leq 0\},$$

where the v^i's are the 2^m vertices of $B_\infty(\frac{1}{2})$.

Proof Since the function $\psi_{(x,y)}(z) := g(x, y+z)$ is convex on $B_\infty(\frac{1}{2})$ for each (x, y), we obtain from the vertex theorem of convex maximization (cf. [21, Corollary 32.3.4])

$$\max_{z\in B_\infty\left(\frac{1}{2}\right)} g(x, y + z) = \max_{i=1,\ldots,2^m} g(x, y + v^i),$$

where the v^i's are the 2^m vertices of $B_\infty(\frac{1}{2})$. Now

$$\begin{aligned} G^- &= \{(x, y) \in B^- : \forall z \in B_\infty(\tfrac{1}{2}) : g(x, y + z) \leq 0\} \\ &= \{(x, y) \in B^- : \max_{z \in B_\infty\left(\frac{1}{2}\right)} g(x, y + z) \leq 0\} \\ &= \{(x, y) \in B^- : \forall i \in \{1, \ldots, 2^m\} : g(x, y + v^i) \leq 0\} \end{aligned}$$

holds. □

In the following, we will need to assume that the set B^- has nonempty interior. For the continuous variables this means that $x_i^l < x_i^u$ holds in all components i. This is not a restriction, since in the case that the lower bound equals the upper bound in one component, we can just fix the corresponding variable. For the integer variables this means that no binary variables (or integer variables which only attain two consecutive integer values) are allowed. This seems to be a significant drawback, but it can be easily tackled by enlarging the set B, resp. B^- in the y-components, which is the topic of Sect. 7. For better readability, the index j is dropped from now on.

4 Finding Polynomial Overestimators

Let $g \in \mathbb{R}[x, y]$. In this section we explain, how we can calculate a polynomial $h \in \mathbb{R}[x, y]$ such that it yields a tight inner approximation of the set G^- via the moment-/sos-hierarchy introduced by Lasserre [9, 12, 13]. Such a polynomial h can be calculated by solving the following polynomial optimization problem

$$\begin{aligned} \inf_{h \in \mathbb{R}[x,y]} \int_{B^-} h(x, y)\, d\lambda(x, y) \\ \text{s.t. } h(x, y) \geq g(x, y + z) \text{ for all } (x, y, z) \in S, \end{aligned} \tag{POP}$$

where

$$S := \{(x, y, z) \in \mathbb{R}^n \times \mathbb{R}^{2m} : (x, y) \in B^-, z \in B_\infty(\tfrac{1}{2})\}.$$

Interpreting (POP), one wants to find a polynomial h with minimal integral value with respect to the Lebesgue measure λ, but which lies above g on the set S. This is equivalent to finding a polynomial h lying above g such that the area between g and h is minimal. Since the objective of (POP) is an integral over the set B^-, we need that B^- is not a null set, that is, the interior of B^- has to be nonempty. Further, the following proposition holds:

Proposition 4.1 *If h is feasible for (POP), then*

$$T_h^- := \{(x, y) \in B^- : h(x, y) \leq 0\} \subseteq G^-.$$

Proof Let $(x, y) \in T_h^-$, then $(x, y) \in B^-$ and $h(x, y) \leq 0$ hold. Since h lies above g on the set S, we have $g(x, y + z) \leq h(x, y) \leq 0$ for all $z \in B_\infty(\frac{1}{2})$, which yields $(x, y) \in G^-$. □

As intended, a solution h of (POP) supplies a subset of the set G^-. The opposite of Proposition 4.1 is not true in general:

Example 4.1 Let $g(y) = (y + 2)(y - 1)$, $B = [-1, 1]$ and $h(y) = y - 1$. Then $G^- = T_h^- = B^- = [-\frac{1}{2}, \frac{1}{2}]$, but for $y' = z' = \frac{1}{2}$, we have $(y', z') \in S$ and $h(y') = -\frac{1}{2} < 0 = g(y' + z')$.

Remark 4.1 If allowing not only polynomials to maximize over, an optimal solution of (POP) is the function J from Lemma 3.2. Since J is only continuous, but no polynomial in general, (POP) does not always attain its minimum. Nonetheless there is some convergence behavior as we will see at the end of this section.

In the case that J is indeed a polynomial and the interior of B^- is nonempty, it is even the unique solution of (POP). This can be seen easily: Let h' be a polynomial feasible for (POP), that is $h' \geq J$ on B^-. If $\int_{B^-} J(x, y)\, d\lambda(x, y) = \int_{B^-} h'(x, y)\, d\lambda(x, y)$, we obtain $J = h'$ (on $\mathbb{R}^n \times \mathbb{R}^m$), since the interior of B^- is nonempty.

Note that the variables of the optimization problem above are the coefficients of the polynomial h: Write $h = \sum_{\alpha,\beta} c_{\alpha,\beta} x^\alpha y^\beta$, where $\alpha \in \mathbb{N}_0^n$, $\beta \in \mathbb{N}_0^m$ are multi-indices and $c_{\alpha,\beta} \in \mathbb{R}$ is the corresponding coefficient, then

$$\int_{B^-} h(x, y)\, d\lambda(x, y) = \sum_{\alpha,\beta} c_{\alpha,\beta} \int_{B^-} x^\alpha y^\beta\, d\lambda(x, y).$$

Since the box B^- is known, the moments can be calculated in advance and hence they are only constants in the optimization problem, whereas the $c_{\alpha,\beta}$'s are the variables over which we want to optimize.

Furthermore, the box S can be written as a semialgebraic set, which is a set described by finitely many polynomial equalities and inequalities. We first define a short notation for describing the bound constraints with polynomials. Let $a, b \in \mathbb{R}$ with $a \leq b$ and define the univariate polynomial

$$\varphi_{a,b}(x) := (x - a)(b - x).$$

Then $\varphi_{a,b}(x) \geq 0$ describes the interval $[a, b]$. We set

$$\begin{aligned}
\varphi_i(x, y, z) &= \varphi_{x_i^l, x_i^u}(x_i) && \text{for } i = 1, \ldots, n, \\
\varphi_i(x, y, z) &= \varphi_{y^l + \frac{1}{2}e, y^u - \frac{1}{2}e}(y_i) && \text{for } i = n + 1, \ldots, n + m, \\
\varphi_i(x, y, z) &= \varphi_{-\frac{1}{2}, \frac{1}{2}}(z_i) && \text{for } i = n + m + 1, \ldots, n + 2m.
\end{aligned}$$

In this way we obtain a polynomial description of S:

$$S = \{(x, y, z) \in \mathbb{R}^n \times \mathbb{R}^{2m} : \forall i \in \{1, \ldots, n + 2m\} : \varphi_i(x, y, z) \geq 0\}.$$

Now (POP) has the general form of a polynomial optimization problem that can be tackled by a hierarchy of sos and moment relaxations introduced by Lasserre [9, 12]. The hierarchy of sos relaxations can be stated as follows:

$$\rho_r = \min_{h,\sigma_i,\theta_j} \int_{B^-} h(x,y)\, d\lambda(x,y) \tag{SOS$_r$}$$

$$\text{s.t. } h(x,y) - g(x,y+z) = \sigma_0(x,y,z) + \sum_{i=1}^{n+2m} \sigma_i(x,y,z)\varphi_i(x,y,z),$$

$$h \in \mathbb{R}[x,y]_{2r}, \sigma_0 \in \Sigma_r[x,y,z], \sigma_i \in \Sigma_{r-1}[x,y,z], i = 1,\dots,n+2m,$$

where $r \in \mathbb{N}$ with $r \geq \lceil \deg(g)/2 \rceil$ denotes the order of the relaxation. In (SOS$_r$) the "$\geq$"-condition from (POP) is changed to a sum of squares constraint bounded by a degree r. The degree of h is bounded by $2r$, but can also be fixed to any arbitrary degree $\leq 2r$. Since the solution h of (SOS$_r$) is also feasible for (POP), Proposition 4.2 follows from Proposition 4.1. It allows the calculation of an approximation of the solution of (POP).

Proposition 4.2 *If h is feasible for (SOS$_r$), then $T_h^- \subseteq G^-$.*

Note that (SOS$_r$) is a simplification of the hierarchy introduced in [12]. The method therein can be applied to a more general semialgebraic set S.

Using the truncated quadratic modules from Sect. 2 for x, y and z instead of x, the sos-constraint in (SOS$_r$) simplifies to

$$h(x,y) - g(x,y+z) \in \mathcal{M}_r(\varphi_1,\dots,\varphi_{n+2m}) \subseteq \mathbb{R}[x,y,z]_{2r}.$$

Naturally, one could also use the linear box constraints for describing the box S. But this description of the semialgebraic set yields a slightly weaker hierarchy as shown by the following lemma.

Lemma 4.1 *Let $a, b \in \mathbb{R}$ with $a < b$. Then*

1. $x - a, b - x \in \mathcal{M}_1(\varphi_{a,b})$,
2. $\varphi_{a,b} \notin \mathcal{M}_1(x-a, b-x)$ *and*
3. $\varphi_{a,b} \in \mathcal{M}_2(x-a, b-x)$.

Proof The first and third statement follow from the following identities:

$$x - a = \frac{1}{b-a}((x-a)^2 + (x-a)(b-x)) \in \mathcal{M}_1(\varphi_{a,b}),$$

$$b - x = \frac{1}{b-a}((-x+b)^2 + (x-a)(b-x)) \in \mathcal{M}_1(\varphi_{a,b}),$$

$$\varphi_{a,b} = \frac{1}{b-a}((b-x)^2(x-a) + (x-a)^2(b-x)) \in \mathcal{M}_2(x-a, b-x).$$

Showing the second statement, we assume that $\varphi_{a,b} \in \mathcal{M}_1(x-a, b-x)$. Then

$$\varphi_{a,b} = \sigma_0 + \sigma_1(x-a) + \sigma_2(b-x)$$

with $\sigma_0 \in \Sigma_1[x]$ and $\sigma_1, \sigma_2 \in \mathbb{R}$. Since $\deg(\varphi_{a,b}) = 2$ and $\sigma_1, \sigma_2 \in \mathbb{R}$, also $\deg(\sigma_0) = 2$ and the leading coefficients of $\varphi_{a,b}$ and σ_0 coincide. But the leading coefficients have different signs, which is a contradiction. □

Using the Riesz functional, moment matrices and localizing matrices introduced in Sect. 2, we can state the hierarchy of moment relaxations (cf. [12]), which are the duals of the semidefinite problems (SOS$_r$):

$$\begin{aligned} \rho_r = \min_{\zeta \in \mathbb{N}_{2r}^{n+2m}} \; & L_\zeta(g(x, y+z))c \\ \text{s.t. } & \mathcal{M}_r(\zeta) \succeq 0 \\ & \mathcal{M}_{r-1}(\varphi_j(x,y,z)\zeta) \succeq 0, \;\; j = 1, \dots, n+2m, \\ & L_\zeta(x^\alpha y^\beta) = \xi_{(\alpha,\beta)}, \;\; \alpha \in \mathbb{N}_{2r}^n, \beta \in \mathbb{N}_{2r}^m \end{aligned} \qquad (\text{MOM}_r)$$

where

$$\xi_{(\alpha,\beta)} := \frac{1}{\lambda(B^-)} \int_{B^-} x^\alpha y^\beta \, d\lambda(x,y)$$

with $\alpha \in \mathbb{N}_{2r}^n$, $\beta \in \mathbb{N}_{2r}^m$ and $\lambda(B^-) = \int_{B^-} 1 \, d\lambda(x,y)$. Recalling the definition of the Riesz functional in (1) $\zeta_{(\alpha,\beta,\gamma)} = L_\zeta(x^\alpha y^\beta z^\gamma)$ holds for all $\alpha \in \mathbb{N}_{2r}^n$, $\beta, \gamma \in \mathbb{N}_{2r}^m$. Thus the last line in (MOM$_r$) fixes the values of the variables $\zeta_{(\alpha,\beta,0)}$ to the values $\xi_{(\alpha,\beta)}$ for all $\alpha \in \mathbb{N}_{2r}^n$, $\beta \in \mathbb{N}_{2r}^m$.

In Theorem 4.1 we state convergence results from [12]: The first part states that the solutions h_r^* of the hierarchy converges in the L_1-norm to the continuous function J. The second one states that, if r is large enough, the set $T_{h_r^*}^-$ coincides with the inner parallel set G^- up to a set of small Lebesgue volume. For proving the convergence, the quadratic module $\mathcal{M}(\varphi_1, \dots, \varphi_{n+2m})$ must be Archimedean, that is, there must exist some $N \in \mathbb{N}$ such that

$$N - \sum_{i=1}^n x_i^2 - \sum_{j=1}^m y_j^2 - \sum_{j=1}^m z_j^2 \in \mathcal{M}(\varphi_1, \dots, \varphi_{n+2m}).$$

This is ensured by the following lemma.

Lemma 4.2 *Let $a, b \in \mathbb{R}$ with $a \leq b$. Then there exists some $N \in \mathbb{N}$ such that*

$$N - x^2 \in \mathcal{M}(\varphi_{a,b})$$

Proof From Lemma 4.1 we obtain $x-a, b-x \in \mathcal{M}(\varphi_{a,b})$. Choose $\widetilde{N} \in \mathbb{N}$ such that $\widetilde{N} \geq \max\{|a|, |b|, |ab|\}$. Then

$$x + \widetilde{N} = x - a + (\widetilde{N} + a) \in \mathcal{M}(\varphi_{a,b}),$$

since $(\widetilde{N} + a) \geq 0$. Similarly, $\widetilde{N} - x, \varphi_{a,b} + (\widetilde{N} + ab) \in \mathcal{M}(\varphi_{a,b})$. Now

$$\begin{aligned}(1 + |a| + |b|)\widetilde{N} - x^2 = \varphi_{a,b} &+ (\widetilde{N} + ab) \\ &+ \frac{|a|}{2}\left((1 + \operatorname{sgn}(a))(\widetilde{N} - x) + (1 - \operatorname{sgn}(a))(x + \widetilde{N})\right) \\ &+ \frac{|b|}{2}\left((1 + \operatorname{sgn}(b))(\widetilde{N} - x) + (1 - \operatorname{sgn}(b))(x + \widetilde{N})\right) \\ &\in \mathcal{M}(\varphi_{a,b}).\end{aligned}$$

The statement follows for $N = \lceil (1 + |a| + |b|)\widetilde{N} \rceil \in \mathbb{N}$. ◻

Theorem 4.1 *[12, Theorem 5, Theorem 3] Let $S := B^- \times B_\infty(\frac{1}{2})$ have nonempty interior. Then there is no duality gap between the semidefinite program (SOS$_r$) and its dual (MOM$_r$). Moreover, (SOS$_r$) (resp. (MOM$_r$)) has an optimal solution $h_r \in \mathbb{R}[x, y]_{2r}$ (resp. $\zeta = (\zeta_{(\alpha,\beta,\gamma)}), (\alpha, \beta, \gamma) \in \mathbb{N}_{2r}^{n+2m}$) and*

$$\lim_{r\to\infty} \int_{B^-} |h_r(x, y) - J(x, y)|\, d\lambda(x, y) = 0,$$

where J is defined as in Lemma 3.2. Further, let G^- have nonempty interior and $\{(x, y) \in B^- : J(x, y) = 0\}$ Lebesgue measure zero. Then

$$\lambda(G^- \setminus T_{h_r}^-) \to 0 \quad as\; r \to \infty, \tag{4}$$

where $\lambda(\cdot)$ is the Lebesgue measure of a Borel set in $\mathbb{R}^n \times \mathbb{R}^m$.

Although Theorem 4.1 suggests that a higher order of the hierarchy yields a better approximation of the inner parallel set G^-, this is only true asymptotically, since the convergence in (4) is not monotone, as the following example illustrates.

Example 4.2 Let $g(y) = y^2 - 1$ and $B = [-1, 1]$. From Example 3.1 or Theorem 3.1 we obtain

$$J(y) = \max_{z\in B_\infty\left(\frac{1}{2}\right)} g(y + z) = \max\{g(y - \tfrac{1}{2}), g(y + \tfrac{1}{2})\}$$

and $G^- = [-\frac{1}{2}, \frac{1}{2}]$. The solution h_1 of SOS$_1$ is the zero polynomial and thus $T_{h_1}^- = G^-$, but the solution of SOS$_2$ is the polynomial

$$h_2(y) \approx 2.73206y^2 - 0.605664$$

with $T_{h_2}^- \approx [-0.471, 0.471] \subsetneq G^-$.

5 The Linear Case

In this section we consider different linearity assumptions for the function g and their consequences for properties of the function J from Lemma 3.2 and the inner parallel set G^-.

Lemma 5.1 *Let $a \in \mathbb{R}[x]_r$, $b_1, \ldots, b_m \in \mathbb{R}[x]$, $\mathfrak{b} := (b_1, \ldots, b_m)$ and*

$$g(x, y) := a(x) + \sum_{j=1}^{m} b_j(x) y_j.$$

Then the following assertions hold.

i) $J(x, y) = g(x, y) + \frac{1}{2}\|\mathfrak{b}(x)\|_1$.

ii) If, in addition, $\mathfrak{b}$ is a vector of constant polynomials (i.e., $b_j \in \mathbb{R}$, $j = 1, \ldots m$), then

$$J(x, y) = g(x, y) + \tfrac{1}{2}\|\mathfrak{b}\|_1 \in \mathbb{R}[x, y]_{\max\{r,1\}}.$$

iii) If, in addition, $a \in \mathbb{R}[x]_1$, say $a(x) = \sum_{i=1}^{n} a_i x_i + g_0$, with $a_i, g_0 \in \mathbb{R}$, $i = 1, \ldots, n$, then J is the linear polynomial

$$J(x, y) = g(x, y) + \tfrac{1}{2}\|\mathfrak{b}\|_1 = \sum_{i=1}^{n} a_i x_i + \sum_{j=1}^{m} b_j y_j + g_0 + \tfrac{1}{2}\|\mathfrak{b}\|_1. \qquad (5)$$

Proof The first assertion is due to

$$\begin{aligned} J(x, y) &= \max_{z \in B_\infty(\frac{1}{2})} \left(a(x) + \sum_{j=1}^{m} b_j(x)(y_j + z_j) \right) \\ &= a(x) + \sum_{j=1}^{m} b_j(x) y_j + \max_{z \in B_\infty(\frac{1}{2})} \sum_{j=1}^{m} b_j(x) z_j \\ &= g(x, y) + \tfrac{1}{2} \max_{z \in B_\infty(1)} \sum_{j=1}^{m} b_j(x) z_j \\ &= g(x, y) + \tfrac{1}{2}\|\mathfrak{b}(x)\|_1, \end{aligned}$$

where the last identity follows since the ℓ_1-norm is the dual norm of the ℓ_∞-norm ([5, A.1.6]).

Under the additional assumption of the second assertion, the functions g and, thus, J lie in $\mathbb{R}[x, y]_{\max\{r,1\}}$, so that it follows from the first assertion. The third assertion is an immediate consequence of the second assertion. □

The closed-form description of the inner parallel set

$$G^- = \{(x, y) \in B^- : g(x, y) + \tfrac{1}{2}\|\mathfrak{b}\|_1 \leq 0\},$$

under the assumption of the third assertion in Lemma 5.1 was already given in [18, Proposition 2.1]. This result can also be derived from Theorem 3.1 and is also respected by (POP):

Lemma 5.2 *Let $g \in \mathbb{R}[x, y]_1$ be a linear polynomial of the form (5) and $S := B^- \times B_\infty(\frac{1}{2})$ defined as above and with nonempty interior. Then $h(x, y) = g(x, y) + \frac{1}{2}\|\mathfrak{b}\|_1$ is the unique optimal solution for (POP).*

Proof From Lemma 5.1 iii), $h = J$ follows. This shows the feasibility of h for (POP). Remark 4.1 yields the optimality and uniqueness. □

Using an inhomogeneous version of Farkas' Lemma [20, Theorem 1.3.4] we show that the first order relaxation SOS_1 is as good as the result from [18].

Theorem 5.1 *(Inhomogeneous Farkas' Lemma) For given $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ suppose that the set $S := \{x \in \mathbb{R}^n : Ax \geq b\}$ is nonempty. Let $c \in \mathbb{R}^n$ and $d \in \mathbb{R}$ such that $c^\top x - d \geq 0$ on S, then there exist $\sigma \in \mathbb{R}^m_{\geq 0}$ and $\nu \in \mathbb{R}_{\geq 0}$ such that*

$$c^\top x - d = \sigma^\top (Ax - b) + \nu.$$

Remark 5.1 Using the definition of the truncated quadratic module from Sect. 2, Farkas' Lemma 5.1 implies that

$$c^\top x - d \in \mathcal{M}_1(a_1^\top x - b_1, \ldots, a_m^\top x - b_m),$$

where $a_1, \ldots, a_m \in \mathbb{R}^n$ denote the rows of A.

Proposition 5.1 *Let $g \in \mathbb{R}[x, y]_1$ be a linear polynomial of the form (5) and $S := B^- \times B_\infty(\frac{1}{2})$ defined as above and with nonempty interior. Then $h(x, y) = g(x, y) + \frac{1}{2}\|\mathfrak{b}\|_1$ is the unique optimal solution of SOS_1.*

Proof From Lemma 5.2 it follows that $h(x, y) - g(x, y + z) \geq 0$ on S. Since the box S is described by linear inequalities, we can conclude

$$h(x, y) - g(x, y + z) \in \mathcal{M}_1(\varphi_1, \ldots, \varphi_{n+2m})$$

with Lemma 4.1, Farkas' Lemma 5.1 and Remark 5.1. Thus, h is feasible for SOS_1. Optimality and uniqueness follow from Lemma 5.2. □

6 The Nonlinear Case

In [19], the approach for dealing with general nonlinear functions is similar to the linear case. The authors look for h of the form $h(x, y) = g(x, y) + \nu$, that is, they want to find $\nu \in \mathbb{R}$ such that

$$T_\nu^- := \{(x, y) \in B^- : g(x, y) + \nu \leq 0\}$$

$$\subseteq \{(x, y) \in B^- : \forall z \in B_\infty(\tfrac{1}{2}) : g(x, y+z) \le 0\} =: G^-.$$

This constant ν stems from a global Lipschitz condition with respect to the variables y uniformly in the variables x for the polynomial g on the set B. Thus, the following assumption shall hold for g:

Assumption 6.1 [19, Assumption 3.1] There exists some $L_\infty \ge 0$ such that for all $x \in \mathbb{R}^n$ and $y_1, y_2 \in \mathbb{R}^m$ with $x^l \le x \le x^u$ and $y^l \le y^1, y^2 \le y^u$, we have

$$|g(x, y^1) - g(x, y^2)| \le L_\infty \|(x, y^1) - (x, y^2)\|_\infty = L_\infty \|y^1 - y^2\|_\infty.$$

Since g is a polynomial and all variables are bounded, such a Lipschitz constant always exists. With $\nu_\infty := L_\infty/2$ one can show the following lemma.

Lemma 6.1 *[19, Lemma 3.1] Under Assumption 6.1, we have $T^-_{\nu_\infty} \subseteq G^-$.*

This fact is also respected by (POP) in the case that g is a polynomial:

Lemma 6.2 *Let $g \in \mathbb{R}[x, y]$ and $S := B^- \times B_\infty(\frac{1}{2})$ defined as above. Then $h(x, y) = g(x, y) + \nu_\infty$ is feasible for (POP).*

Proof Let $(x, y, z) \in S$, then

$$g(x, y+z) - g(x, y) \le |g(x, y+z) - g(x, y)| \le L_\infty \|z\|_\infty = \tfrac{1}{2} L_\infty = \nu_\infty$$

and h is feasible for (POP). □

Setting $h(x, y) = g(x, y) + \nu$ in (SOS_r), we can now calculate upper bounds for

$$\nu^* = \min\{\nu \in \mathbb{R} : L = 2\nu \text{ satisfies Assumption 6.1}\}.$$

In the following example, we show that our method using the moment-/sos-hierarchy yields a better value for ν than the method proposed in [19]. In this example, our value for ν is even optimal. Even better results can be achieved by letting h be an arbitrary polynomial.

Example 6.1 Consider the convex polynomial

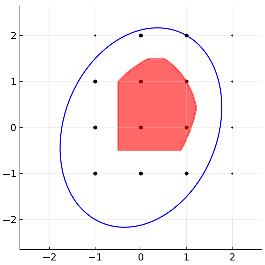
$$g(y) = 3y_1^2 - y_1 y_2 + 2y_2^2 - 9$$

on the box $B = \{y \in \mathbb{R}^2 : -e \le y \le 2e\}$ with e denoting the vector of ones. The Lipschitz constant calculated in [19, Example 4.3, Example 4.4] is $L_\infty = 19$ and therefore $\nu_\infty = 9.5$. Since

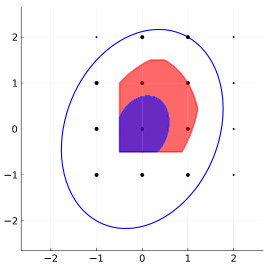
$$3y_1^2 - y_1 y_2 + 2y_2^2 = \frac{1}{8}\left(23y_1^2 + (y_1 - 4y_2)^2\right)$$

is a sum of two squares, it follows that $g(y) + \nu_\infty$ is a strictly positive polynomial and hence $T^-_{\nu_\infty} = \emptyset$. Solving SOS_1, resp. SOS_2, we obtain $\nu_1 \approx 8.074376$, resp. $\nu_2 = 8.$[1]

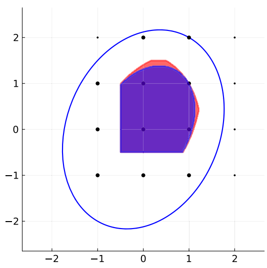
[1] If the linear description for the box constraints is used, SOS_1 will be infeasible (cf. Lemma 4.1).



(a) Inner parallel set



(b) $T^-_{\nu_2}$



(c) $T^-_{h_2}$

Fig. 1 Example 6.1. The blue ellipse is the zero level set of g and the fat black dots are the feasible integer points. The inner parallel set calculated with Theorem 3.1 is shown in red. Its approximations $T^-_{\nu_2}$ (Lipschitz approach) and $T^-_{h_2}$ (quadratic approach) are shown in purple

For both values we obtain that $T^-_{\nu_1}$ and $T^-_{\nu_2}$ are nonempty (see Fig. 1). Also, we cannot find a larger ν than ν_2: For $y = (1.5, -0.5)$ and $z = (0.5, -0.5)$ we have $(y, z) \in S$ and $g(y) + 8 = g(y + z)$.

But our approach is more general than only finding a constant ν. Instead of looking for h of the form $h = g(x, y) + \nu$, we look for an arbitrary polynomial h as long as we fix the degree. For this example we fix the degree of h to 2. Then the solution for SOS_2 is the polynomial

$$\begin{aligned} h_2(y_1, y_2) \approx\ & 5.32525y_1^2 - 2.49835y_1y_2 + 3.61356y_2^2 \\ & - 0.28848y_1 - 0.34123y_2 - 5.79652 \end{aligned}$$

As one can see in Fig. 1, h_2 yields a very large set $T^-_{h_2}$.

7 Enlargements

7.1 Enlargement of Bounds

Up to now, we assumed that the interior of the set B^- has to be nonempty, which excludes problems with binary variables or integer variables, which behave like binaries (i.e., which only attain two consecutive integer values). This drawback can be addressed by enlarging the set B, resp. B^- as shown in [19, Section 4]. We start with a general definition of an enlargement as given in [19].

Definition 7.1 Let M be the feasible set of (MIPOP) and $\widehat{M}$ its relaxed feasible set. Then a set $\widetilde{M}$ is an *enlargement* of the set $\widehat{M}$ if $\widehat{M} \subseteq \widetilde{M}$ and

$$M = \widetilde{M} \cap (\mathbb{R}^n \times \mathbb{Z}^m).$$

The inner parallel set $\widetilde{M}^-$ of $\widetilde{M}$ is called *enlarged inner parallel* set of $\widehat{M}$.

Note that $\widehat{M} \subseteq \widetilde{M}$ implies $\widehat{M}^- \subseteq \widetilde{M}^-$ in Definition 7.1.

The definition of an enlargement is now very general. But since we are not considering the feasible set in total, but every constraint on its own, we are only interested in enlargements of the box B. For this, let $\sigma = (\sigma^l, \sigma^u) \in [0, 1)^m \times [0, 1)^m$, then

$$B_\sigma = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : x^l \le x \le x^u, y^l - \sigma^l \le y \le y^u + \sigma^u\}$$

is an enlargement for B and

$$B_\sigma^- = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : x^l \le x \le x^u, y^l - \sigma^l + \tfrac{1}{2}e \le y \le y^u + \sigma^u - \tfrac{1}{2}e\}$$

its enlarged inner parallel set. Assuming with out loss of generality $x^l < x^u$ and $y^l < y^u$ in all components, this means that the interior of B_σ^- is nonempty for all $\sigma = (\sigma^l, \sigma^u) \in (0, 1)^m \times (0, 1)^m$, even when binary variables appear in (MIPOP). Hence we can solve the polynomial optimization problem (POP) (resp. the approximation (SOS_r)) by replacing the set B_σ^- by B^- (resp. the polynomials describing the corresponding semialgebraic set).

7.2 Enlargement of Constraints

But not only enlargements of the bounds are possible. In the case of purely integer constraints, that is, if $g_j(x, y) = g_j(y) \in \mathbb{Z}[y]$, one can subtract any constant $\tau_j \in [0, 1)$ such that

$$\forall y \in \mathbb{Z}^m : \ g_j(y) \le 0 \Longleftrightarrow g_j(y) - \tau_j \le 0.$$

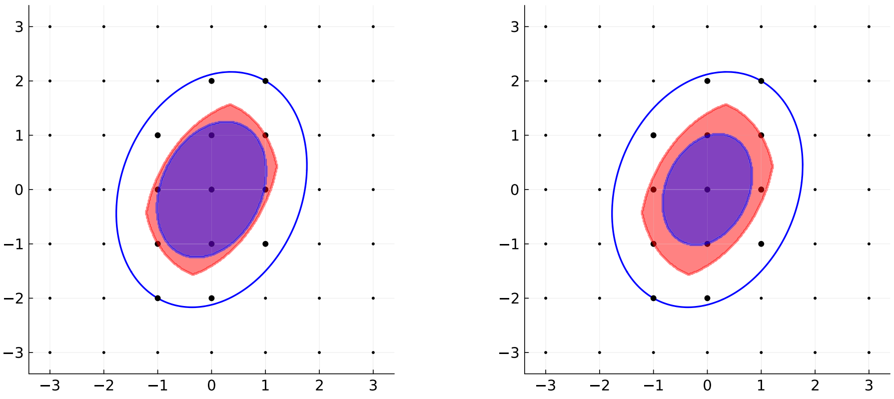
Under some conditions it is even possible to choose a larger τ_j. For more details we refer to [19, Example 4.2].

Note that if we have an optimal solution h_r of (SOS_r) then $h_r - \tau$ is also an optimal solution of (SOS_r) after replacing $g(y)$ with $g(y) - \tau$. Hence we can subtract τ from h_r after solving (SOS_r).

Combining both enlargement ideas, we can solve now

$$\begin{aligned} \min_{(x,y)\in B_\sigma^-} \ & f(x, y) \\ \text{s.t. } & h_j(x, y) - \tau_j \le 0, \ j = 1, \dots, k \end{aligned} \tag{$\text{NLP}_{\sigma,\tau}$}$$

with suitable enlargement parameters $\sigma = (\sigma^l, \sigma^u) \in (0, 1)^m \times (0, 1)^m$ and $\tau \in \mathbb{R}^k_{\ge 0}$ instead of (NLP). We expect that the feasible set $T_{\sigma,\tau}^-$ of ($\text{NLP}_{\sigma,\tau}$) is enlarged compared to the feasible set T^- of (NLP) and we thus find (better) feasible points for (MIPOP), but this is not guaranteed in general. Whereas enlarging the constraints yields the inclusion $T_{\sigma,\tau_1}^- \subseteq T_{\sigma,\tau_2}^-$ for $\tau_1 \le \tau_2$ componentwise, we cannot state anything about possible inclusions if we enlarge the bounds. This is due to the fact that we do not



(a) $T^-_{h_{\delta_1}}$ (b) $T^-_{h_{\delta_2}}$

Fig. 2 The inner parallel set and its approximations with different enlargement parameters from Example 7.1

compute the approximation of the inner parallel set directly, but only computing an approximation by a polynomial description. In some cases enlarging might even have the opposite of the desired effect and the feasible set might shrink, as the following example shows. Thus, enlarging bounds needs to be done carefully.

Example 7.1 Consider the polynomial from Example 6.1

$$g(y) = 3y_1^2 - y_1 y_2 + 2y_2^2 - 9$$

on the box $B = \{y \in \mathbb{R}^2 : -3e \leq y \leq 3e\}$ with e denoting the vector of ones. We now set $\sigma^l = \sigma^u = \delta e$. Then solving SOS$_2$ for $\delta_1 = 0$ and $\delta_2 = 0.99$ yields the quadratic polynomials h_{δ_1} and h_{δ_2} whose sublevel sets are shown in Fig. 2. We see that enlarging the bounds yields a worse approximation of the inner parallel set. This is not surprising, since if the box is larger, then the semialgebraic set S in (POP) is larger, that is, the polynomial inequality needs to be satisfied at more points.

8 An Algorithm for Finding Feasible Points

We can now state our algorithm FRA-SOS (feasible rounding approach with sos-hierarchy) for finding feasible points for mixed-integer polynomial optimization problems (MIPOP).

Bound Tightening

Example 7.1 gives rise to consider a direct improvement for FRA-SOS. The quality of the approximation of the inner parallel set depends in some cases heavily on the size

Algorithm FRA-SOS

Require: (MIPOP), $\sigma = (\sigma^l, \sigma^u) \in \mathbb{R}_+^{2m}$, $\tau \in \mathbb{R}^k$, $r \in \mathbb{N}$,
Ensure: feasible point $(\check{x}, \check{y})$ for (MIPOP) (if possible)
  for each constraint $g_j(x, y)$ **do**
    if no integer variable appears in $g_j(x, y)$ **then**
      $h_j = g_j(x, y)$
    else if integer variables appear linearly in $g_j(x, y)$ **then**
      $h_j = g_j(x, y) + \frac{1}{2}\|\mathfrak{b}\|_1$
    else if any integer variables appears nonlinearly in $g_j(x, y)$ **then**
      compute h_j by solving (SOS_r) with enlargement parameters σ^l and σ^u
    end if
  end for
  Solve ($\text{NLP}_{\sigma,\tau}$)
  if ($\text{NLP}_{\sigma,\tau}$) is feasible **then**
    round any solution (x^*, y^*) of ($\text{NLP}_{\sigma,\tau}$) to $(\check{x}, \check{y})$
    return $(\check{x}, \check{y})$
  end if

of the box B, resp. B^-. If the box B is as small as possible, the approximation of the inner parallel set can be improved. This can be achieved by a well known method form mixed-integer optimization: bound tightening [2]. In common solvers like `SCIP` or `Gurobi` some presolve-procedure is implemented, which can derive tighter bounds for the original optimization problem (e.g., implied bounds or feasibility-based bound tightening). For our numerical tests we would like to investigate this advantage as well. Since we are dealing with mixed-integer polynomial optimization problems, we can use the moment-/sos-hierarchy for this purpose. Probably, this is not very efficient, because we need to solve for each variable two semidefinite programs, but it can be easily replaced by more elaborated methods.

For tightening the lower and upper bounds of the variables in (MIPOP), we solve the following polynomial optimization problems.

$$\begin{aligned} \min_{(x,y)\in\mathbb{R}^n\times\mathbb{R}^m} \quad & \pm v \\ \text{s.t. } & g_j(x, y) \le 0, \;\; j = 1, \dots, k \\ & x^l \le x \le x^u, \\ & y^l \le y \le y^u, \end{aligned}$$

where $v \in (x, y)$ is the variable to be tightened. This can be done by using the classical moment-/sos-hierarchy introduced in [9].

9 Numerical Study

The main purpose of this computational study is to apply the moment-/sos-hierarchy for calculating feasible points for general mixed-integer polynomial nonlinear problem from practice using the concept of granularity developed by Neumann et al. [18, 19]. For this we consider test examples from the MINLPLib [16].

We implement FRA-SOS in `Julia 1.9` using the modeling language `JuMP` [4]. Further we used the software tool `TSSOS` for solving the moment-/sos-hierarchy to exploit sparsity of the mixed-integer polynomial optimization problems [14]. We modified the objective in the source code to solve the problems (SOS_r). For solving the nonlinear problems ($\mathrm{NLP}_{\sigma,\tau}$), we used the (local) solver `Ipopt` [6, 22]. Finally we compared the results with the mixed-integer solver `SCIP` [3]. All tests were carried out on an `Apple M1 Pro` with 32 GB of RAM.

For our numerical study, we set the enlargement parameters $\sigma = \sigma^l = \sigma^u = 1 - 10^{-4}e$ and $\tau_j = 1 - 10^{-4}$ for all j with $g_j(x, y) = g_j(y) \in \mathbb{Z}[y]$.

Sometimes the moment-/sos-hierarchy can cause numerical issues, if high degree monomials are involved [23, Section 5.6]. To avoid these numerical issues, all variables are scaled before solving (SOS_r), such that they are in the interval $[-1, 1]$. This is possible, because all variables are assumed to be bounded. Further, each polynomial is divided through the absolute value of its coefficient with the maximal absolute value. The scaling is also done, if the moment-/sos-hierarchy is used to tighten the bounds of the variables as described in Sect. 8.

In total 108 inequality-constrained polynomial instances from the MINLPLib are considered. All variables have to be bounded or can be bounded efficiently by using the bound tightening from Sect. 8. Most of the instances (70 problems) have only integer variables appearing in linear constraints. For these problems, we can use the approach for linear constraints as seen in Sect. 5. For numerical results, we refer to [19]. The remaining 38 instances have integer variables appearing in nonlinear constraints and are interesting for using the moment-/sos-hierarchy to calculate an approximation of the inner parallel set. In total, we are able to obtain a feasible point for 44 out of 108 instances. Further, we obtain feasible points for 24 out of 38 instances, where integer variables appear nonlinearly.

In Table 1, the results for these 24 instances are shown. The columns of the table read as follows.

- name: name of instance in the MINLPLib,
- variables: (total number of variables, number of integer variables, number of binary variables),
- constraints: (total number of constraints, number of constraints with integer variables, number of constraint with integer variables appearing nonlinearly),
- appr.: approach of calculating the polynomial h in (SOS_r):
  - "h": for each constraint g_j with nonlinear integer variables a polynomial h_j of degree $2r$ is calculated such that $T^-_{h_j}$ approximates G^-_j
  - "$g - h$": for each constraint g_j with nonlinear integer variables a polynomial h_j of degree $2r$ is calculated such that $T^-_{g_j - h_j}$ approximates G^-_j,
  - "$g - \nu$": for each constraint g_j with nonlinear integer variables a constant ν_j is calculated such that $T^-_{g_j - \nu_j}$ approximates G^-_j (cf. Section 6),
- ord.: order r in (SOS_r),
- objective: value $f(\check{x}, \check{y})$ of the feasible point $(\check{x}, \check{y})$ found by FRA-SOS,
- BT/SOS/NLP: solving time in seconds for bound tightening, (SOS_r) and ($\mathrm{NLP}_{\sigma,\tau}$).

Table 1 Comparison of FRA-SOS with SCIP for examples with nonlinear integer variables from the MINLPLib

Name	Variables	Constraints	appr	ord	Objective	BT	SOS	NLP	w/o BT	SCIP	opt
cvxnonsep_normcon20r	(40,10,0)	(21,20,10)	"h"	1	−18.19	0.36	0.04	0.01	**0.05**	0.03	−21.749
cvxnonsep_normcon30r	(60,15,0)	(31,30,15)	"h"	1	−28.43	0.74	0.06	0.01	**0.07**	0.03	−34.24
cvxnonsep_normcon40r	(80,20,0)	(41,40,20)	"h"	1	−25.67	1.38	0.08	0.02	**0.1**	0.04	−32.63
nvs03	(2,2,0)	(2,1,1)	"h"	1	17	0.01	4.0E-03	0.01	**0.01**	2.2E-03	16
nvs07	(3,3,0)	(2,1,1)	"h"	2	252	0.05	0.02	0.01	**0.03**	2.4E-03	4
nvs10	(2,2,0)	(2,2,2)	"h"	1	−308.4	0.02	0.01	0.01	**0.02**	2.5E-03	−310.8
nvs11	(3,3,0)	(3,3,3)	"h"	1	−427.8	0.03	0.02	0.01	**0.03**	3.4E-03	−431
nvs12	(4,4,0)	(4,4,4)	"h"	1	-478	0.04	0.03	0.01	**0.04**	0.01	−481.2
nvs13	(5,5,0)	(5,5,5)	"h"	1	−574.6	0.05	0.07	0.01	**0.08**	0.02	−585.2
nvs17	(6,6,0)	(7,7,7)	"h"	1	−1086.8	0.08	0.01	0.13	*0.14*	0.27	−1100.4
nvs18	(7,7,0)	(6,6,6)	"h"	1	−768.2	0.07	0.09	0.01	**0.1**	0.06	−778.4
nvs19	(8,8,0)	(8,8,8)	"h"	1	−1087.6	0.12	0.2	0.01	*0.21*	0.38	−1098.4
nvs21	(3,2,2)	(2,2,2)	"h"	3	−2.41	0.06	0.16	0.01	0.17	0.02	−5.68
nvs23	(9,9,0)	(9,9,9)	"h"	1	−1098.8	0.14	0.29	0.28	*0.57*	0.81	−1125.2
nvs24	(10,10,0)	(10,10,10)	"h"	1	−1001.4	0.16	0.44	0.3	*0.74*	1.08	−1033.2
prob03	(2,2,0)	(1,1,1)	"h"	1	12	0.87	2.82	0.37	3.19	2.0E-03	10
sonet17v4	(136,136,136)	(2057,17,17)	"$g-h$"	1	1.82E+06	–	5.49	0.4	5.89	0.01	1.18E+06
sonet18v6	(153,153,153)	(2466,18,18)	"$g-h$"	1	6.93E+06	–	7.36	0.24	7.6	0.1	3.39E+06
sonet19v5	(171,171,171)	(2926,19,19)	"$g-h$"	1	4.81E+06	–	10	0.51	10.51	0.12	3.39E+06
sonet20v6	(190,190,190)	(3440,20,20)	"$g-h$"	1	9.25E+06	–	13.26	0.62	13.88	0.15	3.31E+06
sonet21v6	(210,210,210)	(5336,23,23)	"$g-v$"	1	1.07E+07	–	16.74	0.8	17.54	0.2	*
sonet22v4	(231,231,231)	(6096,24,24)	"$g-v$"	1	3.41E+06	–	21.47	0.94	22.41	0.2	*
sonet23v6	(253,253,253)	(4011,21,21)	"$g-h$"	1	1.28E+07	–	28.89	0.89	29.78	0.26	*
sonet24v2	(276,276,276)	(4642,22,22)	"$g-h$"	1	2.15E+07	–	38.73	1.63	40.36	0.29	3.31E+06

– w/o BT: solving time in seconds for FRA-SOS without bound tightening (sum of times for (SOS_r) and ($\text{NLP}_{\sigma,\tau}$)).
– SCIP: time in seconds for SCIP to find a feasible point, which is at least as good as the point found by FRA-SOS.
– opt.: optimal value reported by MINLPLib. If no optimal value is known, this is marked by *.

For most of the problems from the table, we can use the first order of the hierarchy to calculate a feasible point, since the problems are quadratic. Only the two instances `nvs07` and `nvs21` have polynomial constraints of higher order. For the `sonet` instances the approaches "$g - h$" and "$g - \nu$" are reported, since the approach "h" was not successful. The reason for this may be that for the instances involving only binary variables the concept of granularity is not that suited despite of the use of enlargements.

Since most of the problems are small-sized, the times for bound tightening, solving (SOS_r) and ($\text{NLP}_{\sigma,\tau}$) are quite fast. For the larger binary problems, there was no bound tightening carried out.

For 10 out of 24 problems `SCIP` finds faster feasible points that are at least as good as those found by FRA-SOS. For 10 problems, our algorithm is comparable with `SCIP` in time (**bold** in Table 1), if bound tightening is not considered (see discussion in Sect. 8). Finally, although `SCIP` is a very fast mixed-integer solver, FRA-SOS was able to outperform `SCIP` on four of the larger `nvs` instances (*italic* in Table 1). This behavior might indicate that FRA-SOS may be suitable to quickly find good feasible points of practical problems which possess the same structure as the `nvs` instances, that is, problems which have constraints with nonbinary and nonlinear integer variables. To verify this indication, one would naturally need far more examples than provided by the MINLPLib. However, an exhaustive numerical study would go beyond the purpose of the present paper, which is to demonstrate the theoretical potential of our method.

10 Conclusions

In this article, we propose a feasible rounding approach for mixed-integer polynomial optimization problems. The concept of granularity for mixed-integer problems is combined with the moment-/sos-hierarchy from polynomial optimization. Further, we generalize the numerical method from the literature. Instead of calculating only a Lipschitz constant, we can calculate directly polynomial descriptions of approximations of the inner parallel set. Thus, our approach is theoretically at least as good and yields better practical results. Testing our approach on examples from the MINLPLib, we can calculate feasible points for a significantly share of considered problems in our numerical study. Although the number of adequate test problems from the MINLPLib is limited, the numerical results indicate that the approach proposed in this article may be suitable for quickly finding good feasible points for polynomial inequality constrained problems with nonbinary and nonlinear integer variables.

Acknowledgements The first author is funded by Deutsche Bundesstiftung Umwelt (DBU). Further, we would like to express our gratitude to our colleagues Moritz Link and Markus Schweighofer for very fruitful discussions. The authors are grateful to anonymous reviewers for their constructive remarks.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Bank, B., Guddat, J., Klatte, D., Kummer, B., Tammer, K.: Non-Linear Parametric Optimization. Akademie-Verlag, Berlin (1982)
2. Belotti, P., Kirches, C., Leyffer, S., Linderoth, J., Luedtke, J., Mahajan, A.: Mixed-integer nonlinear optimization. Acta Numer. **22**, 1–131 (2013). https://doi.org/10.1017/S0962492913000032
3. Bestuzheva, K., Besançon M., Chen, W-K., Chmiela, A., Donkiewicz, T., van Doornmalen, J., Eifler, L., Gaul, O., Gamrath, G., Gleixner, A., Gottwald, L., Graczyk, C., Halbig, K., Hoen, A., Hojny, C., van der Hulst, R., Koch, T., Lübbecke, M.E., Maher, S.J., Matter, F., Mühmer, E., Müller, B., Pfetsch, M.E. Rehfeldt, D., Schlein, S., Schlösser, F., Serrano, F., Shinano, Y., Sofranac, B., Turner, T., Vigerske, S., Wegscheider, F., Wellner, P., Weninger, D., Witzig, J.: The SCIP Optimization Suite 8.0. Technical Report, Optimization Online (2021). http://www.optimization-online.org/DB_HTML/2021/12/8728.html
4. Bezanson, J., Edelman, A., Karpinski, S., Shah, V.B.: Julia: a fresh approach to numerical computing. SIAM Rev. **59**, 65–98 (2017). https://doi.org/10.1137/141000671
5. Boyd, S.P., Vandenberghe, L.: Convex Optimization. Cambridge University Press (2004)
6. COIN|OR. Computational Infrastructure for Operations Research (2024). https://www.coin-or.org
7. Jeroslow, R.G.: There cannot be any algorithm for integer programming with quadratic constraints. Oper. Res. **21**, 221–224 (1973)
8. Kannan, R., Monma, C. L.: On the computational complexity of integer programming problems. In Optimization and Operations Research, vol. 157 of Lecture Notes in Economic and Mathematical Systems, 161–172 (Springer, Berlin-New York, 1978)
9. Lasserre, J.B.: Global optimization with polynomials and the problem of moments. SIAM J. Optim. **11**, 796–817 (2001). https://doi.org/10.1137/S1052623400366802
10. Lasserre, J. B.: Moments, Positive Polynomials and Their Applications, vol. 1 of Series on Optimization and its Applications (Imperial College Press, 2009)
11. Lasserre, J.B.: An introduction to polynomial and semi-algebraic optimization. In: Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge (2015)
12. Lasserre, J.B.: Tractable approximations of sets defined with quantifiers. Math. Program. **151**, 507–527 (2015)
13. Laurent, M.: Sums of squares, moment matrices and optimization over polynomials. In: Emerging Applications of Algebraic Geometry, pp. 157–270. Springer, New York (2009)
14. Magron, V., Wang, J.: TSSOS: a Julia library to exploit sparsity for large-scale polynomial optimization (2021).https://doi.org/10.48550/arXiv.2103.00915
15. Magron, V., Wang, J.: Sparse Polynomial Optimization—Theory and Practice, vol. 5 of Series on Optimization and its Applications (World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2023). https://doi.org/10.1142/q0382
16. MINLPLib, A.: Library of Mixed-Integer and Continuous Nonlinear Programming Instances (2024). https://www.minlplib.org

17. Neumann, C., Stein, O.: Feasible rounding approaches for equality constrained mixed-integer optimization problems. Optimization **72**, 581–606 (2023). https://doi.org/10.1080/02331934.2021.1981894
18. Neumann, C., Stein, O., Sudermann-Merx, N.: A feasible rounding approach for mixed-integer optimization problems. Comput. Optim. Appl. **72**, 309–337 (2019)
19. Neumann, C., Stein, O., Sudermann-Merx, N.: Granularity in nonlinear mixed-integer optimization. J. Optim. Theory Appl. **184**, 433–465 (2020)
20. Nie, J.: Moment and polynomial optimization. In: MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2023)
21. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)
22. Wächter, A., Biegler, L.T.: On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. Math. Program. **106**, 25–57 (2006). https://doi.org/10.1007/s10107-004-0559-y
23. 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**, 218–242 (2006). https://doi.org/10.1137/050623802

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.