Mathematical Programming 55 (1992) 319-337
North-Holland

The bisection method in higher dimensions

G.R. Wood

Mathematics Department, University of Canterbury, Christchurch, New Zealand

Received 8 June 1988
Revised manuscript received October 1989

Is the familiar bisection method part of some larger scheme? The aim of this paper is to present a natural and useful generalisation of the bisection method to higher dimensions. The algorithm preserves the salient features of the bisection method: it is simple, convergence is assured and linear, and it proceeds via a sequence of brackets whose infinite intersection is the set of points desired. Brackets are unions of similar simplexes. An immediate application is to the global minimisation of a Lipschitz continuous function defined on a compact subset of Euclidean space.

AMS 1980 Subject Classification: 49D37.

Key words: Bisection, simplex, global optimisation, linear convergence, zonotope, tiling.

Introduction

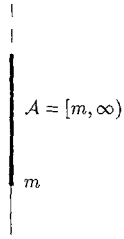
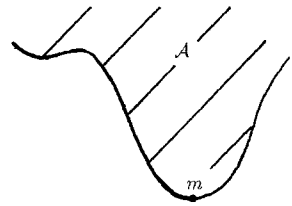
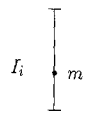
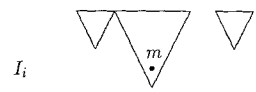
The bisection method appears in diverse roles throughout mathematics. Most commonly today it is associated with root-finding, but it can also be used to show that a continuous function on a closed interval achieves its maximum, or that every infinite subset of a closed interval has a limit point. In each instance the method uses an initial bracketing interval which is successively halved in such a way that the intersection of the sequence of brackets is the point of interest. Necessarily convergence is assured and linear.

The aim of this paper is to present a natural and useful generalisation of the bisection method to higher dimensions (Algorithm 3.1). The need for such a generalisation has been expressed for example in [7, p. 25]. Our starting point is the recognition that the underlying "bisection problem" can be considered to be that of finding all globally minimal points of a closed, cone-ordered subset of Euclidean space. The generalisation of the bisection algorithm presented here begins by bracketing such points in a simplex, and proceeds so that at each iteration the bracket is a union of similar simplexes. The infinite intersection of all such brackets is the required set of global minima. Convergence is again assured and linear. Table 1 presents a quick comparison of the familiar bisection method, viewed in this light, and the generalisation to higher dimensions presented in this paper.

A theoretical account of the algorithm will be presented here. An immediate application of the method is to the global minimisation, over a compact domain, of a Lipschitz continuous function of many variables. Details of this application, together with variations upon the algorithm, have appeared in [10].

Table 1

A comparison of the bisection method and the generalisation to higher dimensions

Bisection ($n = 0$)	Higher dimensional bisection ($n > 0$)
Problem (in $\mathbf{R}^{n+1}$)	
	
Finding the least element of a semi-infinite interval $\mathcal{A}$, a set closed in the usual topology and partial order of $\mathbf{R}$	Finding the least element of $\mathcal{A}$, a set closed in the usual topology and a partial ordering of $\mathbf{R}^{n+1}$
Domain (in $\mathbf{R}^n$)	
A point	A line segment ($n = 1$), a hexagon ($n = 2$), a rhombic dodecahedron ($n = 3$), ...
Bracket (in $\mathbf{R}^{n+1}$)	
An interval I_i, a simplex in $\mathbf{R}$	A union, I_i, of simplexes in $\mathbf{R}^{n+1}$
	
Convergence	
i) $\cap_{i=0}^{\infty}\{I_i\} = \{m\}$	i) $\cap_{i=0}^{\infty}\{I_i\} = \{\text{all global minima of } \mathcal{A}\}$
ii) $\lvert I_{i+1}\rvert = \frac{1}{2}\lvert I_i\rvert$	ii) Variation$(I_{i+1}) \leq \frac{n}{n+1}$ Variation(I_i)

Bisection is reviewed in Section 1 of this paper. In Section 2 the higher dimensional problem is stated and in Section 3 the higher dimensional algorithm presented. The details of the algorithm are discussed in the next three sections: the iterative step is described in Section 4, the convergence results in Section 5, and in Section 6 the initial system is set up. Simple examples are presented in Section 7, followed by discussion in Section 8.

1. Bisection reviewed: The problem and the method

The bisection method is not limited to root-finding. As the examples in the introduction indicate, the method finds wide application throughout mathematics. In this section we isolate the general problem solved by the bisection method, and formulate

it in a new and helpful way. In the next section we shall pose the problem in higher dimensions, and discover a simple extension of the bisection method.

Each of the three applications of the bisection method described in the introduction involved an initial bracket, I_0, containing a point to be found, m. At each stage of the algorithm the mid-point of the current bracket, I_i, is examined, and information available allows the selection of a half-interval I_{i+1} retaining m. This mechanism is pictured conveniently here in Figure 1(a).

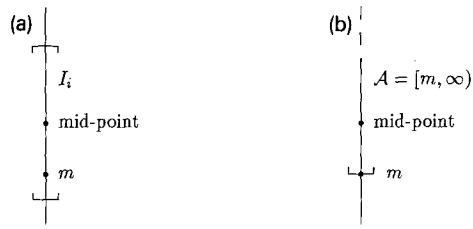


Fig. 1. (a) Traditional view and (b) alternative view of the bisection method. In (a) the point of interest, m, lies in the lower subinterval of I_i. This is equivalent to the interval mid-point lying in $\mathcal{A}$, shown in (b).

We now present a new setting for this old mechanism. At each iteration it suffices to know whether m lies in the lower sub-interval, but note that

$$m \in \text{lower subinterval} \Leftrightarrow \text{interval mid-point} \geqslant m \Leftrightarrow \text{interval mid-point} \in \mathcal{A},$$

where $\mathcal{A} = [m, \infty)$, the closed half line above m, as illustrated in Figure 1(b). That is, knowing whether the lower subinterval contains the minimum point m is equivalent to knowing whether the interval mid-point lies in $\mathcal{A}$. This allows us to isolate those problems amenable to the bisection method in the following way:

The Bisection Problem. Let $\mathcal{A} = [m, \infty) \subseteq \mathbb{R}$, the real line. Suppose that for any x in $\mathbb{R}$ it is possible to decide whether x lies in $\mathcal{A}$. Find m.

In this new setting the familiar algorithm can be phrased as follows:

The Bisection Method.

Initial step: Form an initial bracket, I_0.

Iterative step: For $i \in \mathbb{Z}^+$, the non-negative integers, and x the mid-point of I_i, if $x \in \mathcal{A}$ then I_{i+1} is the lower subinterval, else I_{i+1} is the upper subinterval.

Consequences: (1) For $i \in \mathbb{Z}^+$, $|I_{i+1}| = \frac{1}{2}|I_i|$.

(2) $\bigcap_{i=0}^{\infty} I_i = \{m\}$.

2. Higher dimensions: The problem

Problems suited to the usual bisection method evidently require us to find the least element of a closed half-line, $\mathcal{A} = [m, \infty)$. Amidst proper subsets of the line such half-lines are characterised by the following two closure properties:

(1) $\mathscr{A}$ is topologically closed, and

(2) $\mathscr{A}$ is algebraically closed, with respect to the usual ordering of the real line.

Our search is for the analogues of such sets in higher dimensions, whose minimal points a multidimensional bisection method would locate. Only the algebraic closure presents a problem. Recall that the familiar orderings of Euclidean space are those partial orderings which lead to an ordered vector space structure, and to each such ordering there corresponds a positive cone, the set of all points greater than or equal to the origin (see for example [2, pp. 169–171]). The simplest such cone is that whose base is a simplex: this is our departure point.

Bearing in mind the problem of minimising a real-valued function of n variables, we set up the higher dimensional problem in $\mathbb{R}^{n+1}$, thought of as the cartesian product of a domain $\mathbb{R}^n$ and range $\mathbb{R}$. With this view, the degenerate case $n=0$ corresponds to a 0-dimensional domain, as illustrated in Table 1. Let ∇ be a cone in $\mathbb{R}^{n+1}$ whose base is a simplex. We standardise by assuming that ∇ has apex the origin and axis the $(n+1)$st coordinate axis, and that the cross-section, or *base*, of ∇ orthogonal to the $(n+1)$st axis is a regular simplex. Then ∇ can be used to order $\mathbb{R}^{n+1}$ in the following way: for $x, y \in \mathbb{R}^{n+1}$, we say $x \leqslant y$ if and only if $y \in x+\nabla$. We now define those sets in $\mathbb{R}^{n+1}$ whose minimal points the generalised algorithm will find.

Definition 2.1.

(1) The proper subset $\mathscr{A}$ of $\mathbb{R}^{n+1}$ is termed an *admissible* region if

(a) $\mathscr{A}$ is topologically closed, and

(b) $\mathscr{A}$ is algebraically closed, with respect to the ordering of $\mathbb{R}^{n+1}$ by ∇.

(2) We say $m \in \mathscr{A}$ is *minimal* if $m_{n+1} \leqslant x_{n+1}$ for all $x \in \mathscr{A}$, where x_{n+1} denotes the $(n+1)$st coordinate of x.

For $n>0$, minimal points of $\mathscr{A}$ need not exist. For example, let $\nabla = \{(x, y): y \geqslant |x|\}$ in $\mathbb{R}^2$, and $\mathscr{A}$ be the epigraph of $e^{-x^2/2}$. If we restrict ourselves, however, to admissible regions over a compact subset K of $\mathbb{R}^n$ (the span of the first n coordinate axes of $\mathbb{R}^{n+1}$), that is, regions of the form $(K \times \mathbb{R}) \cap \mathscr{A}$, then it is straightforward to show that minimal points will exist. In future we term K the *domain* of the problem.

We are now in a position to state:

The Multidimensional Bisection Problem. Let $\mathscr{A}$ be an admissible region in $\mathbb{R}^{n+1}$, and K be a compact subset of $\mathbb{R}^n$. Suppose that for any x in $\mathbb{R}^{n+1}$ it is possible to decide whether $x \in \mathscr{A}$. Find the minimal admissible points over K.

We pause briefly to set up some essential notation, required only for the extended problem, when $n>0$. Let $\{e_1, \ldots, e_{n+1}\}$ be the standard basis for $\mathbb{R}^{n+1}$. Let $\{u_1, \ldots, u_{n+1}\}$ comprise unit vectors from the origin to the vertices of a regular simplex, with centroid the origin, in $\mathbb{R}^n$, the span of $\{e_1, \ldots, e_n\}$. Thus $u_1 + \cdots + u_{n+1} = 0$ and $u_k \cdot u_l = -1/n$ for all k, l, $k \neq l$. Given $M>0$, let ∇ be the cone with

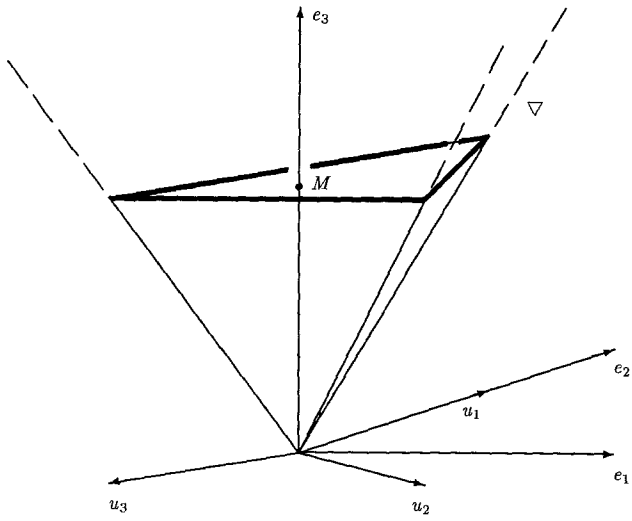


Fig. 2. The ordering cone, ∇.

apex the origin and cross section at height M along the $(n+1)$st axis a translate along this axis of $\text{co}\{u_1, \ldots, u_{n+1}\}$, where co denotes the convex hull. Formally,

$$\nabla = \text{pos}\{u_k + Me_{n+1} : k = 1, \ldots, n+1\},$$

where pos denotes all positive linear combinations. Figure 2 illustrates the situation in the case $n = 2$.

In the sequel we shall find it useful to refer, for example, to a "high" vertex of a simplex: by this we shall mean a vertex of the simplex with largest $(n+1)$st coordinate.

3. Higher dimensions: The algorithm

We now present the algorithm which solves the multidimensional problem in a fashion extending the familiar case. We assume an admissible region $\mathscr{A}$ in $\mathbb{R}^{n+1}$ is given. Shortly we shall see that the algorithm operates naturally over certain domains in $\mathbb{R}^n$: when $n = 1$ this domain is a line segment, for $n = 2$ it is a regular hexagon, while for $n = 3$ it is a regular rhombic dodecahedron — the honeycomb cell. In general a standard domain H will take the form $c + rU$ where $c \in \mathbb{R}^n$, $r \geq 0$ and U is the vector sum $U_1 + \cdots + U_{n+1}$, where U_i is the line segment from 0 to u_i. Conveniently, these regions tile the space in which they lie, so allowing us to use unions of standard domains to approximate arbitrary domains. For now we assume a standard domain H has been specified.

The algorithm begins by forming an *initial system* $\mathscr{S}_0$, a single standard simplex, which brackets all minimal admissible points over H. By means of a pair of processes, *reduction* ($\mathscr{R}$) and *elimination* ($\mathscr{E}$) each iteration decreases the size of the system

(measured by the system *variation*, $V(\mathcal{S})$) in such a way that all minimal admissible points appear as the infinite intersection of the set of system brackets. The main new feature is that beyond the familiar case, when $n=0$, brackets are in general a *union* of standard simplexes. So when $n=1$, a bracket is a union of standard triangles, when $n=2$ a bracket is a union of standard tetrahedra, and so on. With no more ado we state the algorithm.

Algorithm 3.1. Given $\mathcal{A}$ and H:

Initial step: Form the initial system $\mathcal{S}_0$ (Definition 6.1).

Iterative step: For $i \in \mathbb{Z}^+$, let $\mathcal{S}_{i+1} = \mathcal{E}(\mathcal{R}(\mathcal{S}_i))$ (Definitions 4.3 and 4.4).

Consequences: (1) For $i \in \mathbb{Z}^+$, $V(\mathcal{S}_{i+1}) \leq (n/(n+1))V(\mathcal{S}_i)$ (Lemma 5.1).

(2) $\bigcap_{i=0}^{\infty} I_i = \{\text{all minimal admissible points over } H\}$ (where I_i is the ith bracket) (Theorem 5.1).

Stopping rule: Given $\varepsilon > 0$, stop when $V(\mathcal{S}_i) < \varepsilon$.

Outcome: If $\mathcal{S}_i = (\mathcal{T}_i, a_i)$, then for i such that $V(\mathcal{S}_i) < \varepsilon$, a_i is an admissible point such that $a_{i,n+1}$ is within ε of the minimal admissible level.

We explain the algorithm now in three stages. The heart, the iterative step, is examined first. This involves a study of the two processes, reduction and elimination, which diminish the bracket size while retaining the minimal admissible points. Following this we consider the convergence of the algorithm, and finally we show how to set up the initial system.

4. The iterative step

We begin with four definitions. These allow us to describe the brackets of a system, and to measure their size.

Definition 4.1.

(1) A *standard simplex* in $\mathbb{R}^{n+1}$ is a translate of a cap of the cone ∇, of the form

$$T(x, h) = \operatorname{co}\left(\{x\} \cup \left\{x + he_{n+1} + \frac{h}{M} u_k : k = 1, \ldots, n+1\right\}\right)$$

where $x \in \mathbb{R}^{n+1}$ is the apex, and $h \in \mathbb{R}$ the height. By the *top* of $T(x, h)$ we shall mean the facet opposite the apex, x.

(2) A *system of simplexes* (or *system*) $\mathcal{S}$ in $\mathbb{R}^{n+1}$ is

(a) a finite set $\mathcal{T}$ of standard simplexes, plus

(b) an admissible point a in $\mathbb{R}^{n+1}$, lying in a lowest top of the system.

We write $\mathcal{S} = (\mathcal{T}, a)$, with $\mathcal{T} = \{T_j\}$, where $T_j = T(x_j, h_j)$ and j runs through a finite set.

(3) A *uniform system* is a system $\mathcal{S}$ in which all tops lie in the same hyperplane of $\mathbb{R}^{n+1}$. Alternatively, $x_{j,n+1} + h_j = x_{k,n+1} + h_k$ for all T_j, T_k in $\mathcal{T}$, where $x_{j,n+1}$ is the $(n+1)$st coordinate of x_j.

(4) The *variation of a system* $\mathscr{S}$, $V(\mathscr{S})$, is the difference between the highest and lowest points in the system $\mathscr{S}$. That is,

$$V(\mathscr{S}) = \max_j \{x_{j,n+1} + h_j\} - \min_j \{x_{j,n+1}\}.$$

We proceed now to discuss the iterative step. For the moment we take an initial system $\mathscr{S}_0 = (\mathscr{T}_0, a_0)$ for granted. This consists of a single simplex, T_0, bounding $\mathscr{A}$ over H from below. In the sequel we must make the following assumption:

Assumption 4.1. The minimal admissible points of $\mathscr{A}$ over the top of T_0 (precisely, its projection onto $\mathbb{R}^n$) occur over H.

Given $x \in \mathbb{R}^{n+1}$, the least element of $\mathscr{A}$ over x itself we denote by x'. In future we assume x' is obtainable, using if necessary the usual bisection method. For convenience we term this process "evaluation of $\mathscr{A}$ at x". When applying the method to optimisation problems, $\mathscr{A}$ is the epigraph of a function f of n variables, so x' is obtained from an evaluation of f.

The algorithm is built entirely around two critical, but simple, facts. For the x' associated with x in $\mathbb{R}^{n+1}$,

(1) No point higher than x' can be minimal.

(2) No point inside $x' - \nabla$ can be admissible (else that x' is the least element of $\mathscr{A}$ over x is contradicted).

These two facts allow us to strip away upper halfspaces and inverted cones from the initial simplex, while leaving all minimal admissible points. That a base for $-\nabla$ is an n-simplex dually oriented to that forming a base for ∇ will enable us to create a simple algorithm, involving standard simplexes at each stage. We now separately introduce system reduction and system elimination.

4.1. System reduction

To reduce a system we reduce every simplex in the system. Simplex reduction itself is the key to the algorithm: to reduce the simplex $T(x, h)$ we use the evaluation of $\mathscr{A}$ at x to truncate T at height x'_{n+1} and remove the interior of the cone $x' - \nabla$. This leaves at most $n+1$ standard simplexes, with variation reduced by a factor of at least $n/(n+1)$. The case for $n = 2$ is illustrated in Figure 3. Formally, we have:

Definition 4.2 (simplex reduction). Let $T(x, h)$ be a standard simplex. Two cases occur:

(1) Upper reduction (case $h \leq x'_{n+1} - x_{n+1}$). If $h \leq x'_{n+1} - x_{n+1} \leq (n+1)h$ then the reduction of T is (see Figure 3(a)):

$$\mathscr{T} = \left\{ T\left[x + \frac{x'_{n+1} - x_{n+1}}{M(n+1)}(u_k + Me_{n+1}), h - \frac{1}{n+1}(x'_{n+1} - x_{n+1}) \right] : k = 1, \ldots, n+1 \right\},$$

else if $(n+1)h < x'_{n+1} - x_{n+1}$, then $\mathscr{T}$ is the empty set.

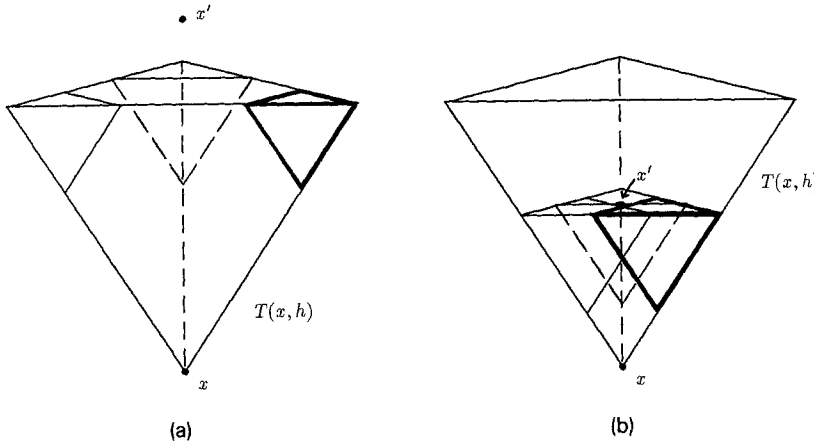


Fig. 3. Simplex reduction. (a) Upper reduction (when $h \leq x'_{n+1} - x_{n+1}$). (b) Lower reduction (when $x'_{n+1} - x_{n+1} < h$).

(2) Lower reduction (case $x'_{n+1} - x_{n+1} < h$). If $0 \leq x'_{n+1} - x_{n+1} < h$ then the reduction of T is (see Figure 3(b)):

$$\mathcal{T} = \left\{ T\left[x' + \frac{x'_{n+1} - x_{n+1}}{M(n+1)}(u_k - Mne_{n+1}), \frac{n}{n+1}(x'_{n+1} - x_{n+1}) \right] : k = 1, \ldots, n+1 \right\},$$

else if $x'_{n+1} - x_{n+1} < 0$ then $\mathcal{T}$ is the empty set.

It is readily shown that $V(\mathcal{T}) \leq (n/(n+1))h$, where $V(\mathcal{T})$ is the variation of the reduction $\mathcal{T}$. For system reduction we now reduce every simplex in the system:

Definition 4.3 (system reduction). The reduction of the uniform system $\mathcal{S} = (\mathcal{T}, a)$ is $\mathcal{R}(\mathcal{S}) = (\mathcal{T}', a')$ where

$$\mathcal{T}' = \bigcup_{T_j \in \mathcal{T}} \mathcal{T}_j, \quad \text{where } \mathcal{T}_j \text{ is the reduction of simplex } T_j,$$

and

$$a' = \begin{cases} a, & \text{if no lower reductions occur,} \\ x' \text{ such that } x'_{n+1} = \min\limits_j \{x'_{j,n+1}\}, & \text{otherwise.} \end{cases}$$

Proposition 4.1. *Let $\mathcal{S}$ be a uniform system, inside the simplex T_0, containing all minimal admissible points over H. Then*

(1) *$\mathcal{R}(\mathcal{S})$ contains all minimal admissible points over H, and*

(2) *$\mathcal{R}(\mathcal{S})$ forms a system.*

Proof. From the definition of $\mathcal{T}'$, (1) will follow if we show that minimal admissible points in any $T_j \in \mathcal{T}$ remain in the $n+1$ simplexes in the reduction of T_j, namely in $\bigcup_{k=1}^{n+1} T_{jk}$, where T_{jk} is the subsimplex of T_j determined by u_k (see Definition 4.2). Let $P^- = \{x \in \mathbb{R}^{n+1} : x_{n+1} \leq x'_{j,n+1}\}$ be the halfspace below the evaluation of $\mathcal{A}$ at x_j. Let F_k be the hyperplane supporting the kth facet of $x'_j - \nabla$ (that is, the facet

opposite the edge of $-\nabla$ determined by $-u_k$), and F_k^+ be the closed half-space not containing $x_j'-\nabla$. Then

$$\bigcup_{k=1}^{n+1} T_{jk} = \bigcup_{k=1}^{n+1} (T_j \cap P^- \cap F_k^+)$$

$$= T_j \cap P^- \cap \bigcup_{k=1}^{n+1} F_k^+$$

$$= T_j \cap P^- \cap (\mathbb{R}^{n+1} \backslash \text{int}(x_j' - \nabla)),$$

where int denotes the interior. Since no point above x_j' can be minimal (here we use Assumption 4.1 that, over the top of the initial simplex, $\mathcal{A}$ assumes its least value on H itself), and no point inside $x_j'-\nabla$ can be admissible, we conclude that no minimal admissible points are shed during simplex reduction.

From (1) we know that $\mathcal{T}'$ is a non-empty set of simplexes, standard by construction. The definition of a' ensures that it is admissible and lies in a lowest top of the reduced system. Thus (2) holds. □

4.2. System elimination

Tops of simplexes in a reduced system will lie at a variety of levels. The elimination process truncates all simplexes at the level of the lowest top, yielding a uniform system.

Definition 4.4 (system elimination). The eliminated system associated with the system $\mathcal{S} = (\mathcal{T}, a)$ is $\mathcal{E}(\mathcal{S}) = (\mathcal{T}', a)$ where $\mathcal{T}' = \{T \cap P^-: T \in \mathcal{T}$ and T is not contained in some other member of $\mathcal{T}\}$ with $P^- = \{x \in \mathbb{R}^{n+1}: x_{n+1} \leq a_{n+1}\}$.

Proposition 4.2. *Let $\mathcal{S}$ be a system, inside the initial simplex T_0, containing all minimal admissible points over H. Then*

(1) *$\mathcal{E}(\mathcal{S})$ contains all minimal admissible points over H, and*

(2) *$\mathcal{E}(\mathcal{S})$ forms a uniform system.*

Proof. Again, using Assumption 4.1, that over the top of T_0 the admissible region $\mathcal{A}$ assumes its least value on H, we know that no point above a can be minimal. Thus (1) follows: all minimal admissible points remain in $\mathcal{E}(\mathcal{S})$. Since a is both admissible and no higher than any top in $\mathcal{S}$, truncation of the system at this level yields a uniform system. □

5. Convergence results

Linear convergence of the variation of our system sequence is guaranteed by the following lemma.

Lemma 5.1. *Let $\mathcal{S}$ be a uniform system inside the initial simplex T_0. Then*

$$V(\mathcal{E}(\mathcal{R}(\mathcal{S}))) \leq \frac{n}{n+1} V(\mathcal{S}).$$

Proof. Let $\mathcal{S} = (\mathcal{T}, a)$, and select the T_j in $\mathcal{T}$ whose reduction contains a point at least as low as any point in any other reduction. Letting $V(\mathcal{T}_j)$ be the variation of $\mathcal{T}_j$, the reduction of T_j, we have

$$V(\mathcal{E}(\mathcal{R}(\mathcal{S}))) \leq V(\mathcal{T}_j) \quad \text{(by choice of } T_j\text{)},$$

$$\leq \frac{n}{n+1} h_j \quad \text{(see remark after Definition 4.2)},$$

$$\leq \frac{n}{n+1} V(\mathcal{S}) \quad (\text{since } T_j \in \mathcal{T}). \qquad \square$$

We now present the main convergence theorem.

Theorem 5.1. *Let $\mathcal{S}_i = (\mathcal{T}_i, a_i)$ be the ith system, and $I_i = \bigcup_{T \in \mathcal{T}_i} T$ the ith bracket, for $i \in \mathbb{Z}^+$, as presented in the multidimensional algorithm. Then the set of all minimal admissible points over H is precisely $\bigcap_{i=0}^{\infty} I_i$.*

Proof. Suppose that m is a minimal admissible point over H. Then $m \in I_0$ $(= T_0)$, at this stage by assumption (but shown later in Proposition 6.1). If $m \in I_i$ for some $i \geq 0$, then m is also in I_{i+1} (Propositions 4.1(1) and 4.2(1)). Thus $m \in \bigcap_{i=0}^{\infty} I_i$.

For the reverse inclusion, suppose that $x \in I_0$ is such that $x_{n+1} \neq l$, the minimum admissible level. Let $d = x_{n+1} - l \neq 0$. Then Lemma 5.1 gives that

$$V(\mathcal{S}_i) \leq \left(\frac{n}{n+1}\right)^i V(\mathcal{S}_0),$$

so we can choose i_0 such that for $i \geq i_0$, $V(\mathcal{S}_i) < |d|$. Thus $x \notin I_i$ for $i \geq i_0$, or $x \notin \bigcap_{i=0}^{\infty} I_i$. Finally, suppose that $x \in I_0$ is such that $x_{n+1} = l$ yet $x \notin \mathcal{A}$. Since $\mathcal{A} \cap I_0$ is compact, the minimum distance from $\mathcal{A} \cap I_0$ to x is positive. Thus a standard simplex of sufficiently small variation containing x would be entirely removed by an upper reduction, so again $x \notin I_i$ for i sufficiently large. $\square$

The outcome of Algorithm 3.1 now follows from Definition 4.1(2), Propositions 4.1(2) and 4.2(2), and Theorem 5.1.

Why, when $n = 0$, does the variation reduction formula of Lemma 5.1 not yield the familiar interval halving formula? The answer lies in our choice of the reduction process, here influenced by the global optimisation problem. We have thought of the boundary of $\mathcal{A}$ as the graph of a function, and assumed that, for each x in $\mathbb{R}^n$, $f(x) = x'$ can be found.

A simpler reduction process is possible, one that precisely parallels what is happening in the case when $n=0$. In such a process a simplex $T(x, h)$ is reduced by examining the mid-point, $x'=x+\frac{1}{2}he_{n+1}$, of the axis of the simplex. If $x'\in\mathscr{A}$, then T is truncated at the level of x', whereas if $x'\notin\mathscr{A}$ then $-\nabla$ is removed at x', in the usual way. The variation of the reduction $\mathscr{T}$ of the single simplex T is now given by

$$V(\mathscr{T})=\begin{cases}\frac{1}{2}V(\mathscr{T}), & \text{if } x'\in\mathscr{A},\\ \frac{1}{2}(1+n/(n+1))V(\mathscr{T}), & \text{if } x'\notin\mathscr{A},\end{cases}$$

while the formula

$$V(\mathscr{E}(\mathscr{R}(\mathscr{S})))\leq\tfrac{1}{2}(1+n/(n+1))V(\mathscr{S})$$

holds for systems, using the proof method of Lemma 5.1. This gives an exact generalisation of the case when $n=0$. In the sequel, however, we return to the original definition of system reduction.

6. The initial system

Our aim is to find a standard simplex which brackets the minimal admissible points of $\mathscr{A}$ over a compact set K. We begin with an informal description of the natural solution, with Figure 4 illustrating the situation when $n=2$. Consider a cube in $\mathbb{R}^{n+1}$ positioned such that a major diagonal is aligned with the $(n+1)$st coordinate axis.

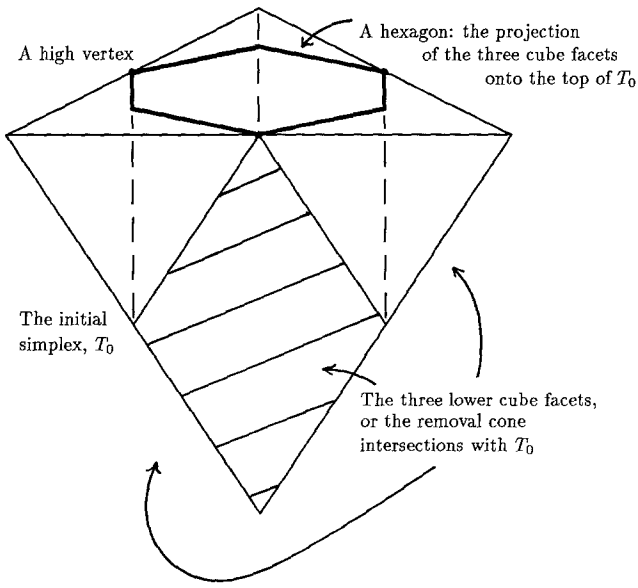


Fig. 4. Construction of the initial simplex when $n=2$.

Consider now the $n+1$ facets of the cube which include the low vertex. By stretching or contracting the cube along the $(n+1)$st coordinate axis, and rotating the cube suitably about this axis, these facets will sit precisely within a standard simplex, T_0, with top passing through the high vertices of these cube facets. Suppose for a moment that the evaluation of $\mathscr{A}$ at these $n+1$ high vertices are equal, and equal to the vertex itself. Then the $n+1$ facets will be exactly the intersections of T_0 with the removal cones at these $n+1$ high vertices. The projection of this set of facets onto the top of T_0 will be the domain covered by T_0. In the event that the evaluations differ, certain removal cones are raised, leading to a bracketing simplex reduced in size.

Thus domains which are projections of such "half-cubes", which prove to be line segments, hexagons, honeycomb cells and so on, lead in a straightforward way to a bracketing standard simplex: to set up the initial simplex over such a region $n+1$ evaluations of $\mathscr{A}$ must be made (see Definition 6.1). As mentioned earlier, these standard domains prove extremely convenient. Generally, given a compact domain K it will suffice to find a standard domain H such that $K \subseteq H$. We shall see, however, in Remark 6.1, that in each dimension the standard domain tiles $\mathbb{R}^n$, so if necessary we can approximate arbitrary domains with a patchwork of standard domains, and apply the algorithm to each in turn.

We proceed now to a formal definition of the initial system and a statement and proof of the bracketing property.

Definition 6.1 (the initial system).

(1) A *standard domain* in $\mathbb{R}^n$ has the form $c+rU$, where $c \in \mathbb{R}^n$, $r \geq 0$ and U is the vector sum $U_1 + \cdots + U_{n+1}$, where U_k is the line segment from 0 to u_k.

(2) For $H = c + rU$, a standard domain in $\mathbb{R}^n$, and admissible region $\mathscr{A}$, the initial system $\mathscr{S}_0 = (\mathscr{T}_0, a_0)$, with $\mathscr{T}_0 = \{T_0 = T(x_0, h_0)\}$, is given by

$$x_0 = c + \frac{1}{M(n+1)} \sum_{k=1}^{n+1} (v'_{k,n+1} - m) u_k + \left(\frac{1}{n+1} \sum_{k=1}^{n+1} v'_{k,n+1} - Mnr \right) e_{n+1},$$

$$h_0 = Mnr - \frac{1}{n+1} \sum_{k=1}^{n+1} (v'_{k,n+1} - m),$$

$$a_0 = c - ru_l + me_{n+1},$$

with

$$v_k = c - ru_k, \quad k = 1, \ldots, n+1, \quad \text{the dual vertices of } H,$$

$$m = \min_k \{v'_{k,n+1}\},$$

l = an index associated with the lowest evaluation (that is, $v'_{l,n+1} = m$).

Proposition 6.1. *$\mathcal{S}_0$ is a uniform system containing all minimal admissible points over H.*

Proof. We build to the general result by considering a sequence of four cases.

Case 1. $c=0$, $r=1$, and $v'_{k,n+1}=0$ for $k=1,\ldots,n+1$. Here $T_0=T(-Mne_{n+1}, Mn)$ and $a_0=-u_1$, so $\mathcal{S}_0$ is certainly a uniform system. In order to show that T_0 bounds $\mathcal{A}$ from below over $H=U$ we show:

(1) The intersection of the kth facet of $v_k-\nabla$ (the removal cone at v_k) with the kth facet of T_0 (that opposite the edge determined by u_k) is the convex set with the 2^n extreme points of the form

$$\sum_{j\in J} u_j-(n-|J|)Me_{n+1},$$

where $J\subseteq\{1,\ldots,n+1\}\backslash\{k\}$ and $|J|$ is the cardinality of J. (This convex set is affinely isomorphic to a cube.)

(2) The projection of this intersection of facets onto $\mathbb{R}^n$ is H_k, the vector sum of $U_1,\ldots,U_{k-1},U_{k+1},\ldots,U_{n+1}$.

Since $\bigcup_{k=1}^{n+1} H_k=H$, and no point above the origin is minimal, and no point inside the removal cones $v_k-\nabla$, $k=1,\ldots,n+1$ is admissible, it follows that T_0 brackets all minimal admissible points over H.

Proof of (1). The kth facet of $-u_k-\nabla$ and the kth facet of T_0 lie in a common hyperplane of $\mathbb{R}^{n+1}$. The base for the kth facet of $-u_k-\nabla$ is dually oriented to that of the kth facet of T_0, so their intersection is affinely isomorphic to an n-cube. It is not hard to check that the extreme points of this intersection have the form

$$-u_k+\sum_{j\in J}(-u_j-Me_{n+1}),$$

where J is any subset of $\{1,\ldots,n+1\}\backslash\{k\}$. Since $\sum_{k=1}^{n+1} u_k=0$, these 2^n extreme points can be alternatively written in the form

$$\sum_{j\in J} u_j-(n-|J|)Me_{n+1}.$$

Proof of (2). A projection of a convex polytope is the convex hull of the projections of the extreme points. Now the projections of the extreme points in (1) are just $\sum_{j\in J} u_j$ in $\mathbb{R}^n$. It is a straightforward exercise to show these are precisely the extreme points of $H_k=U_1+\cdots+U_{k-1}+U_{k+1}+\cdots+U_{n+1}$, so proving (2).

Case 2. $c=0$, $r=1$, $v'_{k,n+1}=m$, for all k. By Case 1, $T(-Mne_{n+1}, Mn)$ would provide an initial simplex for the admissible region $\mathcal{A}-me_{n+1}$, whence $T((m-Mn)e_{n+1}, Mn)$ will suffice for $\mathcal{A}$, together with the admissible point $-u_1+me_{n+1}$.

Case 3. $c=0$, $r=1$, $v'_{k,n+1}$ not all equal. A satisfactory bracket would be $T=T((m-Mn)e_{n+1}, Mn)$. The inequality of the evaluations, however, permits us to reduce T in size. If $v'_{k,n+1}-m>0$ for some k, then the (now raised) kth facet of the removal cone $v'_k-\nabla$ intersects the top of T at a distance $(1/Mn)(v'_{k,n+1}-m)$ from the kth facet of T in the direction of u_k. Since all points inside the removal

cone can be ignored, this means we can advance the centroid of the initial simplex top a distance $(n/(n+1))(1/Mn)(v'_{k,n+1}-m)$ towards u_k. At the same time the apex of the initial simplex is lifted a distance $(1/(n+1))(v'_{k,n+1}-m)$. Putting these movements together for all vertices v_k gives that

$$x_0 = \frac{1}{M(n+1)} \sum_{k=1}^{n+1} (v'_{k,n+1}-m)u_k$$

$$+\left[(m-Mn)e_{n+1}+\frac{1}{n+1}\sum_{k=1}^{n+1}(v'_{k,n+1}-m)e_{n+1}\right]$$

$$=\frac{1}{M(n+1)}\sum_{k=1}^{n+1}(v'_{k,n+1}-m)u_k+\left(\frac{1}{n+1}\sum_{k=1}^{n+1}v'_{k,n+1}-Mn\right)e_{n+1},$$

$$h_0 = Mn-\frac{1}{n+1}\sum_{k=1}^{n+1}(v'_{k,n+1}-m),$$

while $a_0=-u_l+me_{n+1}$ remains an admissible point in the modified initial simplex top.

We now show that this modified initial simplex T_0 still brackets $\mathscr{A}$ over H. The hyperplane F_k in which the kth facet of $v'_k-\nabla$ lies is now a distance $v'_{k,n+1}-m$ above the hyperplane containing the kth facet of T. For this reason $(v'_{k,n+1}-m)e_{n+1}$, when added to $(m-Mn)e_{n+1}$, the apex of T, is strictly inside cone (T), the cone generated by T. Thus

$$\text{cone}(k\text{th facet of } v'_k-\nabla)\cap\text{cone}(T\cap F_k),$$

(an intersection of dually oriented, simplicially-based n-dimensional cones, with apexes at a greater distance than those considered in Case 1), when projected onto $\mathbb{R}^n$, will cover more than H_k. Thus our modified initial simplex T_0 brackets $\mathscr{A}$ over a larger region than does T itself.

Case 4. The general case. The general domain is affinely equivalent to Case 3. We can repeat the previous argument, bearing in mind the translation and scaling factors, to complete the proof. □

We conclude this section by demonstrating the tiling property of the standard domains.

Remark 6.1. By considering an appropriate packing of translates of the $(n+1)$-cube introduced at the beginning of this section, and the associated projections onto $\mathbb{R}^n$, it is possible to give a first principles proof that for each n a standard domain H tiles $\mathbb{R}^n$. Viewing a hexagonal tiling of $\mathbb{R}^2$ as a projection of a 3-cube tiling of $\mathbb{R}^3$ suggests the essential ideas. For pointing out this first principles approach, and many other helpful comments, the author is grateful to Professor Carl Lee.

A very brief proof is available, however, using the zonotope theory of McMullen [4]. An n-zonotope Z is a vector sum of $s>n$ line segments, with affine span of

dimension n. Theorem 2 of [4] states that Z tiles $\mathbb{R}^n$ if and only if its *associated* zonotope $\bar{Z}$ tiles $\mathbb{R}^{s-n}$. Clearly our standard domains H are such zonotopes, with $s = n+1$. Since the linear dependences of H form a one-dimensional subspace of $\mathbb{R}^{n+1}$ it is readily checked that the associated zonotope $\bar{H}$ is a line segment in $\mathbb{R}^1$, which certainly tiles the line. Thus H tiles $\mathbb{R}^n$.

7. Examples

A Fortran 77 code for the algorithm has been developed. A system is stored as a linked list of simplexes, with each simplex held as a centroid-radius pair, (c, r). Here we present two illustrative examples.

Example 1. This example serves to illustrate the mechanism of the algorithm. Consider the function

$$f(x, y) = \max(\sqrt{3}x + y, -2y, y - \sqrt{3}x),$$

whose epigraph is the standard cone we have used in $\mathbb{R}^3$. The global minimum, of 0, is realised at the origin.

(i) Suppose that we use a Lipschitz constant of $M = 1$, and a domain of $H = c + rU$, where $c = (0, 0.5, 0)$ and $r = 1$. Here $u_1 = (0, 1, 0)$, $u_2 = (\frac{1}{2}\sqrt{3}, -\frac{1}{2}, 0)$ and $u_3 = (-\frac{1}{2}\sqrt{3}, -\frac{1}{2}, 0)$. Then the global minimum is reached in the first iteration, as summed up in Table 2.

Table 2
The initial system and first iteration for Example 1, with Lipschitz constant $M = 1$

Iteration	System	
i	$\mathcal{T}_i$	a_i
0	$\{T[(0, 0, 0), 1]\}$	$(0, -0.5, 1)$
1	$\{T[(0, 0, 0), 0]\}$	$(0, 0, 0)$

Figure 5 shows H, and the top of the initial simplex, T_0. The location of T_0 is attributable to the variation in the heights, shown in brackets, at the dual vertices. Since the centroid of the top of the initial simplex is $(0, 0, 0)$, the algorithm finds the global minimum at the first iteration.

(ii) When $M = 2$, a less favourable value, the initial development is shown in Table 3.

The variation of the system becomes less than 10^{-3} after 14 iterations, when the current admissible point is $(0, 0, 0)$ to 3 decimal places. The maximum number of simplexes in any system to this stage is 54.

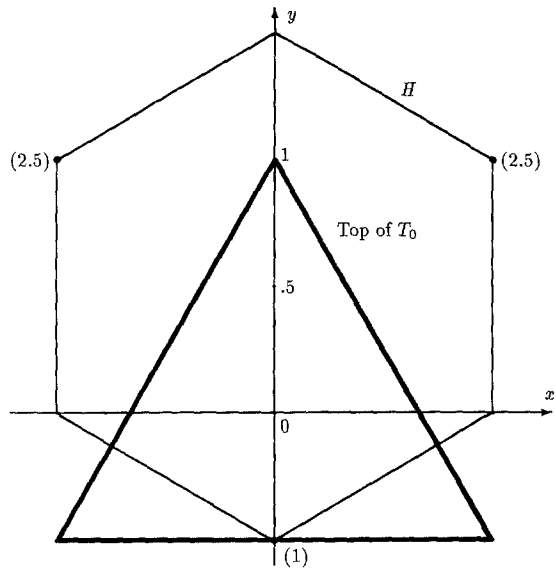


Fig. 5. The hexagonal domain, H, and the top of the initial simplex, T_0, for Example 1, with $M = 1$.

Table 3
The initial system and first iteration for Example 1, with $M = 2$

i	$\mathcal{T}_i$	a_i
0	$\{T[(0, 0.25, -2), 3]\}$	$(0, -0.5, 1)$
1	$\left\{\begin{array}{l} T[(0, 0.625, -1.25), 1.5], \\ T[(0.32476, 0.0625, -1.25), 1.5], \\ T[(-0.32476, 0.0625, -1.25), 1.5] \end{array}\right\}$	$(0, 0.25, 0.25)$
⋮	⋮	⋮

Example 2. Our second example illustrates the main virtue of the algorithm, the ability to home in on the global minimum, by-passing local minima on the way. Consider the function

$$f(x, y) = -e^{-x^2} \sin x + |y|,$$

which has a global minimum of -0.396653 when $x = 0.653273$ and $y = 0$. It has an infinite number of local minima along the x-axis, for example, when $x = 6.361622$. We begin with a hexagonal domain centered on $(10, 10, 0)$ and with radius 20. We use $M = 1$.

Table 4 sums up the progress of the algorithm as it finds the global minimum, using a stopping value of $\varepsilon = 10^{-3}$.

Note that the variation is usually considerably less than two-thirds, $n/(n+1)$, that of the previous variation. The more sharply defined the global minimum, the further the deviation from equality. Also note how the number of simplexes in the system initially builds up to 384, then drops to 63. This drop corresponds to the

Table 4

The course of the iterations for Example 2. Following the reduction step, the number of simplexes in the system is R, while the number in the system after elimination is E

i	R	E	$V(\mathcal{S}_i)$	a_i
0	1	1	33.3	(10.000, −10.000, 10.000)
1	3	3	20.0	(10.000, 6.667, 6.667)
2	9	9	9.892	(1.340, 1.667, 1.505)
3	18	18	6.129	(5.623, −0.806, 0.806)
4	48	48	3.602	(5.623, −0.806, 0.806)
5	132	108	1.959	(25.637, 0.185, 0.185)
6	192	141	0.961	(0.648, −0.238, −0.158)
7	333	264	0.504	(0.839, 0.074, −0.294)
8	384	39	0.257	(0.649, −0.036, −0.361)
9	63	60	0.161	(0.770, −0.009, −0.376)
10	75	69	0.095	(0.756, 0.009, −0.379)
11	105	105	0.058	(0.601, 0.010, −0.384)
12	159	132	0.030	(0.669, 0.000, −0.396)
13	162	162	0.017	(0.669, 0.000, −0.396)
14	285	285	0.010	(0.669, 0.000, −0.396)
15	369	369	0.007	(0.669, 0.000, −0.396)
16	531	456	0.004	(0.650, 0.000, −0.397)
17	663	663	0.002	(0.650, 0.000, −0.397)
18	1029	924	0.001	(0.653, 0.000, −0.397)
19	1302	1287	0.000	(0.651, 0.000, −0.397)

sudden elimination of simplexes around all local, and non-global, minima along the x-axis. The simplex build up at the end heralds the point at which a local algorithm needs to be called.

8. Discussion

The contribution of this paper has been to examine the usual bisection problem, to formulate it in a way which has a clear analogue in higher dimensions, and to develop an algorithm to solve the higher dimensional problem. Since the algorithm can be used to find global minima of Lipschitz continuous functions, there is related work already present in the literature, although it is not presented in the context established here.

Independently and some fifteen years ago, Piyavskii [6] and Shubert [8] discovered a variant of the algorithm for the case where $n = 1$. Their exposition is in terms of a piecewise-linear lower envelope, rather than a bracket comprising a union of similar triangles, with each iteration consisting of a *single* evaluation of $\mathcal{A}$, over the lowest simplex apex. This has the virtue of producing an algorithm optimal in a minimax sense: the evaluations occur in such a way that at any stage the maximum difference between the lowest evaluation and a potential global minimum is least.

In assessing the potential global minima, $\mathcal{A}$ is allowed to range over all epigraphs of Lipschitz continuous functions with Lipschitz constant M, the admissible regions of this paper in the case $n = 1$. As pointed out by Basso [1], however, the sequence of localisations (that is, the sequence of subsets of the domain known to contain the points at which the global minimum is attained) need not converge to this set of global minimisers, the *solution set*. Basso proposes a modification to ensure convergent localisations: in each iteration there must be an evaluation in each disjoint interval in the localisation.

A generalisation of the Piyavskii-Shubert algorithm has already been presented by Mladineo in [5]. Her paper deals with epigraphs of Lipschitz continuous functions of many variables, so ordering is via the cone with spherical base, rather than the cone with simplex base used here. In this paper the algorithm was presented in an extreme form, with *all* simplexes reduced at each iteration, to highlight the simplicity of the underlying ideas, and to emphasise the intimate connection with the bisection method. Piyavskii-Shubert and Basso type variants will also be developed. The Piyavskii-Shubert variant, in which each iteration comprises a lowest evaluation, can be considered to be a discrete analogue of the Mladineo algorithm.

The Mladineo algorithm and the Piyavskii-Shubert variant of the bisection method share the advantage of being deterministic: they have assured convergence to all global minima. In another sense they are complementary. The "continuous" nature of the Mladineo algorithm ensures it has the virtue of minimax optimality with respect to epigraphs of Lipschitz continuous functions, but it also means that much effort must be expended to find the next evaluation point. The Piyavskii-Shubert variant of the bisection method is no longer minimax optimal in this sense, but does have the virtue of giving rise to a simple data structure in which the next evaluation point is readily found, being over the centroid of the largest simplex.

A comment on the unlikely origins of this paper: it was in the context of a problem in probability theory, that of computing the dispersion function of a set of random variables, that the author discovered a variant of the Piyavskii-Shubert algorithm [9] which led to the algorithm presented here.

This section would be incomplete without mention of an alternative generalisation of the bisection method, presented in [3]. This generalisation was motivated by the particular bisection problem of root-finding, rather than the problem common to all applications of the bisection method described here in Section 1. By means of a sequence of simplex halvings this algorithm locates a root of a system of non-linear equations. The algorithm does not address the general higher dimensional bisection problem formulated here in Section 3.

References

[1] Patricio Basso, "Iterative methods for the localization of the global maximum," *SIAM Journal on Numerical Analysis* 19 (1982) 781-792.

[2] Gustave Choquet, *Lectures on Analysis, Vol. 1* (Benjamin, New York, 1969).
[3] A. Eiger, K. Sikorski and F. Stenger, "A bisection method for systems of nonlinear equations," *ACM Transactions on Mathematical Software* 10 (1984) 367-377.
[4] P. McMullen, "Space tiling zonotopes," *Mathematika* 22 (1975) 202-211.
[5] Regina Hunter Mladineo, "An algorithm for finding the global maximum of a multimodal, multivariate function," *Mathematical Programming* 34 (1986) 188-200.
[6] S.A. Piyavskii, "An algorithm for finding the absolute extremum of a function," *USSR Computational Mathematics and Mathematical Physics* 12 (1972) 57-67.
[7] J.E. Dennis Jr. and Robert B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (Prentice-Hall, Englewood Cliffs, NJ, 1983).
[8] Bruno O. Shubert, "A sequential method seeking the global maximum of a function," *SIAM Journal on Numerical Analysis* 9 (1972) 379-388.
[9] G.R. Wood, "On computing the dispersion function," *Journal of Optimization Theory and Applications* 58 (1988) 331-350.
[10] G.R. Wood, "Multidimensional bisection applied to global optimisation," *Computers and Mathematics with Applications* 21 (1991) 161-172.