

## ON DERIVATIVE BASED BOUNDING FOR SIMPLICIAL BRANCH AND BOUND

ELIGIUS M. T. HENDRIX<sup>1,\*</sup>, BOGLARKA G.-TÓTH<sup>2</sup>, FREDERIC MESSINE<sup>3</sup> AND LEOCADIO G. CASADO<sup>4</sup>

**Abstract.** Simplicial based Global Optimization branch and bound methods require tight bounds on the objective function value. Recently, a renewed interest appears on bound calculation based on Interval Arithmetic by Karhbet and Kearfott [*Reliable Comput.* **25** (2017) 53–73] and on exploiting second derivative bounds by Mohand [*RAIRO Oper. Res.* **55** (2021) S2373–S238]. The investigated question here is how partial derivative ranges can be used to provide bounds of the objective function value over the simplex. Moreover, we provide theoretical properties of how this information can be used from a monotonicity perspective to reduce the search space in simplicial branch and bound.

**Mathematics Subject Classification.** 65K05, 90C30, 90C34.

Received December 22, 2020. Accepted May 12, 2021.

### 1. INTRODUCTION

Deterministic global optimization based on the concept of branch and bound looks for the set of global minimum points up to a certain accuracy by rejecting regions where it is guaranteed that the global minimum cannot be found. An overview of the knowledge in the nineties is the handbook [5] of Reiner Horst and Herbert Tuy in 1990. It describes various bounding techniques and partition sets based on cones, boxes and simplices. The latter is rather natural in blending type of problems [2], where the feasible set is given as the standard simplex. The focus of the further study of Reiner Horst has been on the convergence of several subdivision methods. Meanwhile, Paulavičius and Žilinskas elaborated on practical implementations and provide a nice summary in the handbook [10]. Their focus is among others on the simplicial division of the box into  $n!$  simplices as introduced in the seventies by Mike Todd [11] and bounds mainly based on Lipschitz constants.

Recently, there appeared a renewed interest in different ways to provide tight bounds over simplices. In [6], Karhbet and Kearfott take the point of view from Interval Arithmetic, where the paper provides a nice illustration based on 7 instances. The same illustration has been used in a contribution by Ouanes Mohand to

---

*Keywords.* Simplex, branch and bound, derivative.

<sup>1</sup> Universidad de Málaga, Av Cervantes 2, Málaga, Spain and Wageningen University, Hollandseweg 1, Wageningen, The Netherlands.

<sup>2</sup> Computational Optimization Dept., University of Szeged, Árpád tér 2, Szeged, Hungary.

<sup>3</sup> LAPLACE-ENSEEIHT, Toulouse-INP, University of Toulouse, 2 rue Charles Camichel, Toulouse, France.

<sup>4</sup> Informatics Dept., University of Almería (CeiA3), s/n, 04120 La Cañada, Almería, Spain.

\*Corresponding author: [eligiush@uma.es](mailto:eligiush@uma.es)

this journal [9], that studies different ways to use a bound on the second derivative to create objective function bounding over a simplex.

As we have been working on simplicial bounds for some time, our curiosity is how first derivative ranges can be used to create objective function bounds over the simplex and how the different ways compare to earlier results for the 7 instances. The rest of this paper is organised as follows. Section 2 describes the mathematical question and instances we are investigating. Section 3 describes several ways to derive bounds of a function over a simplex based on derivative ranges. Section 4 describes how monotonicity considerations may play a role for practical simplicial branch and bound algorithms. Section 5 describes the findings of our investigation.

## 2. MATHEMATICAL DESCRIPTION OF THE QUESTION

One of the questions in simplicial branch and bound algorithms as outlined in [10] is how to obtain good bounds over the simplicial partition set. Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a twice continuously differentiable function over a simplicial set. Let  $\mathcal{V} = \{v_0, v_1, \dots, v_n\}$  be a set of  $n+1$  affinely independent vertices of  $n$ -simplex  $\Delta = \text{conv}(\mathcal{V})$ . Let the function range be denoted by  $[f^*, \bar{f}] := [\min_{x \in \Delta} f(x), \max_{x \in \Delta} f(x)]$ . Then, the question posed in [6, 9] is to find a sharp enclosure  $\mathbf{F} \supset [f^*, \bar{f}]$ . Considering the minimization of function  $f$ , our question will focus on finding a tight lower bound  $\varphi \leq f^*$ .

We also introduce the minimum enclosing box  $X = [\underline{x}, \bar{x}] \supset \Delta$ , given component-wise by  $\underline{x}_i = \min_{v \in \mathcal{V}} v_i$  and  $\bar{x}_i = \max_{v \in \mathcal{V}} v_i$ . For the notation, we will use as much as possible the index  $i$  for the component and index  $j$  for an element of a set when necessary. The set of vertices of the box  $X$  are denoted by  $\mathcal{W}$ , *i.e.*

$$\mathcal{W} := \{w \in X \mid w_i = \underline{x}_i \text{ or } w_i = \bar{x}_i, i = 1, \dots, n\},$$

and we use  $w$  for any vertex of  $X$ . Our specific question is what happens when the function values at the vertices are available, but also the gradient range defined by  $[\nabla f, \bar{\nabla} f]$ . Karhbet and Kearfott [6] discuss some variants for the enclosure  $\mathbf{F}$  taking derivative bounds over box  $X$  rather than simplex  $\Delta$ . Our focus is on how this information can be used to provide a lower bound  $\varphi \leq f^*$ . We will describe various methods in Section 3 and illustrate them on the instances introduced in [6, 9].

### 2.1. Illustrative instances

In total 7 instances were introduced based on 3 different 2-dimensional functions and varying vertex sets.

- $f_1(x) = x_1^2 + x_2^3$ ,  $\nabla f_1(x) = (2x_1, 3x_2^2)^T$ ,  $H_1(x) = \begin{pmatrix} 2 & 0 \\ 0 & 6x_2 \end{pmatrix}$ , convex for  $x_2 \geq 0$ .
- $f_2(x) = \frac{1}{4}x_1^2 + x_1 + x_2 + \frac{1}{4}x_1x_2 + \frac{1}{4}x_2^3$ ,  $\nabla f_2(x) = (1 + \frac{1}{2}x_1 + \frac{1}{4}x_2, 1 + \frac{1}{4}x_1 + \frac{3}{4}x_2^2)^T$ ,  
 $H_2(x) = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{2}x_2 \end{pmatrix}$ , convex for  $x_2 \geq \frac{1}{12}$ .
- $f_3(x) = \frac{1}{4}x_1^2 + x_1 + x_2 + \frac{1}{4}x_1x_2 + \frac{1}{2}x_2^2$ ,  $\nabla f_3(x) = (1 + \frac{1}{2}x_1 + \frac{1}{4}x_2, 1 + \frac{1}{4}x_1 + x_2)^T$ ,  
 $H_3(x) = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & 1 \end{pmatrix}$ , convex.

It is typical that the introduced polynomial functions in paper [6] which focuses on Interval Arithmetic have fractional coefficients which can be represented exactly in a computer. This means that function evaluations are also exact, when points (vertices) are evaluated with the same property. We can observe this effect when we consider the data in Table 1. Up to the last instance, all vertices are numbers represented exactly in a computer and likewise the minima over the simplices.

We first determined analytically the derivative ranges over the enclosing box  $X$ . One can observe that apart from the first instance, all functions are monotonic over the corresponding simplex  $\Delta = \text{conv}(\mathcal{V})$  given in Table 1. This implies there is no interior minimum point. However, one cannot conclude from the enclosures of

TABLE 1. The 7 instances introduced in [6] with the corresponding minimum function value  $f^*$  and the gradient range over the enclosing box  $X$  and simplex  $\Delta$ .

Inst.	$f$	$f^*$	$V$	$X = [\underline{x}, \bar{x}]$	$[\nabla f, \bar{\nabla} f]_X$	$[\nabla f, \bar{\nabla} f]_\Delta$
1	$f_1$	$-\frac{3}{4}$	$\begin{bmatrix} -1 & \frac{1}{2} & \frac{1}{2} \\ 0 & -1 & 1 \end{bmatrix}$	$\begin{pmatrix} \left[ -1, \frac{1}{2} \right] \\ \left[ -1, 1 \right] \end{pmatrix}$	$\begin{pmatrix} [-2, 1] \\ [0, 3] \end{pmatrix}$	$\begin{pmatrix} [-2, 1] \\ [0, 3] \end{pmatrix}$
2	$f_2$	$-\frac{13}{16}$	$\begin{bmatrix} -1 & \frac{1}{2} & \frac{1}{2} \\ 0 & -1 & 1 \end{bmatrix}$	$\begin{pmatrix} \left[ -1, \frac{1}{2} \right] \\ \left[ -1, 1 \right] \end{pmatrix}$	$\begin{pmatrix} \left[ \frac{1}{4}, \frac{3}{2} \right] \\ \left[ \frac{3}{4}, \frac{15}{8} \right] \end{pmatrix}$	$\begin{pmatrix} \left[ \frac{1}{2}, \frac{3}{2} \right] \\ \left[ \frac{3}{4}, \frac{15}{8} \right] \end{pmatrix}$
3	$f_2$	$-\frac{15}{64}$	$\begin{bmatrix} -\frac{1}{4} & \frac{1}{8} & \frac{1}{8} \\ 0 & -\frac{1}{4} & \frac{1}{4} \end{bmatrix}$	$\begin{pmatrix} \left[ -\frac{1}{4}, \frac{1}{8} \right] \\ \left[ -\frac{1}{4}, \frac{1}{4} \right] \end{pmatrix}$	$\begin{pmatrix} \left[ \frac{13}{16}, \frac{9}{8} \right] \\ \left[ \frac{15}{16}, \frac{69}{64} \right] \end{pmatrix}$	$\begin{pmatrix} \left[ \frac{7}{8}, \frac{9}{8} \right] \\ \left[ \frac{15}{16}, \frac{69}{64} \right] \end{pmatrix}$
4	$f_2$	$-\frac{33}{4}$	$\begin{bmatrix} -2 & 2 & 0 \\ 0 & -3 & 3 \end{bmatrix}$	$\begin{pmatrix} [-2, 2] \\ [-3, 3] \end{pmatrix}$	$\begin{pmatrix} \left[ -\frac{3}{4}, \frac{11}{4} \right] \\ \left[ \frac{1}{2}, \frac{33}{4} \right] \end{pmatrix}$	$\begin{pmatrix} \left[ 0, \frac{7}{4} \right] \\ \left[ \frac{1}{2}, \frac{33}{4} \right] \end{pmatrix}$
5	$f_2$	$-\frac{15}{64}$	$\begin{bmatrix} -\frac{1}{4} & \frac{1}{4} & 0 \\ 0 & -\frac{3}{8} & \frac{3}{8} \end{bmatrix}$	$\begin{pmatrix} \left[ -\frac{1}{4}, \frac{1}{4} \right] \\ \left[ -\frac{3}{8}, \frac{3}{8} \right] \end{pmatrix}$	$\begin{pmatrix} \left[ \frac{25}{32}, \frac{39}{32} \right] \\ \left[ \frac{15}{16}, \frac{299}{256} \right] \end{pmatrix}$	$\begin{pmatrix} \left[ \frac{7}{8}, \frac{35}{32} \right] \\ \left[ \frac{15}{16}, \frac{299}{256} \right] \end{pmatrix}$
6	$f_3$	$-\frac{15}{64}$	$\begin{bmatrix} -\frac{1}{4} & \frac{1}{4} & 0 \\ v_0 & -\frac{3}{8} & \frac{3}{8} \end{bmatrix}$	$\begin{pmatrix} \left[ -\frac{1}{4}, \frac{1}{4} \right] \\ \left[ -\frac{3}{8}, \frac{3}{8} \right] \end{pmatrix}$	$\begin{pmatrix} \left[ \frac{25}{32}, \frac{39}{32} \right] \\ \left[ \frac{9}{16}, \frac{23}{16} \right] \end{pmatrix}$	$\begin{pmatrix} \left[ \frac{7}{8}, \frac{35}{32} \right] \\ \left[ \frac{11}{16}, \frac{11}{8} \right] \end{pmatrix}$
7	$f_3$	$\frac{19}{2}$	$\begin{bmatrix} 1 & 1.2 & 0.6 \\ 3 & 3.4 & 3.8 \end{bmatrix}$	$\begin{pmatrix} [0.6, 1.2] \\ [3, 3.8] \end{pmatrix}$	$\begin{pmatrix} [2.05, 3.45] \\ [4.15, 5.1] \end{pmatrix}$	$\begin{pmatrix} [2.25, 2.45] \\ [4.25, 4.95] \end{pmatrix}$

the gradients that the minimum is attained at a vertex; one can conclude that it can be found on the boundary, which for the 2D cases is an edge of the simplex, see [4]. In a practical branch and bound context, one can reduce the dimension and work with lower dimensional simplices as partition sets. Specifically for all instances, the minimum  $f^*$  is attained at a vertex. We will discuss the theoretical background for these observations in Section 4.

As mentioned in [6], the gradient range is of course tighter over the simplex  $\Delta$  than over the enclosing box  $X \supset \Delta$ . In Table 1, the ranges over the simplex are given, computed in an analytical way.

### 3. VARIOUS BOUNDS BASED ON DERIVATIVE INFORMATION

#### 3.1. 1-norm Lipschitz relaxation

In [10] Chapter 2, an overview is given of the use of lower bounds based on the so-called Lipschitz constant. In the context of differentiable functions, the tightest Lipschitz constant over the simplex is given by

$$L_{\Delta p} = \max_{x \in \Delta} \|\nabla f(x)\|_p \quad (3.1)$$

for different norms represented by the parameter  $p$ . From any vertex  $v \in \mathcal{V}$ , we have the relaxation

$$f(x) \geq f(v) - L_{\Delta p} \|x - v\|_{\frac{p}{p-1}} \quad \forall x \in \Delta, \quad (3.2)$$

where  $\frac{p}{p-1}$  gives the dual norm of  $p$ , e.g.  $\|\cdot\|_1 \leftrightarrow \|\cdot\|_\infty$  or  $\|\cdot\|_2 \leftrightarrow \|\cdot\|_2$ . A lower bound of  $f$  over  $\Delta$  can be found by solving

$$\min_{x \in \Delta} \max_{v \in \mathcal{V}} \left\{ f(v) - L_{\Delta p} \|x - v\|_{\frac{p}{p-1}} \right\}. \quad (3.3)$$

As (3.3) enhances the minimization of the maximum over concave functions, this implies in general solving a global optimization problem. Instead of explicitly solving problem (3.3), one often takes as lower bound the best relaxation over all the vertices, e.g. [3]. Written in terms of 1-norm distance [10], this is

$$\varphi_L := \max_{v \in \mathcal{V}} \left\{ f(v) - L_{\Delta \infty} \max_{u \in \mathcal{V}} \|u - v\|_1 \right\}. \quad (3.4)$$

This is one of the bounds used in [10] to derive simplicial function bounds. Notice that for the 7 instances, the expression in (3.1) to be maximized is a convex function, such that  $L_{\Delta \infty}$  can be found by evaluating the infinite norm of the gradient at the vertices. The corresponding lower bound  $\varphi_L$  is given in Table 2 as decimal number for the 7 instances.

For a box shaped partition set as  $X$ , Meewella and Mayne in [7] use the 1-norm distance, where (3.3) becomes an LP problem. Notice that the Lipschitz constant should then also be taken over  $X$ :

$$L_{X \infty} := \max_{x \in X} \|\nabla f(x)\|_\infty = \max_{x \in X} \max_i |\nabla_i f(x)|. \quad (3.5)$$

We add to the lower bounding LP of [7] the restriction of optimizing over the simplex  $\Delta$ . To capture the 1-norm distance from a box vertex without using the absolute value function, we introduce the vector  $I_w$  for all vertices  $w \in \mathcal{W}$ :

$$I_{iw} = 1 \text{ if } w_i = \underline{x}_i, \text{ else } I_{iw} = -1, \quad i = 1, \dots, n.$$

Using the function values  $f(w)$  on the vertices  $w \in \mathcal{W}$  and the Lipschitz constant  $L_{X \infty}$  over the box (3.5), the lower bound is given by

$$\begin{aligned} \varphi_M := \min z \\ \text{s.t. } z &\geq f(w) - L_{X \infty} I_w^T (y - w), \quad w \in \mathcal{W} \\ y &= \sum_{k=0}^n v_k \lambda_k \\ \sum_{k=0}^n \lambda_k &= 1 \\ \lambda_k &\geq 0. \end{aligned} \quad (3.6)$$

The obtained lower bounds for the 7 instances are given in Table 2. Notice that the determination of the lower bound  $\varphi_M$  requires solving a small LP problem which may imply calling an external routine. The computation of lower bound  $\varphi_L$  will be less time consuming, although the efficiency depends on the implementation of course. One can observe that both lower bounds are far from the minimum and that in all cases  $\varphi_L$  is worse than the more sophisticated  $\varphi_M$ , i.e.  $\varphi_L < \varphi_M$ . This tendency is general, but it is not a mathematical truth. It is possible to design degenerate cases (see Example 3.1) where  $\varphi_L > \varphi_M$ , which requires the gradient range over the simplex  $\Delta$  to be much smaller in absolute value than the gradient range over the box  $X$ .

**Example 3.1.** Consider the function  $f(x) = x_1^2 x_2^2$  over the simplex defined by vertex set  $\{(-1, -1)^T, (2, 0)^T, (0, 2)^T\}$ . The minimum of 0 is attained at a minimizer set with infinitely many elements, where  $x_1 = 0$  or  $x_2 = 0$ . The bounds for this instance are given in Table 3. For this specific case, we have that  $\varphi_L = 1 - 2.37 \times 4 = -8.48$ . Solving the LP (3.6) with the data of this instance provides  $\varphi_M = -39.5$ .

TABLE 2. The minimum  $f^*$  over  $\Delta$  rounded to a decimal number,  $\infty$ -Lipschitz constant over the enclosing box  $X$  and simplex  $\Delta$  and lower bounds  $\varphi_L$  based on (3.4),  $\varphi_M$  based on (3.6).

Inst.	$f^*$	$L_{X\infty}$	$L_{\Delta\infty}$	$\varphi_L$	$\varphi_M$
1	-0.75	3	3	-6.25	-4.625
2	-0.8125	$\frac{15}{8}$	$\frac{15}{8}$	-2.75	-2.5625
3	-0.2344	$\frac{9}{8}$	$\frac{9}{8}$	-0.3125	-0.3066
4	-8.25	$\frac{33}{4}$	$\frac{33}{4}$	-56.25	-38.75
5	-0.2344	$\frac{39}{32}$	$\frac{299}{256}$	-0.7798	-0.3813
6	-0.2344	$\frac{23}{16}$	$\frac{11}{8}$	-0.9297	-0.502
7	9.5	5.1	4.95	6.81	8.7

TABLE 3. Data of example 3.1 instance where  $\varphi_L > \varphi_M$ .

$f$	$f^*$	$X = [\underline{x}, \bar{x}]$	$[\underline{\nabla}f, \bar{\nabla}f]_X$	$[\underline{\nabla}f, \bar{\nabla}f]_{\Delta}$	$L_{X\infty}$	$L_{\Delta\infty}$
$x_1^2 x_2^2$	0	$\begin{pmatrix} [-1, 2] \\ [-1, 2] \end{pmatrix}$	$\begin{pmatrix} [-2, 16] \\ [-2, 16] \end{pmatrix}$	$\begin{pmatrix} [-2, 2.37] \\ [-2, 2.37] \end{pmatrix}$	16	2.37

### 3.2. Using the box gradient range

What we observed in the Lipschitz way of thinking is that the same global derivative bound is taken in all directions. Another way of thinking is to focus on the directional derivative and to use the gradient range explicitly over the enclosing box as suggested in [8] and apply this over the simplex. Given any base point  $b \in X$ , a valid lower bound is given by the Taylor view

$$f(x) \geq f(b) + \underline{(x-b)^T \nabla f} \quad x \in X. \quad (3.7)$$

The difficulty here is that given the gradient range  $[\underline{\nabla}f, \bar{\nabla}f]$ , the lower bound of the directional derivative in direction  $x - b$ ,

$$\underline{(x-b)^T \nabla f} = \sum_{i=1}^n \min\{(x_i - b_i)\underline{\nabla}_i f, (x_i - b_i)\bar{\nabla}_i f\}, \quad (3.8)$$

is not very tight. The complexity of the overestimation becomes less if we know that given base point  $b$  and variable  $x$  the term  $x_i - b_i$  is always positive or always negative. This is the case, if we consider the box vertices  $w \in \mathcal{W}$  as base points like in [7,8]. To extend the idea for a simplicial lower bound, we again introduce directions of interest for each box vertex. Let vectors  $G_w$  be defined by

$$G_{iw} = \underline{\nabla}_i f \text{ if } w_i = \underline{x}_i, \text{ else } G_{iw} = \bar{\nabla}_i f, \quad i = 1, \dots, n.$$

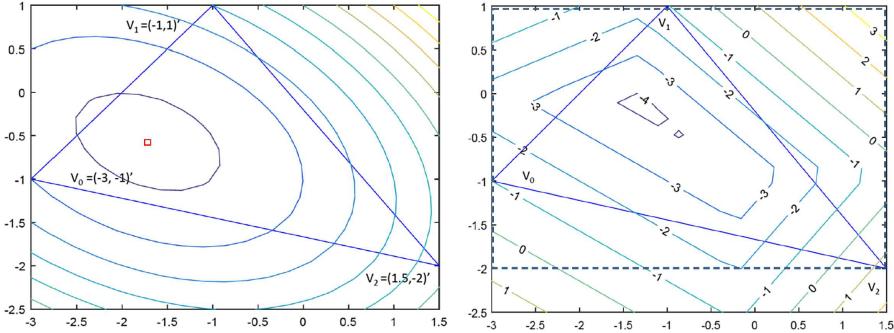


FIGURE 1. At the *left*, contours of  $f_3$  over simplex, minimum point  $(-1.71, -0.57)$ ,  $f^* = -1.14$  as red square. At the *right*, contours of lower bounding function  $z$  of (3.9).

Using a linear relaxation from each box vertex and extending towards the simplex, we obtain the following LP:

$$\begin{aligned}
 \varphi_F := \min z \\
 \text{s.t. } z &\geq f(w) + G_w^T(x - w), \quad w \in \mathcal{W} \\
 x &= \sum_{k=0}^n v_k \lambda_k \\
 \sum_{k=0}^n \lambda_k &= 1 \\
 \lambda_k &\geq 0.
 \end{aligned} \tag{3.9}$$

The underestimation is illustrated in Figure 1 by a simplex around the minimum point of  $f_3$  along with the contours of underestimating function  $z(x) := \max_{w \in \mathcal{W}} \{f(w) + G_w^T(x - w)\}$ . The reached lower bounds for the 7 instances are reported in Table 4. As one can observe, they are always tighter than the Lipschitzian lower bounds. This is not a coincidence.

**Proposition 3.2.** *Consider an instance of  $f^* := \min_{x \in \Delta} f(x)$  with enclosing box  $X \supset \Delta$  and derivative range  $[\nabla f, \bar{\nabla} f]_X$ . The lower bounds  $\varphi_M$  according to (3.6) and  $\varphi_F$  according to (3.9) fulfil  $\varphi_M \leq \varphi_F \leq f^*$ .*

*Proof.* Since both bounds are based on the maximum of linear lower bounding functions,  $\varphi_M$  and  $\varphi_F$  are valid lower bounds. Given that  $L_{X\infty} \geq \max\{\underline{\nabla}_i f, \bar{\nabla}_i f\}$  implies that  $-L_{X\infty} I_{iw} \geq G_{iw}$  for all  $w \in \mathcal{W}, i = 1, \dots, n$ . Thus,  $\forall x \in X$ ,

$$\underline{\nabla} f^T(x - w) = G_w^T(x - w) \geq -L_{X\infty} I_w^T(x - w).$$

Given the maximum is taken over  $w \in \mathcal{W}$  using the same function evaluations  $f(w)$  in both underestimations, we have that  $\varphi_F \geq \varphi_M$ .  $\square$

In [6], several variants are elaborated around what is called the center form, or mean value form. In this concept, one considers for base point  $b$  the center or midpoint  $m = \frac{1}{2}(\underline{x} + \bar{x})$  of the enclosing box in (3.7). As the corresponding lower bounding function  $z(x) = f(m) + (x - m)^T \nabla f$  is concave over the enclosing box, its minimum over the simplex is attained in one of the vertices defining a lower bound

$$\varphi_m := f(m) + \min_{v \in \mathcal{V}} (v - m)^T \nabla f. \tag{3.10}$$

TABLE 4. The minimum  $f^*$  over  $\Delta$  rounded to a decimal number, the best lower bound  $\varphi_K$  found in [6] and the value  $\varphi_O$  found in [9] using second order information. Lower bounds  $\varphi_L$  based on (3.4),  $\varphi_M$  based on (3.6) and  $\varphi_F$  based on (3.9),  $\varphi_m$  according to (3.10) and  $\varphi_c$  based on (3.11). In grey, the sharpest lower bounds are emphasized.

Inst.	$f^*$	$\varphi_K$	$\varphi_O$	$\varphi_L$	$\varphi_M$	$\varphi_F$	$\varphi_m$	$\varphi_c$
1	-0.75	-1	-1.57	-6.25	-4.625	-1.5	-2.1875	-2.5
2	-0.8125	-1.75	-1.215	-2.75	-2.5625	-1.2625	-1.3594	-1.5
3	-0.2344	-0.282	-0.252	-0.3125	-0.3066	-0.2383	-0.2725	-0.2813
4	-8.25	-26.25	-12.521	-56.25	-38.75	-10.9643	-19.25	-19.25
5	-0.2344	-0.274	-0.254	-0.7798	-0.3813	-0.2476	-0.3047	-0.3047
6	-0.2344	-0.297	-0.235	-0.9297	-0.502	-0.3047	-0.3047	-0.3047
7	9.5	9.22	9.5	6.81	8.7	9.46	9.3525	9.3144

Notice that using such a lower bound implies evaluating  $f$  at the midpoint of an enclosing box and not at its vertices. Due to the concave shape of the underestimating function, such lower bound may work well for instances where  $f$  is concave over the simplex in contrast to the convex shaped underestimating function in (3.9).

The distance of the base point to the farthest vertex is smaller and consequently the underestimation may be less when using the centroid  $c := \frac{1}{n+1} \sum_{j=0}^n v_j$ . This provides lower bound

$$\varphi_c := f(c) + \min_{v \in \mathcal{V}} (v - c)^T \nabla f. \quad (3.11)$$

Table 4 shows the evaluation of the center form lower bounds  $\varphi_m$  and  $\varphi_c$  using the directional derivative lower bound (3.8). The corresponding bounds are not tight for the 7 instances, as the instances are mostly convex and one can observe that  $\varphi_m$  and  $\varphi_c$  are less tight than  $\varphi_F$ . Also here, one can construct an instance where  $\varphi_m > \varphi_F$ , although this is not the general tendency.

**Example 3.3.** Consider the one-dimensional function  $f(x) = -|x|$  on the 1-simplex  $[-1, 1]$ . Although the function is not differentiable at 0, the derivative can be understood as sub-gradient, hence the gradient range is  $[-1, 1]$ . The function is concave and has a minimum of  $f^* = -1$  at the two vertices. Now we have that  $\varphi_m = 0 - 1 \times 1 = -1 = f^*$  and  $\varphi_F = -2$ , such that  $\varphi_m > \varphi_F$ . Notice that this function  $f$  fits exactly the underestimating function (3.7), such that the lower bound equals the minimum. A similar phenomenon can be observed in [9], where a careful quadratic lower bounding function is constructed. For instance 7, it even coincides with  $f_3$ , such that the lower bound  $\varphi = f^*$  for this case. Notice that there are some numerical errors, as several reported bounds  $\varphi$  in [9] are even  $\varphi > f^*$ , which is impossible.

Notice also that if we change the example towards the convex variant, *i.e.*  $f(x) = |x|$  on the 1-simplex  $[-1, 1]$ , then  $\varphi_F = 1 - 1 \times 1 = 0 = f^*$  and  $\varphi_m = -1$ , so the instance completely fits model (3.9).

We compare all the discussed lower bounds with the best results from [6, 9] in Table 4. For the used instances, the lower bounds reached by exploiting the derivative range and function evaluations in vertices or midpoints of partition sets are in general tighter than the ones reported in [6], which use less information. For instance 1, [6] shows that a naive interval enclosure provides a tighter bound than using other information. We observe the same in Table 4. Using more information on second derivatives may lead to even sharper bounds as sketched in [9], although also most lower bounds based on derivative ranges reported in Table 4 are tighter.

The instances used all exhibit a form of monotonicity. When the simplices represent a partition set, then usually lower bounding is not necessary, because the monotonicity information can be used to either eliminate the partition set or reduce it to lower dimensional faces. We discuss some theoretical results in Section 4 and illustrate them with the provided instances.

## 4. MONOTONICITY

In case we have monotonicity over a simplicial partition set, we may consider a dimension reduction towards a facet of the simplex which is contained in a face of the feasible area. Dimension reduction is of interest, because it tightens bounds on the function value and its derivatives. We first introduce some notation in Section 4.1. Then Section 4.2 discusses theoretical results on monotonicity questions and Section 4.3 on the directional derivatives. An illustration with the instances is provided in Section 4.4.

### 4.1. Box and simplex faces

Consider box constrained domain  $D := \{x \in \mathbb{R}^n \mid \underline{x}_i \leq x_i \leq \bar{x}_i, i = 1, \dots, n\}$ . We first discuss the concepts of faces, facets and relative interior for this domain. Consider one of the  $3^n - 1$  faces  $b_k$  of box  $D$ . A face is determined by the assignment of the variables (component-wise) represented by index  $i \in \{1, \dots, n\}$  to one of three sets  $\mathcal{L}_k, \mathcal{U}_k$  and  $\mathcal{J}_k$  defined in the following way.

- Indices  $i \in \mathcal{L}_k$  represent variables  $x_i$  on the lower bound  $x_i = \underline{x}_i$ .
- Variable coordinate  $i \in \mathcal{U}_k$  means it takes the upper bound  $x_i = \bar{x}_i$ .
- The free coordinates are given by  $i \in \mathcal{J}_k$  where  $\underline{x}_i \leq x_i \leq \bar{x}_i$ .

The relative dimension (degree of freedom) of box face  $b_k$  is  $m_k = |\mathcal{J}_k|$ . The relative interior of  $b_k$  is given by

$$\text{rint}(b_k) := \{x \in b_k \mid \underline{x}_i < x_i < \bar{x}_i, i \in \mathcal{J}_k\}. \quad (4.1)$$

A facet  $b_\ell$  of face  $b_k$  is a face that has one degree of freedom less; typically it is defined by adding one of the free variables  $i$  to either  $\mathcal{L}_k$  or  $\mathcal{U}_k$ . We will denote the set of facets of face  $b_k$  by  $\mathcal{F}(b_k)$ .

A global minimum point can be in the interior of  $D$ , in one of the vertices of  $D$  or in a relative interior of one of its other faces. The challenge for boundary solutions is that beforehand, we do not know in which (relative interior of) faces the minimum points can be found.

We consider a simplicial partition  $\mathcal{P}$  of  $D$ , *i.e.*  $\cup_{\Delta \in \mathcal{P}} \Delta = D$  and the intersection of two simplices can only consist of a face of both simplices;  $\Delta, \Xi \in \mathcal{P}$  implies  $\Xi \cap \Delta = \partial\Delta \cap \partial\Xi$ . The analysis becomes more complicated when we consider simplices as subsets which are not full dimensional, *i.e.* they can be  $m$ -simplices with  $m < n$ . Moreover, we focus on a further dimension reduction towards facets of the subset  $\Delta$ .

A facet  $F = \text{conv}(\mathcal{V} \setminus \{u\})$  of  $m$ -simplex  $\Delta = \text{conv}(\mathcal{V})$  is given by removing a vertex  $u$  from vertex set  $\mathcal{V}$ . The relative interior of a simplex  $\Delta$  is given by

$$\text{rint}(\Delta) := \left\{ x \in \mathbb{R}^n \mid x = V\lambda, \sum_{j=0}^m \lambda_j = 1; \lambda_j > 0, j = 0, \dots, m \right\}. \quad (4.2)$$

The relative boundary of simplex  $\Delta$  is  $\partial\Delta := \Delta \setminus \text{rint}(\Delta)$ . Notice that the relative boundary consists of the set of facets  $\mathcal{F}(\Delta)$ , that is  $\partial\Delta = \cup_{F \in \mathcal{F}(\Delta)} F$ ,  $|\mathcal{F}(\Delta)| = |\mathcal{V}|$ .

### 4.2. Monotonicity properties

A simplicial branch and bound on the box  $D$  keeps a list  $\Lambda$  of simplices, where the optimum can still be located;  $\min_{x \in D} f(x) = \min_{\Delta \in \Lambda} \min_{x \in \Delta} f(x)$ . The idea is to refine partition sets  $\Delta$  and to eliminate  $\Delta$  if it can be shown that it cannot contain a global minimum point. So far, this paper focused on lower bounding to eliminate partition sets. Our question here is how derivative range  $[\underline{\nabla} f, \bar{\nabla} f]_\Delta$  can be used to help the refinement process. The first observation is close to the concept of a monotonicity test.

**Proposition 4.1.** *Let  $\Delta := \text{conv}(\mathcal{V}) \subset b_k$  with  $|\mathcal{V}| = m_k + 1$  and  $\exists i \in \mathcal{J}_k, 0 \notin [\underline{\nabla}_i f, \bar{\nabla}_i f]_\Delta$ . Then  $\text{rint}(\Delta)$  does not contain a minimum point.*

*Proof.* A necessary condition for a minimum point  $x^*$  in the relative interior of face  $b_k$  is that

$$\forall i \in \mathcal{J}_k, \frac{\partial f}{\partial x_i}(x^*) = 0.$$

The condition of the proposition implies that

$$\forall x \in \text{rint}(\Delta), \frac{\partial f}{\partial x_i}(x) \neq 0,$$

i.e. in each point  $x \in \text{rint}(\Delta)$  there exists an improving direction in component  $i$  such that  $x$  cannot be a minimum point.  $\square$

Basically, for a practical use, this means we can remove the relative interior of  $\Delta$  and refine by storing its facets in  $\Lambda$ . However, if facets  $F \in \mathcal{F}(\Delta)$  are in the relative interior of  $b_k$  we are not interested in keeping them either.

**Corollary 4.2.** *Let  $\Delta := \text{conv}(\mathcal{V}) \subset b_k$  with  $|\mathcal{V}| = m_k + 1$  be a partition set in list  $\Lambda$  and  $\exists i \in \mathcal{J}_k, 0 \notin [\nabla_i f, \bar{\nabla}_i f]_\Delta$ . If  $F \in \mathcal{F}(\Delta), \#b_\ell \in \mathcal{F}(b_k), F \subset b_\ell$  then  $F$  can be removed from consideration in branch and bound list  $\Lambda$ . Consequently, if  $\forall F \in \mathcal{F}(\Delta), \#b_\ell \in \mathcal{F}(b_k), F \subset b_\ell$  then  $\Delta$  can be removed from consideration in branch and bound list  $\Lambda$ .*

The consequence for a practical refinement is that if  $f$  appears monotonic in  $\Delta$  in the sense of Proposition 4.1, then we only have to save those facets in  $\mathcal{F}(\Delta)$  that are included in a facet of  $b_k$ , i.e.  $\exists b_\ell \in \mathcal{F}(b_k)$  such that  $F \in b_\ell$ . It also means that we can eliminate all of  $\Delta$  in some circumstances. This means that  $\Delta$  may have a local optimum on its boundary, but this boundary is also included in other partition sets and can therefore be removed. On the other hand, this says that we still should keep facets  $F$  that are in a facet  $b_\ell$  of  $b_k$ .

There is an interesting case where it can be shown that the relative interior of  $F$  cannot contain a minimum point. For this we introduce the so-called normal vector  $p$  with respect to facet  $F$ . Let  $\Delta = \text{conv}(\mathcal{V})$ ,  $w \in \mathcal{V}$ ,  $\mathcal{W} = \mathcal{V} \setminus \{w\}$ , facet  $F = \text{conv}(\mathcal{W})$ . Consider  $u \in \mathcal{W}$  and construct a set of vectors  $\mathcal{Y} := \{v - u, v \in \mathcal{W}\} \setminus \{0\}$  with corresponding matrix representation  $Y$ . The normal vector of  $F$  in the direction of  $w$  can be expressed by defining (see e.g. [1])

$$p := [I - Y(Y^T Y)^{-1} Y^T](w - u). \quad (4.3)$$

Vector  $p$  points into the same half-space where  $\Delta$  is located. From each point in  $\text{rint}(F)$  direction  $p$  points into  $\text{rint}(\Delta)$ .

**Proposition 4.3.** *Let  $\mathcal{V}$  be a vertex set,  $\Delta := \text{conv}(\mathcal{V}) \in b_k$  with  $|\mathcal{V}| = m_k + 1$  and  $\exists i \in \mathcal{J}_k, (\bar{\nabla}_i f)_\Delta < 0$ . Let  $u \in \mathcal{V}$ ,  $\mathcal{Y} := \mathcal{V} \setminus \{u\}$ , facet  $F := \text{conv}(\mathcal{Y})$ . If  $p_i > 0$ , then  $\text{rint}(F)$  cannot contain a minimum point of  $\min_{x \in \Delta} f(x)$ .*

*Proof.* For any point  $y \in \text{rint}(F)$  we have that coordinate unit vector  $e_i$  is a feasible direction according to  $e_i^T p = p_i > 0$  in which we know that the function value goes down.  $\square$

For the ease of reasoning, Proposition 4.3 is posed in an increasing partial derivative way. The same applies of course the other way around.

**Corollary 4.4.** *Consider the same conditions of Proposition 4.3, but having  $\exists i \in \mathcal{J}_k, (\nabla_i f)_\Delta > 0$ . If  $p_i < 0$ , then  $\text{rint}(F)$  cannot contain a minimum point of  $\min_{x \in \Delta} f(x)$ .*

### 4.3. Directional derivatives

A bound on the directional derivative as in (3.8) can be used as follows.

**Proposition 4.5.** *Let  $\mathcal{V}$  be a vertex set,  $\Delta := \text{conv}(\mathcal{V})$ ,  $u \in \mathcal{V}$  and facet  $F := \text{conv}(\mathcal{V} \setminus \{u\})$ . If  $\exists y \in F$  such that there is a lower bound  $\underline{(u-y)^T \nabla f} \geq 0$  on the directional derivative, i.e.  $\nabla f(x)^T(u-y) \geq 0$  for all  $x \in \Delta$ , then  $F$  contains a minimum point of  $\min_{x \in \Delta} f(x)$ .*

*Proof.* Consider the vertices of  $\mathcal{V}$  ordered such that  $u = v_0$ . Let  $x = \sum_{j=1}^m \lambda_j v_j + \lambda_0 u$  be a minimum point  $x \notin F$ . The point  $y$  can be written as convex combination  $y = \sum_{j=1}^m \mu_j v_j$ . We construct a point  $z$  on  $F$  walking in direction  $y - u$  according to  $z = x - \lambda_0(u-y) = \sum_{j=1}^m (\lambda_j + \lambda_0 \mu_j) v_j$ . Then we have that  $f(z) \leq f(x) - \lambda_0(u-y)^T \nabla f \leq f(x)$ . So, minimum point  $x$  either does not exist, or  $z$  is also minimum point of  $\min_{x \in \Delta} f(x)$  and located on facet  $F$ .  $\square$

This proposition implies an extension of Corollary 4.2. We can do a dimension reduction from  $\Delta$  to  $F$ , or alternatively remove  $\Delta$  from the list.

**Corollary 4.6.** *Let  $\Delta := \text{conv}(\mathcal{V}) \subset b_k$  with  $|\mathcal{V}| = m_k + 1$  be a partition set in list  $\Lambda$  and  $u \in \mathcal{V}$  and  $\exists b_\ell \in \mathcal{F}(b_k)$  such that facet  $F := \text{conv}(\mathcal{V} \setminus \{u\}) \in b_\ell$ . If  $\exists y \in F$  such that  $\forall x \in \Delta$  the directional derivative  $(u-y)^T \nabla f(x) \geq 0$ , then  $\Delta$  can be replaced in  $\Lambda$  by  $F$ .*

**Corollary 4.7.** *Let  $\Delta := \text{conv}(\mathcal{V}) \subset b_k$  with  $|\mathcal{V}| = m_k + 1$  be a partition set in list  $\Lambda$  and  $u \in \mathcal{V}$  and  $\nexists b_\ell \in \mathcal{F}(b_k)$  such that facet  $F := \text{conv}(\mathcal{V} \setminus \{u\}) \in b_\ell$ . If  $\exists y \in F$  such that  $\forall x \in \Delta$  the directional derivative  $(u-y)^T \nabla f(x)^T \geq 0$ , then  $\Delta$  can be removed from  $\Lambda$ .*

A special case of the lowest dimensional faces with a relative interior is an edge.

**Corollary 4.8.** *Let  $\Delta := \text{conv}(\{u, v\})$ . If  $\underline{(u-v)^T \nabla f} \geq 0$  on the directional derivative, then  $v \in \text{argmin}_{x \in \Delta} f(x)$ .*

The use of this theoretical result to generate a test is not easy when only the partial derivative ranges  $[\underline{\nabla_i f}, \bar{\nabla}_i f]_\Delta$  are available, as (3.8) is not very tight. Focusing on the direction  $u - y$  with  $y = \sum_{j=1}^m \lambda_j v_j$ , we can demonstrate there is a (maximum) positive directional derivative solving the LP

$$\begin{aligned}
 & \max \sum_{i \in \mathcal{J}_k} z_i \\
 \text{s. t. } & z_i \leq \underline{\nabla_i f} \sum_{j=1}^m (u_i - v_{ji}) \lambda_j, \quad i \in \mathcal{J}_k \\
 & z_i \leq \bar{\nabla}_i f \sum_{j=1}^m (u_i - v_{ji}) \lambda_j, \quad i \in \mathcal{J}_k \\
 & \sum_{i=1}^m \lambda_i = 1 \\
 & \lambda_i \geq 0
 \end{aligned} \tag{4.4}$$

and checking the result is positive. Notice that for  $j \in \mathcal{L}_k \cup \mathcal{U}_k$ , we have that  $u_j = v_{ij}$ , i.e. the coordinates have the same values and we do not have to take the corresponding partial derivative into account.

TABLE 5. Monotonicity analysis of the 7 instances. Faces left over expressed as vertex set after using theoretical results.

Inst.	Proposition 4.3	Proposition 4.5	Corollary 4.8
1	$\{v_0, v_2\}, \{v_1, v_2\}$	$\{v_0, v_1\}$ using $y = v_1$	
2	$\{v_0, v_1\}, \{v_2\}$	$\{v_0, v_1\}$ using $y = v_1$	
3	$\{v_0, v_1\}, \{v_2\}$	$\{v_0, v_1\}$ using $y = v_1$	
4	$\{v_0, v_1\}, \{v_2\}$	$\{v_0, v_1\}$ using $y = v_0$	
5	$\{v_0, v_1\}, \{v_2\}$	$\{v_0, v_1\}$ using $y = v_0$	
6	$\{v_0, v_1\}, \{v_2\}$	$\{v_0, v_1\}$ using $y = v_0$	
7	$\{v_0, v_2\}, \{v_1\}$	$\{v_0, v_2\}$ using $y = v_0$	$\{v_0\}$

#### 4.4. Monotonicity on the illustrative instances

For the analysis of simplicial partition sets over a box constrained region, it is good to realize that all initial partition sets in the subdivision introduced by Todd [11] have facets in faces of the box. In future implementations, we will make use of that. Here we go step-wise through the consequences of the theoretical results for the 7 instances. A minimum can be attained at a relative interior point of a face, or at a vertex. Notice that all instances attain their minimum at a vertex.

Consider that in fact  $[\nabla f, \bar{\nabla} f]_\Delta$  in Table 1 shows us that all 7 instances have boundary optima, given that  $\nabla_2 f \geq 0$ . For instance 1, this means that it could contain an interior optimum point, but that the optimum is also attained on the boundary. Proposition 4.1 shows that in fact for all instances,  $\Delta$  can be replaced by its facets in reduced dimension in the list of subsets  $\Lambda$  before considering the lower bound over  $\Delta$ .

The next step is to consider what Proposition 4.3, or in our case Corollary 4.4 concludes about which faces we do not have to consider. Focusing on the normal vector (4.3), it can be shown that the (normalized) normal vectors  $p$  among instances 1–3 and among instances 4–6 are the same. A sharp test may conclude that for instance 1, we do not have to consider  $\text{rint}(\text{conv}(\{v_0, v_2\}))$  and for instances 2–6  $\text{rint}(\text{conv}(\{v_0, v_2\}))$  and  $\text{rint}(\text{conv}(\{v_1, v_2\}))$  can be left out of consideration. Due to the numbering of vertices, this is  $\text{rint}(\text{conv}(\{v_0, v_1\}))$  and  $\text{rint}(\text{conv}(\{v_1, v_2\}))$  for instance 7, see Table 5.

The last step is to exploit Proposition 4.5 on the directional derivative. For instances 1–6 the focus on  $u = v_2$  provides the insight, that the minimum is attained at facet  $\text{conv}(\{v_0, v_1\})$ . For instances 1–3 one can take  $y = v_1$  and for instances 4–6 consider  $y = v_0$ . For instance 7, the facet with a minimum is  $\text{conv}(\{v_0, v_2\})$ , so  $u = v_1$  and one can consider  $y = v_0$ .

Continuation of this monotonicity analysis using Corollary 4.8 is only successful for instance 7 to indicate the optimal vertex  $v_0$ . For the other instances, the underestimation of the directional derivative using (3.8) over the edge does not provide a proof of an increasing direction over the edge.

## 5. CONCLUSIONS AND DISCUSSION

The bounding in simplicial branch and bound has gotten a renewed attention recently. Our contribution focuses on the question how the knowledge of the derivative range of either the simplicial partition set or the enclosing box can be used to derive bounds and to perform a monotonicity analysis. We presented several theoretical results and used an earlier set of 7 instances introduced in literature to illustrate them.

From the analysis we learned that bounds are in general sharper when more detailed information is used; in our case gradient ranges are more detailed than a Lipschitz constant. Bounds are in general tighter for the 7 instances than the ones published before in literature. We also remark that all instances have a certain monotonicity. We show that the simplicial sets can be reduced in dimension using the introduced theoretical monotonicity results.

Bounds based on solving an LP in general require more computational time for a partition set. In our experience, sharper bounds pay off in a branch and bound run, as one avoids the management of larger trees. In future work, we use the derived theory for tests in a simplicial branch and bound over a large set of numerical instances. However, it is known that the efficiency of various bound and monotonicity calculations depends on the order in which tests are done and also on the phase the algorithm is in; the beginning has a global search character, whereas in the end when the algorithm approaches the minimum points, that algorithm behaves in another way. Such an experimental investigation, requires a systematic design of experiments to evaluate the bound and monotonicity calculations introduced here.

*Acknowledgements.* This work has been funded by grant RTI2018-095993-B-I00 from the Spanish Ministry.

## REFERENCES

- [1] G. Aparicio, L.G. Casado, E.M.T. Hendrix, B.G.-Toth and I. Garcia, On the minimum number of simplex shapes in longest edge bisection refinement of a regular n-simplex. *Informatica* **26** (2015) 17–32.
- [2] E.M.T. Hendrix, L.G. Casado and I. Garcia, The semi-continuous quadratic mixture design problem: Description and branch-and-bound approach. *Eur. J. Oper. Res.* **191** (2008) 803–815.
- [3] E.M.T. Hendrix and J. Pintér, An application of Lipschitzian global optimization to product design. *J. Global Optim.* **1** (1991) 389–401.
- [4] E.M.T. Hendrix, J.M.G. Salmerón and L.G. Casado, On function monotonicity in simplicial branch and bound. AIP conference proceedings **2070** (2019) 020007–4.
- [5] R. Horst and H. Tuy, Global Optimization (Deterministic Approaches). Springer: Berlin Heidelberg (1990).
- [6] S.D. Karhbet and R.B. Kearfott, Range bounds of functions over simplices, for branch and bound algorithms. *Reliable Comput.* **25** (2017) 53–73. <https://interval.louisiana.edu/reliable-computing-journal>.
- [7] C.C. Meewella and D.Q. Mayne, An algorithm for global optimization of Lipschitz continuous functions. *J. Optim. Theory Appl.* **57** (1988) 307–323.
- [8] F. Messine and J.-L. Lagouanelle, Enclosure methods for multivariate differentiable functions and application to global optimization. *J. Universal Comput. Sci.* **4** (1998) 589–603.
- [9] O. Mohand, Tighter bound functions for nonconvex functions over simplexes. *RAIRO Oper. Res.* **55** (2021) S2373–S238.
- [10] R. Paulavičius and J. Žilinskas, Simplicial Global Optimization. Springer: New York (2014).
- [11] M.J. Todd, The Computation of Fixed Points and Applications. Springer: Heidelberg (1976).