J Glob Optim (2007) 37:215–228
DOI 10.1007/s10898-006-9045-9

ORIGINAL ARTICLE

Comparison between Baumann and admissible simplex forms in interval analysis

Pierre Hansen · Jean-Louis Lagouanelle · Frédéric Messine

Received: 27 September 2005 / Accepted: 11 May 2006 /
Published online: 27 June 2006

Abstract Two ways for bounding n-variables functions over a box, based on interval evaluations of first order derivatives, are compared. The *optimal Baumann form* gives the best lower bound using a center within the box. The *admissible simplex form*, proposed by the two last authors, uses point evaluations at $n + 1$ vertices of the box. We show that the *Baumann center* is within any *admissible simplex* and can be represented as a linear convex combination of its vertices with coefficients equal to the dual variables of the linear program used to compute the corresponding admissible simplex lower bound. This result is applied in a branch-and-bound global optimization and computational results are reported.

Keywords Interval arithmetic · Lower bound · Baumann form · Admissible simplex form · Global optimization.

1 Introduction

Interval analysis [14, 15], has led to many *branch-and-bound algorithms* for global optimization of univariate or multivariate non-linear and non-convex analytical functions, possibly subject to constraints [4, 6–9, 11, 13, 17, 18]. These techniques provide precise enclosures for the optimal value and for one or all optimal solutions with an absolute guarantee, i.e. the errors induced can be bounded to an arbitrary degree specified by the user.

P. Hansen
GERAD and Department of Quantitative Methods in Management,
HEC Montréal, Montréal, Québec, Canada
e-mail: Pierre.Hansen@gerad.ca

J.-L. Lagouanelle · F. Messine (✉)
ENSEEIHT-IRIT CNRS-UMR 5055, Toulouse, France
e-mail: Frederic.Messine@n7.fr

J.-L. Lagouanelle
e-mail: lagouane@irit.fr

This paper focuses on such algorithms [6, 9, 17] and specifically on linear bounding techniques where basic tools are linear centered forms [10] (*automatic differentiation* and interval arithmetic are used for enclosing the gradient of a function). We examine two particular cases of linear bounding, the optimal centered form of Baumann [2, 3] and the admissible simplex form, due to the two last authors [12, 13]. We study links between these two methods. In particular, a property of convexity is used to get a tighter enclosure of the optimum.

The study is restricted to unconstrained minimization problem over a box X for multivariate differentiable functions. Section 2 presents notation and basic tools. The optimal centered form of Baumann is introduced briefly in Sect. 3 and the admissible simplex form is presented in Sect. 4. A property of duality and convexity for multivariate functions is brought to light in Sect. 5. The convexity part of this property is used as an accelerating device. Numerical experiments are gathered and discussed in Sect. 6. Brief conclusions are stated in Sect. 7.

2 Basic tools and notation

The problem definition is given and throughout this paper all the following notation are used:

- $\mathbb{R}$ and $\mathbb{R}^n$ denote the sets of real numbers and real n-vectors.
- $X = (X_1, X_2, \ldots, X_n)$ is a n-dimentional interval in $\mathbb{R}^n$ where $X_i = \{x_i \in \mathbb{R} : a_i \le x_i \le b_i\}$, a_i and b_i have fixed values.
- ∇f denotes the gradient of f.
- Searching for the global minimum of an unconstrained real-valued differentiable function f, our aim is to find (x^*, f^*) with a given precisions where $f^* := f(x^*) = \inf_{x \in X} f(x)$ and x^* is a minimizer. We deal, mainly, with linear lower bounds for the function f.
- The central theoretical result, used in this paper as a basic tool to construct linear under estimations of the function, is the mean value theorem, namely $f(x) = f(c) + (x - c)^t \nabla f(\xi)$ for x and c in X, where ξ is a point between x and c.
- Let $c = (c_1, c_2, \ldots, c_n)$ be a fixed point in X, the *center* of the so-called *linear centered form*. We assume that the partial derivatives of f satisfy the relation $L_k < \partial f(x)/\partial x_k < U_k$ for any x in X. In the sequel, the function f is supposed to be not strictly monotone with respect the kth coordinate direction, hence we examine only the hypothesis

$$(H_0)\text{: } L_k < 0 < U_k \quad \text{for any } k.$$

- Let x be any point such that $a_i \le x_i \le c_i$ for $i \in I \subset \{1, 2, \ldots, n\}$ and $c_j \le x_j \le b_j$ for $j \in J = \{1, 2, \ldots, n\} - I$ then the following linear lower bound holds $f(x) \ge z(x) = f(x) + \sum_{i \in I}(x_i - c_i)U_i + \sum_{j \in J}(x_j - c_j)L_j$.

Many choices are possible for the center c. The most popular one is the center of the box, $c_i = (a_i + b_i)/2$. It has been shown, however, by Baumann [2, 3] that an optimal choice c_B^- exists, for linear centered forms, which gives the best lower bound of the function. An extremal point of X may also be chosen as center; then the linear form is called a *linear boundary value form* (*lbvf*) [12, 13, 15].

The main practical difficulty encountered with centered forms lies in the time consuming computation of enclosures for the partial derivatives. This is overcome by using simultaneously *Automatic differentiation* and *Interval Arithmetic* [1, 2, 9].

Automatic (Arithmetic) differentiation is a numerical chain rule-based technique for evaluating derivatives at any order of a function [5, 16]. If the numerical values of $(u(x), u'(x))$ and $(v(x), v'(x))$ are known, then $(u(x) \star v(x), (u(x) \star v(x))')$ may be computed numerically for any elementary operation $\star$. The derivative of $v(u(x))$ may be computed as well and finally the derivatives of elementary functions such as trigonometric or exponential functions are supposed to be known. Different algorithmic modes may be used but the reverse mode has the advantage of a low computational cost, see for example the book of Griewank [5].

Interval arithmetic uses operators defined over compact sets which are real intervals of $\mathbb{R}$. Let $X = [x^L, x^U] := \{x \in \mathbb{R} : x^L \le x \le x^U\}$ and let $Y = [y^L, y^U]$ be another interval, $Z = X \star Y := \{x \star y : x \in X, y \in Y\}$ where $\star \in \{+, -, \times, \div\}$. Z is an interval since the operations are continuous in Standard Interval Arithmetic [4], because it is assumed that $0 \notin Y$ for the computation of X–Y.

An interval vector $G = (G_1, G_2, \ldots, G_n)$ [1, 14] is assumed to enclose the partial derivatives of the function: $G_i = [L_i, U_i]$ and $L_i \le \partial f(X)/\partial x_i \le U_i$ for $i = 1, 2, \ldots, n$. The components of G are available from automatic differentiation, again interval arithmetic is used for the computations, but any other inclusion function for the range of $\nabla f(x)$ over the box X could be used. Then for any $x \in X$, $f(x) \in f(c) + (x - c)^t G$ and $f(X) \subseteq F_c(X) := f(c) + (X - c)^t G$ when interval operators are used. However our main target is not a tight enclosure of the range $f(X)$ but only a sharp lower bound for the function.

The optimal centered form of Baumann presented in the next section yields the greatest lower bound among all possible centered forms.

3 Optimal centered form of Baumann

Let f be an univariate function defined over $X := [a, b] \subset \mathbb{R}$ with $L \le f'(X) \le U$ and $L < 0 < U$, see (H_0).

The mean value form induces a *concave relaxation* obtained by linear lower bound functions, see Fig. 1 for a geometrical interpretation. Then a lower bound of the function is reached at an extremal point of X:

$$z_c^- = \min\{f(c) + (a - c)U, f(c) + (b - c)L\}.$$

The optimal center c_{B}^- of Baumann [2] is obtained when $f(c)+(a-c)U = f(c)+(b-c)L$ and gives the so-called *optimal (linear) centered form of Baumann*, one can see on Fig. 1 that, in this case, the two points $\mathrm{M} = (a, f(a))$ and $\mathrm{N} = (b, f(b))$ are at the same level, leading to

$$c_{\mathrm{B}}^- = \frac{aU - bL}{U - L}$$

and the corresponding lower bound,

$$z_{\mathrm{B}}^- = f\left(\frac{aU - bL}{U - L}\right) + \frac{(b - a)LU}{U - L}.$$

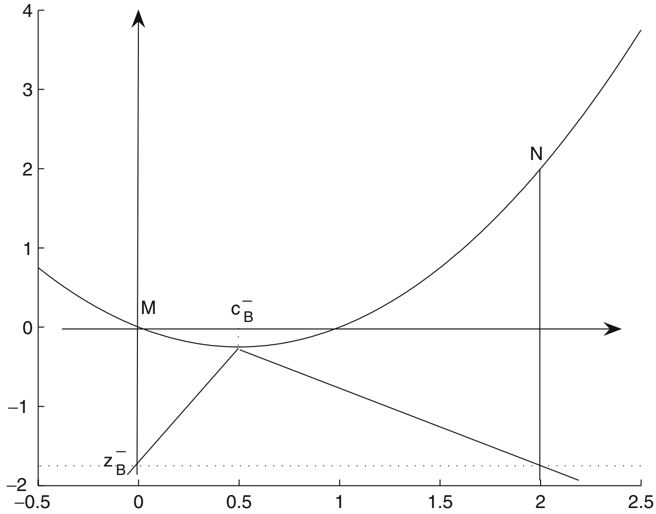


Fig. 1 Example of a Baumann under-estimation

Example 1 For $f(x) = x^2 - x$, $X = [0, 2]$, $\min f(X) = -1/4$ and the lower bound is $z_B^- = -7/4$ with $c_B^- = 1/2$ (see Fig. 1).

The Baumann formula is componentwise separable [3]. This advantage follows from the fact that the coordinates of the optimal center are independent of the function values. Therefore, the optimal Baumann centered form can be easily generalized to the multivariate case.

Let $c_B^- = (c_1^-, c_2^-, \ldots, c_n^-)$ then we get immediately $c_i^- = \frac{a_i U_i - b_i L_i}{U_i - L_i}$ for $i = 1, 2, \ldots, n$ and the corresponding lower bound is given by:

$$z_B^- = f(c_B^-) + \sum_{i=1}^{n} \frac{(b_i - a_i) L_i U_i}{U_i - L_i},$$

which follows from a *linear concave relaxation*.

Example 2 Consider the function $f_1(x_1, x_2) = 1 + (x_1^2 + 2)x_2 + x_1 x_2^2$, with $x_1 \in [1, 2]$ and $x_2 \in [-10, 10]$; one has $z_B^- = -442.58\cdots$ and $c_B^- = (1.2222\cdots, -1.084333\cdots)$ for $f^* = -3.5$.

4 Linear boundary value forms

One assume again that $L_i \le \frac{\partial f}{\partial x_i}(X) \le U_i$ and that (H_0) holds, i.e. $L_i < 0 < U_i$ (for all $i \in \{1, \ldots, n\}$).

4.1 The univariate case

As shown, e.g. in Neumaier's book [15], p60, a useful bicentered form can be built by choosing as centers, the end points a and b of X. This induces a *convex relaxation* of the function f obtained again by linear lower bound functions, the so-called lbvf. A lower bound on the function values for $x \in [a, b]$ is then obtained by intersecting the two lines $y = f(a) + (x - a)L$ and $y = f(b) + (x - b)U$ leading to the point $S = (s^-, z^-_{\mathrm{lbvf}})$, vertex of a convex cone, with coordinates:

$$s^- = \frac{f(a) - f(b)}{U - L} + \frac{(b - a)LU}{U - L}$$

and,

$$z^-_{\mathrm{lbvf}} = \frac{Uf(a) - Lf(b)}{U - L} + \frac{(b - a)LU}{U - L}.$$

Example 3 For $f(x) = x^2 - x$, $X = [0, 2]$, $\min f(X) = -1/4$ and the lower bound is $z^-_{\mathrm{lbvf}} = -1$ with $s^- = 1$ (see Fig. 2).

4.2 The multivariate case

The preceding method may be extended to multivariate functions [13] and was called *admissible simplex form* in this paper. In that case the lower bound follows again from a *convex relaxation,* which is constructed with linear boundary forms. Geometrically this lower bound is reached at a point S obtained by intersecting $n+1$ suitably chosen hyperplanes of $\mathbb{R}^n \times \mathbb{R}$. S is the vertex of a simplicial convex cone. Each hyperplane is the geometrical representation of a linear centered form, the center of which is a vertex S_k of X. Let $I_k \subset N = \{1, 2, \ldots, n\}$ and $J_k = N - I_k$, the equation of the hyperplane Π_k related to the vertex S_k is:

$$z_k(x) = f(S_k) + \sum_{i \in I_k} (x_i - a_i)L_i + \sum_{j \in J_k} (x_j - b_j)U_j.$$

The following relations are satisfied :

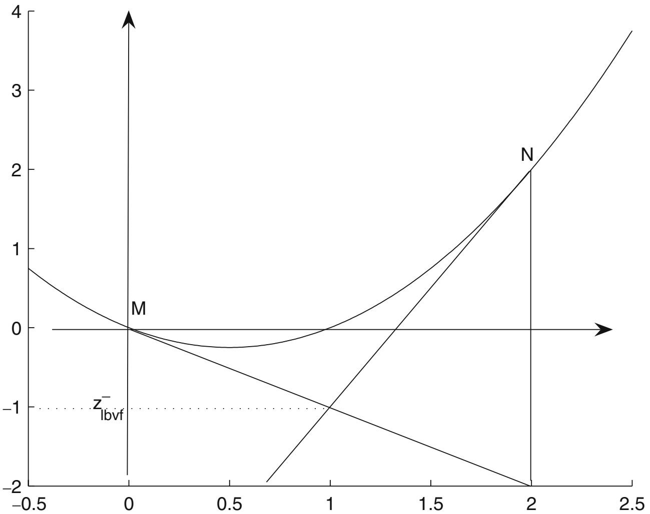


Fig. 2 Example of a lbvf under-estimation

- $z_k(S_k) = f(S_k)$,
- $\forall x \in X$, $z_k(x) \leq f(x)$ and $z_k(x) \leq z_k(S_k)$.

We must find an *admissible set* of vertices $S_0, S_1, \ldots, S_n$, which means that the intersection of their corresponding hyperplane Π_k yields effectively a lower bound of the function over X. For that purpose we use the following method from [12, 13].

Let S_0 be any initial vertex of X, then for every k the vertex S_{k+1} is selected by changing only one coordinate in S_k, according to the following rule:

- Let the $H_l = \mid K_l \mid /(U_l - L_l)$, $l = 1, \ldots, n$, where K_l is the coefficient of x_l in $z_0(x)$, equation related to S_0, and thus $K_l = L_l$ if $l \in I_0$ and $K_l = U_l$ if $l \in J_0$ be ranked in order of decreasing magnitude.
- Assume that $H_{l_1} \geq H_{l_2} \geq \cdots \geq H_{l_n}$; starting from S_0 change the coordinate l_1 to get S_1, then change the coordinate l_2 in S_1 to get S_2 and so on until one gets S_n which is the opposite vertex of S_0 on the box X.

In fact the set of vertices is *admissible* if and only if:

$$\alpha_0 = H_{l_1} \leq 1, \quad \alpha_1 = H_{l_1} - H_{l_2} \geq 0, \ldots, \quad \alpha_{l_{n-1}} = H_{l_{n-1}} - H_{l_n} \geq 0, \quad \alpha_n = H_{l_n} \geq 0.$$

This follows from the optimal solution of the linear program:

$$(P_l) \min z \text{ subject to } (x, z) \in E_k^+, k = 0, 1, \ldots, n$$

where $E_k^+ = \{(x, z) \in I\!R^n \times \mathbb{R} : z \geq z_k(x), x \in X\}$.

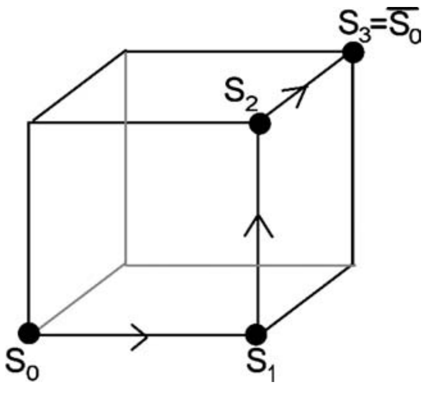
We then get the lower bound:

$$z_{\text{asf}}^- = \sum_{k=0}^{n} \alpha_k f(x_k) + \sum_{i=1}^{n} \frac{(b_i - a_i)L_i U_i}{U_i - L_i}.$$

See [12, 13] for further details, and see Fig. 3 for an example of an admissible path in $I\!R^3$.

Example 4 Looking for the function $f_1(x_1, x_2) = 1 + (x_1^2 + 2)x_2 + x_1 x_2^2$, with $x_1 \in [1, 2]$ and $x_2 \in [-10, 10]$. We obtain $z_{\text{asf}}^- = -314.5957\cdots$ using the following admissible simplex $(1, -10), (1, 10), (-10, 10)$, for $f^* = -3.5$.

Fig. 3 Admissible path in $\mathbb{R}^3$



5 Links between the preceding forms

In this section, we examine which of the two approaches described above gives the best lower bound.

Let assume that $L_i \leq \frac{\partial f}{\partial x_i}(X) \leq U_i$ and $L_i < 0 < U_i$ (for all $i \in \{1, \ldots, n\}$), see (H_0). We consider univariate functions before turning to the general case.

First, equality between z_{B}^- and z_{lbvf}^- implies that:

$$\frac{Uf(a) - Lf(b)}{U - L} = f\left(\frac{aU - bL}{U - L}\right),$$

this means that the point $(c_{\mathrm{B}}^-, f(c_{\mathrm{B}}^-))$ is on the line joining the end points $M(a, f(a))$ and $N(b, f(b))$ of the graph of the function f.

Second, the point c_{B}^- is a convex combination of the end points of X with weights $U/(U-L)$ and $-L/(U-L)$.

Therefore, we get the following property:

Proposition 1 *For any univariate differentiable function f, when* (H_0) *is satisfied, the lower bound resulting from the optimal centered form of Baumann is greater than the lower bound given by the lbvf centered form if and only if*

$$\frac{Uf(a) - Lf(b)}{U - L} < f\left(\frac{aU - bL}{U - L}\right).$$

Proof This is clear from the previous formula. This result may also be found in [8]. □

Remark 1 If f is a convex function over X, lbvf bounding yields a tighter lower bound, and the advantage must be given to Baumann formula in the concave case. Moreover, when f is convex, which is the case around a local minimum, $f(c_{\mathrm{B}}^-)$ may be a good upper bound for $f^* = \inf f(X)$ and then this value may be used in the computation of a middle point test instead of $mid(X) := (a+b)/2$, (see [17]).

Let $\widetilde{f}$ be the current upper bound for the minimum f^*, then its new value could be $\widetilde{f} := \min(\widetilde{f}, f(c_B^-), f(a), f(b))$ and finally $f^* \in [\max\left\{z_{\mathrm{B}}^-, z_{\mathrm{lbvf}}^-\right\}, \widetilde{f}]$.

Example 5 For $f(x) = x^2 - x, X = [0, 2]$ we get $f(c_{\mathrm{B}}^-) = -1/4, z_{\mathrm{B}}^- = -7/4, z_{\mathrm{lbvf}}^- = -1, f(a) = 0, f(b) = 2$ and then $f^\star \in [-1, -1/4]$. Over X, f is a convex function, the linear boundary form gives the best lower bound and $f(c_{\mathrm{B}}^-)$ is a good upper bound for $f^\star$, (see Fig. 4).

Conversely for the opposite (and concave) function $f(x) = x - x^2$, $z_{\mathrm{B}}^- = -9/4$ and $z_{\mathrm{lbvf}}^- = -3$.

For the unidimensional case, these properties have already been pointed out in [1, 14]. We show in the next subsection how they can be extended to multivariate functions.

5.1 Baumann center and admissible sets of vertices

An *admissible set* of vertices $S_0, S_1, \ldots, S_n$, as defined in Sect. 4, is not contained in one hyperplane of $\mathbb{R}^n$. The convex hull of these vertices is a non-empty interior set, simplex of $\mathbb{R}^n$ contained in X, and will be called an *admissible simplex*. The following property is a generalization of the univariate case.

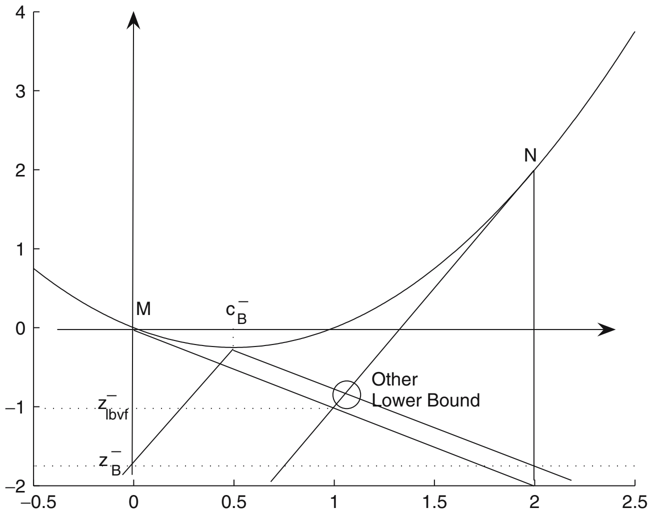


Fig. 4 Comparison between Baumann and lbvf forms on a convex function

Proposition 2 *The optimal center c_{B}^- of Baumann for linear centered forms belongs to the intersection of all the admissible simplexes.*

Proof Let $x_1, x_2, \ldots, x_n$ be the coordinates of S_0, assume that the coordinates are modified in increasing order of indices to get the successive vertices $S_1, S_2, \ldots, S_n$, after a new numbering if necessary. Then x_1 is changed into y_1 to get S_1, x_2 is changed into y_2 to get S_2, and so on.

- First, let us write that $c_{\mathrm{B}}^- = (c_1, c_2, \ldots, c_n)^t$ is a linear combination of the vertices $S_k, c_{\mathrm{B}}^- = \sum_{k=0}^{n} \beta_k S_k$ and that $\sum_{k=0}^{n} \beta_k = 1$.
  Then we must solve the linear system of equations $M\beta = \widetilde{c}$, where $\widetilde{c} = (c_1, c_2, \ldots, c_n, 1)^t$ and M is a square matrix of order $n+1$

$$M = \begin{bmatrix} x_1 & y_1 & y_1 & y_1 & \ldots & y_1 \\ x_2 & x_2 & y_2 & y_2 & \ldots & y_2 \\ x_3 & x_3 & x_3 & y_3 & \ldots & y_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ x_n & x_n & \ldots & \ldots & x_n & y_n \\ 1 & 1 & \ldots & \ldots & 1 & 1 \end{bmatrix}.$$

  This linear system has one and only one solution if and only if $\prod_{k=1}^{n}(y_k - x_k) \neq 0$, which implies that X is a non-degenerated box. We get:

$$\beta_0 = (c_1 - y_1)/(x_1 - y_1),$$
$$\beta_k = (c_{k+1} - y_{k+1})/(x_{k+1} - y_{k+1}) - (c_k - y_k)/(x_k - y_k), \quad k = 1, 2, \ldots, n-1,$$
$$\beta_n = (x_n - c_n)/(x_n - y_n).$$

- Second, let us prove that the linear combination is convex.
  It remains to prove that $\beta_k \in [0, 1]$. But taking the values c_k for the coordinates of c_B^-, we find that $\beta_k = \alpha_k$ and the proof follows directly from Sect. 4. Thus the

coefficients $\alpha_k, k = 0, 1, \ldots, n$ are the barycentric coordinates of $c_{\rm B}^-$ in the simplex $conv(S_0, S_1, \ldots, S_n)$ for any S_0.

This achieves the proof. □

One can deduce a new relation between the two affine lower bound functions.

Proposition 3 *The vector of the barycentric coordinates of $c_{\rm B}^-$ is the optimal solution of the linear program, dual of (P_l) for the corresponding admissible simplex.*

Proof The coefficients α_k defined in Sect. 4 are the marginal costs related to the solution of (P_l) from where we get the lower bound $z_{\rm asf}^-$. □

Then, we see that these two bounding techniques associated, respectively, to a convex and concave relaxation are strongly connected.

5.2 Comparison of the lower bounds

We deal now with the relative efficiencies of the lower bounds $z_{\rm asf}^-$ and $z_{\rm B}^-$ for multivariate functions. Keeping in mind their respective values, we can claim, following the notation above:

Proposition 4 *Let X be a box in $\mathbb{R}^n$, let $S_0, S_1, \ldots, S_n$ be any admissible set of vertices of X defined as above, then for a multivariate differentiable function the lower bound $z_{\rm B}^-$ given by the optimal linear centered form of Baumann is greater than the lower bound $z_{\rm asf}^-$ given by the simplex linear boundary value form over X if and only if*

$$\sum_{k=0}^{n} \alpha_k f(S_k) \leq f(c_{\rm B}^-).$$

Proof Obvious from the analytical expressions of the lower bounds. □

Moreover, the coefficients α_k are the barycentric coordinates of $c_{\rm B}^-$ in the admissible simplex; the convex set defined by the admissible simplex is denoted by $\text{conv}(S_0, S_1, \ldots, S_n)$. Therefore the two lower bounds have the same value when the point $(c_{\rm B}^-, f(c_B^-)) \in \text{conv}(\ (S_k, f(S_k)),\ k = 0, 1, \ldots, n)$, n-simplex in $\mathbb{R}^n \times \mathbb{R}$. This follows from the fact that $c_{\rm B}^- = \sum_{k=0}^n \alpha_k S_k$ with $\sum_{k=0}^n \alpha_k = 1$ and $\alpha_k \in [0, 1]$.

Corollary 1 *If the function f is concave over the box X then $z_{\rm B}^-$ is a greater lower bound than $z_{\rm asf}^-$. Conversely, if the function f is convex (which is the case for a local minimum) the best lower bound is given by $z_{\rm asf}^-$.*

Example 6 Consider once more the function $f_1(x_1, x_2) = 1 + (x_1^2 + 2)x_2 + x_1x_2^2$, with $x_1 \in [1, 2]$ and $x_2 \in [-10, 10]$; one has $f(c_{\rm B}^-) = -1.35141\cdots$ and $\sum_{k=0}^n \alpha_k f(S_k) = 126.6358\cdots$. Therefore, $\sum_{k=0}^n \alpha_k f(S_k) > f(c_{\rm B}^-)$. This implies that the lower bound is most efficient using the asf form $z_{\rm asf}^- = -314.5957\cdots$ (using the admissible simplex $(1, -10), (1, 10), (-10, 10)$), for $f^* = -3.5$.

From the last result we can think that the two formulations may be used simultaneously to get a better efficiency in a branch-and-bound type method for global optimization. For example, if the formula of Baumann yields a bad lower bound, one can deduce that $f(c_{\rm B}^-)$ has a lower value which could give a tight upper bound for f^*, and thus that value of the function may be used in a *compound middle point test*.

6 Application to global optimization

The purpose of this section is to compare the efficiency of different inclusion functions for some test problems of global optimization. The branch-and-bound algorithm we use, is due to Moore–Skelboe and can be found in [17]. This algorithm is modified by setting a new value for the current solution $\tilde{f}$, see the second item in 7(b).

6.1 Algorithm

The modified Moore-Skelboe algorithm is the following:
Algorithm

1. `Set` $X :=$ `the initial domain in which the global minimum is sought for,` $X \subseteq \mathbb{R}^n$.
2. `Set` $\tilde{f} := +\infty$.
3. `Set` $\mathcal{L} := (+\infty, X)$.
4. `Extract from` $\mathcal{L}$ `the box for which the lowest lower bound has been computed.`
5. **Choose a coordinate direction** k `parallel to which the box has the maximum length`
6. **Bisect the selected box** `normal to direction` k, `yielding two sub-boxes` V_1, V_2.
7. `For j:=1 to 2 do`
   (a) **Compute** $v_j :=$ **lower bound of** f **over** V_j **by using one of the presented methods.**
   (b) `If` $\tilde{f} \geq v_j$ `then`
      - `Insert` (v_j, V_j) `in` $\mathcal{L}$.
      - `Set` $\tilde{f} := \min(\tilde{f}, f(m))$, `where` m `is a point in` V_j
      - `If` $\tilde{f}$ `has been modified then remove from` $\mathcal{L}$ `all` (z, Z) `where` $z > \tilde{f}$.
8. `If` $\tilde{f} - \min_{(z,Z)\in\mathcal{L}} z \leq \epsilon$ `then STOP. Else GoTo Step 4.`

ϵ is a positive value which represents the desired accuracy for the global minimum value. The computation of the enclosures of the gradient are performed by using an interval automatic differentiation code in direct mode [9]. The point m in a considered box V_j is generally the middle point c of that box (i.e. $c_i = \frac{(V_j)_i^L + (V_j)_i^U}{2}, \forall i$). Nevertheless the enclosure of the function may be improved choosing a suitable point m, on one hand considering the Baumann center c_{B}^- and on the other hand considering the middle of the interval c' only when the monotonicity of the function is not established:

- $c'_i = \dfrac{(V_j)_i^L + (V_j)_i^U}{2}$, if $L_i U_i < 0$,
- $c'_i = (V_j)_i^L$, if $U_i \leq 0$,
- $c'_i = (V_j)_i^U$, if $L_i \geq 0$,

where L_i and U_i represent the enclosure (lower and upper bounds) of the ith partial derivative of the function considered. In this way, monotonicity is considered for the computation of the upper bound $\tilde{f}$. The classical monotonicity test is not used here in order to show clearly the efficiency of the lower bounds computed by using the different methods. Here, the following methods using an enclosure of the gradient are considered (the natural interval extension of a function does not give generally any solution for the considered following accuracies): Taylor at the order one, Taylor Baumann, a linear boundary value form using an admissible simplex starting from the vertex $S_0 = \left((V_j)_1^L, \ldots, (V_j)_i^L, \ldots, (V_j)_n^L\right)$, and a mixed form using Proposition 4 (the best value between the Taylor Baumann or admissible simplex lower bound).

6.2 Numerical tests

All the tests are performed on a Standard-PC computer with an 1.8 GHz AMD Athlon Processor and 256 Mb RAM using a Fortran-90 compiler. All the computations, even the floating-point operations, are performed using *rounded interval analysis* [14]; this implies that all computations are rigorously performed: no numerical error can occur.

The following considered functions come from the classical literature [13, 17]. f^* represents the optimal value and x^* a corresponding optimizer as denoted in Sect. 2.

- $f_1(x) = 1 + (x_1^2 + 2)x_2 + x_1x_2^2, x_1 \in [1, 2], x_2 \in [-10, 10]\, \epsilon = 10^{-8}, f^* = -3.5, x^* = (2, -1.5)$.
- $f_2(x) = 2x_1^2 - 1.05x_1^4 + x_2^2 - x_1x_2 + \frac{1}{6}x_2^6, \forall x_i \in [-2, 4]\, \epsilon = 10^{-8}, f^* = -239.696, x^* = (4, 1.115)$.
  $-f_2(x) = -2x_1^2 - 1.05x_1^4 + x_2^2 - x_1x_2 + \frac{1}{6}x_2^6, \forall x_i \in [-2, 4]\, \epsilon = 10^{-8}, f^* = -239.696, x^* = (4, 1.115)$.
- $f_3(x) = (x_1 - 2x_2 - 7)^2 + (2x_1 + x_2 - 5)^2, x_1 \in [-2.5, 3.5], x_2 \in [-1.5, 4.5]\, \epsilon = 10^{-8}, f^* = 0.45, x^* = (3.4, -1.5)$.
- $$f_4(x) = \left(1 + ((x_1 + x_2 + 1)^2)(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)\right) \times \left(30 + ((2x_1 - 3x_2)^2)(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)\right),$$
  $\forall x_i \in [-2, 2]\, \epsilon = 10^{-8}, f^* = 3, x^* = (0, -1)$.
- $f_5(x) = (x_1 - 1)(x_1 + 2)(x_2 + 1)(x_2 - 2)x_3^2, \forall x_i \in [-2, 2]\, \epsilon = 10^{-8}, f^* = -36, x^* = (-0.5, -2, 2)$.
  $-f_5(x) = -(x_1 - 1)(x_1 + 2)(x_2 + 1)(x_2 - 2)x_3^2, \forall x_i \in [-2, 2]\, \epsilon = 10^{-8}, f^* = -36, x^* = (-0.5, -2, 2)$.
- $f_6(x) = 4x_1^2 - 2x_1x_2 + 4x_2^2 - 2x_2x_3 + 4x_3^2 - 2x_3x_4 + 4x_4^2 + 2x_1 - x_2 + 3x_3 + 5x_4, x_1 \in [-1, 3], x_2 \in [-10, 10], x_3 \in [1, 4], x_4 \in [-1, 5] \epsilon = 10^{-8}, f^* = 5.7708, x^* = (-0.17, 0.33, 1, -0.375)$.
- $f_7(x) = x_1^2 + x_2^2 - \cos(18x_1)x_1 \sin(18x_2) + x_3 \cos(x_3) + x_1x_2x_3, \forall x_i \in [1, 500]^3\, \epsilon = 10^{-8}, f^* = 1, x^* = (1, 1, 9.148)$.
- $f_8(x) = (x_1 + 10x_2)^2 + (5x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4, x_i \in [-10, 10], \forall i, \epsilon = 10^{-3}, f^* = 0, x^* = (0, 0, 0, 0)$. Function due to Powell.
- $f_9(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4, x_i \in [-1000, 1000], \forall i, \epsilon = 10^{-8}$, the solutions are $f^* = -1.0316, x^* = (0, -0.713)$. Function due to Ratschek.
- $f_{10}(x) = 12x_1^2 - 6.3x_1^4 + x_1^6 + 6x_2(x_2 - x_1), x_i \in [-100, 100], \forall i, \epsilon = 10^{-8}, f^* = 0, x^* = (0, 0)$. Function named Three-Hump.
- $f_{11}(x) = -\frac{1}{(\sqrt{x_1^2 + x_2^2} - 0.3)^2 + 0.01} - \frac{1}{(\sqrt{x_1^2 + x_2^2} - 0.9)^2 + 0.04} + 6, x_1 \in [-1, 2], x_2 \in [0.001, 3]\, \epsilon = 10^{-4}$, the solutions are $f^* = -96.5, x^* = (0, 0.288)$. Function hump.
  $-f_{11}(x) = \frac{1}{(\sqrt{x_1^2 + x_2^2} - 0.3)^2 + 0.01} + \frac{1}{(\sqrt{x_1^2 + x_2^2} - 0.9)^2 + 0.04} - 6, x_1 \in [-1, 2], x_2 \in [0.001, 3]\, \epsilon = 10^{-8}$, the solutions are $f^* = -5.77, x^* = (2, 3)$.

In Table 1, computations of the different lower bounds only are considered. They are denoted by T_1 for the inclusion function based on a centered Taylor expansion at the first order, T_{B} for the inclusion function due to Baumann, AS for the inclusion function which used the admissible simplex method, MIXED is a narrowing method which chooses at each step the best method between T_{B} and AS in a deterministic way. For every technique the classical midpoint test is used considering $f(c)$; where $c_i = \frac{(V_j)_i^L+(V_j)_i^U}{2}$. In all the tables, "Pbs" represents the addressed problems: $f_1, f_2 \ldots$, "Its" is the number of iterations, "time(s)" is the CPU-time in seconds, "$|\mathcal{L}|$" is the number of elements in the list at the end of the program. In Table 3, "Nb AS" and "Nb Baumann" represent the number of time where the admissible simplex versus the Baumann methods are used inside the mixed algorithm (when equality occurs it is noticed in "Nb equality").

Comparing the four methods on Table 1, one can observe that AS is the most efficient. In fact, it produces results close to those of the T_{B} technique excepted for the two hardest cases: f_8 and f_{11}. The mixed method leads to the best results when considering only the number of iterations; that follows from Proposition 4. Unfortunately, each iteration of the mixed algorithm is more expensive in CPU-time than in the other methods and thus, most often the CPU-times for the mixed method are almost equivalent to those of AS.

Considering the first column in Table 3, which denotes how much time each method T_{B} versus AS is used in the MIXED technique—always with the classical midpoint test $f(c)$-, we see that the AS method generally produces the most efficient lower bounds.

In Table 2, the middle point test is changed by the point c' (which takes into account the monotonicity of the considered function) and c_{B}^- which is the Baumann center (computed for z_{B}^-) and an heuristic choice between these two values for the mixed technique. The heuristic is to consider $f(c_{\mathrm{B}}^-)$ when the AS produces the best lower bound and $f(c')$ else. Comparing Tables 1 and 2, we observe impressive gains by considering these most efficient upper bounds: $f(c')$ and/or $f(c_{\mathrm{B}}^-)$ (expected for f_8 and f_{11}). Thus considering efficient upper bounds—which exploit the monotonicity of the

Table 1 Numerical results with $f(c)$

Pbs	T_1			T_{B}			AS			MIXED		
	Its	Time(s)	$\|\mathcal{L}\|$	Its	Time(s)	$\|\mathcal{L}\|$	Its	Time(s)	$\|\mathcal{L}\|$	Its	Time(s)	$\|\mathcal{L}\|$
f_1	167	0.16	34	133	0.11	28	131	0.11	26	132	0.11	28
f_2	87	0.11	23	87	0.11	23	84	0.16	22	86	0.11	23
$-f_2$	116	0.17	27	109	0.11	27	84	0.16	22	110	0.11	29
f_3	124	0.11	26	107	0.11	25	101	0.11	25	101	0.11	25
f_4	5503	2.64	63	3847	1.92	47	3734	1.92	47	3731	2.15	47
f_5	735	0.39	38	364	0.22	56	364	0.33	57	364	0.33	57
$-f_5$	148	0.21	6	127	0.22	6	128	0.22	6	125	0.17	6
f_6	2464	1.27	1079	1585	0.77	697	1136	0.60	511	1183	0.66	531
f_7	13856	4.50	1567	9908	3.46	1564	9778	3.63	1518	9587	3.84	1535
f_8	266973	2649.10	68035	112841	496.2	32847	57023	116.33	14881	57023	116.87	14881
f_9	1222	0.49	24	753	0.27	19	691	0.27	19	687	0.28	19
f_{10}	1270	0.39	10	813	0.28	11	693	0.27	9	689	0.28	9
f_{11}	38048	68.22	14432	26894	42.89	12146	21583	27.91	10621	21583	28.17	10621
$-f_{11}$	265	0.17	1	183	0.11	1	180	0.17	1	178	0.22	1

Table 2 Numerical results with $f(c')$ or $f(c_B^-)$

Pbs	$T_1+f(c')$			$T_B+f(c_B^-)$			AS $+f(c')$			MIXED $+f(c')$ or $f(c_B^-)$		
	Its	Time(s)	$\vert\mathcal{L}\vert$	Its	Time(s)	$\vert\mathcal{L}\vert$	Its	Time(s)	$\vert\mathcal{L}\vert$	Its	Time(s)	$\vert\mathcal{L}\vert$
f_1	95	0.05	1	72	0.05	2	73	0.05	2	71	0.06	2
f_2	41	0.05	0	39	0.05	0	37	0.06	0	37	0.05	0
$-f_2$	64	0.05	1	56	0.06	2	84	0.16	22	53	0.05	0
f_3	71	0.05	0	57	0.06	1	54	0.06	1	54	0.06	1
f_4	5503	2.47	63	3847	1.81	47	3734	1.92	47	3731	1.98	47
f_5	453	0.16	15	221	0.05	7	213	0.05	14	213	0.05	11
$-f_5$	49	0.01	22	25	0.04	5	26	0.01	2	26	0.01	2
f_6	1330	0.55	54	840	0.33	30	597	0.33	36	596	0.33	34
f_7	13856	3.95	1133	9908	2.69	1073	9778	3.19	1027	9587	3.13	1034
f_8	266885	2744.02	67980	112769	492.73	32728	57023	113.92	14880	56987	116.55	14847
f_9	1222	0.22	24	753	0.16	19	691	0.22	19	687	0.21	17
f_{10}	1270	0.32	10	811	0.06	9	693	0.17	9	689	0.11	3
f_{11}	38048	67.44	14432	26894	42.62	12146	21583	27.63	10621	21583	28.23	10621
$-f_{11}$	217	0.05	1	135	0.01	0	132	0.05	0	130	0.01	0

Table 3 Number of computed lower bounds

Pbs	MIXED $+f(c)$			MIXED $+f(c')$ or $f(c_B^-)$		
	Nb AS	Nb Baumann	Nb equality	Nb AS	Nb Baumann	Nb equality
f_1	193	27	44	116	11	15
f_2	85	10	77	52	4	18
$-f_2$	109	31	72	81	5	20
f_3	167	0	35	91	0	17
f_4	6696	763	3	6696	763	3
f_5	508	60	160	332	19	75
$-f_5$	26	15	209	26	15	11
f_6	2024	176	166	332	19	75
f_7	13624	2974	2576	13624	2974	2576
f_8	107818	93	6135	107748	93	6133
f_9	1198	58	118	1198	58	118
f_{10}	1088	108	182	1088	108	182
f_{11}	32250	57	10859	32250	57	10859
$-f_{11}$	155	79	122	155	79	26

function—the convergence of such branch-and-bound algorithms can be considerably improved.

In that case, the AS method produces most often the best lower bounds (see the second column of Table 3).

7 Conclusion

This paper proposes a deterministic way to compare two of the most efficient methods for computing a lower bound of a function over a box. The first method is due to Baumann [2, 3]. The resulting lower bound is calculated using a Taylor expansion at the first order and computing the optimal center of this form. The second method was introduced in [12, 13] and is based on a linear boundary value form.

Proposition 4 allows determination of which one between the two methods T_B and AS is locally the most efficient by only comparing the evaluation of the considered function at the Baumann center with $n + 1$ function values.

These main results on duality and convexity are introduced into a Branch-and-Bound algorithm for global optimization in order to compare these different bounding methods on some numerical tests. Fourteen functions from the literature are minimized. These numerical experiments clearly show that the AS method produces the most efficient bounds (see Table 3). Proposition 4 also leads to a mixed method, called MIXED in the tables of results. This mixed method produces the most efficient lower and upper bounds and consequently the best results when considering the number of iterations only. However, the comparison test of MIXED is expensive in CPU-time and with this last criterion the results are close to those of the AS method.

References

1. Alefeld, G., Herzberger, J.: Introduction to Interval Computations. Academic Press, New York (1983)
2. Baumann, E.: Globale Optimierung Stetig differenzierbarer Funktionen einer variablen. Freiburger Intervall-Berichte 86(6), 1–93 (1986)
3. Baumann, E.: Optimal Centered Forms. BIT 28, 80–87 (1988)
4. Carrizosa, E., Hansen, P., Messine, F.: Improving interval analysis bounds by translations. J. Global Optimization 29(2), 157–172 (2004)
5. Griewank, A.: Evaluating Derivative, Principles and Techniques of Algorithmic Differentiation. SIAM 369 p (2000)
6. Hansen, E.: Global optimization using interval analysis. Marcel Dekker, New-York (1992)
7. Hansen, P., Jaumard, B., Xiong, J.: Decomposition and interval arithmetic applied to global minimization of polynomial and rational functions. J. Global Optimization 3, 421–437 (1992)
8. Hansen, P., Jaumard, B., Xiong, J.: Cord-slope form of Taylor's expansion in univariate global optimization. J. Optimization Theory Appl. **80**(3), 441–464 (1994)
9. Kearfott, R.B.: Rigorous Global Search: Continuous Problems. Kluwer Academic Publishers, Boston (1996)
10. Krawczyk, R., Nickel, K.: The centered form in interval arithmetics: quadratic convergence and inclusion isotonicity. Computing 28, 117–137 (1982)
11. Lagouanelle, J.L.: Algorithme du Cerf-Volant en Optimisation Globale par Intervalles. Internal Report, RT/APO/01/2, ENSEEIHT, Toulouse (2002)
12. Lagouanelle, J.L., Messine, F.: Algorithme d'encadrement de l'optimum global d'une fonction différentiable. Compte Rendu à l'Académie des Sciences, Numerical Analysis, T. 326, Série I (1998)
13. Messine, F., Lagouanelle, J.L.: Enclosure methods for multivariate differentiable functions and application to global optimization. J. Univ. Comput. Sci. **4**(6), 589–603 Springer-Verlag, (1998)
14. Moore, R.E.: Interval Analysis. Prentice-Hall Inc. Englewood Cliffs, NJ (1966)
15. Neumaier, A.: Interval Methods for Systems of Equations. Cambridge University Press, Cambridge (1990)
16. Rall, L.B.: Automatic differentiation, techniques and applications. Lecture Notes in Computer Science, N. 120, Springer-Verlag, Berlin (1981)
17. Ratschek, H., Rokne, J.G.: New Computers Methods for Global Optimization. Wiley, New-York (1988)
18. Vinko, T., Lagouanelle, J.L., Csendes, T.: A New Inclusion Function for Optimization: Kite- The One Dimensional Case. J. Global Optimization, **30**(4), 435–456 (2004)