Theoretical rate of convergence for interval inclusion functions

Daniel Scholz

Abstract   Geometric branch-and-bound methods are commonly used solution algorithms for non-convex global optimization problems in small dimensions, say for problems with up to six or ten variables, and the efficiency of these methods depends on some required lower bounds. For example, in interval branch-and-bound methods various well-known lower bounds are derived from interval inclusion functions. The aim of this work is to analyze the quality of interval inclusion functions from the theoretical point of view making use of a recently introduced and general definition of the rate of convergence in geometric branch-and-bound methods. In particular, we compare the natural interval extension, the centered form, and Baumann's inclusion function. Furthermore, our theoretical findings are justified by detailed numerical studies using the Weber problem on the plane with some negative weights as well as some standard global optimization benchmark problems.

Keywords   Global optimization · Interval analysis · Continuous problems · Approximation algorithms · Geometric branch-and-bound · Facility location problems

1 Introduction

In global optimization, we want to find the global minimum of a given objective function $f : \mathbb{R}^n \to \mathbb{R}$ over a feasible area X, i.e. we want to solve the optimization problem

$$\min_{x \in X} \, f(x).$$

Since these problems are often hard to solve even in relatively small dimensions n, geometric branch-and-bound methods are suitable solution approaches, see e.g. Horst et al. (2000) or Floudas and Pardalos (2009). The main task throughout these algorithms is the calculation of the required lower bounds. Several quite different techniques to do so can be found in the

D. Scholz (✉)
Institut für Numerische und Angewandte Mathematik, Georg-August-Universität Göttingen, Göttingen, Germany
e-mail: dscholz@math.uni-goettingen.de

literature. For instance, commonly used techniques for facility location problems have been introduced in Drezner and Suzuki (2004), Hansen et al. (1985), Plastria (1992), and Blanquero and Carrizosa (2009). Recently, we analyzed the quality of these bounding operations from the theoretical point of view, see Schöbel and Scholz (2010).

Interval analysis is a general framework for calculations with intervals, see e.g. Ratschek and Rokne (1988), which is also a suitable tool for the calculation of lower bounds as outlined in the textbook Hansen (1992) or in Ratschek and Voller (1991). Some empirical studies concerning the quality of these bounds can be found in the literature, see Csallner and Csendes (1996), Tóth and Csendes (2005), and Tóth et al. (2007). Our main interest in the present paper is to extend our results from Schöbel and Scholz (2010) to interval methods. In particular, we will analyze bounding operations derived from interval analysis from the theoretical point of view using the same definition of the rate of convergence as introduced in Schöbel and Scholz (2010). In other words, our definition of the rate of convergence can be used to analyze several quite different bounding procedures known from the literature.

The remainder of the paper is organized as follows. In the next section we give a summary of interval analysis which will be needed throughout this paper. Next, the geometric branch-and-bound method is briefly reviewed in Sect. 3. The main contribution of our work can be found in Sect. 4 where we analyze the calculation of some lower bounds derived from interval analysis from the theoretical point of view. Furthermore, detailed numerical results in Sect. 5 justify our theoretical findings from the empirical point of view. Finally, a brief discussion of our results can be found in Sect. 6.

2 Interval analysis

In this section we summarize some principles of interval analysis as given for example in Neumaier (1990), Ratschek and Rokne (1988), or Hansen (1992). Note that we assume compact intervals throughout this paper.

Notation 1 A (compact) interval X is denoted by

$$X = [a, b] \subset \mathbb{R}$$

with $a \leq b$. Moreover, the left and right endpoints are denoted by $X^L = a$ and $X^R = b$, respectively. If $X^L = X^R = z$ we will sometimes use the short form $X = z = [z, z]$, i.e. $[z, z]$ is equivalent to z.

Next, arithmetic operations between intervals are defined as follows.

Definition 2 Let $X = [a, b]$ and $Y = [c, d]$ be two intervals. Then the ***interval arithmetic*** is given by

$$X \star Y := \{x \star y : x \in X, \ y \in Y\},$$

where $\star$ denotes the addition, multiplication, subtraction, division, minimum, or maximum as long as $x \star y$ is defined for all $x, y \in Y$.

By definition, $X \star Y$ again yields an interval which contains $x \star y$ for all $x \in X$ and $y \in Y$ and which can be computed easily. Apart from interval arithmetic also interval operations are defined as follows.

Definition 3 Let $X = [a, b]$ be an interval. Then the ***interval operation*** is given by

$$op(X) := \{op(x) : x \in X\} = \left[\min_{x \in X} op(x), \max_{x \in X} op(x) \right],$$

where $op : X \to \mathbb{R}$ denotes a continuous function such that $op(X)$ is an interval.

In the following we assume operations such that the interval $op(X)$ can be computed easily. For example, if op is an increasing function we obtain

$$op(X) = [op(a), op(b)]$$

and if op is a decreasing function we obtain

$$op(X) = [op(b), op(a)]$$

for all intervals $X = [a, b]$.

Definition 4 An ***interval function*** $F(X_1, \ldots, X_n)$ is an interval valued function with n intervals as argument using interval arithmetics and interval operations as defined before.

Example 1 An interval function $F(X, Y)$ with two intervals X and Y as argument is

$$F(X, Y) = \exp\left(\frac{X + Y}{Y^2 + [1, 1]} \right) = \exp\left(\frac{X + Y}{Y^2 + 1} \right).$$

For example, we obtain

$$F([0, 2], [-1, 1]) = \exp\left(\frac{[-1, 3]}{[0, 1] + [1, 1]} \right) = \exp([-1, 3]) = [\exp(-1), \exp(3)].$$

Since our main goal in this work is to solve global optimization problems, we now define the natural interval extension.

Definition 5 Let $f(x_1, \ldots, x_n)$ be a fixed representation (see Example 4) of a real valued function with n real numbers as argument using arithmetics and operations such that the corresponding interval arithmetics and interval operations are defined.

Then the ***natural interval extension*** of $f(x_1, \ldots, x_n)$ is given by the interval function $F(X_1, \ldots, X_n)$ where arithmetics and operations are replaced by their corresponding interval arithmetics and interval operations.

Example 2 The natural interval extension of

$$f(x, y) = 4 \cdot x^2 + \frac{\sin(y)}{x^2 + 1}$$

is given by

$$F(X, Y) = [4, 4] \cdot X^2 + \frac{\sin(Y)}{X^2 + [1, 1]} = 4 \cdot X^2 + \frac{\sin(Y)}{X^2 + 1}$$

where X and Y are intervals.

The natural interval extension leads to general lower bounds as required throughout branch-and-bound algorithms. To this end, we need the following statement which can be found in any standard textbook of interval analysis such as Neumaier (1990), Ratschek and Rokne (1988), or Hansen (1992).

Theorem 1 (Fundamental theorem) *Let $F(X_1, \ldots, X_n)$ be the natural interval extension of $f(x_1, \ldots, x_n)$. Then*

$$f(Y_1, \ldots, Y_n) \subseteq F(Y_1, \ldots, Y_n)$$

for all intervals $Y_k \subseteq X_k$ for $k = 1, \ldots, n$, where

$$f(Y_1, \ldots, Y_n) := \{f(x_1, \ldots, x_n) \; : \; x_k \in Y_k \;\; for \; k = 1, \ldots, n\}.$$

Proof See, for instance, Hansen (1992).   □

The following two examples show some properties of natural interval extension $F(X_1, \ldots, X_n)$ which we always should keep in mind.

Example 3 In general we have

$$f(X_1, \ldots, X_n) \neq F(X_1, \ldots, X_n).$$

For example, consider $f(x) = x^2 - 2x$ with natural interval extension

$$F(X) = X^2 - 2 \cdot X$$

and let $Y = [1, 2]$. Since f is monotone increasing on Y, we find

$$f(Y) = [f(1), f(2)] = [-1, 0].$$

But the natural interval extension yields

$$F(Y) = F([1, 2]) = [1, 4] - [2, 4] = [-3, 2].$$

Example 4 Defining

$$f_1(x) = 4(x^2 - x) \quad \text{and} \quad f_2(x) = (2x - 1)^2 - 1,$$

we have $f_1(x) = f_2(x)$ for all $x \in \mathbb{R}$. This is not true for the natural interval extension and intervals. Consider the natural interval extensions

$$F_1(X) = 4 \cdot (X^2 - X) \quad \text{and} \quad F_2(X) = (2 \cdot X - 1)^2 - 1.$$

Then, for $Y = [0, 2]$, we obtain

$$F_1([0, 2]) = [-8, 16] \quad \text{and} \quad F_2([0, 2]) = [-1, 8].$$

Therefore, we always assume a fixed representation of $f(x)$ if we consider the corresponding natural interval extension $F(X)$, see Definition 5.

Moreover, note that $F_2([0, 2]) = f_2([0, 2])$ since $f_2(\frac{1}{2}) = -1$ and $f_2(2) = 8$. Thus, the interval function $F_2(X)$ yields stronger bounds as we will see in the following.

In order to avoid these problems, we introduce the following definition which is of fundamental importance in the following sections.

Definition 6 A fixed representation of a continuous real valued function $f(x_1, \ldots, x_n)$ is said to be a ***single-use expression*** if in its representation every variable x_1 to x_n occurs at most once.

Example 5 Consider

$$f_1(x) = \exp(4(x^2 - x)) \quad \text{and} \quad f_2(x) = \exp((2x - 1)^2 - 1).$$

Then f_2 is a single-use expression while f_1 is not although $f_1(x) = f_2(x)$ for all $x \in \mathbb{R}$. As a second example consider

$$f_1(x, y) = (x + y)^2 + 3 \quad \text{and} \quad f_2(x, y) = x^2 + 2xy + 3 + y^2.$$

Here, only f_1 is a single-use expression although again $f_1(x, y) = f_2(x, y)$ for all $(x, y) \in \mathbb{R}^2$.

The reason for the definition of single-use expressions is the following result.

Theorem 2 *Let $F(X_1, \ldots, X_n)$ be the natural interval extension of a single-use expression $f(x_1, \ldots, x_n)$. Then*

$$f(Y_1, \ldots, Y_n) = F(Y_1, \ldots, Y_n)$$

for all intervals $Y_k \subseteq X_k$ for $k = 1, \ldots, n$.

Proof See, for instance, Neumaier (1990).   □

Remark 1 Let $f(x_1, \ldots, x_n)$ be a single-use expression, let $F(X_1, \ldots, X_n)$ be the corresponding natural interval extension, and consider some intervals $Y_k \subseteq X_k$ for $k = 1, \ldots, n$. Then, from Theorem 2 we know that there exists a $(y_1, \ldots, y_n) \in Y_1 \times \cdots \times Y_n$ such that

$$f(y_1, \ldots, y_n) = F(Y_1, \ldots, Y_n)^L.$$

But note that it is in general not an easy task to find such a $(y_1, \ldots, y_n)$.

For the case of general interval functions, we need the following definition.

Definition 7 An interval function $F(X_1, \ldots, X_n)$ is said to be an ***interval inclusion function*** of $f(x_1, \ldots, x_n)$ if

$$f(Y_1, \ldots, Y_n) \subseteq F(Y_1, \ldots, Y_n)$$

for all intervals $Y_k \subseteq X_k$ for $k = 1, \ldots, n$.

Hence, every interval inclusion function leads to lower bounds on the objective function as required throughout the following prototype algorithm. As an example, the natural interval extension yields an interval inclusion function due to Theorem 1.

3 The prototype algorithm

In order to simplify our notations we need the following definition.

Definition 8 A compact ***box*** X is the Cartesian product of intervals, i.e.

$$X = [a_1, b_1] \times \cdots \times [a_n, b_n].$$

Moreover, the ***diameter*** of a box $X \subset \mathbb{R}^n$ is

$$\delta(X) = \max\{\|x - x'\|_2 \, : \, x, x' \in X\} = \sqrt{(b_1 - a_1)^2 + \cdots + (b_n - a_n)^2}$$

and the ***center*** of a box $X \subset \mathbb{R}^n$ is defined by

$$c(X) = \left(\frac{1}{2}(a_1 + b_1), \ \ldots, \ \frac{1}{2}(a_n + b_n) \right).$$

Consider an objective function $f : \mathbb{R}^n \to \mathbb{R}$ defined on a box $X \subset \mathbb{R}^n$. Then our goal is to minimize f on X. To this end, we suggest a geometric branch-and-bound algorithm which needs a bounding operation as introduced in Schöbel and Scholz (2010).

Notation 9 Let $X \subset \mathbb{R}^n$ be a box and consider $f : X \to \mathbb{R}$. A ***bounding operation*** is a procedure to calculate for any subboxes $Y \subseteq X$ a ***lower bound*** $LB(Y) \in \mathbb{R}$ with

$$LB(Y) \ \leq \ f(x) \quad \text{for all} \quad x \in Y$$

and to specify a point $P(Y) \in Y$.

Note that the choice of $P(Y)$ is of fundamental importance for our theoretical and empirical studies in the following.

For a given bounding operation we can use the following algorithm to find an approximation of the global minimum $x^* \in X$ within an absolute accuracy of $\varepsilon > 0$.

(1) Calculate a lower bound $LB(X)$ for $f(Y)$ and set $UB = f(P(X))$ and $\mathcal{L} = \{X\}$.
(2) Choose a box with the lowest lower bound in $\mathcal{L}$, split it into s congruent smaller boxes $Y_1, \ldots, Y_s$, delete the selected box from $\mathcal{L}$, and add $Y_1, \ldots, Y_s$ to $\mathcal{L}$. Calculate lower bounds $LB(Y_1), \ldots, LB(Y_s)$ and update

$$UB = \min\{UB, f(P(Y_1)), \ldots, f(P(Y_s))\}.$$

Delete all boxes Y from $\mathcal{L}$ with $LB(Y) + \varepsilon \geq UB$.
(3) When there are no boxes left, i.e. $\mathcal{L} = \emptyset$, the algorithm terminates and UB is within the absolute accuracy of ε from the optimum. If there are boxes left, return to step (**2**).

We remark that this algorithm is only appropriate for problems in small dimensions, say for problems with up to six or ten variables. The reason is that the number of considered subboxes throughout the algorithm as well as the time for calculating the required lower bounds increases very fast with increasing number of variables.

In the next section, we discuss several bounding operations which can be easily derived using methods from interval analysis.

4 Interval inclusion functions

In order to analyze bounding operations from the theoretical point of view, we need one more definition which again was introduced in Schöbel and Scholz (2010).

Definition 10 Let $X \subset \mathbb{R}^n$ be a box and $f : X \to \mathbb{R}$. Furthermore, consider the minimization problem

$$\min_{x \in X} f(x).$$

We say a bounding operation has the ***rate of convergence*** $p \in \mathbb{N}$ if there exists a constant $C > 0$ such that

$$f(P(Y)) - LB(Y) \ \leq \ C \cdot \delta(Y)^p \tag{1}$$

for all boxes $Y \subseteq X$.

Although some empirical studies concerning a related concept for the rate of convergence can be found in the literature, see Csallner and Csendes (1996) and Tóth and Csendes (2005), our definition including the specific point $P(Y)$ leads to a concept which can be applied for bounding operations in general and not only for interval inclusion functions. In particular, we calculated the theoretical rate of convergence for the Lipschitzian bounding operation, the d.c. bounding operation, and a general bounding operation in Schöbel and Scholz (2010). Our goal in the following is to calculate the rate of convergence also for some bounding operations given by interval inclusion functions.

4.1 Natural interval bounding operation

Assume that $f(x) = f(x_1, \ldots, x_n)$ consists only of arithmetics and operations such that the corresponding natural interval extension $F(X) = F(X_1, \ldots, F_n)$ exists, see Definition 5, and recall that the superscript ^{L} indicates the left endpoint of an interval. Then, from the fundamental theorem of interval analysis, see Theorem 1, we directly obtain the ***natural interval bounding operation***

$$LB(Y) = F(Y)^L \quad \text{and} \quad P(Y) = c(Y).$$

The natural interval bounding operation can be used in general for global optimization problems. Numerical examples can be found e.g. in Hansen (1992). Moreover, this bounding operation also finds several applications in location problems, see e.g. Fernández et al. (2007) and Tóth et al. (2009).

Before we can present a general statement concerning the rate of convergence of the natural interval bounding operation, we need the following result.

Lemma 3 *Consider a box $X \subset \mathbb{R}^n$, let $f : X \to \mathbb{R}$ be a Lipschitzian function with constant L such that $f(x)$ is a single-use expression, and assume that the natural interval extension $F(X)$ of $f(x)$ exists. Then*

$$\delta(F(Y)) = F(Y)^R - F(Y)^L \; \le \; C \cdot \delta(Y)$$

holds for all boxes $Y \subseteq X$ and the constant $C > 0$ does not depend on Y.

Proof From Theorem 2 we know that

$$f(Y) = F(Y) = [F(Y)^L, \, F(Y)^R]$$

for all subboxes $Y \subseteq X$ since $f(x)$ is a single-use expression. Thus, there are some $\ell, r \in Y$ with

$$f(\ell) = F(Y)^L \quad \text{and} \quad f(r) = F(Y)^R.$$

Therefore, we find

$$F(Y)^R - F(Y)^L = f(r) - f(\ell) \; \le \; L \cdot \|r - \ell\|_2 \; \le \; L \cdot \delta(Y)$$

and the statement is shown.   □

Lemma 4 *Consider a box $X \subset \mathbb{R}^n$, let $m_1, \ldots, m_s : X \to \mathbb{R}$ be Lipschitzian functions such that $m_1(x)$ to $m_s(x)$ are single-use expressions, and assume that the natural interval extensions $M_1(X)$ to $M_s(X)$ exist. Furthermore, consider*

$$f(x) = h(m_1(x), \ldots, m_s(x)),$$

where $h : \mathbb{R}^s \to \mathbb{R}$ is a Lipschitzian function with constant L such that $h(x)$ is a single-use expression and assume that the natural interval extension $H(X)$ exists. Then

$$\delta(F(Y)) = F(Y)^R - F(Y)^L \leq C \cdot \delta(Y)$$

holds for all boxes $Y \subseteq X$ and the constant $C > 0$ does not depend on Y. Moreover, $F(X)$ is the natural interval extension of $f(x)$.

Proof For any $Y \subseteq X$, we find some

$$\ell = (\ell_1, \ldots, \ell_s), \ r = (r_1, \ldots, r_s) \in M_1(Y) \times \cdots \times M_s(Y)$$

such that

$$H(M_1(Y), \ldots, M_s(Y))^L = h(\ell) \quad \text{and} \quad H(M_1(Y), \ldots, M_s(Y))^R = h(r)$$

since $h(x)$ is a single-use expression. Hence, using Lemma 3 we obtain

$$
\begin{aligned}
\delta(F(Y)) &= \delta(H(M_1(Y), \ldots, M_s(Y))) \\
&= H(M_1(Y), \ldots, M_s(Y))^R - H(M_1(Y), \ldots, M_s(Y))^L \\
&= h(r) - h(\ell) \leq L \cdot \|r - \ell\|_2 \\
&= L \cdot \sqrt{(r_1 - \ell_1)^2 + \cdots + (r_s - \ell_s)^2} \\
&\leq L \cdot \sqrt{\sum_{k=1}^{s} \left(M_k(Y)^R - M_k(Y)^L\right)^2} \\
&\leq L \cdot \sqrt{\sum_{k=1}^{s} (C_k \cdot \delta(Y))^2} \leq L \cdot \sqrt{s} \cdot C_{max} \cdot \delta(Y)
\end{aligned}
$$

with $C_{max} = \max\{C_1, \ldots, C_s\}$. □

Finally, we can prove the following result.

Theorem 5 *Consider a box $X \subset \mathbb{R}^n$, let $m_1, \ldots, m_s : X \to \mathbb{R}$ be Lipschitzian functions such that $m_1(x)$ to $m_s(x)$ are single-use expressions, and assume that the natural interval extensions $M_1(X)$ to $M_s(X)$ exist. Furthermore, consider*

$$f(x) = h(m_1(x), \ldots, m_s(x))$$

where $h : \mathbb{R}^s \to \mathbb{R}$ is a Lipschitzian function with constant L such that $h(x)$ is a single-use expression and assume that the natural interval extension $H(X)$ exists.

Then the natural interval bounding operation has a rate of convergence of $p = 1$.

Proof Using the previous result, we directly obtain

$$f(P(Y)) - LB(Y) \leq F(Y)^R - F(Y)^L \leq C \cdot \delta(Y)$$

which shows the rate of convergence of $p = 1$. □

To sum up, we found a general family of functions such that the natural interval bounding operation has a rate of convergence of $p = 1$. We remark that almost all commonly used objective functions in location theory as well as standard global optimization benchmark functions, see Sect. 5.2, can be expressed in the required form. However, this is not true in general as the following counterexample shows.

Example 6 Consider the one-dimensional and differentiable objective function

$$f(x) = \sqrt{\sin(7x) + \cos(3x) + 2}.$$

This function should be written as $f(x) = h(m_1(x), m_2(x))$ with

$$m_1(x) = \sin(7x) + 1, \quad m_2(x) = \cos(3x) + 1, \quad \text{and} \quad h(x_1, x_2) = \sqrt{x_1 + x_2}.$$

Thus, we know that $m_1(x) \geq 0$ and $m_2(x) \geq 0$ for all $x \in \mathbb{R}$ and m_1 and m_2 are Lipschitzian functions, but h is not Lipschitzian on $[0, \infty)^2$. However, in our numerical studies in Sect. 5.2 it can be seen that even for this function we obtain an empirical rate of convergence of one, see Table 4.

The next example shows that we cannot expect a rate of convergence higher than $p = 1$ for the natural interval bounding operation.

Example 7 Consider $f(x) = x^2$ and the box $X = [0, 2]$. Furthermore, define the sequence $Y_\mu = [1 - \mu, 1 + \mu]$ for $0 < \mu < 1$ with $c(Y_\mu) = 1$. Then the natural interval bounding operation yields

$$LB(Y_\mu) = (Y_\mu^2)^L = (1 - \mu)^2 = 1 - 2\mu + \mu^2.$$

Thus, we find

$$\frac{f(P(Y_\mu)) - LB(Y_\mu)}{\delta(Y_\mu)^2} = \frac{1 - (1 - 2\mu + \mu^2)}{4\mu^2} = \frac{2\mu - \mu^2}{4\mu^2} = \frac{1}{2\mu} - \frac{1}{4}$$

which is unbounded for $\mu \to 0$. Hence, we cannot find a constant $C > 0$ such that

$$f(P(Y_\mu)) - LB(Y_\mu) \leq C \cdot \delta(Y)^2$$

for all $\mu > 0$.

4.2 Centered interval bounding operation

Apart from the natural interval bounding operation we can use the natural interval extension of the Taylor expansion as follows, see also Hansen (1992).

Consider $X \subset \mathbb{R}^n$ and assume that $f : X \to \mathbb{R}$ is continuously differentiable. Then, for all subboxes $Y \subseteq X$ and $c = (c_1, \ldots, c_n) = c(Y)$ we know that

$$f(x) = f(c) + \nabla f(\vartheta(x))^T \cdot (x - c) \quad \text{for all} \quad x \in Y$$

and $\vartheta(x) \in Y$. Using again the fundamental theorem of interval analysis, this Taylor expansion leads to

$$f(Y) = f(Y_1, \ldots, Y_n) \subseteq f(c) + \sum_{k=1}^{n} G_k(Y) \cdot (Y_k - c_k),$$

where $G_k(X)$ is the natural interval extension of

$$g_k(x) := \frac{\partial f}{\partial x_k}(x) \quad \text{for} \quad k = 1, \ldots, n.$$

Finally, define $z = (z_1, \ldots, z_n) \in Y$ with $z_k \in \{Y_k^L, Y_k^R\}$ such that

$$(G_k(Y) \cdot (Y_k - c_k))^L = (G_k(Y) \cdot (z_k - c_k))^L$$

for $k = 1, \ldots, n$. Hence, we constructed the ***centered interval bounding operation***

$$LB(Y) = f(c) + \sum_{k=1}^{n} (G_k(Y) \cdot (Y_k - c_k))^L \quad \text{and} \quad P(Y) = z.$$

Theorem 6 *Let $X \subset \mathbb{R}^n$ and consider a continuously differentiable function $f : X \to \mathbb{R}$ such that the natural interval extensions of*

$$g_k(x) = \frac{\partial f}{\partial x_k}(x) \quad \text{for} \quad k = 1, \ldots, n$$

exist and satisfy the conditions given in Lemma 4.

Then the centered interval bounding operation has a rate of convergence of $p = 2$.

Proof For all $Y \subseteq X$ the first order Taylor expansion of f yields

$$f(x) = f(c) + \nabla f(\vartheta(x))^T \cdot (x - c)$$

with $\vartheta(x) \in Y$ for all $x \in Y$. Next, for all $k = 1, \ldots, n$ find $w_k \in G_k(Y)$ such that

$$(G_k(Y) \cdot (Y_k - c_k))^L = w_k \cdot (z_k - c_k),$$

where $z_k \in Y_k$ as defined before. Note that these values exist since $x \cdot y$ is a single-use expression. Finally, define

$$u = (u_1, \ldots, u_n) = \nabla f(\vartheta(z)) \in G_1(Y) \times \cdots \times G_n(Y).$$

Thus, Lemma 4 yields

$$f(P(Y)) - LB(Y) = \nabla f(\vartheta(z))^T \cdot (z - c) - \sum_{k=1}^{n} (G_k(Y) \cdot (Y_k - c_k))^L$$

$$= u^T \cdot (z - c) - \sum_{k=1}^{n} w_k \cdot (z_k - c_k)$$

$$= \sum_{k=1}^{n} (u_k - w_k) \cdot (z_k - c_k)$$

$$\leq \sum_{k=1}^{n} |u_k - w_k| \cdot |z_k - c_k|$$

$$\leq \sum_{k=1}^{n} \delta(G_k(Y)) \cdot \frac{1}{2} \delta(Y) \leq \sum_{k=1}^{n} C_k \cdot \delta(Y) \cdot \frac{1}{2} \delta(Y)$$

$$= \left(\frac{1}{2} \cdot \sum_{k=1}^{n} C_k \right) \cdot \delta(Y)^2$$

and the rate of convergence of $p = 2$ is shown.   □

Example 8 Consider $f(x) = x^3$, the box $X = [0, 2]$, and the sequence $Y_\mu = [1 - \mu, 1 + \mu]$ for $0 < \mu < 1$ with $c = c(Y_\mu) = 1$. Then we find

$$f(c) + F'(Y_\mu) \cdot (Y_\mu - c) = 1 + 3 \cdot [1 - \mu, 1 + \mu]^2 \cdot [-\mu, \mu].$$

Thus, the centered interval bounding operation yields

$$LB(Y_\mu) = 1 + 3 \cdot \left([(1-\mu)^2, (1+\mu)^2] \cdot [-\mu, \mu]\right)^L$$
$$= 1 + 3 \cdot (1+\mu)^2 \cdot (-\mu) = 1 - 3\mu - 6\mu^2 - 3\mu^3$$

and we obtain $P(Y_\mu) = z = (1-\mu)$. This leads to

$$\frac{f(P(Y_\mu)) - LB(Y_\mu)}{\delta(Y_\mu)^3} = \frac{(1-\mu)^3 - (1 - 3\mu - 6\mu^2 - 3\mu^3)}{8\mu^3}$$
$$= \frac{(1 - 3\mu + 3\mu^2 - \mu^3) - (1 - 3\mu - 6\mu^2 - 3\mu^3)}{8\mu^3}$$
$$= \frac{9\mu^2 + 2\mu^3}{8\mu^3} = \frac{9}{8\mu} + \frac{1}{4}$$

which is again unbounded for $\mu \to 0$. Therefore, the centered interval bounding operation does not have a rate of convergence of $p \geq 3$.

4.3 Baumann's interval bounding operation

In the previous subsection, we used the first order Taylor expansion of f at $c = c(Y)$. But note that for both for the calculation of the bounding operation and for the rate of convergence of $p = 2$ we can use the Taylor expansion of f at any point $y \in Y$. The idea of Baumann (1988) was to use a specific point $b = b(Y) \in Y$ instead of $c = c(Y) \in Y$ as follows.

Consider again a box $X \subset \mathbb{R}^n$ and assume that $f : X \to \mathbb{R}$ is continuously differentiable. Denote by $G_k(X)$ again the natural interval extensions of

$$g_k(x) := \frac{\partial f}{\partial x_k}(x) \quad \text{for} \quad k = 1, \ldots, n.$$

Then we define $b = (b_1, \ldots, b_n) \in Y$ by

$$b_k = \begin{cases} (\delta(G_k(Y))^{-1} \cdot \left(G_k(Y)^R \cdot Y_k^L - G_k(Y)^L \cdot Y_k^R\right) & \text{if} \quad 0 \in G_k(Y) \\ Y_k^L & \text{if} \quad G_k(Y)^L \geq 0 \\ Y_k^R & \text{if} \quad G_k(Y)^R \leq 0 \end{cases}$$

for $k = 1, \ldots, n$. As before, we obtain

$$f(Y) = f(Y_1, \ldots, Y_n) \subseteq f(b) + \sum_{k=1}^{n} G_k(Y) \cdot (Y_k - b_k).$$

Defining $z = (z_1, \ldots, z_n) \in Y$ with $z_k \in \{Y_k^L, Y_k^R\}$ such that

$$(G_k(Y) \cdot (Y_k - b_k))^L = (G_k(Y) \cdot (z_k - b_k))^L$$

for $k = 1, \ldots, n$ we have **Baumann's interval bounding operation**

$$LB(Y) = f(b) + \sum_{k=1}^{n}(G_k(Y) \cdot (Y_k - b_k))^L \quad \text{and} \quad P(Y) = z.$$

Analogous to the centered interval bounding operation, we find quadratic convergence.

Corollary 7 *Let $X \subset \mathbb{R}^n$ and consider a continuously differentiable function $f : X \to \mathbb{R}$ such that the natural interval extensions of*

$$g_k(x) = \frac{\partial f}{\partial x_k}(x) \quad for \quad k = 1, \ldots, n$$

exist and satisfy the conditions given in Lemma 4.

Then Baumann's interval bounding operation has a rate of convergence of $p = 2$.

Proof The proof is similar to Theorem 6 and therefore omitted here. □

Baumann (1988) has shown the specific choice of b which leads to the ***optimal centered form*** in the following sense. For all $y = (y_1, \ldots, y_n) \in Y$ we have

$$f(y) + \sum_{k=1}^{n} (G_k(Y) \cdot (Y_k - y_k))^L \leq f(b) + \sum_{k=1}^{n} (G_k(Y) \cdot (Y_k - b_k))^L.$$

Thus, Baumann's interval bounding operation yields the best lower bound among all possible centered forms. Hence, in our numerical results presented in the next section we expect a smaller constant C compared to the centered interval bounding operation.

We remark that in the literature some related results about quadratic convergence for Baumann's form can be found. For instance, under certain conditions Chuba and Miller (1972) and Krawczyk and Nickel (1982) proved that

$$\delta \left(f(b) + \sum_{k=1}^{n} G_k(Y) \cdot (Y_k - b_k) \right) - \delta(f(Y)) \leq C \cdot \delta(Y)^2.$$

But note that our result is stronger than this one since we in general do not have any information about $\delta(f(Y))$ in the left hand side of this inequality.

5 Numerical results

Our goal in this section is to measure the rate of convergence from the empirical point of view. To this end, changing the inequality into an equality in Equation (1) and applying the natural logarithm yields

$$\log(f(P(Y)) - LB(Y)) = \log(C) + p \cdot \log(\delta(Y)).$$

For a given test function f, we can calculate the left hand side of this expression and $\log(\delta(Y))$ for some selected boxes $Y \subseteq X$. We then obtain the empirical rate of convergence p and $\log(C)$ by linear regression. This strategy was also used in Tóth and Csendes (2005) and Tóth et al. (2007) to analyze related concepts and in Schöbel and Scholz (2010) to evaluate Lipschitzian and d.c. bounding operations.

5.1 The Weber problem on the plane

In order to present some first numerical experiences, we consider the Weber problem on the plane with positive and negative weights, see Tuy et al. (1995) or Drezner and Suzuki (2004). Assume a given set of s demand points $a_1, \ldots, a_s \in \mathbb{R}^2$ and weights $w_1, \ldots, w_s \in \mathbb{R}$. Then the Weber problem is to minimize the objective function

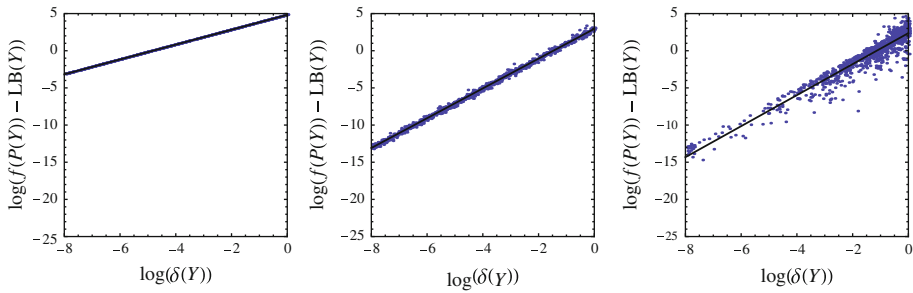$$f(x) = \sum_{k=1}^{s} w_k \cdot \|a_k - x\|_2.$$

Fig. 1 Results for the empirical rate of convergence using randomly selected boxes. Natural interval bounding operation (*left*), centered interval bounding operation (*middle*), and Baumann's interval bounding operation (*right*)

Table 1 Results for the empirical rate of convergence using randomly selected boxes

Bounding operation	C	p
Natural interval bounding operation	124.01	1.00
Centered interval bounding operation	19.06	2.00
Baumann's interval bounding operation	9.32	2.04

For this function, we can directly apply the natural interval bounding operation. But since the objective function is not differentiable at the demand points $a_1, \ldots, a_s$, we cannot directly apply the centered interval bounding operation and Baumann's interval bounding operation. Therefore, for any box $Y \subseteq X$ we consider

$$\tilde{f}(x) = \sum_{\substack{k=1 \\ a_k \notin Y}}^{s} w_k \cdot \|a_k - x\|_2 + \sum_{\substack{k=1 \\ a_k \in Y, \, w_k < 0}}^{s} w_k \cdot d_k^{max}(Y),$$

where $d_k^{max}(Y) = \max\{\|a_k - x\|_2 : x \in Y\}$. Note that $\tilde{f}$ is calculated for every subbox Y separately. Hence, $\tilde{f}(x) \le f(x)$ for all $x \in Y$ and $\tilde{f}$ is differentiable for all $x \in Y$. Moreover, $f = \tilde{f}$ if $a_k \notin Y$ for $k = 1, \ldots, s$ or if $\delta(Y) = 0$.

In our numerical experiences, we generated a fixed problem instance with $s = 100$ demand points uniformly distributed in $X = [-10, 10] \times [-10, 10]$ and weights uniformly distributed in $[-4, 6]$.

In a first study, we randomly selected 1,000 boxes $Y \subseteq X$ with different widths such that $f(P(Y)) - LB(Y) \ne 0$ for all three bounding operations since $\log(0)$ does not exist. Figure 1 shows the results where linear regressions yield the values of C and p summarized in Table 1.

As can be seen, the empirical rates of convergence agree fairly well with our theoretical results. Note that the constant C is smaller for Baumann's interval bounding operation compared to the centered interval bounding operation as expected from Sect. 4.3. Furthermore, the larger deviation for Baumann's interval bounding operation can be explained since for randomly selected boxes we are very often in a case where we obtain an exact lower bound. Thus, especially for smaller boxes it is very likely to find $f(P(Y)) - LB(Y) = 0$, see above.

Since the natural interval bounding operation has a rate of convergence of only $p = 1$, the branch-and-bound method as described in Sect. 3 is too slow if we are using the natural
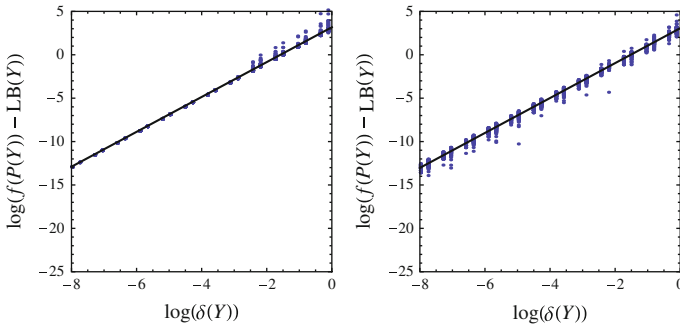
Fig. 2 Results for the empirical rate of convergence solving one particular problem instance. Centered interval bounding operation (*left*) and Baumann's interval bounding operation (*right*)

Table 2 Results for the empirical rate of convergence solving one particular problem instance

Bounding operation	Iterations	C	p
Centered interval bounding operation	2,116	28.96	2.01
Baumann's interval bounding operation	1,363	21.41	2.01

Table 3 Minimum, maximum, and average number of iterations for the branch-and-bound algorithm for 10 different Weber problem instances

Bounding operation	Min	Max	Ave.
Centered interval bounding operation	1,373	2,150	1,781.2
Baumann's interval bounding operation	545	1,660	1,130.3

interval bounding operation. Therefore, we only consider the other two bounding operations for the remainder of this subsection.

In a second study, we again used the same problem instance but we now considered all boxes which occur while running the branch-and-bound method with $\varepsilon = 10^{-10}$. Our results are illustrated in Fig. 2 and Table 2.

As before, we obtained a rate of convergence of approximately $p = 2$ for both bounding operations and the constant C is again smaller for Baumann's interval bounding operation compared to the centered interval bounding operation.

In a third study, we analyzed the number of iterations needed throughout the branch-and-bound algorithm for 10 different randomly generated problem instances as before with $s = 100$ demand points. Our results can be found in Table 3.

Although both bounding operations have a rate of convergence of $p = 2$, the number of iterations is much smaller for Baumann's interval bounding operation. This fact can be explained by the smaller constant C as discussed in our second study. We remark that similar results can be found for other values of s.

5.2 Global optimization benchmark problems

For some further numerical results, we solved some standard global optimization benchmark problems in different dimensions n, see Hansen (1992) and Schwefel (1981). A detailed list of the test functions used in the following can be found in the Appendix.

For all functions, we calculated the empirical rate of convergence as described before by linear regressions. Furthermore, we applied the branch-and-bound algorithm using the initial

Table 4 Numerical results for several standard global optimization test problems with $n \leq 2$

Bounding operation	Iterations	Run time (s)	C	p
Example 6 with $n = 1$ and $\varepsilon = 10^{-10}$				
Natural interval	270,502	653.25	1.28	1.01
Centered interval	218	0.00	4.56	2.03
Baumann's interval	119	0.00	4.16	2.15
Six Hump Camel with $n = 2$ and $\varepsilon = 10^{-10}$				
Natural interval	n.a.	n.a.	4,026.63	0.99
Centered interval	2,003	0.12	1,553.56	2.02
Baumann's interval	821	0.03	10.05	2.03
Levy 3 with $n = 2$ and $\varepsilon = 10^{-10}$				
Natural interval	n.a.	n.a.	87.63	1.02
Centered interval	8,039	0.40	314.80	2.03
Baumann's interval	4,273	0.16	225.38	2.17
Levy 5 with $n = 2$ and $\varepsilon = 10^{-10}$				
Natural interval	n.a.	n.a.	89.78	1.01
Centered interval	2,893	0.28	314.97	2.02
Baumann's interval	2,278	0.17	191.57	2.18
Levy 13 with $n = 2$ and $\varepsilon = 10^{-10}$				
Natural interval	n.a.	n.a.	42.71	0.97
Centered interval	18,758	0.95	510.79	2.00
Baumann's interval	13,717	0.58	432.21	2.22
Schwefel 25 with $n = 2$ and $\varepsilon = 10^{-10}$				
Natural interval	169	0.00	39.85	0.99
Centered interval	1,077	0.04	6.58	2.00
Baumann's interval	320	0.00	2.01	2.01
Fermat-Weber problem with $n = 2$ and $\varepsilon = 10^{-10}$				
Natural interval	n.a.	n.a.	124.12	1.00
Centered interval	2,116	0.26	19.21	2.01
Baumann's interval	1,363	0.17	8.88	2.03

box $X = [-10, 10]^n$ and $\varepsilon = 10^{-10}$ for all problems. Note that each selected box throughout the algorithm was bisected perpendicular to the direction of the maximum width component in two subboxes.

Our numerical results can be found in Tables 4 and 5. Therein, we reported the empirical rates of convergence calculated from 1,000 randomly selected boxes $Y \subset X$ as before as well as the run time and the number of iterations needed to solve the problems. Note that n.a. stands for not available since the algorithm was stopped if the number of iterations exceeded one million and the run time was >1 hour.

For all benchmark problems, the empirical rate of convergence is again similar to our theoretical studies. Moreover, for almost all problems the lower bounds obtained from the natural interval bounding operation were not sharp enough such that we could solve the problems efficiently. However, note that for two benchmark problems, namely for *Schwefel 25* and for

Table 5 Numerical results for several standard global optimization test problems with $n \geq 3$

Bounding operation	Iterations	Run time (s)	C	p
Rosenbrock with $n = 3$ and $\varepsilon = 10^{-10}$				
Natural interval	179	0.00	$3.54 \cdot 10^5$	0.99
Centered interval	133,015	10.25	$0.80 \cdot 10^5$	1.99
Baumann's interval	29,185	0.74	170.09	2.02
Fermat-Weber problem with $n = 3$ and $\varepsilon = 10^{-10}$				
Natural interval	n.a.	n.a.	95.31	1.00
Centered interval	39,426	8.88	9.11	2.00
Baumann's interval	23,844	4.67	2.34	1.99
Fermat-Weber problem with $n = 4$ and $\varepsilon = 10^{-10}$				
Natural interval	n.a.	n.a.	115.06	1.00
Centered interval	238,719	134.96	8.95	2.00
Baumann's interval	96,452	33.85	1.98	2.03
Shekel 5 with $n = 4$ and $\varepsilon = 10^{-10}$				
Natural interval	n.a.	n.a.	$20.22 \cdot 10^{-4}$	1.01
Centered interval	18,885	1.00	$2.93 \cdot 10^{-4}$	2.00
Baumann's interval	3,057	0.12	$0.76 \cdot 10^{-4}$	1.81
Shekel 7 with $n = 4$ and $\varepsilon = 10^{-10}$				
Natural interval	n.a.	n.a.	$27.16 \cdot 10^{-4}$	1.01
Centered interval	19,325	1.02	$4.02 \cdot 10^{-4}$	2.00
Baumann's interval	3,006	0.23	$1.32 \cdot 10^{-4}$	1.91
Shekel 10 with $n = 4$ and $\varepsilon = 10^{-10}$				
Natural interval	n.a.	n.a.	$37.72 \cdot 10^{-4}$	1.00
Centered interval	19,566	1.37	$5.56 \cdot 10^{-4}$	1.99
Baumann's interval	3,062	0.22	$2.25 \cdot 10^{-4}$	1.89

Rosenbrock, the fastest way to solve the problem was to use the natural interval bounding operation.

Finally, we remark that similar results can be obtained for different values of ε. However, for same larger values of ε, say $\varepsilon = 10^{-2}$ or $\varepsilon = 10^{-1}$, the natural interval bounding operation fares better than the other two bounding operations since for larger boxes we often obtain the sharpest bounds using the natural interval bounding operation.

6 Discussion

In this paper, we analyzed bounding operations derived from interval analysis from the theoretical point of view. We found out that the natural interval extension yields a bounding operation with a rate of convergence of $p = 1$ while the centered interval bounding operation as well as Baumann's interval bounding operation have a rate of convergence of $p = 2$. Our theoretical results are also verified by various numerical experiences using the Weber problem on the plane and some standard global optimization benchmark problems.

We remark that we cannot generalize the centered interval bounding operation to a bounding operation with an arbitrary rate of convergence using higher order Taylor extensions since we are not in a position to determine the point $P(Y) = z$ as required. For a general bounding operation, we refer to Schöbel and Scholz (2010).

Moreover, note that the empirical rates of convergence as calculated in Tóth and Csendes (2005) and Tóth et al. (2007) seem to be very close to our theoretical results if we ignore boxes with $f(P(Y)) - LB(Y) = 0$. Therefore, our work now explains the empirical studies reported e.g. in Tóth and Csendes (2005) from the theoretical point of view.

Appendix

In the following, we give a detailed list of test functions used for our numerical results in Sect. 5.2. All these functions are standard global optimization benchmark problems and can be found e.g. in Hansen (1992) and Schwefel (1981).

(1) *Six Hump Camel* with $n = 2$:

$$f(x_1, x_2) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1 x_2 - 4x_2^2 + 4x_2^4.$$

(2) *Levy 3* with $n = 2$:

$$f(x_1, x_2) = \sum_{k=1}^{5} (k \cdot \cos((k-1)x_1 + k)) \cdot \sum_{k=1}^{5} (k \cdot \cos((k+1)x_2 + k)).$$

(3) *Levy 5* with $n = 2$:

$$f(x_1, x_2) = \sum_{k=1}^{5} (k \cdot \cos((k-1)x_1 + k)) \cdot \sum_{k=1}^{5} (k \cdot \cos((k+1)x_2 + k))$$
$$+ (x_1 + 1.42513)^2 + (x_2 + 0.80032)^2.$$

(4) *Levy 13* with $n = 2$:

$$f(x_1, x_2) = \sin^2(3\pi x_1) + (x_1 - 1)^2(1 + \sin^2(3\pi x_2)) + (x_2 - 1)(1 + \sin^2(2\pi x_2)).$$

(5) *Schwefel 25* with $n = 2$:

$$f(x_1, x_2) = (x_1 + 2x_2 - 7)^2 + (2x_1 + x_2 - 5)^2.$$

(6) *Fermat-Weber problem* with $n \in \{2, 3, 4\}$:

$$f(x) = \sum_{k=1}^{100} w_k \cdot \|x - a_k\|_2$$

with $a_k \in [-10, 10]^n$ and $w_k \in [-4, 6]$, see also Sect. 5.1.

(7) *Rosenbrock* with $n = 3$:

$$f(x_1, x_2, x_3) = 100(x_2 - x_1^2)^2 + (x_1 - 1)^2 + 100(x_3 - x_2^2)^2 + (x_2 - 1)^2.$$

(8) *Shekel 5*, *Shekel 7*, and *Shekel 10* with $n = 4$:

$$(x) = -\sum_{k=1}^{m} \frac{1}{\|A_k - x\|_2^2 + \frac{1}{10}c_k}$$

with $m = 5$, $m = 7$, and $m = 10$, respectively, and

$$A = \begin{pmatrix} 4 & 4 & 4 & 4 \\ 1 & 1 & 1 & 1 \\ 8 & 8 & 8 & 8 \\ 6 & 6 & 6 & 6 \\ 3 & 7 & 3 & 7 \\ 2 & 9 & 2 & 9 \\ 5 & 5 & 3 & 3 \\ 8 & 1 & 8 & 1 \\ 6 & 2 & 6 & 2 \\ 7 & 3.6 & 7 & 3.6 \end{pmatrix} \quad \text{and} \quad c = (1, 2, 2, 4, 4, 6, 3, 7, 5, 5).$$

References

Baumann, E.: Optimal centered forms. BIT Numer. Math. **28**, 80–87 (1988)

Blanquero, R., Carrizosa, E.: Continuous location problems and big triangle small triangle: constructing better bounds. J. Global Optim. **45**, 389–402 (2009)

Chuba, W., Miller, W.: Quadratic convergence in interval arithmetic. Part I. BIT Numer. Math. **12**, 284–290 (1972)

Csallner, A.E., Csendes, T.: The convergence speed of interval methods for global optimization. Comput. Math. Appl. **31**, 173–178 (1996)

Drezner, Z., Suzuki, A.: The big triangle small triangle method for the solution of nonconvex facility location problems. Oper. Res. **52**, 128–135 (2004)

Fernández, J., Pelegrín, B., Plastria, F., Tóth, B.: Solving a Huff-like competitive location and design model for profit maximization in the plane. Eur. J. Oper. Res. **179**, 1274–1287 (2007)

Floudas, C.A., Pardalos, P.M.: Encyclopedia of Optimization. 2nd edn. Springer, New York (2009)

Hansen, E.: Global Optimization Using Interval Analysis. 1st edn. Marcel Dekker, New York (1992)

Hansen, P., Peeters, D., Richard, D., Thisse, J.F.: The minisum and minimax location problems revisited. Oper. Res. **33**, 1251–1265 (1985)

Horst, R., Pardalos, P.M., Thoai, N.V.: Introduction to Global Optimization. 2nd edn. Springer, Berlin (2000)

Krawczyk, R., Nickel, K.: Die zentrische Form in der Intervallarithmetik, ihre quadratische Konvergenz und ihre Inklusionsisotonie. Computing **28**, 117–137 (1982)

Neumaier, A.: Interval Methods for Systems of Equations. 1st edn. Cambridge University Press, Cambridge (1990)

Plastria, F.: GBSSS: the generalized big square small square method for planar single-facility location. Eur. J. Oper. Res. **62**, 163–174 (1992)

Ratschek, H., Rokne, J.: New Computer Methods for Global Optimization. 1st edn. Ellis Horwood, Chichester, England (1988)

Ratschek, H., Voller, R.L.: What can interval analysis do for global optimization?. J. Global Optim. **1**, 111–130 (1991)

Schöbel, A., Scholz, D.: The theoretical and empirical rate of convergence for geometric branch-and-bound methods. J. Global Optim. **48**, 473–495 (2010)

Schwefel, H.P.: Numerical Optimization of Computer Models. 1st edn. Wiley, New York (1981)

Tóth, B., Csendes, T.: Empirical investigation of the convergence speed of inclusion functions in a global optimization context. Reliab. Comput. **11**, 253–273 (2005)

Tóth, B., Fernández, J., Csendes, T.: Empirical convergence speed of inclusion functions for facility location problems. J. Comput. Appl. Math. **199**, 384–389 (2007)

Tóth, B., Fernández, J., Pelegrín, B., Plastria, F.: Sequential versus simultaneous approach in the location and design of two new facilities using planar Huff-like models. Comput. Oper. Res. **36**, 1393–1405 (2009)

Tuy, H., Al-Khayyal, F., Zhou, F.: A D.C. optimization method for single facility location problems. J. Global Optim. **7**, 209–227 (1995)