Global Optimization Simplex Bisection Revisited Based on Considerations by Reiner Horst*

Eligius M.T. Hendrix[1], Leocadio G. Casado[2], and Paula Amaral[3]

[1] Arquitectura de Computadores, Universidad de Málaga and Logistics and Operations Research, Wageningen University
eligius@uma.es
[2] Dpt. de Arquitectura de Computadores y Electronica, Universidad de Almería
leo@ual.es
[3] Department of Mathematics and CMA, Universidade Nova de Lisboa
paca@campus.fct.unl.pt

Abstract. In this paper, the use of non-optimality spheres in a simplicial branch and bound (B&B) algorithm is investigated. In this context, some considerations regarding the use of bisection on the longest edge in relation with ideas of Reiner Horst are reminded. Three arguments highlight the merits of bisection of simplicial subsets in B&B schemes.

Keywords: Global Optimization, simplicial partition, branch and bound, bisection.

1 Introduction

This work is dedicated to Reiner Horst, who encouraged and inspired the study of Global Optimization branch and bound (B&B) methods. In his last (2010) contribution titled "Bisection by global optimization revisited" [8], some considerations were elaborated regarding the use of simplices in branch and bound. Reiner ideas on simplicial partitioning, developed in discussion with his co-workers Micheal Nast and Nguyen Van Thoai, are summarized in the book [11] and elaborated and experimented in the thesis of Ulrich Raber [17]. The main issue in [8] is that "bisection is not optimal". It is clear that optimality depends on the objective under consideration, and we would like to stress that Reiner had a wider view on the use of simplices than B&B only, namely the typical lower dimensional tessalation in physics and the use of triangulations for finding roots of mappings. This paper focuses on several aspects for which bisecting the longest edge in simplicial branch and bound in Global Optimization may be convenient.

* This paper has been supported by The Spanish Ministry of Science and Innovation (project TIN2008-01117) and Junta de Andalucía (P11-TIC-7176), in part financed by the European Regional Development Fund (ERDF) and by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through PEst-OE/MAT/UI0297/2011 (CMA). Eligius M.T. Hendrix is a fellow of the Spanish "Ramon y Cajal" contract program, co-financed by the European Social Fund.

B. Murgante et al. (Eds.): ICCSA 2012, Part III, LNCS 7335, pp. 159–173, 2012.

Previous experience regarding the use of B&B on the unit simplex in applications in mixture design for multinational Unilever, [4], showed that the study of Reiner [7], on splitting the unit simplex by bisection, leads typically to edge lengths of $1, \sqrt{3}/2, \sqrt{2}/2$ and $1/2$. Moreover, it was shown that, implicitly, this technique leads to samples points over an equidistant grid when using an ε accuracy in the decision space [4]. In the above mentioned application, the practical importance of this design property has connections with robustness considerations, in the sense that finding an acceptable design means that all points in its environment are feasible. Bisecting the longest edge gives relatively 'round' partition sets. For running a B&B tree to the bottom, where simplices have at most a size of ε, this feature is convenient. Using radial splitting over the centroid, as suggested in [7], leads to needle shaped subsimplices. Deviating from the midpoint requires keeping track of ε robustness. Concluding *Bisecting over the middle of the longest edge can be convenient for ε robustness considerations.*

A second aspect has to do with implementation issues in B&B. Instead of questioning how many small subsets the B&B search may lead to at the bottom of the tree, Reiner showed his always optimistic perspective of following a subset to be split iteratively from top to bottom of the tree to see how fast it converges to a singleton. In [8] he repeats the proposition that can also be found in [7] and with an extensive proof in [11] that after splitting an n-simplex n times, the longest edge will be shorter than $\sqrt{3}/2$. He also reminded that Beaker Kearfott already published a similar result in 1978. In [4] it is shown that after splitting all edges i.e. going $n(n+1)/2$ deeper in the tree, then the size is at most $1/2$. Up to about $n = 9$ this is even a sharper bound. However, seen from the worst case (pessimistic) B&B perspective this is not encouraging; the number of simplices to be evaluated is astronomically high.

After obtaining in practice millions of subsets to be stored in RAM, there are two optional directions. From a theoretical point of view, this means looking for sharper (and more elegant) bounds. From a practical perspective, it implies designing a convenient way to store and manage the search tree, allowing sorting, easy look up, and workload distribution over several processors. This is a second reason why using the midpoint of the longest edge can be convenient; the same (evaluated) point appears several times as vertex of sub-simplices, allowing repeated re-use of its information without the need to evaluate it. Concluding, *bisecting over the middle of the longest edge can be convenient for storing a B&B tree structure with subsets linked to evaluated vertices.*

The use of simplices in B&B to solve GO problems is common [12,16]. The idea can be found in the work of Ulrich Raber [16], also in the work of Julius and Antanas Žilinskas [20] and Remigijus Paulavičius, J. Žilinskas and Andreas Grothey [14]. Specifically, in [19] the idea of using more elegant partitioning than bisection is discussed. However, these proposals are applicable only to low dimensional cases. Theoretically, it is known that convergence is guaranteed. As discussed, bisection may be a good basis from the computer science perspective, despite it is not efficient from a bounding perspective. In the sequel, a third aspect related to the practical use in B&B is outlined. To address this issue, a question is posed: how can simplicial partition sets and bisection be used to have early pruning of nodes? This means to develop methods that detect subspaces which cannot contain a global optimizer in an early phase.

Regarding node pruning, we focus on covering methods, based on bounds on first (Lipschitz constant) and second derivative. For higher dimensions, greater than 1, Reiner Horst mentioned in [11] that Lipschitz optimization "does not look very practical". His focus was rather on B&B; in [10] a B&B view on covering methods is presented. Our study deals with linking the two concepts using so-called non-optimality spheres.

We describe the idea of covering algorithms that can basically also be found in the books of Reiner Horst [9,11] with the aid of simple examples and figures in Sect. 2. In Sect. 3, the concept of non-optimality spheres is presented and a B&B algorithm is given in Sect. 4 based on simplicial partition sets. We discuss how to infer simplicial partition set covering by non-optimality spheres in Sect. 5. We numerically illustrate the concept of using bisection in such a procedure in Sect. 6. This is followed by conclusions in Sect. 7.

2 Covering Algorithms

The generic box-constrained GO problem consists in finding the global minimum f^* of a real valued n-dimensional function $f : S \rightarrow \mathbb{R},\ S \subset \mathbb{R}^n$, and the corresponding set S^* of global minimum points, where S is a box, i.e.:

$$f^* = f(x^*) = \min_{x \in S} f(x),\ x^* \in S^* \ . \tag{1}$$

Covering methods approach this problem by defining iteratively a covering function $\varphi_k(x) \leq f(x)$, where a minimum point of $\varphi_k(x) \leq f(x)$ over S is then used as the next iterate x_{k+1}. A basic method with this property is due to Piyavski and Shubert [5,15,18], who published in parallel about an algorithm where the so-called saw-tooth cover is based on information about the Lipschitz constant. The knowledge of a scalar L is assumed such that

$$\mid f(x_1) - f(x_2) \mid \leq L\|x_1 - x_2\| \ \forall x_1, x_2 \in S \ . \tag{2}$$

For evaluated points $x_1, \ldots, x_i, \ldots, x_k$, with function values $f_1, \ldots, f_i, \ldots, f_k$ the covering function is defined by

$$\varphi_k(x) = \max_{1 \leq i \leq k} (f_i - L\|x - x_i\|) \ . \tag{3}$$

By keeping track of the best function value as upper bound of the global minimum $U = \min_{1 \leq i \leq k} f_i$, one can show that the algorithm evaluating iteratively the minimum of φ_k leads to a guaranteed approximation of the optimum with accuracy δ, when using $U - \min_x \varphi(x) < \delta$ as stopping criterion.

The real challenge is the application of this concept for dimensions higher than 1, given the scepticism of some authors, Reiner Horst included in [11] that a direct application looks impractical. A continuing report on achievements on covering methods in Russian is due to the work of Yuri Evtushenko (e.g. [6]) during 40 years. A description of the corresponding algorithm and minimization of φ_k is also described in [13].

A second base for covering algorithms is due to the work of Breiman and Cutler, [2] when using a bound K on the second derivative, such that $K \geq -f''(x), x \in S$ or more general (in higher dimensions) on an overestimate of the negative of the minimum eigenvalue of the Hessian, such that

$$f(x) \geq f(x_1) + \nabla f^T(x)(x - x_1) - \frac{1}{2}K\|x - x_1\|^2 \; \forall x, x_1 \in S \ . \tag{4}$$

Analogously to (3), the corresponding covering function is given by

$$\varphi_k(x) = \max_{1 \leq i \leq k} (f_i + \nabla f_i^T(x - x_i) - \frac{1}{2}K\|x - x_i\|^2) \ . \tag{5}$$

The algorithm of Breiman-Cutler takes iteratively a minimum point of φ_k as next iterate. The original article [2] describes also the approach for the multivariate case where it is necessary to find the minimum of intersecting parabolics leading to polytope shaped regions that are similar to Voronoi diagrams. The method is very elegant, but also very elaborative, as it requires storing information on all evaluated points, intersecting planes and resulting vertices of the polytopes.

Baritompa in [1] showed how (2) and (4) can be relaxed by focussing on the behavior around global optimum x^*, f^*. Let M and K be values such that $f(x) \leq f^* + M\|x - x^*\|, \forall x \in S$ and $f(x) \leq f^* + \frac{1}{2}K\|x - x^*\|^2, \forall x \in S$. So, it is

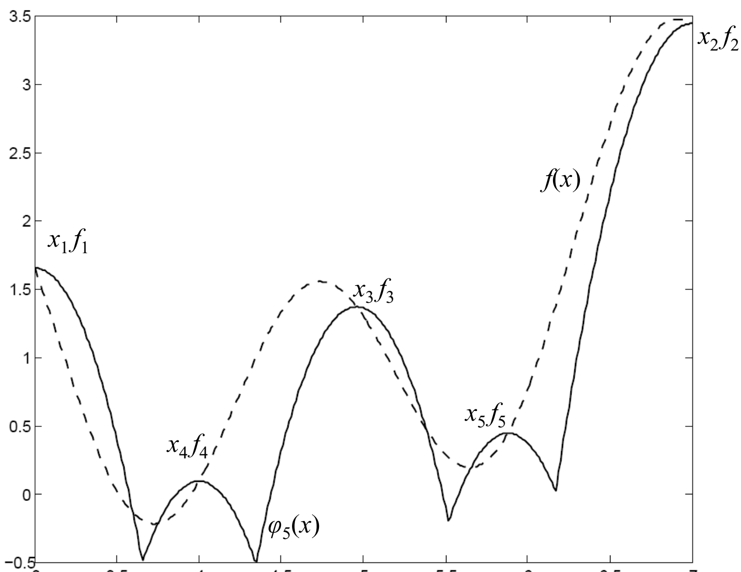

Fig. 1. Iterate is a minimum of (7) for $f(x) = sin(x) + sin(3x) + ln(x)$

not necessary to have a global overestimate of neither Lipschitz constant, nor of the second derivative (or the negative of the minimum eigenvalue of the Hessian in higher dimensions), K. Then one can take as cover

$$\varphi_k(x) = \max_{1\leq i\leq k}\{f_i - M\|x - x_i\|\} \tag{6}$$

or alternatively

$$\varphi_k(x) = \max_{1\leq i\leq k}\{f_i - \frac{1}{2}K\|x - x_i\|^2\}\ . \tag{7}$$

An interesting aspect is that φ is not necessarily an underestimating function of f, but it neither cuts away a global minimum point.

Example 1. Consider function $f(x) = sin(x) + sin(3x) + ln(x)$ on the interval $X = [3, 7]$. We take $K = 10$, as the maximum value of the second derivative is reached close to x^*. Fig. 1 depicts iterates corresponding to a minimum point of (7). Function φ_k is not a lower bounding function, but neither cuts away the global minimum. The minimum point of φ_k is a lower bound for the minimum of f.

The example illustrates the deterministic view of Reiner Horst, where the global minimum can be obtained with a guarantee given certain information. Usually one refers to a bound on derivatives, but the assumptions of Baritompa around the global minimum point do not require the function to be differentiable neither continuous over the whole domain. On the other hand, practically information on M or K is required, which may be as hard to obtain as solving the original problem.

Our interest is the multivariate variants of using (6) and (7) in simplicial branch and bound. Finding iteratively the minimum point of $\varphi_k(x)$ may be a tedious job. However, from a B&B perspective, it is not necessary to know exactly the minimum of φ. Our focus is on the potential of so-called non-optimality spheres, close to the covering concepts of [6] and the concept of infeasibility spheres in [4]. The purpose is to come to simplicial B&B based algorithms applying (6) and (7) in order to illustrate the usefulness of bisection.

3 Non-optimality Spheres

As the name suggests, non-optimality spheres are spheres that are guaranteed not to contain an optimal solution. We start describing how non-optimality spheres can be derived from sample points and global value information. Next, a specific B&B algorithm which uses simplicial partition sets is presented.

Consider sample points $x_1, \ldots, x_i, \ldots, x_k$, with $f_1, \ldots, f_i, \ldots, f_k$ as function values, and U the best function value found, $U = \min_{1\leq i\leq k} f_i$. For a value of M such that

$$f(x) \leq f^* + M\|x - x^*\|, \forall x \in S, \forall x^* \in S^*, \tag{8}$$

a non-optimality sphere BM_i centered at x_i with radius r_i is given by

$$BM_i = \{x \in S|\ \|x - x_i\| < (r_i = \frac{f_i - U}{M})\}\ . \tag{9}$$

For a value K with

$$f(x) \leq f^* + \frac{1}{2}K\|x - x^*\|^2, \forall x \in S, \forall x^* \in S^* \tag{10}$$

a non-optimality sphere is given by

$$BK_i = \{x \in S| \; \|x - x_i\|^2 < (r_i^2 = 2\frac{f_i - U}{K})\} \; . \tag{11}$$

First notice that in (8) and (10) necessarily $M > 0$ and $K > 0$. The definition of the non-optimality sphere radius is obtained by a simple manipulation of (8) and (10), bounding f^* by U and replacing x and $f(x)$ respectively by x_i and f_i. Comparing BK_i as in (11) with the sphere that could be obtained using a similar procedure and the Breiman-Cutler assumption (4), the difference is that the center of the sphere is shifted towards $x_i + \frac{1}{K}\nabla f_i$ and in the radius definition in (11), one should take instead of f_i the value of the top of the parabola $f_i + \frac{1}{2K}\nabla f_i^T \nabla f_i$.

Notice that for a current best point where $f_i = U$, the non-optimality sphere is empty, i.e. it could be an optimum point. Moreover, if the upper bound U goes down during the iterations, the spheres are getting bigger. The fact that the area of the non-optimality spheres can be left out of further consideration is given in the following theorems.

Theorem 1. *Non-optimality sphere BM_i does not contain a global minimum point $x^* \in S^*$.*

Proof. Proof by contradiction. By definition of M, $f(x_i) \leq f^* + M\|x_i - x^*\|$, such that $f^* \geq f_i - M\|x_i - x^*\|$. Let $x^* \in B_i$. Substitution of definition (9) gives

$$f^* \geq f_i - L\|x_i - x^*\| > f_i - f_i + U = U, \tag{12}$$

which contradicts U being an upper bound of f^*. □

For the parabolic non-optimality sphere this is given as follows.

Theorem 2. *Non-optimality sphere BK_i does not contain a global minimum point $x^* \in S^*$.*

Proof. Proof by contradiction. By definition of K, $f(x_i) \leq f^* + \frac{1}{2}K\|x_i - x^*\|^2$, such that $f^* \geq f_i - \frac{1}{2}K\|x_i - x^*\|^2$. Let $x^* \in BK_i$. Substitution of definition (11) gives

$$f^* \geq f_i - \frac{1}{2}K\|x_i - x^*\|^2 > f_i - f_i + U = U, \tag{13}$$

which contradicts with U being an upper bound of f^*. □

In general, a non-optimality sphere may be completely covered by another one depending on the values of M and K. This is interesting from algorithmic perspective, as the covered sample point and its sphere apparently do not add any information to the search. However, if M is a strict overestimate of the Lipschitz constant

$$M > \frac{|f_2 - f_1|}{\|x_2 - x_1\|} \forall x_1, x_2 \in S, \tag{14}$$

then one sphere cannot be covered by another one.

Theorem 3. *Let M be an overestimate (14), BM_i be defined by (9) and $x_1, x_2 \in S$ with function values f_1, f_2. Then neither $BM_1 \subset BM_2$ nor $BM_2 \subset BM_1$.*

Proof. A sphere BM_2 of radius r_2 and center x_2 contains a sphere BM_1 with radius r_1 and center x_1 if

$$r_2 \geq r_1 + \|x_2 - x_1\|.$$

W.l.o.g. let $f_2 > f_1$, such that $r_2 > r_1$. Then $BM_2 \subset BM_1$ is not possible, as $r_1 \geq r_2 + \|x_2 - x_1\|$ is not possible. Furthermore, from (9) we have

$$r_2 - r_1 = \frac{f_2 - U}{M} - \frac{f_1 - U}{M} = \frac{f_2 - f_1}{M}.$$

Now using (14) we obtain

$$r_2 - r_1 < \frac{f_2 - f_1}{|f_2 - f_1|}\|x_2 - x_1\| = \|x_2 - x_1\|,$$

so

$$r_2 < r_1 + \|x_2 - x_1\|.$$

So neither $BM_2 \subset BM_1$ nor $BM_1 \subset BM_2$. □

Example 2. Consider the six-hump camel-back function:

$$f(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4 \tag{15}$$

taking as feasible area $S = [-2, 2] \times [-2, 2]$. It has 6 local optimum points two of which describe the set of global optimum solutions. All vertices of S and 16 more generated sample points x_i are evaluated. The maximum eigenvalue of the Hessian goes up to 184. In [2], experiments are done with $K = 9$, as the lower bounding is based on the most negative eigenvalue. In the illustration, a value of $K = 60$ is used.

The resulting Emmentaler set $S \setminus \cap BK_i$, where the optimum still can be located, is drawn in Fig. 2. Similar figures can be made using $M = 38$ as valid upper bound in the determination of BM_i. The spheres close to the vertices of S are relatively big, because the highest function values are attained there. Only 19 spheres are drawn because the one that corresponds to $\min_i f_i$ is empty; a cross marks its center.

As such, the described set is difficult to work with. However, an alternative is to link covering algorithms to B&B as Reiner Horst did in [10]. Specifically, we apply the spheres in a B&B framework where n-simplices are used as suggested by Reiner in [8].

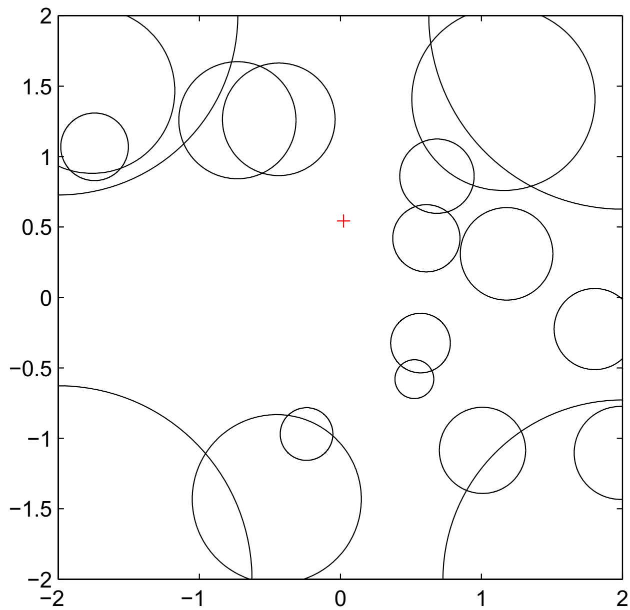

Fig. 2. Emmentaler set for six-hump camel-back evaluated at vertices and 16 additional sample points

4 B&B Simplicial Covering Algorithm

The use of value information on M and K aims at guaranteeing that in the end the best sampled point has a function value that differs less than a predefined accuracy δ from f^*; $f^* \leq U \leq f^* + \delta$. This target can be reached by having a dense sampling, e.g. a grid. If a value for M is given and there is for every sampling point, another sampling point such that the distance in between them is at most $\varepsilon = 2\frac{\delta}{L}$, then any $x \in S$ is closer than $\frac{1}{2}\varepsilon$ from a sampled point, such that, $f(x) > U - \frac{1}{2}M\varepsilon = U - \delta$, $\forall x \in S$. In this case U is a δ-accuracy optimal solution.

Similarly, if a value for K is given, one can sample up to an accuracy of points being $\varepsilon = \sqrt{\frac{8\delta}{K}}$ apart to guarantee $f(x) > U - \frac{1}{2}K(\frac{1}{2}\varepsilon)^2 = U - \delta$. The essential of branch and bound is not to sample everywhere dense, but to remove areas where it has been proven that the optimum cannot be located. In the B&B method, the set is subsequently partitioned into more and more refined parts (branching) over which bounds of an objective function value, and in this case, non-optimality spheres can be determined. Parts completely covered by the spheres are deleted (pruning), since these parts of the domain cannot contain optimum solutions.

Algorithm 1. B&B algorithm.

Inputs: - S: box constrained feasible area
- f: objective function
- δ: accuracy
- K: parabolic parameter

Output: - best proven solution x^U

Funct B & B Algorithm

1. $\varepsilon := \sqrt{\frac{8\delta}{K}}$, $\Lambda := \{C_1, \ldots, C_p\}$ as first partition of S
2. **for** sample points $x_i \in C_j, C_j \in \Lambda$ EvaluateVertex(x_i)
3. **for** simplices $C_j \in \Lambda$ EvaluateSimplex(C_j)
4. **while** $\Lambda \neq \emptyset$
5. Take one subset C from list Λ according to a selection rule. Subdivide C into two new subsets $Cnew_1$ and $Cnew_2$ by splitting over the longest edge, generating new point x_k.
6. EvaluateVertex(x_k),
7. EvaluateSimplex($Cnew_1$), EvaluateSimplex($Cnew_2$)
8. return x^U

A possible algorithm based on bisection is outlined (see Algorithm 1). The method starts with a partitioning of set S into simplices $C_1, \ldots, C_p$ to be stored as first elements of a list Λ of subsets (partition sets) and stops when the list Λ is empty. We also store the generated sample points x_i and their function value f_i on which the radius of the non-optimality spheres is based. Finally we keep track of points, that are proven to have a function value which differs less than δ from the global minimum f^*.

A generated subset C_k is not stored in Λ, if it can be proven that it is covered. In Sect. 5, results on proving coverage are discussed. Moreover, partition sets smaller in size than ε are discarded. The branching concerns the further refinement of the partition. This means that one of the subsets is selected to be split into new subsets. A selection rule determines the subset to be split next.

As discussed before based on the considerations in [7], an advantage of bisection splitting along the longest edge is due to the shape of the partition sets. The length of the longest edge is at most twice the size of the shortest edge. Therefore the sets can never get a needle shape.

Algorithm 2. Evaluate subset, decide to put on list based on cover

Funct EvaluateSimplex (C); global Λ, U, ε

1. **if** $size(C) > \varepsilon$
2. Cover check of C by $\cup BK_i$
3. **if** C not proven to be covered
4. store C in Λ

Algorithm 3. Evaluate a point and update global information.

Funct EvaluateVertex (x); global Λ, U, x^U

1. Determine $f(x)$ either from stored points or evaluate
2. **if** $f(x) < U$
3. $U := f(x)$ and $x^U := x$ *Update global information*
4. Update all BK_i and remove all $C_k \in \Lambda$ that are covered *Pruning*

5 Check on Covering a Simplex by Spheres

The question if simplex C is covered by spheres $B_i = \{x| \, \|x - v_i\| \leq r_i\}$ centered at its vertices v_i has been dealt with extensively in [3]. Notice that the vertices v_i are a subgroup of the evaluated points; $\{v_i \in C\} \subset \{x_1, \ldots, x_k\}$. Even the question of covering the simplex by spheres at the vertices is not for each instance easy to verify. The following three rules are useful:

1. check first if one of the spheres alone covers C, i.e. $\max_j \|v_j - v_i\| < r_i$.
2. if an interior point $x \in C$ is covered by the intersection of spheres $x \in \cap_{v_i \in C} B_i$, all the simplex is covered, i.e. $C \subset \cup_{v_i \in C} B_i$. One can try a weighted average of the vertices.
3. the best point to check is the so called θ-point where $\|\theta - v_i\|^2 - r_i^2 = \|\theta - v_j\|^2 - r_j^2, \forall v_i, v_j \in C$. Even if θ is not interior, but covered, the whole simplex C is covered.

In the algorithmic context, the first rule is the easiest to check and should be tried first. The determination of the θ-point requires solving a set of n linear equalities. Consider the vertices $v_1, \ldots, v_{n+1}$ of C. Equating

$$(\theta - v_1)^T(\theta - v_1) - r_1^2 = (\theta - v_i)^T(\theta - v_i) - r_i^2, \quad i = 2, \ldots, n+1 \tag{16}$$

and bringing the terms with θ to the left hand side gives

$$2(v_i - v_1)^T\theta = r_1^2 - r_i^2 + v_i^T v_i - v_1^T v_1, \quad i = 2, \ldots, n+1 \ . \tag{17}$$

Example 3. Consider the following three spheres in 2-dimensional space:

$$v_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 5 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 3 \\ 6 \end{pmatrix}, r_1^2 = 4, r_2^2 = 3, r_3^2 = 1.$$

Point $\theta = \begin{pmatrix} 2.6 \\ 2.7 \end{pmatrix}$ can be determined equating the two planes (17) between v_1 and v_2 and between v_1 and v_3, see Fig. 3. The corresponding solution has equal values $\|\theta - v_j\|^2 - r_j^2 = 10.05$ for the three vertices, v_1, v_2 and v_3.

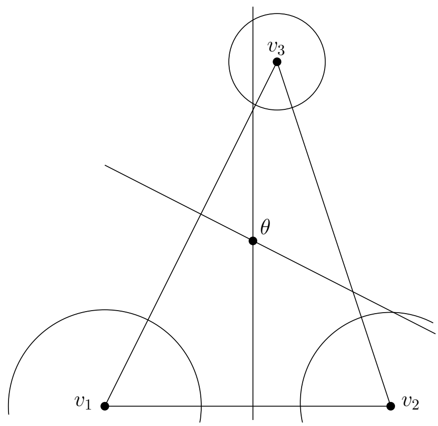

Fig. 3. Determination of the θ-point

It is interesting that the θ-point is closely related to the vertices that Breiman-Cutler keep track of in their algorithm. In fact, in [3] it is shown that if θ is interior with respect to C, then it is a global minimum point of ψ_C, defined similar as φ_k in (7), where one only considers the vertices of C. To be more precise:

$$\psi(x) = \max_{v_i \in C}\{\|x - v_i\|^2 - r_i^2\} \ . \tag{18}$$

Using (11), where $r_i^2 = 2\frac{f_i - U}{K}$, we can redefine

$$\psi_C(x) = \max_{v_i \in C}\{f_i - \frac{1}{2}K\|x - v_i\|^2\} \ . \tag{19}$$

Notice that $\psi_C(x) \leq \varphi_k(x)$, because $\{v_i \in C\} \subset \{x_1, \ldots, x_k\}$. As has been shown in [3], $l(C) := \psi(\theta) \leq \min_{x \in C} \psi(x)$ is a lower bound of ψ_C over C. In that sense, $l(C)$ is also a lower bound of φ_k over C. The consequence of this theoretical results is that to check the cover, θ can be computed to find $l(C)$. If $l(C) > U$, then C cannot contain an optimum solution. For the underestimate based on M (8), one can also redefine the function ψ of (18). However, in that case also the ratio between radii (r_i/r_j) depends on the best function value found, U. That means, that also the θ-point depends on U. So, one can construct a similar test, but if an update of the global upper bound U has been found, the ratio changes, such that the θ-point is shifted.

Now getting back to the main question of our research that deals with the use of bisecting the longest edge by the midpoint. In the empirical work of running the algorithms, we found that if the θ-point and therefore minimum point of ψ, has more tendency to be inside the simplex under consideration, than in the case of needle shaped simplices. Concluding, *bisecting over the middle of the longest edge can be convenient for checking the cover of a simplex by non-optimality spheres centered at its vertices.*

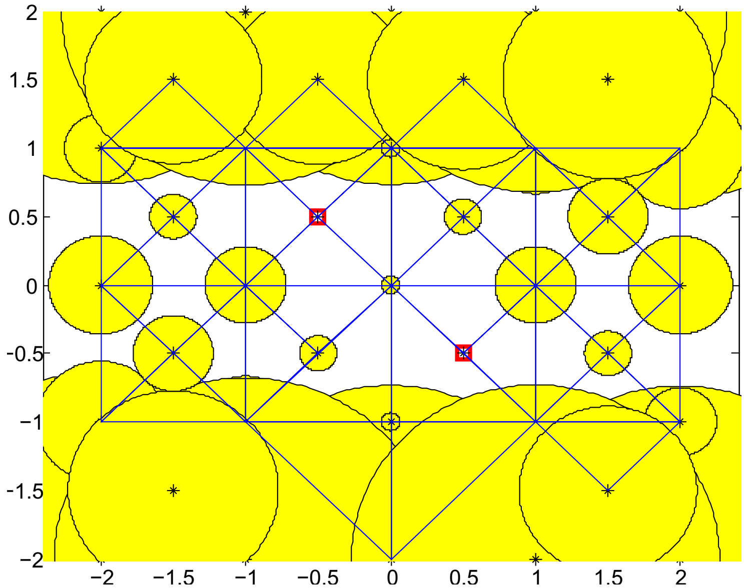

Fig. 4. Progress of a simple covering algorithm on six-hump after 50 iterations. The left over simplices, the evaluated points and their corresponding non-optimality spheres are depicted. The best points found are given by a small square.

If an inspected simplex C is not covered by the spheres at its vertices, it may still be covered completely by spheres centered at other points in $\{x_1, \ldots, x_k\} \setminus \{v_i \in C\}$. To check this individually, one can run over the list of evaluated points $x_1, \ldots, x_j, \ldots x_k$ and check whether

$$\max_{v_i \in C} \|x_j - v_i\|^2 < r_j^2 \quad . \tag{20}$$

Intuitively, a sphere has more tendency to cover a 'round' simplex than a needle shaped one. In the illustration, (20) is used on the bisected partition sets.

6 Numerical Illustration

How does the development of using non-optimality spheres in simplicial B&B look like? The presented algorithm is rather generic as many details can be filled in. A simple illustration is given without any pretention to outperform other covering based algorithms. The algorithm was applied to the six-hump camel-back function, where an accuracy of $\delta = 0.0001$ and $K = 60$ were used. The only used cover check is the validation of (20) for all evaluated points and a breadth-first-search selection was applied. A list of ns simplices is maintained and the number $ndel$ of deleted simplices and number nf of function evaluations is measured during the iterations it in Table 1.

The algorithm converges after 622 iterations returning the global minimum points. Notice that at each iteration two simplices are evaluated and that about half of them are not put on the list in the first place. Figure 4 sketches the progress after 50 iterations. Proceeding, U is updated and consequently spheres increase. Figure 5 shows the state after 150 iterations. These figures also show well that many simplices are covered by a set of non-optimality spheres, but not by a single one, so test (20) is quite rough. The main interpretation of the illustration is that in fact more 'round' simplices have an earlier covering by individual spheres than needle shaped simplices, advocating the use of bisection as division rule.

Table 1. Progress of the B&B algorithm on six-hump, $\delta = 0.0001$ and $K = 60$

it	50	200	400	622
ns	36	96	19	0
$ndel$	16	106	383	624
nf	39	129	285	454
U	-.13	-.98	-1.02	-1.03

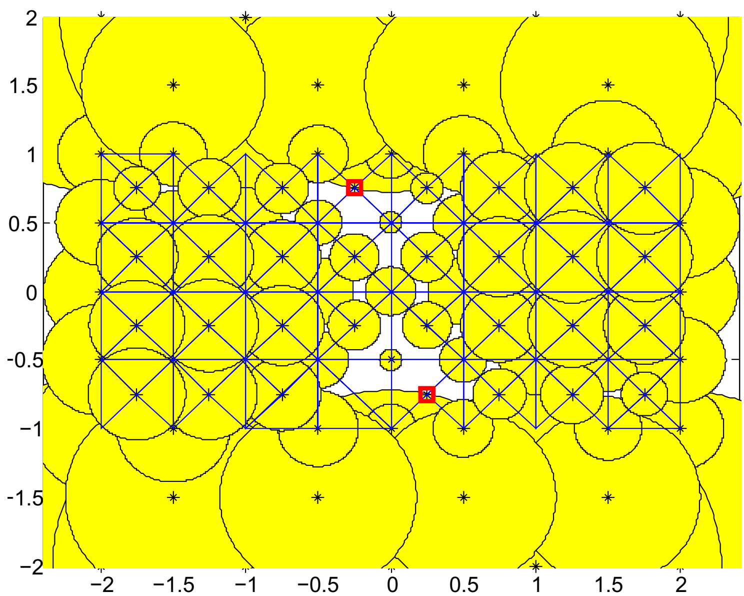

Fig. 5. Progress of a simple covering algorithm on six-hump after 150 iterations. The left over simplices, the evaluated points and their corresponding non-optimality spheres are depicted. The best points found are given by a small square.

7 Conclusions and Future Work

In a recent publication Reiner Horst stated that bisection is "not optimal" referring to volume considerations and convergence rates within a branch and bound tree.

In this paper we discuss several aspects for which the use of bisection in simplicial B&B may be convenient due to the feature of leading to relatively 'round' partition sets, and implicitly sampling over an equidistant grid. Bisecting the longest edge over the midpoint appears to be convenient for

- Robustness considerations searching for feasible ε-spheres;
- storage issues in branch and bound trees;
- checking the cover of a simplex by non-optimality spheres.

The first two observations follow from experience solving practical design problems by B&B. To elaborate the latter, a generic B&B algorithm has been outlined and several properties regarding non-optimality spheres in this context have been elaborated. As reported, relatively 'round' simplices appear to be convenient. Like Reiner, we also looked into other ways to divide simplices. Although convenient in low dimensional applications, regular equilateral subsivisions cannot be extended to higher dimensional branch and bound methods. Further research can even be focussed on overlapping equilateral subdivisions.

References

1. Baritompa, W.P.: Customizing methods for global optimization, a geometric viewpoint. Journal of Global Optimization 3, 193–212 (1993)
2. Breiman, L., Cutler, A.: A deterministic algorithm for global optimization. Mathematical Programming 58, 179–199 (1993)
3. Casado, L.G., García, I., Tóth, B.G., Hendrix, E.M.T.: On determining the cover of a simplex by spheres centered at its vertices. Journal of Global Optimization 50, 654–655 (2011)
4. Casado, L.G., Hendrix, E.M.T., García, I.: Infeasibility spheres for finding robust solutions of blending problems with quadratic constraints. Journal of Global Optimization 39, 557–593 (2007)
5. Danilin, Y., Piyavski, S.A.: An algorithm for finding the absolute minimum. Theory of Optimal Decisions 2, 25–37 (1967) (in Russian)
6. Evtushenko, Y., Posypkin, M.: Coverings for global optimization of partial-integer nonlinear problems. Doklady Mathematics 83, 1–4 (2011)
7. Horst, R.: On generalized bisection of n-simplices. Mathematics of Computation 66(218), 691–698 (1997)
8. Horst, R.: Bisection by global optimization revisited. Journal of Optimization Theory and Applications 144, 501–510 (2010)
9. Horst, R., Pardalos, P.M., Thoai, N.V.: Introduction to Global Optimization, Nonconvex Optimization and its Applications, vol. 3. Kluwer Academic Publishers, Dordrecht (1995)
10. Horst, R., Tuy, H.: On the convergence of global methods in multiextremal optimization. Journal of Optimization Theory and Applications 54, 253–271 (1987)

11. Horst, R., Tuy, H.: Global Optimization (Deterministic Approaches). Springer, Berlin (1990)
12. Locatelli, M., Raber, U.: On convergence of the simplicial branch-and-bound algorithm based on ω-subdivisions. J. Optim. Theory Appl. 107, 69–79 (2000)
13. Mladineo, R.H.: An algorithm for finding the global maximum of a multimodal multivariate function. Mathematical Programming 34, 188–200 (1986)
14. Paulavičius, R., Žilinskas, J., Grothey, A.: Investigation of selection strategies in branch and bound algorithm with simplicial partitions and combination of lipschitz bounds. Optimization Letters 4, 173–183 (2010)
15. Piyavski, S.A.: An algorithm for finding the absolute extremum of a function. USSR Computational Mathematics and Mathematical Physics 12, 57–67 (1972) (in Russian)
16. Raber, U.: A simplicial branch-and-bound method for solving nonconvex all-quadratic programs. Journal of Global Optimization 13, 417–432 (1998)
17. Raber, U.: Nonconvex All-Quadratic Global Optimization Problems: Solution Methods, Application and Related Topics. Ph.D. thesis, Trier University (1999)
18. Shubert, B.O.: A sequential method seeking the global maximum of a function. SIAM Journal of Numerical Analysis 9, 379–388 (1972)
19. Zilinskas, A., Clausen, J.: Subdivision, sampling, and initialization strategies for simplicial branch and bound in global optimization. International Journal of Computers and Mathematics with Applications 44, 943–955 (2002)
20. Zilinskas, A., Zilinskas, J.: Global optimization based on a statistical model and simplicial partitioning. International Journal of Computers and Mathematics with Applications 44, 957–967 (2002)