On Polytopal Branch and Bound with Monotonicity

E. M. T. Hendrix[1(✉)], L. G. Casado[2], B. G.-Tóth[3], and F. Messine[4]

[1] Computer Architecture, Universidad de Málaga, 29080 Málaga, Spain
eligius@uma.es
[2] Informatics, Almería University, (CeiA3), Almería, Spain
leo@ual.es
[3] Informatics, University of Szeged, Szeged, Hungary
boglarka@inf.szte.hu
[4] LAPLACE, ENSEEIHT, University of Toulouse, Toulouse, France
frederic.messine@laplace.univ-tlse.fr

Abstract. In the field of Interval Arithmetic Branch and Bound (B&B) monotonicity detection and its use for dimension reduction and eliminations of partition sets has a long tradition. Recent investigation of extending ideas to simplicial B&B has shown that the relative success depends on the dimension of the simplex, i.e. the dimension of its affine hull. Linear constraints may provide a feasible area which is a polytope in a lower dimension due to equality constraints. We investigate the question how to deal with subsets defined by polytopes. We study how to construct a converging branch and bound, how to define bounds and how to refine and how to exploit monotonicity when bounds are provided by the interval hull.

Keywords: Global Optimization · polytope · branch and bound · interval

1 Introduction

Interval Arithmetic based branch and bound has a long tradition in monotonicity considerations to remove subsets, see [7,12]. The idea is that one can remove interior boxes where the function is monotone over the box and reduce the dimension with respect to a monotone component when a box facet is at the boundary of the search space. The idea to use the interval hull of a simplex to derive lower bounds has been discussed in [11]. We did not only extend this idea for deriving function bounds in [4,9], but we focused further on exploiting the derivative bounds of the interval hull to construct monotonicity tests in simplicial B&B [1,2].

This work has been funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe", Grant PID2021-123278OB-I00 of the Spanish ministry of Science and Innovation.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
O. Gervasi et al. (Eds.): ICCSA 2024 Workshops, LNCS 14816, pp. 397–414, 2024.
https://doi.org/10.1007/978-3-031-65223-3_27

In recent numerical experiments comparing interval B&B which has quite some acceleration devices with simplicial B&B, we noticed that mainly for a feasible set which is on a lower dimensional hyperplane, simplicial B&B may provide some advantage in computation time [6]. We found that this is mainly due to handling linear constraints, as simple linear constraints actually define a polytope for a box region or simplicial feasible set. Our research question is whether it is possible to extend the analysis on lower bounding and monotonicity for a B&B defined by polytopal partition sets. We have several sub-questions. For instance, we would like to keep relatively "round" partition sets that fit well their interval hull to keep tight bounds on the derivatives. How can this be promoted and how can monotonicity be handled? How can we generate tight bounds?

To investigate these question, Sect. 2 introduces mathematical notation and Sect. 3 provides properties of monotonicity over polytopal sets. Moreover, we discuss some special cases together with LP and MIP models to find monotone directions. Section 4 describes a specifically designed algorithm. Section 5 investigates several ways to refine the polytopal subsets, while Sect. 6 summarises our findings.

2 Mathematical Notation

We consider the minimization of a continuously differentiable function $f : \mathbb{R}^d \rightarrow \mathbb{R}$, over a polytope P as feasible set. The 'polytope' as a convex hull of a finite number of points, is also called convex polytope in literature. Our main motivation is the use of linear equality constraints intersecting a box or simplex, which typically leads to polytope dimension $p \leq d$, i.e. in a lower dimensional space than that of the objective function f. The exact dimension of a polytope, like any subset of $\mathbb{R}^d$ is defined as the dimension of its affine hull: $\dim(P) := \dim(\text{aff}(P))$. Notice that also the concepts of relative interior $\text{rint}(P)$ and relative boundary apply. Let $P := \text{conv}(\mathcal{P})$ be defined by a set of q vectors that serve as vertices

$$\mathcal{P} := \{w_1, \ldots, w_q\} \subset \mathbb{R}^d.$$

We assume the vertices to be extreme points of P. The idea is to find or enclose all global minimum points of

$$\min f(x), x \in P. \tag{1}$$

Our experiments in [6] with interval B&B and simplicial B&B codes taught us that using partition sets based on simplices is more effective when the polytope dimension is low with respect to d, whereas partition sets based on boxes (intervals) can exploit better the acceleration devices developed for them when in fact $p = d$. Our main research question is whether we could simply use polytopal partition sets in a direct way. How can they be subdivided effectively and how can monotonicity over them be exploited?

In this concept, the partition set (subset) S is a polytope $S := \text{conv}(V)$, where $V := \{v_1, \ldots, v_m\} \subset \mathbb{R}^d$ is a set of extreme points of S. An interesting

aspect, is that the dimension $n := \dim(S)$ is not necessarily d. We will see why this is relevant. Notice that we can use the barycenter coordinates to describe partition set S:

$$S = \{x = \sum_j \lambda_j v_j, \lambda_j \geq 0, j = 1, \ldots, m, \sum_{j=1}^{m} \lambda_j = 1\} \tag{2}$$

and also its relative interior

$$\text{rint}(S) = \{x = \sum_j \lambda_j v_j, \lambda_j > 0, j = 1, \ldots, m, \sum_{j=1}^{m} \lambda_j = 1\}. \tag{3}$$

The centroid of S is given by $c := \frac{1}{m} \sum_{j=1}^{m} v_j$ and its interval hull $\square S := [\underline{x}, \overline{x}]$, where $\underline{x}_i := \min_{v \in V} v_i$ and $\overline{x}_i := \max_{v \in V} v_i$. In general for a polytope, the minimum is attained at one of its faces. The advantage of a polytopal branch and bound is that the faces of the feasible set are covered by the set of polytopal partition sets. This is not necessarily the case for interval based B&B and difficult to achieve for a simplicial B&B, if the polytope is not partitioned well into simplices.

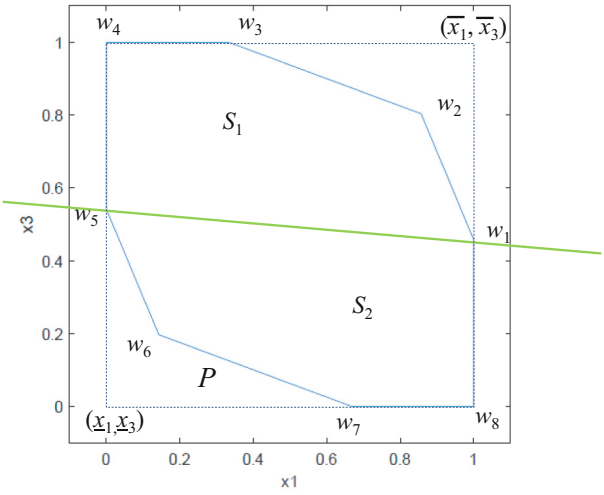

Fig. 1. 2D image from a 6 dimensional polytopal (octagon) feasible set with a partition into two polytopal subsets.

Example 1. Consider a 2-polytope in 6-dimensional space with 8 vertices $\mathcal{P}$ (an octagon) according to the vertex matrix

$$W := \begin{pmatrix} 1.00 & 0.86 & 0.33 & 0 & 0 & 0.14 & 0.67 & 1.00 \\ 0.67 & 1.00 & 1.00 & 0.86 & 0.33 & 0 & 0 & 0.14 \\ 0.46 & 0.80 & 1.00 & 1.00 & 0.54 & 0.20 & 0 & 0 \\ 1.00 & 1.00 & 0.54 & 0.20 & 0 & 0 & 0.46 & 0.80 \\ 0.31 & 0.58 & 0.89 & 1.00 & 0.69 & 0.42 & 0.11 & 0 \\ 0.11 & 0 & 0.31 & 0.58 & 0.89 & 1.00 & 0.69 & 0.42 \end{pmatrix}.$$

It may be clear from the data that the corresponding interval enclosure $\square P$ is given by $[\underline{x}, \overline{x}] = [0, 1]^6$ and the centroid by $c = 0.5(1, 1, 1, 1, 1, 1)^T$. The feasible set has been drawn in Fig. 1 in the coordinate plane x_1, x_3. For the illustration of the concept, we consider the quadratic objective $f(x) = \frac{1}{2}\|x - t\|_2^2$ with $t = (0.2, 0.8, 0.8, 0.3, 0.8, 0.8)^T$. This means that the minimum can be found in the relative interior of P. As we are dealing with a 2-polytope, the corresponding face graph [5] consists only of the vertices, edges and the complete feasible set as depicted in Fig. 2. For comparison reason, we also depicted the face graph of a 3D unit box in Fig. 5. There are several ways to partition the feasible set into polytopal subsets. We will discuss several options in this paper. For the example, we split P into S_1, S_2 by a hyperplane through the vertices w_1 and w_5. Later, we will discuss this choice.

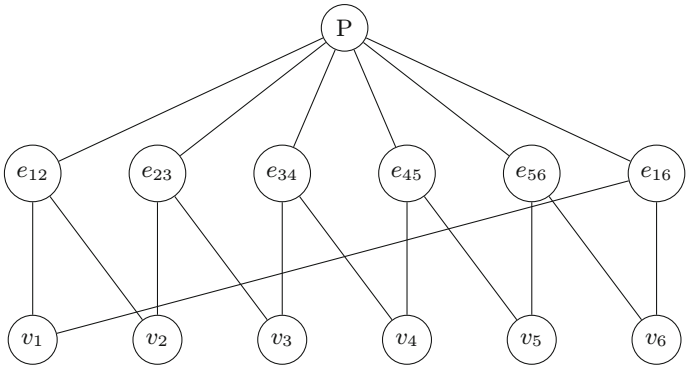

Fig. 2. Face graph of 2D polytopal feasible set

3 Mathematical Properties on Monotonicity

We first focus on mathematical properties with respect to polytopal subsets in Sect. 3.1 and then discuss several ways to find monotonic directions in Sect. 3.2.

3.1 Mathematical Properties on Monotonicity

Consider two points $x, y \in S, x \neq y$ that define a direction $r = x - y$ over polytope S.

Definition 1. *If* $\forall x \in S \ \frac{\partial}{\partial \mu} f(x + \mu r) > 0$ *for* $\mu = 0$*, direction* r *is called monotonically increasing. Monotonically decreasing and non-decreasing is defined similarly.*

As we consider continuously differentiable functions, the monotonicity can be written as $r^T \nabla f(x) > 0$, [8]. In a branch and bound context, we are mainly dealing with lower bound $\underline{r^T \nabla f(S)}$ and upper bound $\overline{r^T \nabla f(S)}$ over the subset S.

Although this is not a necessary consideration for the theoretical properties, in earlier studies we used the interval hull $\square S := [\underline{x}, \overline{x}]$ of the partition set S in order to exploit an enclosure G of the gradient $\nabla f(x) \in G := [\underline{G}, \overline{G}], \forall x \in S$, i.e. $[\underline{\nabla f(S)}, \overline{\nabla f(S)}] \subseteq G$. Interval vector G can be calculated by Interval Automatic Differentiation over the interval hull $\square S$ of subset S, see [13,14]. Considering directional vector r as the difference between two points in S, then the corresponding directional derivative $r^T \nabla f(x)$ is also included in the inner product

$$r^T G = \left[\underline{r^T G}, \overline{r^T G}\right] = \left[\sum_{i=1}^{d} \min\{r_i \underline{G}_i, r_i \overline{G}_i\}, \sum_{i=1}^{d} \max\{r_i \underline{G}_i, r_i \overline{G}_i\}\right]. \quad (4)$$

In [2], we focus on the question how to find monotonic directions in an efficient way in simplicial branch and bound.

The monotonicity is based on directional derivative bounds. This means that if $0 \notin [\underline{r^T \nabla f(S)}, \overline{r^T \nabla f(S)}]$, then r is a monotonic direction. The most generic impact is that, like for simplices [1], the interior of the polytopal partition set cannot contain an optimum.

Proposition 1. *Let* $S \subseteq P$ *be a partition set. If* $\exists\ x, y \in S$*, such that direction* $r = x - y$ *has corresponding directional derivative bounds* $0 \notin [\underline{r^T \nabla f(S)}, \overline{r^T \nabla f(S)}]$ *then* $\operatorname{rint}(S)$ *does not contain a global minimum point of* (1)*.*

Proof. Consider $z \in \operatorname{rint}(S)$. As z is in the relative interior, there exists a feasible direction r in which lower function values can be found, i.e. $\exists \varepsilon \in \mathbb{R}^+$ small enough, such that $z + \varepsilon r \in S$ and $f(z + \varepsilon r) < f(z)$. So z cannot be a minimum point of f. □

The elaboration for an algorithm depends on the choice of the direction r and the way to compute it.

Corollary 1. *Let* $S \subseteq P$ *be a partition set in a branch and bound algorithm. If the conditions of Proposition 1 apply and* $S \in \operatorname{rint}(P)$*, then* S *can be rejected.*

The argument is that the relative boundary of S may contain a global minimum point, but the same point is enclosed in the relative boundary of another partition set. The main consequence of Proposition 1 is that S can be replaced by the collection of its facets, which basically have a lower polytopal dimension and less vertices. This is mainly a theoretical observation. Keeping track of the facets of a polytope is a complex issue. One typical observation following Corollary 1 is the following.

Corollary 2. *Let F be a facet of $S \subseteq P$ in a branch and bound algorithm. If the conditions of Proposition 1 apply to F and $\mathrm{rint}(F) \subseteq \mathrm{rint}(P)$, then F can be rejected.*

The argument is that the relative boundary of F may contain a global minimum point, but the same point is enclosed in the relative boundary of another facet.

We now focus on the direction from a vertex of V pointing towards the relative interior, $r = y - v, v \in V$.

Proposition 2. *Let $S \subseteq P$ be partition set and $y \in \mathrm{rint}(S)$, such that for a vertex v of S, we have $\overline{(y-v)^T \nabla f(S)} < 0$. Then for all facets F of S which include vertex v, $\mathrm{rint}(F)$ cannot contain a global optimum point.*

Proof. Let $z \in \mathrm{rint}(F)$. The function value $f(z + \varepsilon r)$ is lower than $f(z)$ for $\varepsilon > 0$, because r is a decent direction as long as $z + \varepsilon r \in S$. This means, we have to show that $r = y - v$ is a feasible direction. For this, we should find a value for $\varepsilon \in \mathbb{R}^+$ small enough to show $z + \varepsilon(y - v) \in S$. Without loss of generality, number the vertices of S as $v_1, \ldots v_m$ with $v_1 = v$. Notice that F is a convex hull of a subset of those vertices. According to the definition of relative interior, we have that $y = \sum_1^m \mu_j v_j$ with $\mu_j > 0$. Let $z = \sum_1^m \lambda_j v_j$ with $\lambda_1 > 0$ be a point in the interior of a facet which contains v_1. This means that $x = z + \varepsilon r = (\lambda_1 + \varepsilon(\mu_1 - 1))v_1 + \sum_2^m (\lambda_j + \mu_j \varepsilon) v_j$ is a feasible point for $\varepsilon > \frac{\lambda_1}{1-\mu_1}$ with a function value $f(x)$ lower than in point z. So $z \in \mathrm{rint}(F)$ cannot be a minimum point of f. □

Corollary 3. *Let F be a facet of $S \subseteq P$ in a branch and bound algorithm including vertex v. If the conditions of Proposition 2 apply, then F can be rejected.*

The argument is that the relative boundary of F may contain a global minimum point, but the same point is enclosed in the relative boundary of another facet.

Example 2. For the illustration of the concept, consider the elements in Fig. 3. It shows a partition set S_2 with five vertices (re-)numbered as $v_1, \ldots, v_5$. If, for the orange direction $x - c$ we have $0 \notin [(x-c)^T \nabla f(S_2), \overline{(x-c)^T \nabla f(S_2)}]$, i.e. it is a monotonic direction, the interior of S_2 cannot contain an optimum. As pointed out by Corollary 2, the green faces need not be considered anymore. In [6], those are classified as non-border.

Now consider the direction $r = y - v_1$ and assume $\overline{(y-v)^T \nabla f(S_2)} < 0$ (blue arrow). We can not only conclude that vertex v_1 is not optimal, but according to Corollary 3, we can also leave out the facets $\mathrm{conv}(\{v_1, v_2\})$ and $\mathrm{conv}(\{v_1, v_5\})$ from consideration.

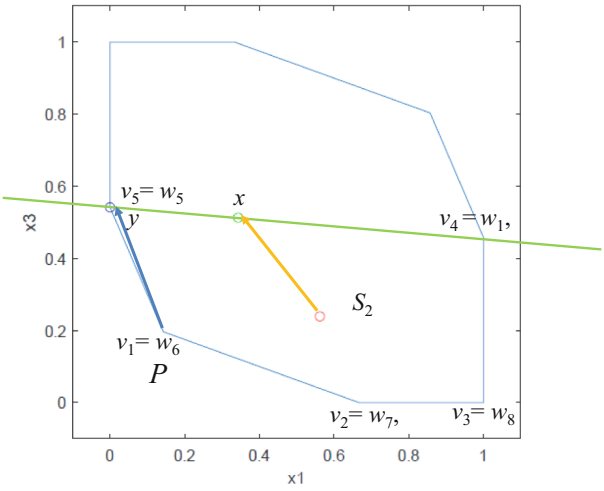

Fig. 3. 2D image from a 6 dimensional subset S_2 with renumbered vertices. The red dot is the centroid of S_2, the green dot the solution x of LP (6) and the blue dot the solution of LP (7) taking v_1 (w_6) as base. (Color figure online)

In the following section, we discuss a way to find the monotonic directions.

3.2 Finding Negative Directional Derivatives

We extend the earlier analysis of [3] for simplicial partition sets to polytopal partition sets. The analysis uses an enclosure G of the gradient based on interval enclosure $\square S$ of S. Notice again that the dimension d may be bigger than the polytopal dimension. This provides overestimation of $[\underline{\nabla f(S)}, \overline{\nabla f(S)}] \subset G$. A necessary, but not sufficient condition for the existence of a monotone direction is to have $0 \notin G$.

Example 3. Consider the 2-polytope in 6-dimensional introduced in Example 1. Due to the quadratic objective, the gradient $\nabla f(x) = (x - t)$ is linear. Bounds on the derivatives could be based on the 64 vertices of P, but also on the 8 vertices according to $\underline{\nabla f(P)}_i = \min_j(w_{ij} - t_i)$ and $\overline{\nabla f(P)}_i = \max_j(w_{ij} - t_i)$.

For the bound values we can write $[\underline{\nabla f(P)}, \overline{\nabla f(P)}]$ as the interval $[-0.2, 0.8] \times [-0.8, 0.2] \times [-0.8, 0.2] \times [-0.3, 0.7] \times [-0.8, 0.2] \times [-0.8, 0.2]$ which clearly includes 0. The idea is that partitioning leads to a tightening of the bounds. In our case, consider the partitioning in $S_1 := \text{conv}(\{w_1, w_2, w_3, w_4, w_5\})$ and $S_2 := \text{conv}(\{w_1, w_5, w_6, w_7, w_8\})$. This tightens $[\underline{\nabla f(S_1)}, \overline{\nabla f(S_1)}]$ towards $[-0.2, 0.8] \times [-0.5, 0.2] \times [-0.3, 0.2] \times [-0.3, 0.7] \times [-0.5, 0.2] \times [-0.8, 0.1]$ and $[\underline{\nabla f(S_2)}, \overline{\nabla f(S_2)}]$ is given by the interval $G :=$ $[-0.2, 0.8] \times [-0.8, -0.13] \times [-0.8, -0.26] \times [-0.3, 0.7] \times [-0.8, -0.1] \times [-0.7, 0.2]$ which does not include 0. This indicates, that f can be monotonic over S_2.

To prove that a monotonic direction exists, we need to find a direction $r = y - x$, with $x, y \in S$ corresponding to a negative upper bound of the directional derivative

$$0 > \overline{r^T G} = \sum_{i=1}^{d} \max\{r_i \underline{G}_i, r_i \overline{G}_i\}. \tag{5}$$

Finding such a direction can be done by looking for the steepest monotone direction $\min_r \overline{r^T G}$. Consider the terms $z_i = \min\{d_i \underline{G}_i, d_i \overline{G}_i\}$. This means we can write $\overline{r^T G} = \sum_{i=1}^{n} z_i$. If we fix one of the points $x \in S$ in $r = x - y$, then the term $z_i(y) = \max\{(x_i - y_i)\underline{G}_i, (x_i - y_i)\overline{G}_i\}$ is a convex function in x, as it is the maximum of two affine functions. Therefore, the upper bound on the directional derivative $\overline{g}(x) := \overline{(x-y)^T G}$ is a convex function being the sum of convex terms. We will show how an LP problem can be formulated to find a minimum of $\overline{g}(x)$.

Looking for the existence of a negative value of $\overline{r^T G}$ for some direction r, we can fix y to the centroid in the directional vector, i.e. $r = x - c$. The minimisation of convex piece-wise linear function $\overline{g}(x)$ over $x \in S$ can be formulated as the following LP problem.

$$\begin{aligned}
& \min \sum_{i=1}^{d} z_i \\
\text{s.t.} \quad & z_i \geq (x_i - c_i)\underline{G}_i, \ \ i = 1, \ldots, d \\
& z_i \geq (x_i - c_i)\overline{G}_i, \ \ i = 1, \ldots, d \\
& \sum_{j=1}^{m} \lambda_j = 1 \\
& y = \sum_{j=1}^{m} \lambda_j v_j \\
& \lambda_j \geq 0, j = 1, \ldots, m
\end{aligned} \tag{6}$$

If there is no monotone direction, then $x = c$ and $\sum_{i=1}^{d} z_i = 0$.

Example 4. Using the gradient enclosure G of subset S_2 we can now illustrate LP (6). We renumber the vertices $V_2 := \{v_1, v_2, v_3, v_4, v_5\}$. Given the gradient enclosure G and the centroid c of S_2, the result of LP (6) is a negative upper bound of $\sum_{i=1}^{d} z_i = -0.05$ for the vector $x = (0.34, 0.45, 0.51, 0.34, 0.56, 0.62)^T$ of which the components (x_1, x_3) are depicted as a point in Fig. 3. This means that f has been shown to be monotonic on S_2.

Following the line of reasoning of the example, according to Proposition 1, we can conclude that simplex S can be replaced by its facets after having shown a monotonic direction exists. Following Corollary 2, we can conclude for the example that the green facet (edge) between w_2 and v_5 can be rejected. This means for the example that S_2 can be replaced by its facets e_{18}, e_{78}, e_{67} and e_{56}. Figure 1 suggests that using Proposition 2, the edges which include w_7 and w_8

can also be removed as it seems that the direction $(x - c)$ is feasible. However, this is just a visual observation and not a mathematical rigorous proof. To show this mathematically, we have to find a feasible direction $r = y - v$ for which $\overline{r^T \nabla f(S_2)} < 0$ and have $y \in \text{rint}(S)$.

To answer the question whether such a direction exists, we can solve the following LP for a specific vertex v minimising the upper bound of a directional derivative. Consider now the re-ordered vertex set $V := \{v_1, v_2, \ldots, v_m\}$. Focusing on the direction $y - v_1$ with $y = \sum_{j=1}^{m} \lambda_j v_j$, we can demonstrate whether or not there is a (minimum) negative directional derivative by solving the LP

$$\begin{aligned} & \min \sum_{i=1}^{d} z_i \\ \text{s.t.} \quad & z_i \geq \underline{G}_i(y_i - v_{1i}), \ i = 1, \ldots, d \\ & z_i \geq \overline{G}_i(y_i - v_{1i}), \ i = 1, \ldots, d \\ & \sum_{j=1}^{m} \lambda_j = 1 \\ & y = \sum_{j=1}^{m} \lambda_j v_j \\ & \lambda_j \geq \delta, j = 1, \ldots, m, \end{aligned} \tag{7}$$

where $\delta > 0$ is a small accuracy, e.g. $\delta = 0.1^{-6}$. If the result is negative, we have proven we can reject faces that include vertex v_1 from consideration. If such a direction does not exist the solution will be $y = v_1$ and the objective function value is $\sum_{i=1}^{d} z_i = 0$.

Example 5. Using the vertex set $V_2 := \{v_1, v_5, v_6, v_7, v_8\}$ we will illustrate LP (7) taking in the ordered set as first vertex v_1. It feels that the result for w_7 and w_8 is more obvious due to the graphical observation in Fig. 3. Running LP (7) with $\delta = 10^{-6}$ gives an objective function value of $\sum_{i=1}^{d} z_i = -0.058$ and $y = (0.00, 0.33, 0.54, 0.0, 0.69, 0.89)^T$. To be more precise $y_1 = 2.7 * 10^{-6}$, yielding an interior point of S_2. This illustrates that for this example all of subset S_2 can be removed from consideration.

4 Lower Bounds in a Polytopal Subset and the Algorithm

Refining the partition can in general be done in many ways like using bisection or multi-section. Using a radial partition or a hyperplane to split parts. The idea of following a box (rectangular) shape seems attractive from a bounding point of view as distances within the partition set diminish and possibly underestimation of the objective function value is tighter. A generic pseudo-code of a B&B method based on bisection using monotonicity is provided in Algorithm 1. We sketch various ways to generate lower bounds in this section and discuss several options for the refinement in Sect. 5.

Algorithm 1. Polytopal branch and bound($f,\ P,\ \alpha$)

Require:
f: the d dimensional continuously differentiable objective function.
P: feasible set $P = \text{conv}(\{w_1, \ldots, w_q\})$
α: accuracy

Store P in Λ ▷ Set of partition sets
$X = \{w_1, \ldots, w_q\}$ ▷ List of points used as vertices
Evaluate and store $f(x)$ for all $x \in X$
$\overline{f} = \min_{x \in X} f(x)$ ▷ Incumbent value
$\underline{f} = \underline{f}(P)$
while $\overline{f} - \underline{f} > \alpha$ **do**
  Extract S from Λ ▷ Retrieve S from Λ with lowest value for $\underline{f}(S)$
  Evaluate gradient enclosure G of S
  if $0 \in G$ **then** Divide = true ▷ $0 \notin G$ is a necessary not sufficient condition
  else ▷ Trying to exploit monotonicity
    Divide = CheckMonotonicity($f,\ S,\ G, \Lambda$) ▷ If monotone store subsets in Λ
  if Divide **then**
    $\{S_1, S_2, \ldots, S_p\}$ = Refine(S) evaluating points x ▷ Refinement rule
    if $x \notin X$ **then**
      Evaluate and store $f(x)$ and add x to X
      if $f(x) < \overline{f}$ **then**
        $\overline{f} = f(x)$
        Λ = CutOff(Λ) ▷ Remove subsets $S \in \Lambda$ with $\underline{f}(S) > \overline{f}$
    for each new subset S_j **do**
      Evaluate $\underline{f}(S_j)$
      if $\underline{f}(S_j) \leq \overline{f}$ **then** ▷ RangeUp test
        Store S_j in Λ ▷ Store S_j and its bounds in Λ
  $\underline{f} = \min_{S \in \Lambda} \underline{f}(S)$
return $\cup_{S \in \Lambda} S$ and $[\underline{f}, \overline{f}]$

An overview of concepts for lower bounding in B&B is provided in the book [8]. More specifically, we discussed lower bounds on a simplicial feasible set based on bounds on the derivative G in [9]. The basis is provided by the Taylor view provided a (or some) base point $b \in S$

$$f(x) \geq f(b) + \underline{(x-b)^T \nabla f(b)} \qquad x \in S. \tag{8}$$

That study used several small full dimensional instances from [11] and showed how an LP using the function values in the vertices of the enclosing box gave most effective bounds. With the example in mind, this looks very unattractive as soon as the dimension of P is far less than d. Specifically, in the example case, the function value and also the derivatives are far higher in absolute value than that at the vertices. Therefore, we will use a small illustration of a lower bound from [9] based on a center form that uses only the centroid of P as base point, although it is less sophisticated than an LP. Given a gradient range $[\underline{\nabla f(S)}, \overline{\nabla f(S)}]$, the lower bound of the directional derivative in direction $x - c$,

$$\underline{(x-c)^T \nabla f(S)} = \sum_{i=1}^{n} \min\{(x_i - c_i)\underline{\nabla_i f(S)}, (x_i - c_i)\overline{\nabla_i f(S)}\}. \tag{9}$$

As (9) is concave in x, its minimum is taken in one of the vertices. Notice that this requires evaluation of the objective function at the centroid:

$$\varphi_c(S) := f(c) + \min_{v \in V} \underline{(v-c)^T \nabla f(S)}. \tag{10}$$

Example 6. The function value in the centroid of P is $f(c) = 0.245$. Using $[\underline{\nabla f(P)}, \overline{\nabla f(P)}]$, the lower bound yields $\varphi_c(P) = -1.52$. After the partition we have the function value at the centroid of S_1 being $f(c) = 0.155$. Actually, this value improves the incumbent best value $\overline{f} := \min_{w \in W} f(w) = f(w_3) = 0.204$. The gradient bounds $[\underline{\nabla f(S_1)}, \overline{\nabla f(S_1)}]$ tighten and also the maximum distance to the centroid gets smaller yielding a lower bound of $\varphi_c(S_1) = -1.12$. For the partition S_2, we have $f(c) = 0.53$ and using gradient bounds $[\underline{\nabla f(S_2)}, \overline{\nabla f(S_2)}]$ the lower bound yields a value of $\varphi_c(S_2) = -1.05$.

The example illustrates how the gradient bounds tighten and also the lower bound increases when bisection is applied. However, notice that the centroid form (10) does not guarantee that the lower bound increases, as the function values in the centroid do not increase with a refinement.

5 Ideas for Refining Subsets

When lower bounding and monotonicity checks are not sufficient to remove a subset, we have to divide it into two (bisection) or more subsets. There are several possibly conflicting criteria to keep in mind:

1. Preferably a subset should be "round" to keep bounds tight
2. We like polytopal subsets to have as few vertices as possible
3. Division cuts should preferably follow coordinate planes for getting tight interval hull bounds on subsets
4. It should be easy to keep track of the faces of the subset for the dimension reduction in case of monotonicity.

5.1 Radial Refinement

One of the refinements promoted by [10] is that of using a radial refinement along the centroid instead of bisecting the subset. The concept is sketched for the 2D example in Fig. 4. The procedure is to connect the centroid of the subset to all faces. It may be clear, that this way of refinement does not promote "round" subsets as the longest edge is not made shorter. For the example, the number of vertices is drastically reduced and one could proceed with a simplicial B&B based on longest edge bisection (LEB).

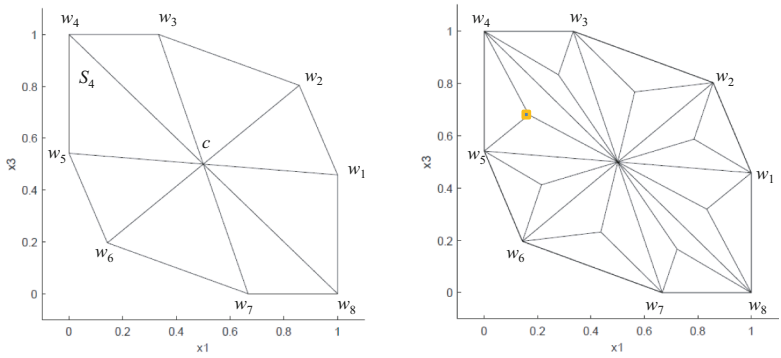

Fig. 4. Radial refinement of the 2D polytope

Example 7. For the used example, we have the radial refinement as sketched in Fig. 4 This refinement fits the lower bounding calculation of Sect. 4. Indeed, one would find a new incumbent, as the centroid of subset conv$\{w_4, w_5, c\}$) has a function value of 0.05. However, all lower bounds do not exceed the value of -0.18, such that no subset can be eliminated due to a cut-off and further refinement is necessary to the right part of the figure which contains 24 subsets.

It is important to recognise, that the convergence towards simplicial subsets is not guaranteed, as one of the facets may contain vertices that do not imply a set of affine independent points. An extreme case is the unit box and its face graph as sketched in Fig. 5.

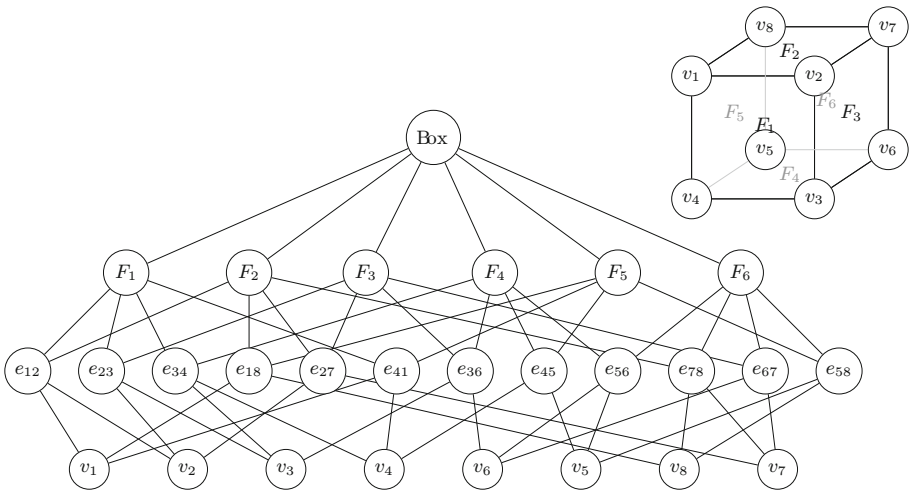

Fig. 5. Face graph of 3D box as polytopal feasible set

5.2 Component Wise Refinement

Interval arithmetic algorithms rely on bisecting (or multi-secting) an enclosing box (interval) along the longest edge. This certainly avoid needle shaped subsets. As sketched in Fig. 6, bisecting along component c_i of the centroid of the enclosing box implies finding the facets of S that include c_i and cutting them along the edges. This generates more vertices. The component wise width is halved after d bisections.

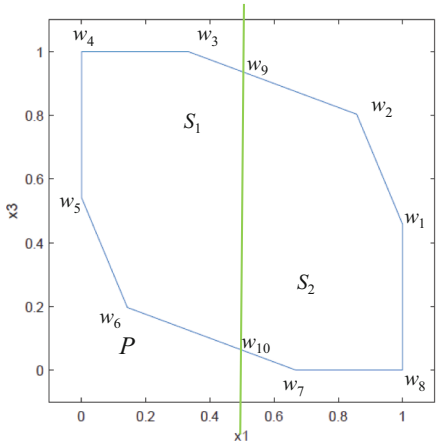

Fig. 6. Component wise bisection along c_1

Example 8. For the used example, facets $\text{conv}(\{w_2, w_3\})$ and $\text{conv}(\{w_6, w_7\})$ include $c_1 = 0.5$ and are split along new vertices $w_9 = (0.5, 1, 0.94, 0, 69, 0, 79, 0.21)^T$ and $w_{10} = (0.5, 0, 0.06, 0.31, 0.21, 0.79)^T$ respectively generating two subsets, $S_1 = \text{conv}(\{w_3, w_4, w_5, w_6, w_9, v_{10}\})$ and $S_2 = \text{conv}(\{w_1, w_2, w_7, w_8, w_9, w_{10}\})$. The evaluated vertices can be renumbered within the subsets of course. Actually for the generic case, it is not known how many new vertices are generated with a box bisection.

5.3 Vertex Wise Bisection Reducing Number of Vertices

An alternative which focuses on vertex reduction is to use n vertices that are not in the same facet and of which one of the components is as close to the centroid as possible to force volume reduction. The idea to select vertices is that in the end, we end up with subsets with $m = n + 1$ vertices and from there can use simplicial bisection of the longest edge. There is still some difficulty with degeneracy though; more than n vertices can lay on the separating hyperplane leading to non-simplicial facets.

Let $U := \{u_1, u_2, \ldots, u_\ell\}$ be groups of n affine independent vertices $u_j \subset V, |u_j| = n$ that are not in a facet of S. The subset S is then split into S_1, S_2

using the normal vector p of u_j (see, [9]) and the hyperplane $x^T p = p^T v_\ell, v_\ell \in u_j$. Consider the simplex conv(u), within subset conv(V), a vertex in the selected subset $v_\ell \subset u$ and a vertex of S outside of u, $v \in V \setminus u$. Construct a set of vectors $\mathcal{Y} := \{y - v_\ell, y \in u \setminus \{v_\ell\}\}$ with corresponding matrix representation Y. The normal vector of conv(u) in the direction of v can be expressed by defining

$$p := \left[I - Y(Y^T Y)^{-1} Y^T\right](v - v_\ell). \tag{11}$$

For finding the best subset u we can take as selection criterion to find vertices relatively close to the centroid with respect to one of the components

$$\min_{u \in U} \min_{i=1,\ldots,d} \max_{v \in u} \frac{|v_i - c_i|}{\overline{x}_i - \underline{x}_i},$$

where $\overline{x}_i - \underline{x}_i$ is the component width of the enclosing box of S.

Example 9. This rule is illustrated in Fig. 7 for subset S_1 assuming that S_2 is rejected due to monotonicity. Initially $u = \{v_1, v_5\}$ had been selected to bisect S into S_1, S_2. For $S_1, U = \{\{v_1, v_3\}, \{v_1, v_4\}, \{v_2, v_4\}, \{v_2, v_5\}, \{v_3, v_5\}\}$. The plane through v_1, v_4 is relatively closest (value 0.157 of the criterion) to c_2 (component 2). Further bisection is done through v_1, v_3 which splits component $i = 6$ best. After these bisections, S_1 has been partitioned into simplices. It is not directly clear from the x_1, x_3 projection that this is the best way to proceed, such that we depict in the figure also the x_2, x_6 projection. From the final partition, a simplicial B&B can be started, where it is easier to keep track of the facets.

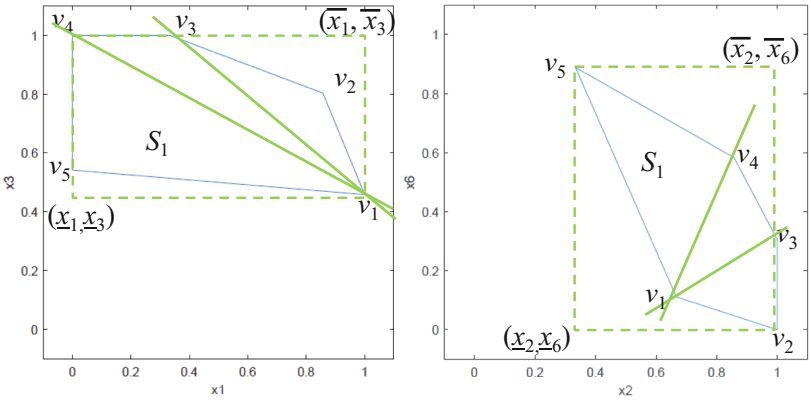

Fig. 7. Bisection along vertex hyperplanes close to c. The same content left and right, other 2D projection.

For the illustration of the calculation of the normal vector which defines the bisection along v_1, v_4, we can use reference vectors $v = v_5$ and $v_\ell = v_1$. This provides us with normal vector $p = (-0.10, -0.51, -0.41, -0.27, -0.24, 0.35)^T$, which is perpendicular to the plane through v_1, v_4.

If degeneracy occurs, more than n vertices are located in the separating hyperplane. This is illustrated in Fig. 8 for splitting the unit box. This means that the best group u is not unique. However, we can still define the corresponding hyperplane and normal vector.

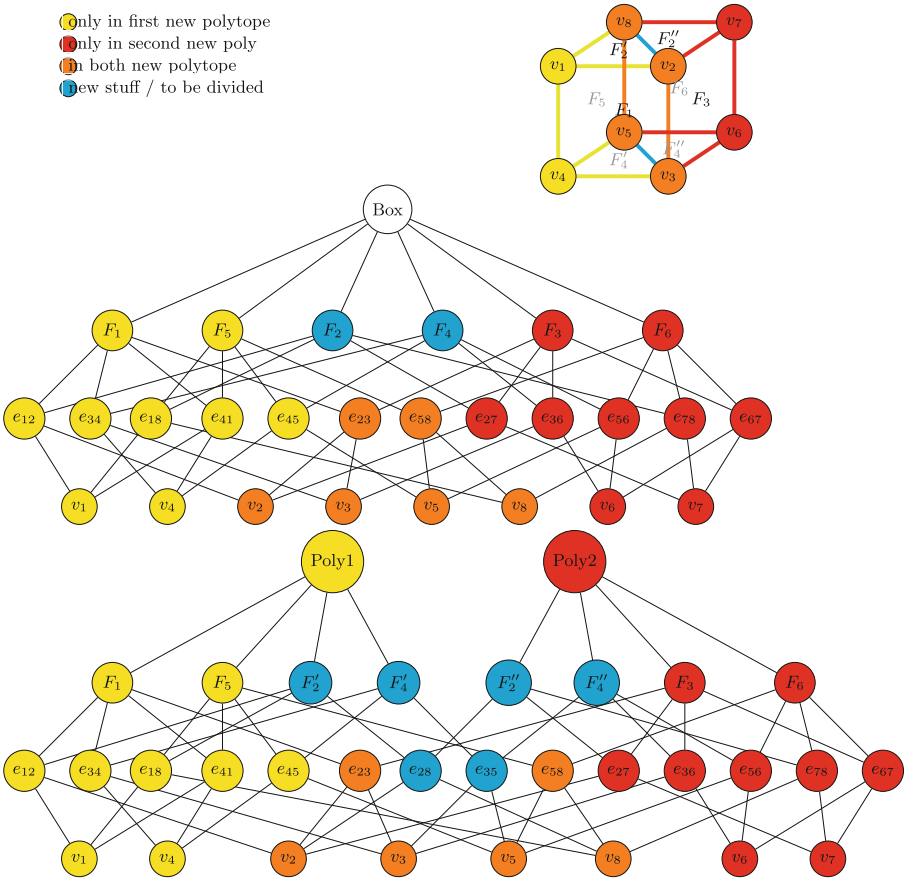

Fig. 8. Face graph of bisected 3D box

5.4 Bisecting Coordinates Along One Vertex

Using only vertices has as characteristic that we do not follow the coordinate wise splitting to promote sharp enclosures. Moreover, focusing on reduction of the width of the enclosing box does not take care of scaling with respect to the gradients. One target is to try to promote monotonicity to have dimension reduction. We can do so by taking one one hand the width of $\overline{G}_i - \underline{G}_i$ into account and on the other hand look for a vector coordinate closest to to the centroid

$$\min_{v \in V} \min_{i=1,\ldots,d} \frac{|v_i - c_i|}{\overline{G}_i - \underline{G}_i}$$

and use $x_i = v_i$ as bisection hyperplane. This does not require using the normal vector expression (11) and generates at most $n-1$ new vertices in each iteration. This procedure is an intermediate between coordinate wise bisection and using vertices. It also converges to a simplicial B&B.

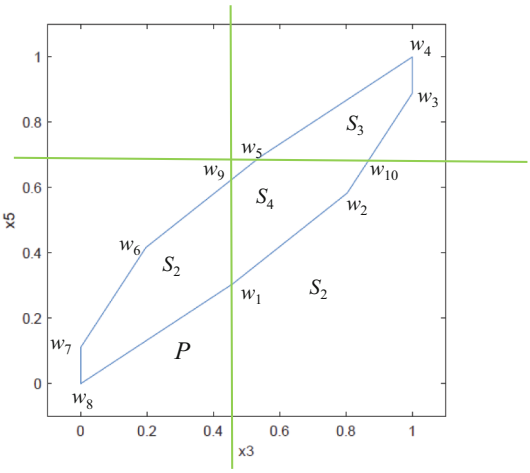

Fig. 9. Bisection along one vertex along a coordinate close to c.

Example 10. This rule is illustrated in Fig. 9. First of all, the criterion gives that we can bisect using the coordinate $i = 3$ of vertex v_1. This is similar as the bisection along vertices into S_1 and S_2, such that S_2 provides a monotonicity and can be removed. The example is very symmetric and there are two other candidates that could have been used with the same criterion value. The difference with a hyperplane supported by vertices is that we now generate a new vertex w_9; in total $n - 1 = 1$ vertices. We now focus on the resulting subset S_1 to be split into S_3 and S_4. There is less symmetry in the gradient bounds and the criterion provides a unique bisection along component $i = 5$ of w_5 generating new vertex w_{10}. To illustrate this, we adapted the projection in the figure. Although not visible, $0 \notin [\underline{\nabla} f(S_4), \overline{\nabla} f(S_4)]$. However, this desired effect is still not sufficient to find a monotonic direction according to (6). It is interesting to see that the number of vertices of the subsets is decreasing when we go deeper into the B&B tree.

6 Conclusions

Our question deals with the use of polytopes as feasible set and partition sets in global optimization branch and bound. The findings of monotonicity of simplicial B&B can be extended, although keeping track of facets to be used as subsets is a bigger challenge. Ideas of lower bounding can be based on vertex evaluations,

but also on vertices of an enclosing box and the centroid. What is the best way to refine the partition sets requires a systematic investigation. One option is to split along vertex hyperplanes reducing the number of vertices and converge to simplicial B&B. Another option is to follow interval bisection rules which may provide tighter bounds, but this results in an increase in the number of vertices for the subsets. Another rule is to combine the two ways of providing a bisection. The idea is to promote convergence to a simplicial B&B for which effective bounding has been defined. These theoretical findings provoke further investigation into the effectiveness of rules in a computational setting.

References

1. Casado, L.G., G.-Tóth, B., Hendrix, E.M.T., Messine, F.: On monotonicity detection in simplicial branch and bound over a simplex. In: Gervasi, O., Murgante, B., Misra, S., Rocha, A.M.A.C., Garau, C. (eds.) ICCSA 2022. LNCS, vol. 13378, pp. 113–126. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-10562-3_9
2. Casado, L.G., G.-Tóth, B., Hendrix, E.M.T., Messine, F.: Local search versus linear programming to detect monotonicity in simplicial branch and bound. J. Glob. Optim. (2023)
3. Casado, L.G., G.-Tóth, B., Messine, F., Hendrix, E.M.T.: Directional derivative bounds and border facets in simplicial B&B monotonicity tests. In: SCAN 2020: The 19th International Symposium on Scientific Computing, Computer Arithmetic and Verified Numerical Computations, pp. 18–19. Szeged (Hungary) (2021)
4. G.-Tóth, B., Casado, L.G., Hendrix, E.M.T., Messine, F.: On new methods to construct lower bounds in simplicial branch and bound based on interval arithmetic. J. Glob. Optim. **80**(4), 779–804 (2021)
5. G.-Tóth, B., Hendrix, E.M.T., Casado, L.G.: On monotonicity and search strategies in face based copositivity detection algorithms. Central Eur. J. Oper. Res. **30**, 1071–1092 (2021)
6. G.-Tóth, B., Hendrix, E.M.T., Messine, F., Casado, L.G.: On dealing with minima at the border of a simplicial feasible area in simplicial branch and bound. J. Optim. Theory Appl. (2024). https://doi.org/10.1007/s10957-024-02480-9
7. Hansen, E., Walster, W.: Global Optimization Using Interval Analysis, 2nd edn. Marcel Dekker Inc., New York (2004)
8. Hendrix, E.M.T., B.-Tóth, G.: Introduction to Nonlinear and Global Optimization. Springer, New York (2010). https://doi.org/10.1007/978-0-387-88670-1
9. Hendrix, E.M.T., G.-Tóth, B., Messine, F., Casado, L.G.: On derivative based bounding for simplicial branch and bound. RAIRO **55**(3), 2023–2034 (2021)
10. Horst, R.: On generalized bisection of n-simplices. Math. Comput. **66**(218), 691–699 (1997)
11. Karhbet, S.D., Kearfott, R.B.: Range bounds of functions over simplices, for branch and bound algorithms. Reliable Comput. **25**, 53–73 (2017). Reliable-Computing-journal
12. Kearfott, R.B.: An interval branch and bound algorithm for bound constrained optimization problems. J. Global Optim. **2**(3), 259–280 (1992)

13. Moore, R.E., Kearfott, R.B., Cloud, M.J.: Introduction to Interval Analysis. Society for Industrial and Applied Mathematics, USA (2009)
14. Rall, L.B.: Examples of software for automatic differentiation and generation of Taylor coefficients. In: Rall, L.B. (ed.) Automatic Differentiation: Techniques and Applications. LNCS, vol. 120, pp. 54–90. Springer, Heidelberg (1981). https://doi.org/10.1007/3-540-10861-0_5