



Convergent upper bounds in global minimization with nonlinear equality constraints

Christian Füllner¹ · Peter Kirst¹ · Oliver Stein¹

Received: 6 November 2018 / Accepted: 12 March 2020
© The Author(s) 2020

Abstract

We address the problem of determining convergent upper bounds in continuous non-convex global minimization of box-constrained problems with equality constraints. These upper bounds are important for the termination of spatial branch-and-bound algorithms. Our method is based on the theorem of Miranda which helps to ensure the existence of feasible points in certain boxes. Then, the computation of upper bounds at the objective function over those boxes yields an upper bound for the globally minimal value. A proof of convergence is given under mild assumptions. An extension of our approach to problems including inequality constraints is possible.

Keywords Branch-and-bound · Convergent upper bounds · Feasibility verification · Theorem of Miranda

Mathematics Subject Classification 90C26

1 Introduction

In this article we propose a method to determine convergent valid upper bounds for the globally optimal value v^* of nonconvex minimization problems of the form

$$P(B) : \min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad h_j(x) = 0, \quad j \in J, \quad x \in B, \quad (1)$$

✉ Christian Füllner
christian.fuellner@kit.edu

Peter Kirst
peter-kirst@web.de

Oliver Stein
stein@kit.edu

¹ Institute of Operations Research, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

where all functions f and h_j , $j \in J = \{1, \dots, m\}$, with $m \in \mathbb{N}$ and $m \leq n$, are at least continuously differentiable on some open set containing the box B . As usual in continuous optimization we are mainly interested in the case $m < n$. Yet, our results also hold for $m = n$. We do not assume any convexity properties for $P(B)$ so that, in particular, the functions h_j , $j \in J$, are not necessarily linear.

Inequalities of the form $g_i(x) \leq 0$, $i \in I$, may be treated by introducing slack variables y_i , $i \in I$, and rewriting the constraints by

$$g_i(x) + y_i^2 = 0$$

where appropriate box constraints for the slack variables can be determined using interval arithmetic [33], for instance. We point out that in general in nonlinear programming this approach is known to have some drawbacks which will be discussed in the setting of spatial branch-and-bound algorithms in Sect. 6.

The box B is defined by

$$B = \{x \in \mathbb{R}^n \mid \underline{b} \leq x \leq \bar{b}\}$$

with $\underline{b}, \bar{b} \in \mathbb{R}^n$, $\underline{b} < \bar{b}$, and all inequalities meant componentwise. We assume all feasible points to strictly satisfy the box constraints. Although this is in contrast to standard literature assumptions in global optimization, in practice usually this should be fulfilled. Otherwise, it can be achieved by introducing additional, slightly larger box constraints and handling the original box constraints as standard inequality constraints, for example by using the aforementioned slack variable approach.

Additionally, we make the following assumption:

Assumption 1.1 For any globally minimal point x^* of $P(B)$ the gradients $\nabla h_j(x^*)$ of all functions h_j , $j \in J$, are linearly independent.

Assumption 1.1 means that each globally minimal point satisfies the Linear Independence Constraint Qualification (LICQ). Note that this assumption is mild in the sense that it has been proven to generically hold everywhere in the feasible set $M(B)$ [17]. The latter means, in particular, that in case of its violation it may be expected to hold at least under small perturbations of the problem data. Let us remark that in nonlinear optimization the assumption of LICQ even in all locally minimal points is standard for convergence proofs. This means, in particular, that common upper bounding procedures in spatial branch-and-bound methods which rely on the local solution of NLP subproblems implicitly use this or related assumptions as well. In this sense, Assumption 1.1 is not restrictive.

For a problem of type $P(B)$ and some predefined tolerance $\varepsilon > 0$ a typical aim in global optimization is to determine an ε -optimal feasible point $x^* \in M(B)$, that is, in addition to feasibility it satisfies $f(x^*) \leq f(x) + \varepsilon$ for all $x \in M(B)$. One of the state-of-the-art approaches to globally solve problems of type $P(B)$ in this sense is to apply spatial branch-and-bound algorithms. In such methods, the problem is iteratively branched into subproblems of the form

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad x \in M(X)$$

with $M(X) = \{x \in X \mid h_j(x) = 0, j \in J\}$ and X a sub-box of B . Then, for those subproblems upper and lower bounds are constructed to approximate the globally minimal value v^* and exclude boxes from the search space.

The application of spatial branch-and-bound algorithms in global optimization goes back to Falk in [12]. Since then, various extensions, improvements and related methods have been proposed, such as branch-and-reduce [37,38], symbolic branch-and-bound [41,42], branch-and-contract [46] or branch-and-cut [44]. Several state-of-the-art global optimization solvers are based on implementations of branch-and-bound algorithms, see for example BARON [39], COUENNE [8], ANTIGONE [30] and LINDOGLOBAL [25]. For extensive reviews on deterministic global optimization we refer to [13,14] and [16]. The latter in particular covers convergence results.

Typically, within a spatial branch-and-bound algorithm, lower bounds are obtained by special bounding procedures. We will discuss some of them in Sect. 2. Upper bounds, on the contrary, can be obtained by evaluating the objective function f at feasible points of $P(B)$, either directly or indirectly by applying local solvers. However, as the problem $P(B)$ is in general nonconvex, finding a feasible point may be as difficult as solving the problem itself. Therefore, there is no guarantee for the occurrence of feasible points within current branch-and-bound algorithms [21]. This is crucial, though, since a branch-and-bound algorithm requires convergent valid upper bounds in the termination criterion. To cope with this, in the literature sometimes it is suggested to accept ε_f -feasible solutions [13], but [21] shows that instead of upper bounds this may generate arbitrarily bad lower bounds. We shall discuss this issue in more detail in Sect. 2.2.

Several upper bound generation techniques proposed in literature suffer from the above mentioned theoretical drawbacks, although they provide sufficiently good upper bounds *on average* for practical applications. As we will discuss more thoroughly in Sect. 2, this also holds for the most common approach to obtain an upper bound by solving the nonconvex problem $P(B)$ locally as proposed in [6]. In fact, we are not aware of any proof of convergence for upper bounds to the true globally minimal value v^* for this or similar approaches.

In this article, we address this issue and propose a method to determine convergent valid upper bounds for the true globally minimal value v^* . Thus, our research focuses on the theoretical convergence proof. In practice, the proposed method may be used to complement existing upper bounding procedures to ensure convergence. We remark that, regarding upper bounds, by the brief term *convergent* we will mean convergence to the optimal value v^* throughout this paper.

To theoretically ensure termination of a spatial branch-and-bound algorithm, it is essential to either determine a specific feasible point or, at least, to verify the existence of such a point in some sub-box X of B . In the latter case, an upper bound for the objective function on X may be used as a valid upper bound for v^* . In this paper, we propose a method to calculate convergent valid upper bounds for v^* based on this approach.

At least in the absence of rounding errors, by using our upper bounding procedure within a standard branch-and-bound algorithm with convergent lower bounding procedure, it should be possible to obtain a global optimization method that terminates

after a finite number of iterations, since the overall lower bound as well as the overall upper bound converge to the globally minimal value. Note that, even without rounding errors, this does not hold for other mentioned upper bounding procedures such as acceptance of ε_f -feasible points or by using local solvers, since it is not clear whether sufficiently good and correct upper bounds are found at all. In contrast, for commonly used lower bounding procedures such proofs of convergence are available, and so the combination of these methods with our new upper bounding procedure leads to convergent branch-and-bound methods where lower as well as upper bounds converge to the globally minimal value, at least in the absence of rounding errors. In order to obtain a rigorous global branch-and-bound solver, finite precision arithmetic has to be taken into consideration during the implementation, which we have not done so far. Our new upper bounding procedure may be seen as the mathematical foundation of such a rigorous global solver. This will be discussed in more detail in Sect. 6.

In the literature, the problem of feasibility verification has been mainly addressed by heuristics, which are not considered in this article. In contrast, our focus is on deterministic methods. In context of rigorous upper bound determination, such approaches are introduced by Kearfott [19,20] and Domes and Neumaier [9]. They determine approximately feasible points with conventional nonlinear solvers and then construct boxes around such points, for which the existence of a feasible point can be proven via interval Newton methods. Yet, as $n - m$ variables have to be fixed using heuristic approaches and some parameters have to be carefully chosen, verification of feasible points is not guaranteed in general. A deterministic method to calculate convergent upper bounds for v^* and ensure the termination of branch-and-bound algorithms for purely inequality constrained problems is presented in [21]. In the context of semi-infinite programming a deterministic upper bounding procedure is derived in [31,32] by exploiting a certain Slater condition.

The idea of feasibility verification in our method is based on a theorem of Miranda in [28] which extends the intermediate value theorem from the one-dimensional to the multidimensional case with several constraints in some sense. Instead of single points, here facets of boxes have to satisfy certain sign conditions to assure the existence of a feasible point. We will show that under mild assumptions and transformations, the conditions of this generalization are guaranteed to be satisfied in a branch-and-bound algorithm after a finite number of steps. Thus, valid upper bounds for v^* can be calculated in a deterministic way and, as the boxes gradually become smaller leading to convergent upper bounds, termination of the branch-and-bound algorithm is assured.

We start by presenting the general spatial branch-and-bound framework for global optimization and its shortcomings with respect to upper bounds for nonconvex problems in Sect. 2. In Sect. 3 we introduce the theorem of Miranda, before we use it to develop our algorithm in Sect. 4. Preliminary numerical results are presented in Sect. 5, and Sect. 6 concludes the paper with some final remarks.

The notation in this paper is standard. The gradient of a function h_j , $j \in J$, is denoted by ∇h_j and the Jacobian of the vector-valued function $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ with $h = (h_1, \dots, h_m)$ by $\nabla h = (\nabla h_1, \dots, \nabla h_m)$ or $Dh = (\nabla h)^\top$. For a box $X = [\underline{x}, \bar{x}] := \{x \in \mathbb{R}^n \mid \underline{x} \leq x \leq \bar{x}\} \subseteq \mathbb{R}^n$ we denote the midpoint by $\text{mid}(X) := (\bar{x} + \underline{x})/2$ and the diagonal length with $\text{diag}(X) := \|\bar{x} - \underline{x}\|_2$.

2 Difficulties in branch-and-bound algorithms

In this section we present a general spatial branch-and-bound framework to solve global minimization problems and then discuss its main difficulties in the presence of nonconvex equality constraints. The framework and our notation are based on [21]. It is primarily designed to approximate the minimal value v^* of problem $P(B)$.

2.1 The spatial branch-and-bound framework

The main principle of the branch-and-bound framework is to branch on the original problem $P(B)$ iteratively, yielding subproblems on sub-boxes. Using these subproblems, convergent upper and lower bounds for the global minimal value v^* are calculated to approximate the globally minimal value v^* . Moreover, the bounds can be used to exclude subproblems from further examination if they cannot contain a globally minimal point of the original problem. In this subsection, we will present the framework in a very generic way. Afterwards, we will discuss procedures to determine lower and upper bounds.

The main steps within the general framework are the following.

Branching In each iteration k the current box X^k is divided into two sub-boxes $X^{k,1}$ and $X^{k,2}$, for example along the midpoint of a longest edge. In the literature also different, and more efficient, branching techniques for spatial branch-and-bound algorithms are proposed, for example in [8] in the setting of mixed-integer variables.

Bounding For both sub-boxes $X^{k,l}, l \in \{1, 2\}$, a lower bounding procedure is used to calculate a lower bound $\ell(X^{k,l})$ for the globally minimal value v_k^l of P restricted to $X^{k,l}$. If the latter problem is infeasible, we formally set $\ell(X^{k,l}) := +\infty$. For boxes $X^{k,l}$ which do not only satisfy $\ell(X^{k,l}) < +\infty$ but also $\ell(X^{k,l}) \leq u_{k-1}$ for some known overall upper bound u_{k-1} of v^* (see below), the pair $(X^{k,l}, \ell(X^{k,l}))$ is added to a list \mathcal{L} of subproblems which are potentially needed to approximate v^* .

Then, some upper bounding procedure is used to determine a valid upper bound \bar{u}_k^l for v^* . The value u_k is defined as the smallest known upper bound for v^* . It is initialized as $u_0 = +\infty$ and updated in each iteration using \bar{u}_k^1 and \bar{u}_k^2 , if possible. Commonly \bar{u}_k^l is determined by evaluating f at some feasible point $x^{k,l} \in M(X^{k,l})$, if such a trial point exists, and the point corresponding to the best known upper bound u_k is stored as the best known feasible point so far.

More generally, in our approach \bar{u}_k^l is determined as an upper bound for f on some box which is only known to contain a point from $M(X^{k,l})$, but the explicit knowledge of this point is not necessarily required. Hence, instead of storing the best known trial point so far, one could store the best known box so far. However, since the proposed algorithm focuses on the approximation of the optimal value v^* , we do not state these possibilities to approximate minimal points in the below pseudo code of Algorithm 1, but we refer to the subsequent discussion for further details.

Fathoming If \mathcal{L} contains an element $(\tilde{X}, \ell(\tilde{X}))$ with $\ell(\tilde{X}) > u_k$, this element can be removed from the list, since the lower bound for the minimal value on the correspond-

ing box is larger than the best known upper bound for v^* . Hence, \tilde{X} cannot contain a globally minimal point.

Update of Lower Bound From all lower bounds $\ell(\tilde{X})$ for the global minimal value of $P(B)$ restricted to boxes \tilde{X} a valid lower bound for v^* can be constructed by $\ell_k := \min \{ \ell(\tilde{X}) \mid (\tilde{X}, \ell(\tilde{X})) \in \mathcal{L} \}$.

Stopping Criterion The algorithm terminates if $u_k - \ell_k \leq \varepsilon$ holds for a prescribed tolerance $\varepsilon > 0$. Otherwise, a new iteration is started in which a box \tilde{X} corresponding to ℓ_k is branched.

Algorithm 1 summarizes the procedure of the branch-and-bound method. Upon termination, the case $\mathcal{L} = \emptyset$ yields a certificate for infeasibility of P . Otherwise, the value v^* lies in the interval $[\ell_k, u_k]$ of length ε . Moreover, if as mentioned above in each iteration the best known feasible point so far, or the best known box, are stored, upon termination the point is ε -minimal or the box contains an ε -minimal point, respectively.

2.2 Bounding and difficulties in presence of nonconvexities

The aim of Algorithm 1 is to solve nonconvex problems of type $P(B)$. The algorithm requires convergent lower and upper bounds for the globally minimal value v^* to terminate. However, as the feasible set of $P(B)$ is nonconvex, difficulties may occur with regard to the upper bounds, as we will show.

Typically, lower bounds at v^* are obtained by special bounding procedures. Among those are procedures based on Lipschitz constants [34], interval arithmetic [33] and its enhancements, such as centered forms [22] or optimal centered forms [7]. Another approach is to obtain lower bounds from convex relaxations. This approach is fol-

Algorithm 1 Conceptual Branch-and-Bound Method

Input: Problem $P(B)$.

- 1: **Initialization:** Choose a tolerance $\varepsilon > 0$. Set the initial lower bound for v^* to $\ell_0 = -\infty$, the upper bound to $u_0 = +\infty$ and the list to $\mathcal{L} = \{(B, \ell_0)\}$. Set the iteration counter to $k = 0$.
 - 2: **repeat**
 - 3: Increase the iteration counter k by one.
 - 4: Choose a pair $(X^k, \ell(X^k))$ from \mathcal{L} with smallest lower bound and remove it from \mathcal{L} .
 - 5: Divide X^k along the midpoint of a longest edge into $X^{k,1}$ and $X^{k,2}$.
 - 6: **for** $l = 1, 2$ **do**
 - 7: Determine a lower bound $\ell(X^{k,l})$ for the minimal value of $P(B)$ on $X^{k,l}$ using some lower bounding procedure, with $\ell(X^{k,l}) = +\infty$ for an infeasible problem.
 - 8: If $\ell(X^{k,l}) < +\infty$ and $\ell(X^{k,l}) \leq u_{k-1}$, add the pair $(X^{k,l}, \ell(X^{k,l}))$ to \mathcal{L} .
 - 9: Determine an upper bound \bar{u}_k^l for v^* using some upper bounding technique based on $X^{k,l}$. If no valid upper bound can be determined set $\bar{u}_k^l = +\infty$.
 - 10: Calculate the best upper bound for v^* as $u_k = \min \{ u_{k-1}, \bar{u}_k^1, \bar{u}_k^2 \}$.
 - 11: Remove all pairs $(\tilde{X}, \ell(\tilde{X}))$ with $\ell(\tilde{X}) > u_k$ from \mathcal{L} .
 - 12: **if** $\mathcal{L} \neq \emptyset$ **then**
 - 13: Update the lower bound for v^* to $\ell_k = \min \{ \ell(\tilde{X}) \mid (\tilde{X}, \ell(\tilde{X})) \in \mathcal{L} \}$.
 - 14: **until** $u_k - \ell_k \leq \varepsilon$ or $\mathcal{L} = \emptyset$.
-

lowed in the α BB-method, which has been originally proposed in [6] and extended in [1–5,26]. Convex relaxations are also used in COUENNE, where symbolic reformulations [27,42], determination of convex underestimators [43] and bound tightening techniques [8] are combined. Moreover, lower bounds can be derived by exploiting duality as proposed in [10,11] or by using piecewise linear approximations of multi-dimensional functions [15,29,35]. For several of those techniques it is proven that the determined lower bounds converge to v^* for decreasing box sizes, as they occur in Algorithm 1.

Upper bounds are commonly obtained by evaluating the objective function f at feasible points of $P(B)$, either explicitly or implicitly by applying local solvers. However, as the problem $P(B)$ is in general nonconvex, it may be hard to determine feasible points. Thus, it cannot be assured that within Algorithm 1, feasible points with sufficiently good objective function value are found and the current upper bound u_k is updated. This also implies that the upper bounds are not guaranteed to converge for decreasing box sizes. Therefore, Algorithm 1 is not guaranteed to terminate.

To cope with this, in the literature sometimes it is suggested to accept ε_f -feasible solutions, that is, for a given tolerance $\varepsilon_f > 0$ points $x \in B$ with

$$\max_{j \in J} |h_j(x)| \leq \varepsilon_f,$$

are considered to be feasible (see for example [13]). The idea behind this is that the small distance between approximately feasible points and the feasible set leads to a small distance in function values as well. However, the function value of an ε_f -feasible solution is not guaranteed to be a valid upper bound for v^* , as exemplarily shown in [21] for purely inequality constrained problems. In fact, for any ε_f examples can be created where the objective value of an ε_f -feasible point is arbitrarily much smaller than v^* . This feasibility tolerance should not be confused with a termination tolerance or a machine precision, since even in the absence of rounding errors and even for a small value of ε_f this approach can lead to arbitrary large errors, depending on the problem at hand.

Despite providing sufficiently good upper bounds *on average* in practical applications, upper bound generation techniques proposed in literature suffer from the above mentioned theoretical drawbacks. The naive approach to obtain upper bounds using the optimal point of the convex relaxation after checking its feasibility, for example proposed in [26], is not guaranteed to work, as such points might never be feasible in the course of the algorithm. For box- and inequality-constrained problems in [21] an example is given, for which the α BB method with such an upper bound procedure does not terminate, since u_0 is never updated.

The current state-of-the-art approach to solve the nonconvex optimization problem $P(B)$ (or some restricted subproblem $P(X)$) locally, as proposed in [6], has several drawbacks: Firstly, it is computationally expensive. Liberti and Maculan even identify it as the most expensive step in current branch-and-bound implementations [24]. Secondly, for nonconvex feasible sets it is not guaranteed that the local solver will eventually find a feasible point of $P(B)$. Thus, usually ε_f -feasible solutions are accepted,

see for example [6,13]. However, as stated before, in such a case it is not ensured that valid upper bounds for the true globally minimal value v^* are determined.

Hence, neither of the two approaches is suitable to guarantee the termination of Algorithm 1. In the following we will present a method which determines a valid upper bound for v^* for a sufficiently small box X based on verifying the existence of a feasible point of the problem $P(B)$. If a feasible point in X is verified, an upper bound for v^* does not have to be determined by evaluating the objective function f at a specific feasible point, but it suffices to apply an upper bounding procedure to the whole box X .

The obtained upper bounds converge for decreasing box sizes and, thus, incorporating our method into Algorithm 1 is sufficient to ensure termination, i.e. convergence to v^* . This can be proven using standard arguments from global optimization [16]. The main idea is that in a branch-and-bound algorithm the boxes eventually become sufficiently small and, hence, upper bounds for v^* can be determined in a finite number of steps. Feasibility verification will also succeed for any sufficiently small box containing a globally minimal point of $P(B)$, and with smaller such boxes the upper bounds for v^* gradually become better and converge to v^* . Contrary to lower bounding procedures, we are not aware of any convergent upper bounding procedure proposed so far.

3 Miranda's theorem

Before presenting our algorithm to calculate convergent upper bounds, we will first introduce the main idea that our approach is based on. As stated before, our algorithm is designed to verify the existence of a feasible point within some sub-box $X \subseteq B$, this means a point $\hat{x} \in X$ satisfying $h_j(\hat{x}) = 0$ for all $j \in J$. If the existence of a feasible point in a sub-box X is verified, an upper bound calculated on this whole box can serve as a valid upper bound for v^* .

The main idea of feasibility verification is based on *Miranda's Theorem*, sometimes also referred to as *Poincaré–Miranda Theorem*. It was first proven by Carlo Miranda in 1940, see [28], and in some sense can be thought of as generalizing Bolzano's intermediate value theorem to the multidimensional case. It states that if not single points but whole facets

$$\bar{X}_i := \{x \in X \mid x_i = \bar{x}_i\}, \quad \underline{X}_i := \{x \in X \mid x_i = \underline{x}_i\}$$

of a box X , with $i = 1, \dots, n$, satisfy certain sign conditions, the existence of a feasible point in X is guaranteed. Formally, this is stated in the following result.

Theorem 3.1 (Miranda's Theorem, from [28]) *Let $X \subseteq \mathbb{R}^n$ be a box and $h = (h_1, \dots, h_n) : X \rightarrow \mathbb{R}^n$ a continuous function. Moreover, let h satisfy the conditions*

$$\begin{aligned} h_i(x) &\leq 0, & \forall x \in \underline{X}_i, \\ h_i(x) &\geq 0, & \forall x \in \bar{X}_i, \end{aligned} \tag{2}$$

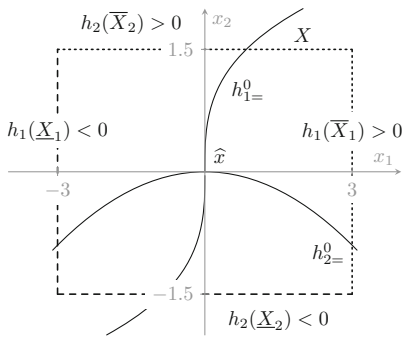
for all $i = 1, \dots, n$. Then $h(x) = 0$ has a solution in X .

Proofs of this theorem can be found in [23,45] among others. The number of facets of a n -dimensional box is $2n$ and, hence, can be iterated over at low computational costs. Figure 1a shows an example where all conditions of Theorem 3.1 are satisfied.

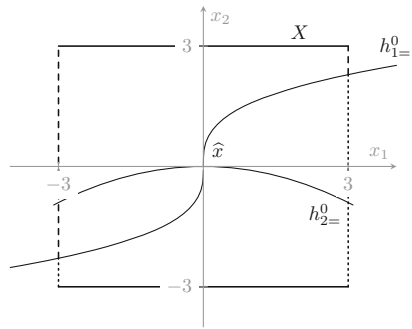
Note that by index changes arbitrary allocations of facet pairs and functions are possible in the conditions (2) and that the sign conditions for the facets could also be reversed to guarantee a zero of h in X . However, later on we will transform the system in such a way that it is reasonable to use the sign conditions as stated above.

Theorem 3.1 can be extended so that it can be applied not only for the same number of equality constraints m and variables n , but more general for $m \leq n$ and, thus, for problems we are interested in.

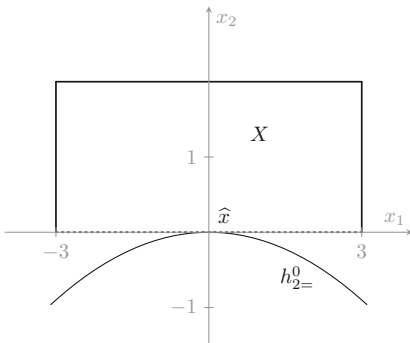
Theorem 3.2 (Extended Miranda’s Theorem) *Let $X \subseteq \mathbb{R}^n$ be a box and $h = (h_1, \dots, h_m) : X \rightarrow \mathbb{R}^m$ a continuous function on X with $m \leq n$. Moreover, let there exist a set of indices $S := \{s_1, \dots, s_m\}$ with $S \subseteq \{1, \dots, n\}$ so that h satisfies*



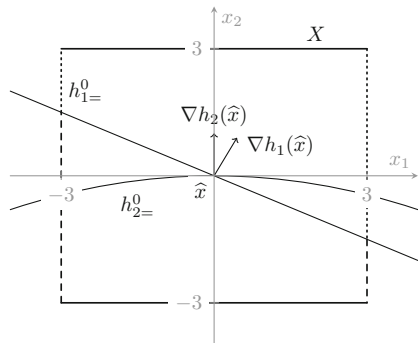
(a) Conditions (3) are satisfied.



(b) Box too large for verification of the existence of a feasible point.



(c) Feasible point on a facet of X .



(d) Example with one gradient in \hat{x} not having unit direction.

Fig. 1 Illustration of Theorem 3.2 and its difficulties

the conditions

$$\begin{aligned} h_j(x) &\leq 0, \quad \forall x \in \underline{X}_{s_j}, \\ h_j(x) &\geq 0, \quad \forall x \in \overline{X}_{s_j}, \end{aligned} \quad (3)$$

for all $j \in J$. Then $h(x) = 0$ has a solution in X .

Proof Without loss of generality we assume that $S = \{1, \dots, m\}$, so that

$$\begin{aligned} h_j(x) &\leq 0, \quad \forall x \in \underline{X}_j, \\ h_j(x) &\geq 0, \quad \forall x \in \overline{X}_j, \end{aligned} \quad (4)$$

holds for all $j \in J$. For arbitrary $\widehat{x}_\ell \in X_\ell$, with $\ell \in \{m+1, \dots, n\}$, we define $\widehat{h} = (\widehat{h}_1, \dots, \widehat{h}_m) : X_1 \times \dots \times X_m \rightarrow \mathbb{R}^m$ with

$$\widehat{h}(x_1, \dots, x_m) := h(x_1, \dots, x_m, \widehat{x}_{m+1}, \dots, \widehat{x}_n).$$

With the conditions in (4) for all $x \in X_1 \times \dots \times X_m$ and $j \in J$ we have

$$\widehat{h}_j(x_1, \dots, x_{j-1}, x_j, x_{j+1}, \dots, x_m) \leq 0, \quad \widehat{h}_j(x_1, \dots, x_{j-1}, \bar{x}_j, x_{j+1}, \dots, x_m) \geq 0,$$

so that with Theorem 3.1 it follows that \widehat{h} has a zero $(\widehat{x}_1, \dots, \widehat{x}_m)$ in $X_1 \times \dots \times X_m$.

From the definition of \widehat{h} it follows $h(\widehat{x}_1, \dots, \widehat{x}_m, \widehat{x}_{m+1}, \dots, \widehat{x}_n) = 0$ and since $(\widehat{x}_1, \dots, \widehat{x}_n) \in X$, the function h has a zero in X . \square

Theorem 3.2 provides a sufficient, but not necessary condition for the existence of feasible points in X . One can think of simple examples of feasible points within boxes for which the conditions (3) are not satisfied. Additionally, even if they are, it is not clear in advance which function h_j , $j \in J$, does satisfy the sign conditions on which pair of facets $\{\underline{X}_j, \overline{X}_j\}$, $i = 1, \dots, n$. Therefore it is not possible to exclude facets from examination or fix variables without excluding possible feasible points from the search space.

For an upper bounding procedure that ensures termination of spatial branch-and-bound algorithms in global optimization we shall overcome these difficulties in the remaining part of this article. We will start by discussing different scenarios, in which the conditions (3) are not satisfied even though a feasible point exists in X .

Case 1 (*The box X is too large*) Despite containing a feasible point the conditions (3) in Theorem 3.2 might not be satisfied for a box X if it is too large, as illustrated exemplarily in Fig. 1b. While for h_2 the conditions are satisfied, they are not for h_1 , even though the point $\widehat{x} = (0, 0) \in X$ is feasible. However, if the box is shrunk sufficiently around \widehat{x} , the existence of a feasible point can be verified by using Theorem 3.2.

Case 2 (*Feasible point on a facet of X*) Further difficulties might occur if a feasible point $\widehat{x} \in X$ is located on a facet of X . In this case, the conditions (3) might still be satisfied for sufficiently small boxes, but this is not guaranteed, as shown in Fig. 1c.

The function h_2 takes solely positive values on \underline{X}_2 . Even if the box X is shrunk with \hat{x} remaining feasible, h_2 will never take negative values on this facet. In Sect. 4.3 we will present an approach which makes it still possible to verify the existence of feasible points using Theorem 3.2 in such a case.

Case 3 (*Gradients in feasible point not directed towards the facets*) The conditions (3) in Theorem 3.2 might also not be satisfied if the gradients $\nabla h_j(\hat{x})$, $j \in J$, in a feasible point \hat{x} have no unit direction and, hence, are not directed towards the facets \bar{X}_j . This is shown in Fig. 1d for the function h_1 . It is not guaranteed that this difficulty vanishes for any type of smaller boxes around \hat{x} . As within a branch-and-bound algorithm usually the rule to construct new boxes is predetermined, this case requires a more specific approach, which we will present in Sect. 4.4.

4 Convergent upper bounds for the globally minimal value

In this section we will develop the announced algorithm to calculate convergent valid upper bounds for the globally minimal value v^* of $P(B)$ based on Theorem 3.2.

A straightforward idea might be to consider every box that occurs during the branch-and-bound algorithm and to check whether the conditions of Theorem 3.2 are fulfilled. If this is the case, a valid upper bound at the globally minimal value could be obtained by computing an upper bound at the objective function over the whole box, for instance by applying interval arithmetic or any other convergent bounding procedure.

However, as already discussed at the end of the previous section, in Theorem 3.2 only a sufficient, but not necessary condition for verifying the existence of a feasible point is provided. Therefore, this approach does not necessarily need to lead to a convergent procedure. In this section, we will show how the boxes obtained by standard branch-and-bound algorithms can be altered so that the conditions are fulfilled, provided a sufficiently small box contains a feasible point.

In order to achieve this we start by considering an artificial and very simple case where

- the box midpoint equals the feasible point \hat{x} ,
- the gradient $\nabla h_j(\hat{x})$ is the j -th unit vector for all $j \in J$ and
- the box X is sufficiently small.

Then it can be ensured that the conditions of Theorem 3.2 are satisfied, as we will prove in Sect. 4.1. Afterwards in Sects. 4.2, 4.3 and 4.4, we will sequentially generalize our results to more realistic sequences of boxes that occur in spatial branch-and-bound algorithms by addressing the difficulties mentioned in Sect. 3, and finally state our algorithm for the general case.

4.1 Feasibility verification in a simple case

In this subsection we consider the simple case where the gradient of h_j , $j \in J$, in a feasible point \hat{x} equals the j -th unit vector e_j , that is, $\nabla h_j(\hat{x}) = e_j$ for all $j \in J$. Moreover, we assume the midpoint $\text{mid}(X)$ of a box X to be feasible, so we have $\hat{x} = \text{mid}(X)$ with $h_j(\hat{x}) = 0$ for all $j \in J$.

We will show that in this specific case, for a sufficiently small box X , the conditions of Theorem 3.2 are guaranteed to be satisfied and hence, the existence of a feasible point can be verified. The intuition behind this is that starting from \widehat{x} the function h_j will take positive values in direction e_j and any direction towards \overline{X}_j for sufficiently small steps and for all $j \in J$. On the other hand, h_j will take negative values in direction $-e_j$ and any other direction towards \underline{X}_j for sufficiently small steps.

To derive the above result, we will analyze *exhaustive sequences of boxes*, that is, sequences $(X^k)_{k \in \mathbb{N}}$ of non-empty boxes $X^k \subseteq B$ with $X^{k+1} \subseteq X^k$ (*nestedness*) and $\lim_{k \rightarrow \infty} \text{diag}(X^k) = 0$. Additionally, we assume that the feasible point \widehat{x} is midpoint of each element X^k . We will show that there is a $\widehat{k} \in \mathbb{N}$ so that for all $k \geq \widehat{k}$ Miranda’s conditions are satisfied. To make this work, in addition, we have to ensure that boxes do not become increasingly deformed. Therefore we will only consider *non-deformed* exhaustive sequences of boxes, i.e. sequences for which the maximum ratio of the length of box edges is bounded above by a constant $\bar{t} < \infty$ for all boxes X^k .

Definition 4.1 (*Non-Deformed Sequence of Boxes*) Let $(X^k)_{k \in \mathbb{N}}$ be a sequence of boxes with the maximum ratio of the length of box edges

$$t^k := \frac{\max_{i=1, \dots, n} (\bar{x}_i^k - \underline{x}_i^k)}{\min_{i=1, \dots, n} (\bar{x}_i^k - \underline{x}_i^k)},$$

bounded above by a constant $\bar{t} < \infty$. Then we call $(X^k)_{k \in \mathbb{N}}$ a *non-deformed sequence of boxes*.

The following lemma shows that this additional requirement does not exclude the typical branching rule of standard branch-and-bound algorithms in global optimization to divide a box along a longest edge. We believe that this may also be shown for other branching rules along the lines of this proof.

Lemma 4.2 Let $(X^k)_{k \in \mathbb{N}}$ be a sequence of boxes where X^{k+1} is constructed by dividing X^k along a longest edge. Then $(X^k)_{k \in \mathbb{N}}$ is non-deformed.

Proof For simplicity we denote the lengths of the edges $i = 1, \dots, n$ by $\Delta x_i^k := \bar{x}_i^k - \underline{x}_i^k$. For $n = 1$ the ratio t^k is trivially 1 and thus, bounded above for all $k \in \mathbb{N}$. For $n > 1$ we distinguish two cases:

First, let $t^k > 2$. As a longest edge is bisected we obtain

$$t^{k+1} = \frac{\max_{i=1, \dots, n} \Delta x_i^{k+1}}{\min_{i=1, \dots, n} \Delta x_i^{k+1}} \leq \frac{\max_{i=1, \dots, n} \Delta x_i^k}{\min_{i=1, \dots, n} \Delta x_i^{k+1}} \stackrel{t^k \geq 2}{=} \frac{\max_{i=1, \dots, n} \Delta x_i^k}{\min_{i=1, \dots, n} \Delta x_i^k} = t^k.$$

Now, let $t^k \leq 2$. Then, after dividing a longest edge the length of the shortest edge is exactly half the length of the previously longest edge. That is, we have $\min_{i=1, \dots, n} \Delta x_i^{k+1} = \frac{1}{2} \cdot \max_{i=1, \dots, n} \Delta x_i^k$. With this we obtain

$$t^{k+1} = \frac{\max_{i=1, \dots, n} \Delta x_i^{k+1}}{\min_{i=1, \dots, n} \Delta x_i^{k+1}} = 2 \cdot \frac{\max_{i=1, \dots, n} \Delta x_i^{k+1}}{\max_{i=1, \dots, n} \Delta x_i^k} \leq 2 \cdot \frac{\max_{i=1, \dots, n} \Delta x_i^k}{\max_{i=1, \dots, n} \Delta x_i^k} = 2,$$

where the second equality follows from the bisection of a longest edge.

Summarizing, the ratio t^k is bounded above by $\bar{t} := \max \{t^1, 2\} < \infty$ for all $k \in \mathbb{N}$. □

For non-deformed exhaustive sequences of boxes we can now show that there is a $\hat{k} \in \mathbb{N}$ so that for all $k \geq \hat{k}$ the conditions of Theorem 3.2 are satisfied. This is stated formally in the following result.

Theorem 4.3 *Let $(X^k)_{k \in \mathbb{N}}$ be a non-deformed exhaustive sequence of boxes with $\hat{x} := \text{mid}(X^k)$ for all $k \in \mathbb{N}$ and \hat{x} feasible. Furthermore, let $h : \mathbb{R}^n \mapsto \mathbb{R}^m$, $h \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ and let the gradients satisfy the conditions $\nabla h_j(\hat{x}) = e_j$ for all $j \in J$. Then, there is some $\hat{k} \in \mathbb{N}$ so that the conditions (3) of Theorem 3.2 are satisfied for all $k \in \mathbb{N}$ with $k \geq \hat{k}$.*

Proof Consider an arbitrary element $X^k \in (X^k)_{k \in \mathbb{N}}$. Without loss of generality we examine constraint h_1 . Let \tilde{x}^k be a maximum point of h_1 on \underline{X}_1^k . Using Taylor’s Theorem with $\lim_{\tilde{x}^k \rightarrow \hat{x}} \omega(\tilde{x}^k) = \omega(\hat{x}) = 0$ we obtain

$$h_1(\tilde{x}^k) = h_1(\hat{x}) + \left\langle \nabla h_1(\hat{x}), \tilde{x}^k - \hat{x} \right\rangle + \omega(\tilde{x}^k) \cdot \|\tilde{x}^k - \hat{x}\|.$$

With the feasibility of \hat{x} , $\nabla h_1 = e_1$ and $\tilde{x}^k \in \underline{X}_1^k$ it follows

$$h_1(\tilde{x}^k) = -\frac{1}{2}(\tilde{x}_1^k - \underline{x}_1^k) + \omega(\tilde{x}^k) \cdot \|\tilde{x}^k - \hat{x}\|.$$

With the box property $\|x - \text{mid}(X)\|_2 \leq \frac{1}{2} \text{diag}(X)$ for all $x \in X$ we obtain

$$h_1(\tilde{x}^k) \leq -\frac{1}{2}(\tilde{x}_1^k - \underline{x}_1^k) + \frac{1}{2}\omega(\tilde{x}^k) \text{diag}(X^k).$$

Using the equivalence of norms and that $(X^k)_{k \in \mathbb{N}}$ is non-deformed we have

$$\begin{aligned} \tilde{x}_1^k - \underline{x}_1^k &\geq \min_{i=1, \dots, n} (\tilde{x}_i^k - \underline{x}_i^k) = \frac{1}{t^k} \max_{i=1, \dots, n} (\tilde{x}_i^k - \underline{x}_i^k) = \frac{1}{t^k} \|\tilde{x}^k - \underline{x}^k\|_\infty \\ &\geq \frac{1}{t^k} \frac{\text{diag}(X^k)}{\sqrt{n}} \geq \frac{1}{\bar{t}} \frac{1}{\sqrt{n}} \text{diag}(X^k) \end{aligned}$$

and hence,

$$h_1(\tilde{x}^k) \leq -\frac{1}{2\sqrt{n}\bar{t}} \text{diag}(X^k) + \frac{1}{2}\omega(\tilde{x}^k) \text{diag}(X^k).$$

Since for k approaching infinity we have $\lim_{k \rightarrow \infty} \text{diag}(X^k) = 0$ but also $\tilde{x}^k \rightarrow \hat{x}$ and hence $\omega(\tilde{x}^k) \rightarrow \omega(\hat{x}) = 0$, the first term on the right hand side determines the sign

of the whole expression. With $n, \bar{t} > 0$, it is always negative, so that there is some $k_1^1 \in \mathbb{N}$ so that for all $k \geq k_1^1$ and all $x \in \underline{X}_1^k$ we have

$$h_1(x) \leq \max_{x \in \underline{X}_1^k} h_1(x) = h_1(\check{x}^k) < 0.$$

Analogously, with \check{x}^k the minimum point of h_1 on \overline{X}_1^k there is some $k_1^2 \in \mathbb{N}$ so that for all $k \geq k_1^2$ and all $x \in \overline{X}_1^k$ we have

$$h_1(x) \geq \min_{x \in \overline{X}_1^k} h_1(x) = h_1(\check{x}^k) > 0.$$

We proceed analogously for the equality constraints h_2, \dots, h_m . Thus, for all $k \geq \widehat{k} := \max_{\ell=1,2} \max_{j=1,\dots,m} \{k_j^\ell\}$ the assertion follows. \square

Consequently, for this specific case an algorithm based on Theorem 3.2 is able to verify the existence of a feasible point within some box X^k and to determine a valid upper bound for v^* .

4.2 A first extension of the simple case

In Sect. 4.1 we assume the midpoint $\text{mid}(X^k)$ of each box X^k in $(X^k)_{k \in \mathbb{N}}$ to be feasible. Now we relax this assumption and generalize our results to exhaustive sequences of boxes for which the feasible point \widehat{x} only becomes midpoint asymptotically, but still $\nabla h_j(\widehat{x}) = e_j$ holds for all $j \in J$. We will show that for sufficiently small boxes, i.e. after a finite number of steps within a branch-and-bound algorithm, still the conditions (3) in Theorem 3.2 are guaranteed to be satisfied. To begin with, we introduce the concept of *asymptotically centered sequences of boxes*.

Definition 4.4 (*Asymptotically Centered Sequence of Boxes*) Let $(X^k)_{k \in \mathbb{N}}$ be a sequence of boxes with $\widehat{x} \in X^k$ for all $k \in \mathbb{N}$, and

$$\lim_{k \rightarrow \infty} \frac{\widehat{x}_i - x_i^k}{\bar{x}_i^k - x_i^k} = \frac{1}{2}, \quad i = 1, \dots, n. \quad (5)$$

Then, we call the sequence $(X^k)_{k \in \mathbb{N}}$ *asymptotically centered on \widehat{x}* .

Definition 4.4 states that a sequence of boxes is asymptotically centered on a point \widehat{x} if this point asymptotically becomes the midpoint of elements of the sequence. Note that for this property it is not sufficient that the midpoint $\text{mid}(X^k)$ of X^k solely converges to \widehat{x} for k approaching infinity.

Although \widehat{x} only becomes midpoint of X^k asymptotically, the following theorem proves that for a non-deformed and exhaustive sequence of boxes $(X^k)_{k \in \mathbb{N}}$, which is asymptotically centered on a feasible point $\widehat{x} \in X^k$ for all $k \in \mathbb{N}$, there still is a $\widehat{k} \in \mathbb{N}$ so that for all $k \geq \widehat{k}$ the conditions of Theorem 3.2 are satisfied.

Theorem 4.5 *Let $(X^k)_{k \in \mathbb{N}}$ be a non-deformed sequence of boxes asymptotically centered on a feasible point \hat{x} of $P(B)$ and with $\lim_{k \rightarrow \infty} \text{diag}(X^k) = 0$. Furthermore, let $h : \mathbb{R}^n \mapsto \mathbb{R}^m$, $h \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ and the gradients satisfy the conditions $\nabla h_j(\hat{x}) = e_j$ for all $j \in J$. Then, there is some $\hat{k} \in \mathbb{N}$ so that the conditions (3) of Theorem 3.2 are satisfied for all $k \in \mathbb{N}$ with $k \geq \hat{k}$.*

Proof We follow the same approach as in the proof of Theorem 4.3 and obtain

$$h_1(\tilde{x}^k) = -(\hat{x}_1 - \tilde{x}_1^k) + \frac{1}{2}\omega(\tilde{x}^k) \text{diag}(X^k).$$

Introducing the ratio

$$s(X^k) := \frac{\hat{x}_1 - \tilde{x}_1^k}{\tilde{x}_1^k - \underline{x}_1^k} = \frac{\hat{x}_1 - \underline{x}_1^k}{\tilde{x}_1^k - \underline{x}_1^k} \in [0, 1]$$

we can express the first term in terms of a box edge. Then, using norm properties and the non-deformity of $(X^k)_{k \in \mathbb{N}}$ it follows

$$h_1(\tilde{x}^k) \leq -\frac{s(X^k)}{\sqrt{n} \, \bar{t}} \text{diag}(X^k) + \frac{1}{2}\omega(\tilde{x}^k) \text{diag}(X^k).$$

Since we assume X^k to be an element of a sequence of boxes asymptotically centered on \hat{x} , $s(X^k)$ converges to $\frac{1}{2}$ and thus, $-\frac{s(X^k)}{\sqrt{n} \, \bar{t}}$ converges to a non-zero constant when $\text{diag}(X^k)$ approaches 0. Hence, there is some $k_1^1 \in \mathbb{N}$, so that for all $k \geq k_1^1$, $k \in \mathbb{N}$, and for all $x \in \underline{X}_1^k$ we have

$$h_1(x) \leq \max_{x \in \underline{X}_1^k} h_1(x) = h_1(\tilde{x}^k) < 0.$$

For \bar{X}_1^k and h_2, \dots, h_m we proceed analogously. With the same line of argumentation as for Theorem 4.3 the assertion follows. □

4.3 A further extension of the simple case

In Sect. 3 we demonstrated that the conditions (3) of Theorem 3.2 may not be satisfied if a feasible point is located on a facet of X . In this subsection we will address this challenge and present an approach to guarantee that upper bounds to v^* can be determined in a finite number of steps within a branch-and-bound algorithm.

The main idea to handle such situations is to artificially extend a box X to a box \tilde{X} so that all feasible points in X are interior points of \tilde{X} . Thereby, for each exhaustive sequence of boxes $(X^k)_{k \in \mathbb{N}}$ with $\hat{x} \in X^k$ we obtain an additional sequence $(\tilde{X}^k)_{k \in \mathbb{N}}$ with $X^k \subseteq \tilde{X}^k$. This sequence is not guaranteed to be nested any longer, but this is no restriction to our results. Additionally, we construct this sequence such that it is

asymptotically centered on \widehat{x} . Then, using Theorem 3.2 the existence of a feasible point in some box \widetilde{X}^k can be verified, as shown in the previous subsection. Generally, the verification of a feasible point in \widetilde{X}^k does not guarantee the existence of a feasible point in X^k . However, we will demonstrate that it is sufficient to determine valid upper bounds for v^* .

The following lemma provides a method to construct a sequence with the desired properties for boxes with edge lengths smaller than 1. This is sufficient for our purposes, as we are only interested in convergence properties for small boxes.

Lemma 4.6 (Construction of a Centered Sequence of Super-Boxes) *Let $(X^k)_{k \in \mathbb{N}}$ be an exhaustive sequence of boxes with $\bar{x}_i^k - \underline{x}_i^k < 1$ for all $i = 1, \dots, n$ and $\widehat{x} \in X^k$ for all $k \in \mathbb{N}$. Furthermore, let $(\widetilde{X}^k)_{k \in \mathbb{N}}$ be a sequence of boxes with $\widetilde{X}^k = [\underline{x}^k, \bar{x}^k]$, $\text{mid}(\widetilde{X}^k) = \text{mid}(X^k)$ and $\widetilde{x}_i^k - \underline{x}_i^k := \sqrt{\bar{x}_i^k - \underline{x}_i^k}$ for all $i = 1, \dots, n$ and $k \in \mathbb{N}$. Then, $X^k \subseteq \widetilde{X}^k$ for each $k \in \mathbb{N}$ and the sequence $(\widetilde{X}^k)_{k \in \mathbb{N}}$ is centered on \widehat{x} .*

Proof With $\text{mid}(\widetilde{X}^k) = \text{mid}(X^k)$, $\widetilde{x}_i^k - \underline{x}_i^k = \sqrt{\bar{x}_i^k - \underline{x}_i^k}$ and $\bar{x}_i^k - \underline{x}_i^k < 1$ we have $X^k \subseteq \widetilde{X}^k$ for all $k \in \mathbb{N}$. To prove the second assertion, we will show that $\lim_{k \rightarrow \infty} \frac{\widehat{x}_i - \underline{x}_i^k}{\widetilde{x}_i^k - \underline{x}_i^k} = \frac{1}{2}$ for an arbitrary $i \in \{1, \dots, n\}$.

We start with the case $\widehat{x}_i \geq \text{mid}(X_i^k)$. Using the definition of \widetilde{X}^k and basic box properties it follows

$$\begin{aligned} s(\widetilde{X}^k) &:= \frac{\widehat{x}_i - \underline{x}_i^k}{\widetilde{x}_i^k - \underline{x}_i^k} = \frac{\widehat{x}_i - \text{mid}(X_i^k) + \text{mid}(X_i^k) - \underline{x}_i^k}{\widetilde{x}_i^k - \underline{x}_i^k} = \frac{\widehat{x}_i - \text{mid}(X_i^k)}{\widetilde{x}_i^k - \underline{x}_i^k} + \frac{1}{2} \\ &\leq \frac{1}{2} \frac{\bar{x}_i^k - \underline{x}_i^k}{\widetilde{x}_i^k - \underline{x}_i^k} + \frac{1}{2} = \frac{1}{2} \frac{\bar{x}_i^k - \underline{x}_i^k}{\sqrt{\bar{x}_i^k - \underline{x}_i^k}} + \frac{1}{2} = \frac{1}{2} \sqrt{\bar{x}_i^k - \underline{x}_i^k} + \frac{1}{2}. \end{aligned}$$

For an exhaustive sequence of boxes $(X^k)_{k \in \mathbb{N}}$ we have $\lim_{k \rightarrow \infty} \bar{x}^k - \underline{x}^k = 0$. Hence, we obtain $\lim_{k \rightarrow \infty} s(\widetilde{X}^k) \leq \frac{1}{2}$. Moreover, with $\widehat{x}_i \geq \text{mid}(X_i^k)$ we have

$$s(\widetilde{X}^k) \geq \frac{\text{mid}(X_i^k) - \underline{x}_i^k}{\widetilde{x}_i^k - \underline{x}_i^k} = \frac{1}{2},$$

and hence the assertion.

For the case $\widehat{x}_i < \text{mid}(X_i^k)$ with the same line of argumentation we obtain $s(\widetilde{X}^k) \geq -\frac{1}{2} \sqrt{\bar{x}_i^k - \underline{x}_i^k} + \frac{1}{2}$ and

$$s(\widetilde{X}^k) \leq \frac{\text{mid}(X_i^k) - \underline{x}_i^k}{\widetilde{x}_i^k - \underline{x}_i^k} = \frac{1}{2},$$

from which the assertion follows for k approaching infinity. □

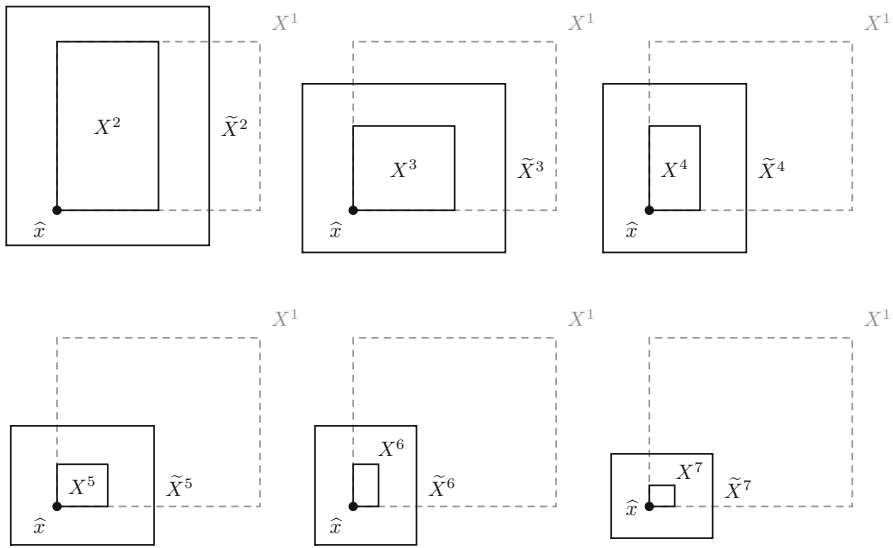


Fig. 2 A sequence of boxes $(X^k)_{k \in \mathbb{N}}$ and the sequence $(\tilde{X}^k)_{k \in \mathbb{N}}$ constructed using Lemma 4.6. The latter is asymptotically centered on \hat{x}

Figure 2 illustrates the proven result for an exemplary sequence of boxes X^k where consecutive boxes are constructed by division along a longest edge.

Using this construction we obtain the following result with regard to the verification of the existence of feasible points:

Corollary 4.7 *Let $(X^k)_{k \in \mathbb{N}}$ be a non-deformed exhaustive sequence of boxes with $\bar{x}_i^k - \underline{x}_i^k < 1$ for all $i = 1, \dots, n$ and $k \in \mathbb{N}$, $\hat{x} \in X^k$ for all $k \in \mathbb{N}$ and \hat{x} feasible. Furthermore, let $h : \mathbb{R}^n \mapsto \mathbb{R}^m$, $h \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ and the gradients satisfy the conditions $\nabla h_j(\hat{x}) = e_j$ for all $j \in J$. Then, there is some $\hat{k} \in \mathbb{N}$ so that for all $k \in \mathbb{N}$ with $k \geq \hat{k}$ the conditions (3) of Theorem 3.2 are satisfied for the boxes \tilde{X}^k constructed as in Lemma 4.6.*

Proof With Lemma 4.6 it follows that the constructed sequence of boxes $(\tilde{X}^k)_{k \in \mathbb{N}}$ is asymptotically centered on \hat{x} . For a sufficiently small box X^k also \tilde{X}^k becomes sufficiently small, since $\text{diag}(X^k) \rightarrow 0$ implies $\text{diag}(\tilde{X}^k) \rightarrow 0$. Consequently, there is some \hat{k} so that all assumptions of Theorem 4.5 are satisfied and the assertion follows. \square

Therefore, if X^k contains a feasible point and X^k is sufficiently small, it is guaranteed that with the construction of \tilde{X}^k the existence of this point can be verified using Theorem 3.2, as $X^k \subseteq \tilde{X}^k$. In a branch-and-bound algorithm with additional construction of boxes \tilde{X}^k this is the case after a finite number of steps. On the contrary, it is also possible that the conditions (3) in Theorem 3.2 are satisfied due to a point $\hat{x} \in \tilde{X}^k$ with $h(\hat{x}) = 0$ but $\hat{x} \notin X^k$ or even $\hat{x} \notin B$. We will discuss this aspect in detail later on and show that it is no strong restriction to our approach.

4.4 The general case

In this subsection we further generalize our results from the previous subsections by dropping the assumption $\nabla h_j(\hat{x}) = e_j$ for all $j \in J$. In this case it is possible, but not guaranteed that the conditions (3) of Theorem 3.2 are satisfied for sufficiently small boxes. The main idea to still guarantee the verification of the existence of feasible points based on this theorem, is to transform the system in such a way that at least asymptotically all conditions from Corollary 4.7 are satisfied. In particular, that all gradients of constraints in a feasible point have unit direction at least asymptotically.

First, we will introduce our transformation approach in Sect. 4.4.1, afterwards in Sect. 4.4.2 we will present our algorithm and prove that within a branch-and-bound context it determines valid upper bounds for v^* after a finite number of steps.

4.4.1 Transformation of the system

As before, we will first construct a sequence of boxes $(\tilde{X}^k)_{k \in \mathbb{N}}$ with midpoints $\text{mid}(X^k)$ which is asymptotically centered on \hat{x} . Then, to transform the system in the desired manner, for each $k \in \mathbb{N}$ we will apply a coordinate transformation from the standard basis of \mathbb{R}^n to another basis \mathcal{B}_k .

Such a coordinate transformation can be described with a matrix A_k whose columns are the basis vectors of \mathcal{B}_k . Then, we obtain $x \mapsto y^k = A_k^{-1}x$ for the transformation of a point x in the original coordinates to a point y in the new coordinates and $y^k \mapsto x = A_k y^k$ for the reverse transformation from the new to the original coordinates. In the following, y^k will always denote the image of a point x under A_k^{-1} and \hat{y}^k that of a feasible point \hat{x} . The functions $h_j, j \in J$, can be expressed in the new coordinates by

$$h_j(x) = h_j(A_k y^k) = (h_j \circ A_k)(y^k) =: \tilde{h}_j^k(y^k). \tag{6}$$

Using the chain rule it follows $\nabla \tilde{h}_j^k(\hat{y}^k) = A_k^T \nabla h_j(\hat{x})$ for the gradient of \tilde{h}_j^k in the transformed feasible point \hat{y}^k .

Our aim is to determine A_k in such a way that the gradients of \tilde{h}_j^k in the new coordinates have unit direction at a feasible point \hat{y}^k . As the feasible points and their gradients are not yet known, we will use the box midpoints instead. With $(\tilde{X}^k)_{k \in \mathbb{N}}$ asymptotically centered on \hat{x} , asymptotically the feasible point \hat{x} becomes box midpoint and the coordinate transformation will be conducted in \hat{x} . This will prove to be sufficient to verify the existence of a feasible point in a finite number of steps in a branch-and-bound algorithm with additional construction of boxes \tilde{X}^k .

We want to choose the (n, n) -matrix A_k^T such that

$$\begin{aligned} A_k^T \left(\nabla h_1(\text{mid}(X^k)), \dots, \nabla h_m(\text{mid}(X^k)) \right) &= (e_1, \dots, e_m) \\ \Leftrightarrow A_k^T \nabla h(\text{mid}(X^k)) &= I_{n,m} \\ \Leftrightarrow Dh(\text{mid}(X^k)) A_k &= I_{m,n}. \end{aligned}$$

If $Dh(\text{mid}(X^k))$ is quadratic and the gradients $\nabla h_j(\text{mid}(X^k))$ are linearly independent, using the inverse of $Dh(\text{mid}(X^k))$ we obtain

$$A_k = (Dh(\text{mid}(X^k)))^{-1}, \quad A_k^{-1} = Dh(\text{mid}(X^k)).$$

If $Dh(\text{mid}(X^k))$ is not quadratic, since $m < n$, we will extend the matrix to a basis of \mathbb{R}^n by adding a basis of its null space $\text{null}(Dh(\text{mid}(X^k)))$ and then proceed analogously. Denoting the extended matrix with $\overline{Dh(m(X^k))}$ we obtain

$$A_k = (\overline{Dh(\widehat{x})})^{-1} \quad A_k^{-1} = \overline{Dh(\widehat{x})}.$$

With Assumption 1.1, linearly independent gradients are guaranteed asymptotically for sequences of boxes asymptotically centered on a globally minimal point x^* of $P(B)$. Moreover, as the following lemma shows, this condition is also satisfied after a finite number of steps within a branch-and-bound algorithm with additional construction of boxes \widetilde{X}^k .

Lemma 4.8 *Let $(\widetilde{X}^k)_{k \in \mathbb{N}}$ be a sequence of boxes centered on a globally minimal point x^* of $P(B)$ and with $\lim_{k \rightarrow \infty} \text{diag}(\widetilde{X}^k) = 0$, and let Assumption 1.1 hold. Then, there is a $\widetilde{k} \in \mathbb{N}$ so that for all $k \geq \widetilde{k}$ the coordinate transformation using $A_k = (\overline{Dh(\widehat{x})})^{-1}$ can be applied in $\text{mid}(X^k)$.*

Proof With Assumption 1.1 we have linearly independent gradients of $h_j, j \in J$ in x^* , and hence, for sequences $(\widetilde{X}^k)_{k \in \mathbb{N}}$ asymptotically. Additionally, all functions $h_j, j \in J$, are presumed to be continuously differentiable, so linear independence will also be guaranteed for midpoints sufficiently close to x^* . Within a branch-and-bound algorithm such midpoints occur after a finite number of steps. \square

Based on this lemma, in the following we will restrict our results to sequences of boxes asymptotically centered on globally minimal points although they also hold for any sequence asymptotically centered on a feasible point \widehat{x} with linearly independent gradients $\nabla h_j(\widehat{x})$.

As intended, with the above transformation asymptotically we obtain unit directed gradients in the transformed globally minimal point y^{*k} , as Lemma 4.9 shows. In its proof we shall need that the matrices $A_k^{-1} = \overline{Dh(\text{mid}(X^k))}, k \in \mathbb{N}$, form a convergent sequence. While this is clear in the case $m = n$, for $m < n$ the bases of the null spaces of the matrices $Dh(\text{mid}(X^k)), k \in \mathbb{N}$, need to be chosen carefully. One such possibility is to collect m linearly independent columns of $Dh(\text{mid}(X^k))$ in a matrix R^k and write $Dh(\text{mid}(X^k)) = (R^k, S^k)$ where, without loss of generality, we assume that one can choose the first m columns. Then the columns of the matrices $\begin{pmatrix} -(R^k)^{-1} S^k \\ I_{n-m, n-m} \end{pmatrix}$ form convergent bases of the null spaces of $Dh(\text{mid}(X^k)), k \in \mathbb{N}$.

Lemma 4.9 *Let $(\tilde{X}^k)_{k \in \mathbb{N}}$ be a sequence of boxes centered on a globally minimal point x^* and with $\lim_{k \rightarrow \infty} \text{diag}(\tilde{X}^k) = 0$, and let Assumption 1.1 hold. Then for all $j \in J$ we have*

$$\lim_{k \rightarrow \infty} \nabla \tilde{h}_j^k(\text{mid}(\tilde{Y}^k)) = \nabla \tilde{h}_j^{x^*}(y^{*k}) = e_j,$$

where $\tilde{h}_j^{x^*}$ refers to a transformation conducted in x^* .

Proof Since $(\tilde{X}^k)_{k \in \mathbb{N}}$ is centered on x^* we have $\lim_{k \rightarrow \infty} \text{mid}(X^k) = x^*$. With Lemma 4.8 there is a $\tilde{k} \in \mathbb{N}$ so that for all $k \geq \tilde{k}$ the coordinate transformation is applicable. For all such k , with $A_k^{-1} = Dh(\text{mid}(X^k))$ and h continuously differentiable we obtain $\lim_{k \rightarrow \infty} A_k^{-1} = A_{x^*}^{-1}$, where A_{x^*} refers to a transformation conducted in x^* . By construction of the transformation we have

$$\nabla \tilde{h}_j^k(\text{mid}(\tilde{Y}^k)) = A_k^T \nabla h_j(\text{mid}(X^k)) = e_j.$$

Since inverting a matrix is a continuous transformation and h is continuously differentiable, introducing limits for all $j \in J$ it follows

$$\begin{aligned} \lim_{k \rightarrow \infty} \nabla \tilde{h}_j^k(\text{mid}(\tilde{Y}^k)) &= \lim_{k \rightarrow \infty} A_k^T \lim_{k \rightarrow \infty} \nabla h(\text{mid}(X^k)) \\ &= A_{x^*}^T \nabla h(x^*) = \nabla \tilde{h}_j^{x^*}(y^{*k}) = e_j. \end{aligned}$$

□

So applying the coordinate transformation we obtain a sequence of functions $(\tilde{h}^k)_{k \in \mathbb{N}}$ whose elements at least asymptotically satisfy the conditions from Corollary 4.7. Yet, the box \tilde{X}^k is not necessarily transformed to another box, but to a n -parallelepiped \tilde{Y}^k . To cope with that, for each $k \in \mathbb{N}$ we will overestimate \tilde{Y}^k with a box \tilde{Y}^{*k} and then prove that all conditions from Corollary 4.7 are fulfilled.

To overestimate a n -parallelepiped \tilde{Y}^k with a box we will construct its minimum bounding box \tilde{Y}^{*k} by determining the maximum and minimum entries for each component $i = 1, \dots, n$ among all vertices of \tilde{Y}^k and then taking them as component-wise interval boundaries. Iterating over all vertices of \tilde{Y}^k is expensive from a computational point of view, but the maxima and minima can also be determined by solving $2n$ linear optimization problems.

To derive these linear optimization problems and some following results, we will make use of the edge vectors of a box X^k , which we define by $d_i^{X^k} := (0, \dots, \bar{x}_i^k - \underline{x}_i^k, \dots, 0)^T$ for all $i = 1, \dots, n$, with edge length $\|d_i^{X^k}\|_2 = \bar{x}_i^k - \underline{x}_i^k$. Any point \tilde{x} in X^k can be addressed by

$$\tilde{x} = \underline{x} + \sum_{i=1}^n \alpha_i d_i^{X^k}$$

with appropriate $\alpha_i \in [0, 1]$ for all $i = 1, \dots, n$. Analogously, we denote the edge vectors of an n -parallelepiped \tilde{Y}^k by $d_i^{\tilde{Y}^k} := A_k^{-1} d_i^{X^k}$. Any point \tilde{y}_k within the n -parallelepiped can be addressed by

$$\tilde{y}_k = \underline{y}_k + \sum_{i=1}^n t_i d_i^{\tilde{Y}^k} = A_k^{-1} \underline{x} + \sum_{i=1}^n t_i A_k^{-1} d_i^{\tilde{X}^k}$$

with appropriate $t_i \in [0, 1]$ for all $i = 1, \dots, n$.

Using these definitions, for each $i = 1, \dots, n$ we have to solve the linear optimization problems

$$P_{\tilde{Y}^k, i, \max} : \max_{t \in \mathbb{R}^n} y_{k, i} \quad \text{s.t.} \quad y_k = \underline{y}_k + \sum_{l=1}^n t_l d_l^{\tilde{Y}^k}, \quad t_l \in [0, 1] \quad \forall l = 1, \dots, n, \quad (7)$$

and

$$P_{\tilde{Y}^k, i, \min} : \min_{t \in \mathbb{R}^n} y_{k, i} \quad \text{s.t.} \quad y_k = \underline{y}_k + \sum_{l=1}^n t_l d_l^{\tilde{Y}^k}, \quad t_l \in [0, 1] \quad \forall l = 1, \dots, n. \quad (8)$$

Constructing the boxes \tilde{Y}^k for each $k \geq \tilde{k}$ we obtain a sequence of boxes $(\tilde{Y}^k)_{k \in \mathbb{N}}$. In order to apply our results from Corollary 4.7 we now have to check whether this sequence has all the required properties.

First, we can state that the transformed midpoint $\text{mid}(\tilde{Y}^k) := A_k^{-1} \text{mid}(X^k)$ of the box \tilde{X}^k is also midpoint of the n -parallelepiped \tilde{Y}^k , i.e. $\text{mid}(\tilde{Y}^k) = \underline{y}_k + \sum_{i=1}^n \frac{1}{2} d_i^{\tilde{Y}^k}$, which follows from the linearity of the coordinate transformation. As constructing the box \tilde{Y}^k does not change the componentwise boundaries, $\text{mid}(\tilde{Y}^k)$ is also midpoint of the box \tilde{Y}^k , and the previous results from this subsection remain valid. In addition, we have to show that $(\tilde{Y}^k)_{k \in \mathbb{N}}$ is a non-deformed sequence of boxes asymptotically centered on the globally minimal point y^{*k} in the new coordinates and satisfies $\lim_{k \rightarrow \infty} \text{diag}(\tilde{Y}^k) = 0$. The following lemma proves these properties.

Lemma 4.10 *Let $(\tilde{X}^k)_{k \in \mathbb{N}}$ be a non-deformed sequence of boxes asymptotically centered on a globally minimal point x^* and with $\lim_{k \rightarrow \infty} \text{diag}(\tilde{X}^k) = 0$, and let Assumption 1.1 hold. Then the sequence $(\tilde{Y}^k)_{k \in \mathbb{N}}$ of boxes*

- (a) *satisfies $\lim_{k \rightarrow \infty} \text{diag}(\tilde{Y}^k) = 0$,*
- (b) *is asymptotically centered on*

$$y^* = \lim_{k \rightarrow \infty} y^{*k} = \lim_{k \rightarrow \infty} A_k^{-1} x^* = A_{x^*}^{-1} x^*,$$

- (c) *is non-deformed.*

Proof To prove *assertion (a)* we use the row-sum norm, i.e. the matrix norm induced by $\|\cdot\|_\infty$. With the compatibility of the row-sum norm of A_k^{-1} we have

$$\|A_k^{-1}x\|_\infty \leq \|A_k^{-1}\|_\infty \|x\|_\infty$$

for all $x \in \mathbb{R}^n, k \in \mathbb{N}$. In particular, for all $x', x'' \in \tilde{X}^k$ it follows

$$\|A_k^{-1}(x' - x'')\|_\infty \leq \|A_k^{-1}\|_\infty \|x' - x''\|_\infty \leq \|A_k^{-1}\|_\infty \max_{i=1, \dots, n} (\tilde{x}_i^k - \underline{x}_i^k).$$

With $(\tilde{X}^k)_{k \in \mathbb{N}}$ asymptotically centered on x^* we have $\lim_{k \rightarrow \infty} \text{mid}(X^k) = x^*$ and with h being continuously differentiable it follows that $\lim_{k \rightarrow \infty} A_k^{-1} = A_{x^*}^{-1}$. With $\lim_{k \rightarrow \infty} \text{diag}(\tilde{X}^k) = 0$ we also have $\lim_{k \rightarrow \infty} \max_{i=1, \dots, n} (\tilde{x}_i^k - \underline{x}_i^k) = 0$. Moreover, $\|A_{x^*}^{-1}\|_\infty$ is a real number. As both limit values exist, we obtain

$$\begin{aligned} 0 &\leq \lim_{k \rightarrow \infty} \|A_k^{-1}(x' - x'')\|_\infty \\ &\leq \lim_{k \rightarrow \infty} \|A_k^{-1}\|_\infty \cdot \max_{i=1, \dots, n} (\tilde{x}_i^k - \underline{x}_i^k) \\ &= \lim_{k \rightarrow \infty} \|A_k^{-1}\|_\infty \cdot \lim_{k \rightarrow \infty} \max_{i=1, \dots, n} (\tilde{x}_i^k - \underline{x}_i^k) = 0. \end{aligned}$$

Hence, for arbitrary $x', x'' \in \tilde{X}^k$ all components of $A_k^{-1}(x' - x'')$ converge to 0. As \tilde{Y}^{nk} is constructed from these component entries, also all edge lengths of \tilde{Y}^{nk} converge to zero and thus, also $\text{diag}(\tilde{Y}^{nk})$.

Now we turn towards *assertion (b)*. The globally minimal point x^* can be described as a linear combination of all edge vectors of \tilde{X}^k starting from \underline{x}^k , i.e.

$$x^* = \underline{x}^k + \alpha_1^k d_1^{\tilde{X}^k} + \alpha_2^k d_2^{\tilde{X}^k} + \dots + \alpha_n^k d_n^{\tilde{X}^k}$$

with $\alpha_i \in \mathbb{R}$ for all $i = 1, \dots, n$ and $k \in \mathbb{N}$. As x^* asymptotically becomes midpoint of the elements of $(\tilde{X}^k)_{k \in \mathbb{N}}$, we have $\lim_{k \rightarrow \infty} \alpha_i^k = \frac{1}{2}$ for all $i = 1, \dots, n$.

Since the coordinate transformation is linear, it follows

$$y^{*k} = \underline{y}_k + \alpha_1^k d_1^{\tilde{Y}^k} + \alpha_2^k d_2^{\tilde{Y}^k} + \dots + \alpha_n^k d_n^{\tilde{Y}^k}. \tag{9}$$

Hence, with $\lim_{k \rightarrow \infty} \alpha_i^k = \frac{1}{2}$ the feasible point \hat{y}^k in the new coordinates also becomes midpoint of \tilde{Y}^k and, thus, \tilde{Y}^{nk} asymptotically.

To show *assertion (c)* again we use the row-sum norm of A_k^{-1} . Similarly to (a), for all $x', x'' \in \tilde{X}^k$ we obtain

$$\|A_k^{-1}(x' - x'')\|_\infty \leq \|A_k^{-1}\|_\infty \max_{i=1, \dots, n} \|d_i^{\tilde{X}^k}\|_\infty. \tag{10}$$

Analogously, with the induced norm of A_k for all $x', x'' \in \tilde{X}^k$ it follows

$$\|A_k^{-1}(x' - x'')\|_\infty \geq \frac{1}{\|A_k\|_\infty} \min_{i=1, \dots, n} \|d_i^{\tilde{X}^k}\|_\infty. \tag{11}$$

The Eqs. (10) and (11) provide the maximum and minimum stretching under the coordinate transformation. These factors also bound the edge lengths of the new box \tilde{Y}^{rk} . Hence, for the maximum ratio of the length of box edges we obtain

$$t^{\tilde{Y}^{rk}} \leq \frac{\|A_k^{-1}\|_\infty \max_{i=1, \dots, n} \|d_i^{X,k}\|_\infty}{\frac{1}{\|A_k\|_\infty} \min_{i=1, \dots, n} \|d_i^{X,k}\|_\infty} = \kappa(A_k^{-1}) t^k \leq \kappa(A_k^{-1}) \bar{t}, \tag{12}$$

where in the last step we used that $(\tilde{X}^k)_{k \in \mathbb{N}}$ is a non-deformed sequence. Thus, the ratio of the edge lengths of \tilde{Y}^{rk} is bounded above by a constant, κ , containing the condition number of A_k^{-1} .

Since the gradients of h in x^* are linearly independent, because of Assumption 1.1, $A_{x^*}^{-1}$ is a regular matrix. For regular matrices the condition number is bounded. Hence, for k approaching infinity the ratio $t^{\tilde{Y}^{rk}}$ is bounded by a finite constant. Thus, the boxes do not become increasingly deformed. \square

Lemmas 4.9 and 4.10 show that using the, possibly extended, Jacobian of h in $\text{mid}(X^k)$ as the transformation matrix A_k^{-1} for all $k \geq \tilde{k}$ the system is transformed in a way that all requirements of Corollary 4.7 are met at least asymptotically. Therefore, we obtain the following result with regard to Theorem 3.2.

Theorem 4.11 *Let $(X^k)_{k \in \mathbb{N}}$ be a non-deformed exhaustive sequence of boxes with $\bar{x}_i^k - \underline{x}_i^k < 1$ for all $i = 1, \dots, n$ and $k \in \mathbb{N}$, $x^* \in X^k$ for all $k \in \mathbb{N}$ and x^* globally minimal point of $P(B)$. Let $(\tilde{X}^k)_{k \in \mathbb{N}}$ be a sequence of boxes with $\tilde{X}^k = [\underline{x}^k, \bar{x}^k]$, $\text{mid}(\tilde{X}^k) = \text{mid}(X^k)$ and $\tilde{x}_i^k - \underline{x}_i^k := \sqrt{\bar{x}_i^k - \underline{x}_i^k}$ for all $i = 1, \dots, n$ and $k \in \mathbb{N}$. Furthermore, let $h : \mathbb{R}^n \mapsto \mathbb{R}^m$, $h \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ and let Assumption 1.1 hold. Then, there is some $\hat{k} \in \mathbb{N}$ so that for all $k \in \mathbb{N}$ with $k \geq \hat{k}$ the conditions (3) of Theorem 3.2 are satisfied for the boxes \tilde{Y}^{rk} obtained by applying the aforementioned transformation of the system.*

Proof With the extension of X^k for each $k \in \mathbb{N}$ according to Lemma 4.6 we obtain a sequence $(\tilde{X})_{k \in \mathbb{N}}$ of boxes which is asymptotically centered on the globally minimal point x^* . According to Lemma 4.8 there is a $\tilde{k} \in \mathbb{N}$ so that for all $k \geq \tilde{k}$ the coordinate transformation, and afterwards the extension of \tilde{Y}^k to a box \tilde{Y}^{rk} can be applied. Thereby we obtain a sequence of boxes $(\tilde{Y}^{rk})_{k \in \mathbb{N}}$ that is non-deformed, asymptotically centered on y^* and satisfies $\lim_{k \rightarrow \infty} \text{diag}(Y^{rk}) = 0$, as shown in Lemma 4.10, and a sequence of functions $(\tilde{h}_j^k)_{k \in \mathbb{N}}$ in the new coordinates. Using Lemma 4.9 it follows $\lim_{k \rightarrow \infty} \nabla \tilde{h}_j^k(y^{*k}) = e_j$. Hence, asymptotically, all assumptions from Corollary 4.7 are met, so that we can conclude that \tilde{h}^k asymptotically satisfies the sign conditions of Theorem 3.2 on \tilde{Y}^{rk} . As \tilde{h}^k is continuously differentiable, there is some \hat{k} so that for all $k \geq \hat{k}$ the sign conditions are satisfied as well. \square

4.4.2 Determining upper bounds in the transformed system

With the previous results we are now able to present our method to calculate valid upper bounds for the globally minimal value v^* .

In a first phase, the box is extended to cope with feasible points on facets. If the box is too large for our extension approach, we set $\tilde{X} = X$. In a second phase, if possible, the system is transformed such that at least asymptotically the gradients $\nabla \tilde{h}_j$, $j \in J$, at a transformed globally minimal point y^* have unit direction. We note again that the transformation can be also applied for sufficiently small boxes in a sequence of boxes asymptotically centered on a feasible point \hat{x} with linearly independent gradients $\nabla h_j(\hat{x})$. In phase three, Theorem 3.2 is used to check for feasibility. Theorem 4.11 states that after a finite number of steps within a branch-and-bound algorithm, its conditions are satisfied and the existence of a zero in \tilde{Y}^{1k} can be verified. However, as checking

$$\tilde{h}_j^k(y) \leq 0, \quad \forall y \in \underline{\tilde{Y}_j^{1k}} \quad \text{and} \quad \tilde{h}_j^k(y) \geq 0, \quad \forall y \in \overline{\tilde{Y}_j^{1k}},$$

for all $j \in J$ is not an easy task, we will use quadratically convergent upper bounding procedures, e.g. centered forms [7,22], to determine upper and lower bounds on the facets instead.

For all $j \in J$ we will determine an upper bound $u_j(\overline{\tilde{Y}_j^{1k}})$ for the maximal value of \tilde{h}_j^k on the facet $\underline{\tilde{Y}_j^{1k}}$ and a lower bound $\ell_j(\overline{\tilde{Y}_j^{1k}})$ for the minimal value of \tilde{h}_j^k on the facet $\overline{\tilde{Y}_j^{1k}}$. Using such bounds is sufficient to verify the existence of feasible points, since $u_j(\overline{\tilde{Y}_j^{1k}}) \leq 0$ and $\ell_j(\overline{\tilde{Y}_j^{1k}}) \geq 0$ for all $j \in J$ imply that all conditions of Theorem 3.2 are satisfied.

In general, the verification of a zero of \tilde{h}^k in a box \tilde{Y}^{1k} does neither guarantee the existence of a feasible point in the original box X^k nor the existence of a feasible point of $P(B)$ at all, since by extending X^k to \tilde{X}^k and \tilde{Y}^k to \tilde{Y}^{1k} the considered set of points is enlarged. Applying the coordinate transformation, for each feasible point $\hat{x} \in X^k \subseteq B$ we obtain a feasible point $\hat{y}^k \in \tilde{Y}^{1k}$ in the new coordinates. Hence, if X^k is sufficiently small, the existence of a feasible point is verified correctly. On the contrary, it is also possible that the conditions (3) in Theorem 3.2 are satisfied due to a point $\hat{y} \in \tilde{Y}^{1k}$ with $\tilde{h}_j^k(\hat{y}^k) = 0$ for all $j \in J$ without a corresponding point $\hat{x} \in X^k$ in the original coordinates. While for $\hat{x} \notin B$ the existence of a feasible point of $P(B)$ cannot be concluded then, if we have $\hat{x} \notin X^k$ but $\hat{x} \in B$, still a feasible point of the original problem $P(B)$ is verified. We will show that for a sufficiently small box X^k the latter condition is guaranteed to be satisfied so that for such a box our algorithm verifies the existence of a feasible point of the problem $P(B)$. Then, an upper bound for the objective function on \tilde{Y}^{1k} may serve as a new upper bound for v^* .

In practice, it is not clear in advance whether a box is sufficiently small for a detected zero to be located in B . As the zero is not determined explicitly, to check for this, we transform the whole box \tilde{Y}^{1k} to an n -parallelepiped \tilde{X}^{1k} in the original coordinates and then extend it to a minimum bounding box \tilde{X}^{1k} with the same ideas as in the previous

subsection. Then, we can check whether $\tilde{X}''^k \subseteq B$ which implies $\hat{x} \in B$. The box \tilde{X}''^k is also used to save the best-known approximation of a globally minimal point in the original coordinates within a branch-and-bound algorithm. However, to calculate an upper bound for v^* that is as tight as possible, the upper bound should be determined on \tilde{Y}''^k in the new coordinates.

Algorithm 2

Input: Box $X \subseteq B \subseteq \mathbb{R}^n$, equality constraints $h_j(x) = 0, j \in J$, objective function f

Phase 1 - Box extension

- 1: **if** $\tilde{x}_i^* - \underline{x}_i^* < 1$ **then**
- 2: Construct \tilde{X} with $\text{mid}(\tilde{X}) = \text{mid}(X)$ and $\tilde{x}_i - \underline{x}_i := \sqrt{\tilde{x}_i - \underline{x}_i}$ for all $i = 1, \dots, n$.
- 3: **else**
- 4: Set $\tilde{X} := X$.

Phase 2 - System transformation

- 5: Extend $Dh(\text{mid}(\tilde{X}))$ to a matrix A^{-1} of basis vectors in \mathbb{R}^n with $\text{null}(Dh(\text{mid}(\tilde{X})))$.
- 6: If possible, apply a coordinate transformation using A^{-1} , which transforms \tilde{X} to \tilde{Y} .
- 7: Determine a minimum bounding box \tilde{Y}' for \tilde{Y} .

Phase 3 - Feasibility verification and upper bound for v^*

Phase 3a - Zero verification for \tilde{Y}'

- 8: **for** $j \in J$ **do**
- 9: Use a quadratically convergent bounding procedure to determine a
 - lower bound ℓ_j for $\min_{y \in \tilde{Y}'_j} \tilde{h}_j(y)$,
 - upper bound u_j for $\max_{y \in \tilde{Y}'_j} \tilde{h}_j(y)$.
- 10: **if** $u_j \leq 0$ and $\ell_j \geq 0$ for all $j \in J$ **then**
- 11: \tilde{Y}' contains a point \hat{y} with $\tilde{h}_j(\hat{y}) = 0$ for all $j \in J$.

Phase 3b - Retransformation to the original coordinates

- 12: Apply a transformation using matrix A , which transforms \tilde{Y}' to \tilde{X}' .
- 13: Determine a minimum bounding box \tilde{X}'' for \tilde{X}' .

Phase 3c - Feasibility verification and upper bound determination

- 14: **if** $\tilde{X}'' \subseteq B$ **then**
- 15: The verified point \hat{y} is also feasible for B .
- 16: Determine an upper bound u_f for f on \tilde{Y}' .

Output: Upper bound u_f for v^* .

The following corollary formally proves that Algorithm 2 determines an upper bound for v^* after a finite number of steps.

Corollary 4.12 *Let $(X^k)_{k \in \mathbb{N}}$ be a non-deformed exhaustive sequence of boxes with $\tilde{x}_i^k - \underline{x}_i^k < 1$ for all $i = 1, \dots, n$ and $k \in \mathbb{N}$, $X^k \subseteq B$, $x^* \in X^k$ for all $k \in \mathbb{N}$ and x^* a globally minimal point of $P(B)$. Furthermore, let $h : \mathbb{R}^n \mapsto \mathbb{R}^m, h \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ and let Assumption 1.1 hold. Then, there is some $\hat{k} \in \mathbb{N}$ so that for all $k \in \mathbb{N}$ with $k \geq \hat{k}$ Algorithm 2 verifies the existence of a feasible point of $P(B)$ and determines a valid upper bound u_f for v^* .*

Proof With Theorem 4.11 there is a k^1 so that the conditions (3) of Theorem 3.2 are satisfied for all $k \geq k^1$. Hence, with \tilde{y}^k the maximum point on $\underline{\tilde{Y}}_j^{k^1}$ for some arbitrary $j \in J$ we have

$$\tilde{h}_j^k(y) \leq \tilde{h}_j^k(\tilde{y}^k) < 0, \quad \forall y \in \underline{\tilde{Y}}_j^{k^1}.$$

Using a quadratically convergent bounding procedure on $\underline{\tilde{Y}}_j^{k^1}$ we obtain

$$u_j(\underline{\tilde{Y}}_j^{k^1}) \leq \tilde{h}_j^k(\tilde{y}^k) + \delta \operatorname{diag}(\underline{\tilde{Y}}_j^{k^1})^2 \leq \tilde{h}_j^k(\tilde{y}^k) + \delta \operatorname{diag}(\tilde{Y}^k)^2.$$

Since the first term is negative for sufficiently large k and converges to zero linearly while the second one converges to zero quadratically, there is some $k^2 \in \mathbb{N}$ so that for all $k \geq k^2$ and for all $j \in J$ we have

$$u_j(\underline{\tilde{Y}}_j^{k^1}) < 0.$$

The proof for the facet $\overline{\tilde{Y}}_j^{k^1}$ follows completely analogously.

In the retransformation phase we obtain a sequence of boxes $(\tilde{X}''^k)_{k \in \mathbb{N}}$ in the original coordinates. Analogously to our ideas in Sect. 4.4 we can conclude that $\lim_{k \rightarrow \infty} \operatorname{diag}(\tilde{X}''^k) = 0$. Additionally, we assume all feasible points of $P(B)$ to strictly satisfy the box constraints. Hence, for any box containing a feasible point there is some $k^3 \in \mathbb{N}$ so that $\tilde{X}''^k \subseteq B$ for all $k \geq k^3$.

So in total, there is some $\hat{k} = \max\{k^1, k^2, k^3\}$ so that Algorithm 2 verifies the existence of a feasible point of $P(B)$ in \tilde{Y}^k for all $k \geq \hat{k}$ and a valid upper bound u_j for v^* is determined. \square

5 Preliminary numerical results

In this subsection we present preliminary numerical results for Algorithm 2. The aim of these results is to demonstrate the effectiveness of the algorithm to detect the existence of feasible points of $P(B)$ and, thus, to complement our theoretical results. The results also indicate that our approach could be used within a branch-and-bound framework to calculate convergent upper bounds for v^* .

For the purpose of demonstration, we construct (sequences of) boxes as they typically occur within a branch-and-bound algorithm by division along the longest edge. Then, we verify the existence of feasible points on the obtained boxes. If the existence of a feasible point is verified for a box, the algorithm determines an upper bound for v^* and terminates. Otherwise, the box is divided along a longest edge, added to a list and the next box is chosen from the list based on the FIFO principle, as long as the list is not empty.

Our current implementation is not rigorous in the sense that it does not take rounding errors into consideration. To the best of our knowledge, currently there is no

commercial branch-and-bound solver in global optimization that ensures that all computations are performed in such a rigorous manner so that the user can be sure that the lower bound is actually less or equal than the globally minimal value, although mathematically the proofs for the lower bounds are available for exact arithmetic. So far, we follow this practice also for our new upper bounding procedure and do not consider rounding errors in our implementation, although the proofs for the case of exact arithmetic are now available in this paper. A more detailed discussion on this subject can be found at the end of this article.

5.1 Implementation details

Algorithm 2 constitutes the basis of the implementation, but some further additions are made to improve its efficiency and the quality of the upper bounds. Figure 3 illustrates the feasibility verification checks and transformations within Algorithm 2 together with these additions.

- Bounds $u(X)$ and $\ell(X)$ for h_j , $j \in J$, are calculated on the current box X to exclude it from further examination, if it is guaranteed not to contain a feasible point, that is, $u(X) < 0$ or $\ell(X) > 0$ hold.
- Despite the better convergence rate of the former, centered forms as well as interval arithmetic are applied as bounding procedures and the tighter bounds are used in each case.
- We use a cascaded system of feasibility verifications: First we check the sign conditions of Theorem 3.2 for X . If they are satisfied, a feasible point is verified in X and there is no need for a box extension or coordinate transformation. If they are not satisfied, we extend X to \tilde{X} and check the sign conditions for the extended box. If they are not satisfied, we apply the system transformation and check the conditions for \tilde{Y}' . The reason behind this approach is that we obtain tighter upper bounds if we verify the existence of a feasible point in a smaller box. Therefore, it is reasonable to extend the boxes only if necessary.

As shown in Corollary 4.12, Algorithm 2 guarantees that the conditions (3) of Theorem 3.2 are satisfied for sufficiently small boxes containing feasible points. After applying the coordinate transformation it is clear which function h_j , $j \in J$ is allocated to which box facet with respect to the sign conditions because their gradients have unit direction. However, a reasonable allocation beforehand might lead to a less rough overestimation of \tilde{X} in the new coordinates and increase the probability that no transformation is required at all, as the conditions (3) might be satisfied for the boxes X or \tilde{X} already. Therefore, we test two different strategies of allocating m of the n facets of a box to the m functions h_j , $j \in J$. *Allocation strategy I* simply allocates the j -th function to the j -th facets of X . *Allocation strategy II* determines the allocation iteratively based on the angle of the m gradients ∇h_j , $j \in J$, with the n unit vectors.

The method is implemented in MATLAB R2018b and executed on an Intel Core i5-6300 U (2.4 GHz) processor with 12 GB RAM. In addition to standard MATLAB commands, the INTLAB toolbox V7.1 for reliable computing is used to realize the calculation of bounds, the determination of the box \tilde{Y}^* and interval computations in general [36].

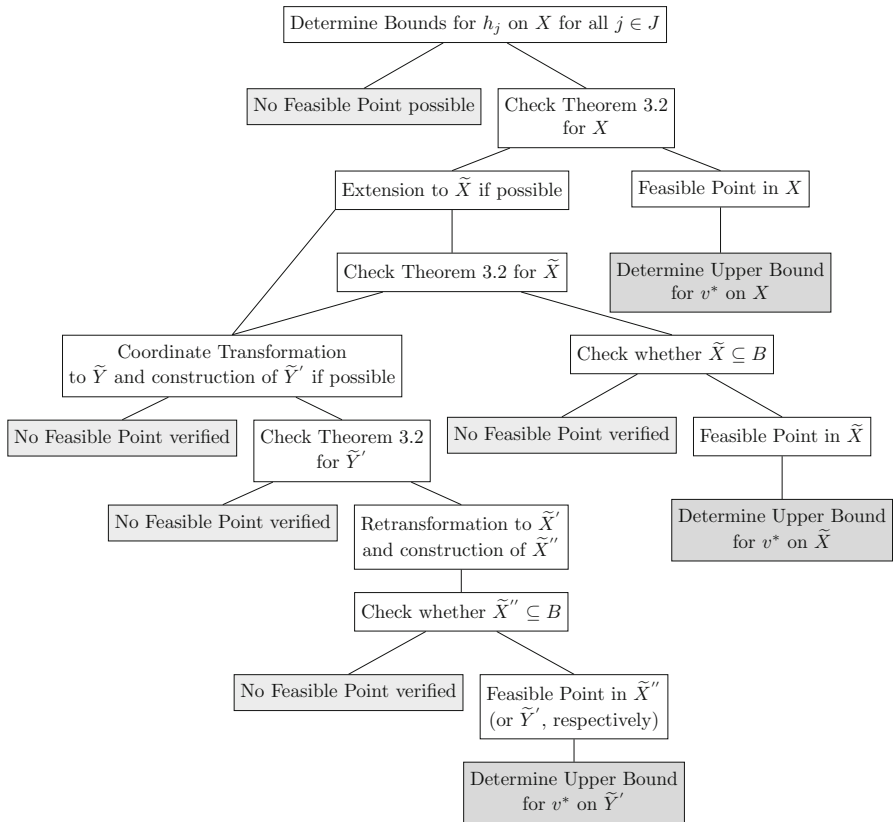


Fig. 3 Possible outcomes of one iteration of the implemented method

5.2 Results for test problems

Our implementation is tested for a selection of problems from the COCONUT benchmark [40]. All selected problems are also examined by Domes and Neumaier in [9]. In the case that a test problem did not include any box constraints, these were added to the problem to make it applicable to our purposes. We chose $B = [-1000, 1000]^n$ as the initial box. The problems *aljazzaf* and *bt13* were slightly modified for the feasible points to strictly satisfy the box constraints. The examined problems and the main results are summarized in Table 1.

For all but two problems the algorithm successfully verifies the existence of feasible points. For 26 of 35 problems the algorithm terminates successfully for both allocation strategies, as indicated by *yes* in Table 1. For the specific problem *parabola* the following example illustrates the application of our method in detail.

Table 1 Summary of test problems analyzed

Problem	n	m	Strategy I		Strategy II	
			Verified	Iterations	Verified	Iterations
bt1	2	1	Yes	153	Yes	153
extrasim	2	1	Yes	15	Yes	1
hs006	2	1	Yes	5	Yes	5
maratos	2	1	Yes	153	Yes	153
tame	2	1	Yes	53	Yes	53
try-b	2	1	Yes	41	Yes	41
gottfr	2	2	Yes	462	Yes	462
a	2	2	Yes	530	Yes	530
booth	2	2	Yes	110	Yes	110
himmelbc	2	2	Yes	17	Yes	17
hs008	2	2	Yes	155	Yes	155
hycpir	2	2	Yes	127	Yes	127
mickey	2	2	Yes	204	Yes	204
parabola	2	2	Yes	23	Yes	23
supersim	2	2	Yes	92	Yes	90
hs028	3	1	Yes	62	Yes	3
aljazzaf	3	1	No		Yes*	
hs061	3	2	Yes	51	Yes	11
rediff3	3	3	Yes	2822	Yes	2822
zangwil3	3	3	Yes	59	Yes	59
bronstein	3	3	Yes	1637	Yes	1637
clo1	3	3	Yes	100	Yes	100
czaporgeddes	3	3	Yes	6719	Yes	6936
eqlin	3	3	Yes	61	Yes	61
hong2	3	3	Yes*		No	
mathews	3	3	Yes	7493	Yes	7493
hs042	4	2	Yes*		Yes*	
kear3	4	4	No		No	
powell	4	4	No		No	
b1	4	4	Yes	765	Yes	8499
eiger	4	4	Yes	952	Yes	1141
hong1	4	4	Yes*		Yes*	
lorentz	4	4	Yes*	8924	Yes*	9073
bt13	5	1	No		Yes	6707
vrahatis	9	9	Yes	1023	Yes	1023

Example 5.1 The problem consists of the following constraints

$$\begin{aligned}h_1(x) &:= x_1^2 - x_2 = 0, \\h_2(x) &:= x_1^2 + x_2^2 - 2 = 0, \\x \in B &:= [0, 10] \times [0, 2].\end{aligned}$$

The only feasible point of this problem is $\hat{x} = (1, 1)$.

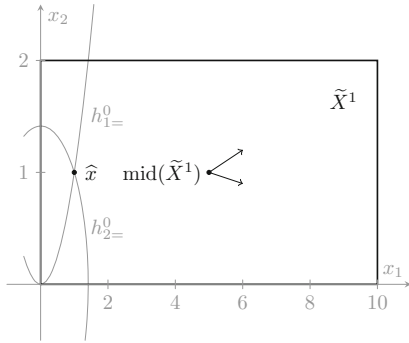
In the first iteration the box $X^1 = B$ is too large for our extension approach, so we set $\tilde{X}^1 = X^1$. The box-midpoint is $\text{mid}(X^1) = (5, 1)$. The gradients $\nabla h_1(\text{mid}(X^1))$ and $\nabla h_2(\text{mid}(X^1))$ do not have unit direction, as can be seen in Fig. 4a. Then, the coordinate transformation is applied in Algorithm 1, such that the gradients ∇h_1 and ∇h_2 in the transformed midpoint do have unit direction, see Fig. 4b. However, this is still not the case in the image \hat{y}_1 of the feasible point \hat{x} in the new coordinates. As the level curves of \tilde{h}_1^1 and \tilde{h}_2^1 in Fig. 4b show, the conditions of Theorem 3.2 are not satisfied. Moreover, we can see that in this iteration the minimum bounding box \tilde{Y}^1 only provides a rough approximation of \tilde{Y}^1 .

However, after 23 iterations our method terminates with the verification of the existence of a feasible point. In this iteration we have $X^{23} = [0.9374, 1.2501] \times [1, 1.5]$, $\tilde{X}^{23} \approx [0.8142, 1.3733] \times [0.8964, 1.6036]$ and $\tilde{Y}^{23} \approx [0.1175, 2.1076] \times [4.0222, 7.0129]$. The original and the transformed system for this iteration are visualized in Fig. 4c and d. The parallelepiped is much less skewed than in the first iteration. The gradients in the transformed feasible point \hat{y}^{23} are approximately equal to the gradients in $\text{mid}(\tilde{Y}^{23})$. Moreover, the constructed box \tilde{Y}^{23} is sufficiently small. Therefore, the existence of a zero of \tilde{h}^{23} is verified using Theorem 3.2. Since \tilde{Y}^{23} transformed back to the original coordinates is a subset of B , a feasible point for the original problem is verified and an upper bound for v^* can be determined.

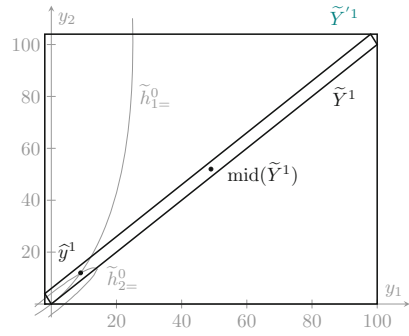
For some problems, starting with the initial box constraints (*bt1*, *hong1*, *hong2*, *aljazaf*, *hs042*, *lorenz*) no verification of a feasible point was possible within a time frame of 30 minutes. This might be explained by the combination of large initial boxes, no fathoming in our implementation and using the FIFO principle instead of choosing boxes based on bounds for v^* . Starting with a sufficiently small initial box containing a feasible point, however, our algorithm was able to verify the existence of such a point for all of these problems at least with one of the two allocation strategies (see *yes** in Table 1). Hence, valid upper bounds for v^* could be determined. These results validate the results proven in Sect. 4.

For allocation strategy I in 28 cases the existence of a feasible point is verified after a coordinate transformation. This underlines the significance of gradient directions towards facets. However, our results also indicate a trade-off between the allocation and the coordinate transformation. If the functions h_j , $j \in J$, and the facets are allocated in a more reasonable way, as done with allocation strategy II, in 12 cases the existence of a feasible point can be verified for box X^k or \tilde{X}^k already. Therefore, the computationally expensive transformation and further box extensions are not required at all.

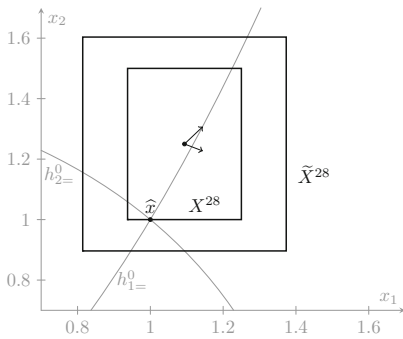
Compared to numerical results in the literature, for some of the problems our algorithm performs better than the majority of the methods analyzed in [9] (*bt1*, *extrasim*,



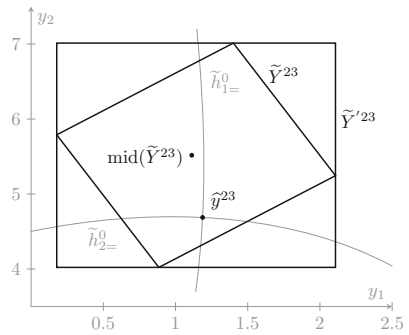
(a) The original system in the first iteration.



(b) The transformed system in the first iteration.



(c) The original system in iteration 23.



(d) The transformed system in iteration 23.

Fig. 4 Illustration of Algorithm 2 applied to *Parabola*

hs006, hs061, maratos, tame, try-b). For two problems, *hs028* and *aljazzaf*, our method is even the only one to successfully verify the existence of a feasible point and provide a valid upper bound for v^* . The only two problems, for which the methods presented in [9] clearly perform better are *kear2* and *powell*, for which the existence of a feasible point cannot be verified with our algorithm, as Assumption 1.1 is violated.

6 Final remarks

In this article, a new method to compute convergent valid upper bounds in spatial branch-and-bound methods is proposed. The numerical results show that our method performs well on standard test problems in global optimization. However, there are still some issues that have to be addressed.

For example, our approach is described for equality constrained problems that do not involve inequalities beside the box constraints. Although, in theory, additional inequalities can be reformulated as equality constraints by using the slack variable approach

as described in the introduction, this reformulation is known to have some drawbacks as explained in [18]. In order to see this in the context of spatial branch-and-bound methods we consider a problem of the form $P(B)$ with additional inequality constraints $g_i(x) \leq 0$, $i \in I$. Using the slack variable approach this can be reformulated as $g_i(x) + y_i^2 = 0$, $i \in I$, together with the additional box constraints $-\bar{c}_i \leq y_i \leq \bar{c}_i$ where \bar{c}_i , $i \in I$, are chosen appropriately, for instance by applying interval arithmetic. Assuming the existence of a globally minimal point x^* with $g_i(x^*) < 0$, $i \in I$, for the original problem, there is also a globally minimal point $(x^*, y^*)^\top$ for the reformulated problem that provides the same globally minimal value. Here, every y_i^* can be either $\sqrt{-g_i(x^*)}$ or $-\sqrt{-g_i(x^*)}$ so that there are $2^{|I|}$ globally minimal points for the reformulated problem. Thus, we see that the number of globally optimal points may grow exponentially in the number of inequality constraints, which is undesirable as this slows down spatial branch-and-bound methods considerably.

A straightforward possibility to circumvent this difficulty in many practical applications might be to introduce the box constraints $-\gamma \leq y_i \leq \bar{c}_i$ with a small $\gamma > 0$ instead of the aforementioned ones. However, even if it cuts off many superfluous globally minimal points introduced during the reformulation step, it might be difficult to determine a reasonable value for γ . Moreover, for problems with active inequalities in the globally minimal point x^* this introduces nearly optimal points that provide very good, although not globally minimal, values. These points can be expected to be difficult to handle for global optimization solvers. Note that letting $\gamma = 0$ is not a valid option as strict box constraints are required by our new upper bounding technique. In summary, the integration of inequality constraints into our new upper bounding procedure needs to be examined in more detail and is left for future research.

Another issue we want to address is our current transformation approach as described in Sect. 4.4.1 that might lead to significantly enlarged boxes and, additionally, leads to $2n$ linear optimization problems that have to be solved as well as the inversion of a matrix. As an alternative, one could try to work with the non-transformed system. However, this would require a different extension of Miranda's Theorem that copes with parallelepipeds instead of boxes. This is also left for future research.

Moreover, so far our implementation of the new upper bounding procedure is not included into a branch-and-bound framework. Instead of this, in the current paper the new method is applied to exhaustive sequences of boxes that also occur during the solution of these problems by standard algorithms in global optimization. This approach enables us to examine our upper bounding procedure isolatedly and in more detail without side effects produced by the remaining solution procedure. However, especially if combined with an improved approach to handle inequality constraints, our aim is to integrate this procedure into a complete spatial branch-and-bound algorithm. This is also left for future research.

By using our new upper bounding procedure we believe that it is possible to develop a rigorous global branch-and-bound solver. In this article we have shown that, at least in exact arithmetic, our new upper bounding procedure converges to the globally minimal value. While this cannot be ensured for other commonly used upper bounding procedures, this is also true for standard lower bounding procedures in global opti-

mization and, thus, the combination of both may serve as starting point for a rigorous implementation.

However, for a rigorous global branch-and-bound solver it is important to distinguish mathematical convergence results from implementation details. During implementation of numerical methods the additional difficulty of finite precision arithmetic arises since on a computer a given real number can be represented only within a certain accuracy. There are at least two possibilities to cope with this issue. Firstly, so-called arbitrary precision arithmetic libraries could be used that allow higher precision at the cost of longer computation times and higher memory usage. Secondly, current processors allow to change rounding modes of floating point numbers. By exploiting this, it is possible to develop advanced interval arithmetic libraries that ensure that the actual value of an arithmetic expression is contained in an interval, where lower as well as upper bounds can be represented by floating point numbers.

Using one of the aforementioned techniques we believe that a rigorous global solver can be implemented. However, the implementation of such a method is beyond the scope of this paper.

Acknowledgements Open Access funding provided by Projekt DEAL. The authors are grateful to the anonymous referees and the editor as well as to Mirjam Dür for their precise and substantial remarks on earlier versions of this manuscript.

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. Adjiman, C.S., Floudas, C.A.: Rigorous convex underestimators for general twice-differentiable problems. *J. Glob. Optim.* **9**(1), 23–40 (1996)
2. Adjiman, C.S., Androulakis, I.P., Maranas, C.D., Floudas, C.A.: A global optimization method, α bb, for process design. *Comput. Chem. Eng.* **20**, 419–424 (1996)
3. Adjiman, C.S., Androulakis, I.P., Floudas, C.A.: Global optimization of MINLP problems in process synthesis and design. *Comput. Chem. Eng.* **21**, 445–450 (1997)
4. Adjiman, C.S., Androulakis, I.P., Floudas, C.A.: A global optimization method, α bb, for general twice-differentiable constrained NLPs: II. Implementation and computational results. *Comput. Chem. Eng.* **22**(9), 1159–1179 (1998)
5. Adjiman, C.S., Dallwig, S., Floudas, C.A., Neumaier, A.: A global optimization method, α bb, for general twice-differentiable constrained NLPs: I. Theoretical advances. *Comput. Chem. Eng.* **22**(9), 1137–1158 (1998)
6. Androulakis, I.P., Maranas, C.D., Floudas, C.A.: α bb: a global optimization method for general constrained nonconvex problems. *J. Glob. Optim.* **7**(4), 337–363 (1995)
7. Baumann, E.: Optimal centered forms. *BIT Numer. Math.* **28**(1), 80–87 (1988)
8. Belotti, P., Lee, J., Liberti, L., Margot, F., Wächter, A.: Branching and bounds tightening techniques for non-convex MINLP. *Optim. Methods Softw.* **24**, 597–634 (2009)
9. Domes, F., Neumaier, A.: Rigorous verification of feasibility. *J. Glob. Optim.* **61**(2), 255–278 (2015)

10. Dür, M.: Dual bounding procedures lead to convergent branch-and-bound algorithms. *Math. Program.* **91**, 117–125 (2001)
11. Dür, M.: A class of problems where dual bounds beat underestimation bounds. *J. Glob. Optim.* **22**, 49–57 (2002)
12. Falk, J.E., Soland, R.M.: An algorithm for separable nonconvex programming problems. *Manag. Sci.* **15**, 550–569 (1969)
13. Floudas, C.A.: *Deterministic Global Optimization. Theory, Methods, and Applications.* Kluwer, Dordrecht (2000)
14. Floudas, C.A., Gounaris, C.E.: A review of recent advances in global optimization. *J. Glob. Optim.* **45**, 3–38 (2009)
15. Geissler, B., Martin, A., Morsi, A., Schewe, L.: Using piecewise linear functions for solving MINLPs. In: Lee, J., Leyffer, S. (eds.) *Mixed Integer Nonlinear Programming*, pp. 287–314. Springer, Berlin (2012)
16. Horst, R., Tuy, H.: *Global Optimization. Deterministic Approaches.* Springer, Berlin (1996)
17. Jongen, H.Th., Jonker, P., Twilt, F.: *Nonlinear Optimization in Finite Dimensions.* Kluwer, Dordrecht (2000)
18. Jongen, H.Th., Stein, O.: On the complexity of equalizing inequalities. *J. Glob. Optim.* **27**, 367–374 (2003)
19. Kearfott, R.B.: On proving existence of feasible points in equality constrained optimization problems. *Math. Program.* **83**, 89–100 (1998)
20. Kearfott, R.B.: On rigorous upper bounds to a global optimum. *J. Glob. Optim.* **59**(2–3), 459–476 (2014)
21. Kirst, P., Stein, O., Steuermann, P.: Deterministic upper bounds for spatial branch-and-bound methods in global minimization with nonconvex constraints. *TOP* **23**, 591–616 (2015)
22. Krawczyk, R., Nickel, K.: Die zentrische form in der Intervallarithmetik, ihre quadratische Konvergenz und Inklusionsisotonie. *Computing* **28**, 117–137 (1982)
23. Kulpa, W.: The Poincaré–Miranda theorem. *Am. Math. Mon.* **104**(6), 545–550 (1997)
24. Liberti, L., Maculan, N.: *Global Optimization: From Theory to Implementation. Nonconvex Optimization and its Applications.* Kluwer Academic Publishers, Dordrecht (2006)
25. Lin, Y., Schrage, L.: The global solver in the LINDO API. *Optim. Methods Softw.* **24**, 657–668 (2009)
26. Maranas, C.D., Floudas, C.A.: Global optimization in generalized geometric programming. *Comput. Chem. Eng.* **21**(4), 351–369 (1997)
27. McCormick, G.P.: Computability of global solutions to factorable nonconvex programs: part I—convex underestimating problems. *Math. Program.* **10**(1), 147–175 (1976)
28. Miranda, C.: Un’ osservazione su una teorema di Brouwer. *Boll. Unione Mat. Ital.* **3**, 527 (1940)
29. Misener, R., Floudas, C.A.: Piecewise-linear approximations of multidimensional functions. *J. Optim. Theory Appl.* **145**, 120–147 (2010)
30. Misener, R., Floudas, C.A.: ANTIGONE: algorithms for continuous/integer global optimization of nonlinear equations. *J. Glob. Optim.* **59**, 503–526 (2014)
31. Mitsos, A.: Global optimization of semi-infinite programs via restriction of the right-hand side. *Optimization* **60**, 1291–1308 (2011)
32. Mitsos, A., Tsoukalas, A.: Global optimization of generalized semi-infinite programs via restriction of the right hand side. *J. Glob. Optim.* **61**, 1–17 (2015)
33. Neumaier, A.: *Interval Methods for Systems of Equations.* Cambridge University Press, Cambridge (1990)
34. Pintér, J.: Branch-and-Bound algorithms for solving global optimization problems with Lipschitzian structure. *Optimization* **19**(1), 101–110 (1988)
35. Rebennack, S., Kallrath, J.: Continuous piecewise linear delta-approximations for bivariate and multivariate functions. *J. Optim. Theory Appl.* **167**(1), 102–117 (2015)
36. Rump, S.M.: INTLAB—INTerval LABoratory. In: Csendes, T. (ed.) *Developments in Reliable Computing*, pp. 77–104. Kluwer Academic Publishers, Dordrecht (1999)
37. Ryoo, H.S., Sahinidis, N.V.: Global optimization of nonconvex NLPs and MINLPs with applications in process design. *Comput. Chem. Eng.* **19**(5), 551–566 (1995)
38. Ryoo, H.S., Sahinidis, N.V.: A branch-and-reduce approach to global optimization. *J. Glob. Optim.* **8**(2), 107–138 (1996)
39. Sahinidis, N.V.: BARON: a general purpose global optimization software package. *J. Glob. Optim.* **8**, 201–205 (1996)

40. Shcherbina, O., Neumaier, A., Sam-Haroud, D., Vu, X.-H., Nguyen, T.-V.: Benchmarking global optimization and constraint satisfaction codes. In: Blicek, Ch., Jermann, Ch., Neumaier, A. (eds.) *Global Optimization and Constraint Satisfaction*, pp. 211–222. Springer, Berlin (2003)
41. Smith, E.M.B., Pantelides, C.C.: Global optimisation of nonconvex MINLPs. *Comput. Chem. Eng.* **21**, 791–796 (1997)
42. Smith, E.M.B., Pantelides, C.C.: A symbolic reformulation/spatial branch-and-bound algorithm for the global optimisation of nonconvex MINLPs. *Comput. Chem. Eng.* **23**, 457–478 (1999)
43. Tawarmalani, M., Sahinidis, N.V.: Global optimization of mixed-integer nonlinear programs: a theoretical and computational study. *Math. Program.* **99**(3), 563–591 (2004)
44. Tawarmalani, M., Sahinidis, N.V.: A polyhedral branch-and-cut approach to global optimization. *Math. Program. B* **103**, 225–249 (2005)
45. Vrahatis, M.: A short proof and a generalization of Miranda’s existence theorem. *Proc. Am. Math. Soc.* **107**(3), 701–703 (1989)
46. Zamora, J.M., Grossmann, I.E.: A branch and contract algorithm for problems with concave univariate, bilinear and linear fractional terms. *J. Glob. Optim.* **14**, 217–249 (1999)

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