On Solving a D.C. Programming Problem by a Sequence of Linear Programs

R. HORST[1], T.Q. PHONG[2], Ng. V. THOAI[3], and J. de VRIES[1]
[1]*Fachbereich IV – Mathematik, Universität Trier, Trier, West Germany*; [2]*Institute of Technology, Danang, Vietnam*; [3]*Institute of Mathematics, Hanoi, Vietnam*

(Received: 28 May 1990; accepted: 4 December 1990)

Abstract. We are dealing with a numerical method for solving the problem of minimizing a difference of two convex functions (a d.c. function) over a closed convex set in $\mathbb{R}^n$. This algorithm combines a new prismatic branch and bound technique with polyhedral outer approximation in such a way that only linear programming problems have to be solved.

Key words. Nonlinear programming, global optimization, d.c. programming, branch and bound, outer approximation, prismatic partition

1. Introduction

In this paper we consider the multiextremal global optimization problem

$$\begin{aligned} &\text{glob min}(f(x) - g(x)) \\ &\text{s.t. } h_j(x) \leq 0 \quad (j = 1, \ldots, J) \end{aligned} \tag{P}$$

where f, g, h_j $(j = 1, \ldots, J)$ are finite convex functions on $\mathbb{R}^n$. Problem (P) is frequently called a d.c. optimization problem, where d.c. is an abbreviation for the difference of two convex functions. The formulation of (P) indicates that we are interested in finding a global minimum of the objective function $(f(x) - g(x))$ over the feasible set

$$D := \{X \in \mathbb{R}^n : h_j(x) \leq 0 \quad (j = 1, \ldots, J)\} .$$

We assume that D is compact (which implies of course that a global solution of (P) exists whenever D is nonempty), and that a feasible point is known in advance. Problem (P) is of interest from a practical as well as from a theoretical viewpoint. From the theoretical viewpoint we should like to mention that the class of d.c. functions, i.e., functions that can be represented as difference of two convex functions, enjoys a remarkable stability with respect to operations frequently encountered in optimization. For example, the class of d.c. functions is closed under operations such as sum, multiplication, multiplication with a scalar, forming maximum and minimum of a finite number of functions, etc. Moreover, we known that every locally d.c. function, i.e., every function that is d.c. in a

Parts of this research were accomplished while the third author was visiting the University of Trier, Germany, as a fellow of the Alexander von Humboldt foundation.

neighbourhood of a point, is also d.c. in the whole space. From this it can easily be deduced, for example, that every C^2-function is d.c., and also that continuous piecewise linear functions are d.c. Since, clearly, polynomials are C^2-functions, we see from the well-known approximation theorem of Weierstraß that, in principle, every global optimization problem of minimizing a continuous function over a compact convex set can be approximated (with respect to the sup-norm) as closely as desired by a sequence of d.c. problems. Surveys of d.c. functions are included, e.g., in Hiriart-Urruty (1985), Horst and Tuy (1990).

Of course, the main concern when using these properties is how to construct a d.c. representation of a function which is known to be a d.c. function but not given in d.c. form. Although this problem of finding appropriate d.c. representations is not yet solved for broad classes of d.c. functions, it should be noted that in many applications a d.c. representation of the functions involved is either given or can easily be found. For example, in many econometric applications we encounter a situation where the objective function is of the form $f(y) - g(z)$, f and g convex, which reflects, for example, the fact that in some activities the unit cost increases when the scale of activity (y) is enlarged (diseconomies of scale), whereas in other activities the unit cost decreases when the scale of activity (z) is enlarged (economies of scale). In other applications, the objective function is supposed to be differentiable and separable, i.e. the sum of function of one real variable; and each of these univariate functions is S-shaped, i.e., it is concave (or convex) in the interval $[a, \tilde{x}]$ and convex (or concave) in the interval $[\tilde{x}, b]$, where $[a, b]$ is the interval of interest for the given function and $\tilde{x} \in (a, b)$. It has been shown in Tuy (1987a) that such a univariate S-shaped function can easily be represented in d.c. form (cf. also Horst (1990), Horst and Tuy (1990)). Another example is the optimization of indefinite quadratic forms x^TAx ($A \in \mathbb{R}^{n\times n}$, symmetric) for which it is well known that there are several ways of finding positive semidefinite matrices F, G such that $x^TAx = x^TFx - x^TGx$ (cf., e.g., Pardalos *et al.* (1987), Pardalos and Rosen (1987)).

D.C. programming problems arise frequently also in engineering and physics (e.g., Giannessi *et al.* (1979), Heron and Sermange (1982), Horst and Tuy (1990), Nguyen and Strodiot (1988), Polak and Vincentelli (1979), Toland (1979), Tuy (1986), (1987a), Vidigal and Director (1982)). Particularly in engineering design we encounter optimization problems with lower C^2-objective functions which have been shown to be d.c. (e.g., Hiriart-Urruty (1985), Horst and Tuy (1990), Polak (1987), Thach (1988)). Another example of mainly theoretical interest that, however, shows how large the class of d.c. problems is, is provided by the fact that the square $d_M^2(x)$ of the distance to any closed set $M \subset \mathbb{R}^n$ is d.c. (see Hiriart-Urruty (1985) and references therein). For a survey of d.c. problems, see Horst and Tuy (1990).

Recent advances in deterministic global optimization (see Horst (1990), Horst and Tuy (1990)) have led to several algorithms for solving problem (P). Note first that, by introducing an additional variable t, problem (P) can be transformed into the equivalent concave minimization problem

$$\begin{aligned} &\text{glob}\min(t - g(x)) \\ &\text{s.t. } f(x) \leq t \\ &\quad h_j(x) \leq 0 \quad (j = 1, \ldots, J). \end{aligned} \tag{CP}$$

Let

$$\tilde{D} := \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq t, x \in D\}$$

denote the feasible set of (CP).

Likewise, problem (P) can be converted into an equivalent convex minimization problem subject to an additional reverse convex constraint. This problem is of the form

$$\begin{aligned} &\text{glob}\min(f(x) - t) \\ &\text{s.t. } g(x) \geq t \\ &\quad h_j(x) \leq 0 \quad (j = 1, \ldots, J). \end{aligned} \tag{RC}$$

A third transformation of (P) which involves two additional variables yields the so called "canonical d.c. program"

$$\begin{aligned} &\text{glob}\min c^T z \\ &\text{s.t. } \tilde{h}_1(z) \leq 0, \\ &\quad \tilde{h}_2(z) \geq 0, \end{aligned} \tag{CDC}$$

where $c \in \mathbb{R}^{n+2}$; $\tilde{h}_1, \tilde{h}_2 : \mathbb{R}^{n+2} \to \mathbb{R}$ convex (cf. Horst and Tuy (1990), Tuy (1986)). The known algorithms for solving (CP), (RC) and (CDC), respectively are mainly either branch and bound methods or outer approximation procedures (cutting plane algorithms). Very recently also combinations of branch and bound with outer approximation have been proposed to solve the concave minimization problem (CP) and the canonical d.c. program (CDC).

Branch and bound methods to solve (CP) are proposed in Horst (1976, 1980, 1986), Tuy *et al.* (1985). Generalizations of some of these approaches to reverse convex problems and canonical d.c. programs are given in Horst (1988). A general theory of branch and bound methods in global optimization can be found in Horst (1986, 1988, 1989), Horst and Tuy (1987), Tuy and Horst (1988), cf. also Horst and Tuy (1990). Outer approximation methods for (CP) are developed in Hoffman (1981), Horst, Thoai and Tuy (1987, 1989), Thieu *et al.* (1983), Tuy (1983), cf. also Horst and Tuy (1990). Problems (RC) and (CDC) are treated by outer approximation techniques in Horst and Tuy (1990), Thoai (1988), Tuy (1986), Tuy (1987), Tuy (1987a), Tuy and Thuong (1988). A general theory of outer approximation in global optimization is presented in Horst, Thoai and Tuy (1987, 1989). A detailed discussion of both types of approaches that includes many additional material on specially structured subclasses and on applications can be found in the monograph Horst and Tuy (1990).

All of these pure branch and bound or pure outer approximation algorithms are numerically expensive. Outer approximation methods require, in each iteration, the determination of all new vertices of a polytope P generated from a polytope P by an intersection of P with a closed halfspace (a cut). Although the originally proposed methods of Hoffman (1981), Thieu *et al.* (1983) for finding all of these new vertices have been improved recently by Horst *et al.* (1987), this part of the outer approximation methods remains computationally very costly (cf. Horst and Tuy (1989), (1990)).

Branch and bound methods use conical or simplicial partitions unless additional structure can be exploited such as, for example, separability of the objective function which suggests to consider rectangular partitions (e.g., Horst and Tuy (1990), Pardalos *et al.* (1987), Pardalos and Rosen (1987)). The computationally most expensive part in these methods usually consists in the calculation of lower bounds. For example, the computation of the lower bounds in the conical branch and bound algorithm of Tuy *et al.* (1985) and is variants of Horst (1986) (see also Horst and Thoai (1989)) requires in each step the calculation of the vertices of a polytope which is the intersection of a polyhedral cone with $k \cdot (n+1)$ cutting halfspaces, where n is the dimension of the feasible set and $k \in \mathbb{N}$ is a number to be specified in the algorithm. The simplicial procedure for solving (CP) proposed in Horst (1976, 1981) mainly requires in each step the solution of two nonlinear convex minimization problems. Extensions of this approach to d.c. programming problems face the additional problem of detecting infeasible partition sets, and, whenever the chosen bounding operation is computationally cheap in these extensions, then – as a rule – the bounds are not very efficient (e.g., Horst and Tuy (1990) and references therein).

A first successful attempt to overcome the difficulties mentioned above is the cone splitting/outer approximation algorithm of Horst, Thoai and Benson (1991) for solving the concave minimization problem (for a simplicial variant, see Benson and Horst (1991); for a specialization to the case when the objective is separable, see Benson (1991)). This algorithm that, in the sequel, will be called H–T–B algorithm combines branch and bound elements with an outer approximation procedure in such a way that only linear programs and one-dimensional convex minimization problems (line searches) have to be solved in each iteration. First numerical experiments indicate that this approach is considerably more efficient than previous ones (cf. Horst, Thoai and Benson (1991)). A generalization of this concave minimization approach to the canonical d.c. program is proposed in Horst, Phong and Thaoi (1991). The H–T–B algorithm, however, does not seem to be appropriate for solving problem (CP). This algorithm requires knowledge of an interior point y^0 (hence full dimensionality) and boundedness of the feasible set, and it operates with partitions consisting of cones emanating from y^0. The feasible set of (CP), however, which is the epigraph of the convex function $f(x)$ over D, is not bounded. Even if one would find it worthwhile to compactify this set, for example, at the expense of carrying out a first iteration of a branch and bound procedure for solving max $f(D)$ which yields an upper bound of f on D, the specific shape of the resulting truncated epigraph would not suggest using the conical partitions of the H–T–B approach.

Likewise, a transformation of (P) into the form (CDC) and application of the procedure in Horst, Phong and Thoai (1991) does not seem to be appropriate. One reason is the increase in the number of variables by two; a second one is that the approach of Horst, Phong and Thoai (1991) seems to be not much better than pure branch and bound procedures because the efficient concave minimization bounding operation used in Horst, Thaoi and Benson (1991) cannot be carried over to problem (CDC). Finally, it should be noted that all *conical* branch and bound/outer approximation algorithms suffer from the drawback that the LP-subproblems become more and more ill-conditioned as the cones shrink to a ray. Therefore, implementations of these procedures delete "thin" cones which amounts to sacrificing accuracy.

It is the purpose of the present paper to propose a new branch and bound/outer approximation algorithm for solving the d.c. problem (P). This approach, though designed in a similar spirit as the H–T–B algorithm, has several advantages over an application of the latter to (CP) resp. of the algorithm of Horst, Phong and Thoai (1991) to (CDC). For example, in contrast to the H–T–B approach which requires full dimensionality of D, no requirement on the dimension of D is needed. Moreover, it will not be necessary to compactify the feasible set of (CP).

The main advantages, however, regard numerical efficiency: only linear subprograms and no univariate convex minimization problems have to be solved. Furthermore, a mechanism for the deletion of infeasible pattern sets is applied that is more efficient then the deletion-by-infeasibility rules proposed previously (cf. Horst (1988, 1989), Horst and Tuy (1990)). Finally, the algorithm uses prismatic partitions which fit much better to the shape of the feasible set $\tilde{D}$ of (CP) than conical partitions.

The paper is organized as follows. An outline of the method is presented in the next section. Section 3 deals with basic operations of the algorithm, and Section 4 contains its detailed description and proves convergence properties. Some illustrative numerical examples are given in the final section.

2. Outline of the Method

We consider problem (CP), and we set

$$h(x) = \max\{h_j(x) : j = 1, \dots, J\}$$

which is clearly convex in $\mathbb{R}^n$.

A brief outline of the algorithm is as follows.

Given an n-simplex S which contains the feasible set $D \subset \mathbb{R}^n$ of (P), a *prism* $T = T(S) \subset \mathbb{R}^{n+1}$ is defined by

$$T = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : x \in S\} \tag{1}$$

The prism T has $n + 1$ edges that are vertical lines (i.e., lines parallel to the t-axis) which pass through the $n + 1$ vertices of S, respectively). Several ways to construct a simplex $S \supset D$ are described in the next section.

For the bounding operation on S (resp. T), we consider a polyhedral convex set

$$P = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : x \in S, t \geq \tilde{t}\}\ , \tag{2}$$

where $\tilde{t}$ is a real number satisfying

$$\tilde{t} \leq \min\{f(x) : x \in D\}$$

(determining $\tilde{t}$ amounts to solving a convex minimization problem which can be done by any standard nonlinear programming technique).

From the above we see that $T \supset P \supset \tilde{D}$, where $\tilde{D} = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : x \in D, f(x) - t \leq 0\}$ is the feasible set of problem (CP). A lower bound β of the function $t - g(x)$ over $\tilde{D}$ is determined by solving a certain linear program which will be introduced below. Let $(\bar{v}, \bar{t})$ be a point in P satisfying.

$$\bar{v} \in S,\ \bar{t} - g(\bar{v}) \leq t - g(x)\ \forall (x, t) \in \tilde{D}\ .$$

If $(\bar{v}, \bar{t})$ is in $\tilde{D}$, then $(\bar{v}, \bar{t})$ is of course an optimal solution of (CP). Otherwise, we cut off a part of the set $P \setminus \tilde{D}$ and (or) construct a (more and more refined) partition of T.

Suppose that a simplicial partition of S is at hand, i.e., we have

$$S = \bigcup_{i=1}^{p} S_i\ , \tag{3}$$

where $p \geq 2$; the sets S_i are n-simplices ($i = 1, \ldots, p$), and each pair of simplices S_i, S_j ($i \neq j$) intersects at most in common boundary points (cf. Section 3 for various ways to construct such a partition). Then

$$T = \bigcup_{i=1}^{p} T_i\ , \tag{4}$$

where

$$T_i = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : x \in S_i\} \tag{5}$$

is a natural prismatic partition of T induced by the simplicial partition (3). For each prism T_i of the partition (4), a lower bound $\beta(T_i)$ for $t - g(x)$ over $T_i \cap P$ is calculated, and

$$\beta_1 = \min\{\beta(T_i) : i = 1, \ldots, p\}$$

is a lower bound of the objective function in (CP).

If throughout the further subdivision of prisms and the bounding operation some feasible points in $\tilde{D}$ are found, then let (x^k, t_k) be the best feasible point obtained so far, i.e., the feasible point with smallest objective function value, and let $\alpha_k = t_k - g(x^k)$. If $\alpha_k - \beta_k = 0$, then (x^k, t_k) is an optimal solution of (CP). Otherwise, we delete prisms that do not contain any feasible solution which is better than (x^k, t_k), and we choose one of the remaining prisms associated with the smallest lower bound β_k for further subdivision and go to the next iteration of the branch and bound procedure.

Clearly, the previous procedure involves four basic operations:

(a) the subdivision process: in each iteration some prism is divided into a finite number of subprisms;

(b) the bound estimation: for each prism generated throughout the algorithm, a lower bound for the objective function $t - g(x)$ over the part of the feasible set contained in this prism is computed;

(c) the construction of cutting planes: throughout the algorithm a sequence of polyhedral convex sets $P_0, P_1, \dots$ is constructed such that $P_0 \supset P_j \supset \cdots \supset \tilde{D}$. Each set P_j is generated by using a cutting plane to cut off a part of P_{j-1};

(d) the deletion rule: at each iteration we have to delete certain prisms that do not contain any feasible solution which is better than the one obtained so far.

In the next section we shall describe the basic operations in detail.

3. Basic Operations

3.1. CONSTRUCTION OF A FIRST PRISM

Throughout this paper by a prism T we always mean a polyhedral prism generated by an n-simplex $S \subset \mathbb{R}^n$, i.e.,

$$T = T(S) = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : x \in S\} . \tag{6}$$

For the initial simplex $S \supset D$ (which yields the initial prism $T \supset \tilde{D}$), we consider two often occurring cases.

CASE 1. Lower bounds $\underline{e}_j$ of the variables x_j are given, i.e.,

$$\underline{e}_j \leq x_j \ (j = 1, \dots, n) , \tag{7}$$

where $\underline{e}_j \in \mathbb{R}$ $(j = 1, \dots, n)$,

A simplex $S \supset D$ can then be defined by

$$S = \left\{ x \in \mathbb{R}^n : \underline{e}_j \leq x_j,\ j = 1, \dots, n,\ \sum_{j=1}^{n} x_j \leq \gamma \right\} , \tag{8}$$

where

$$\gamma = \max\left\{ \sum_{j=1}^{n} x_j : x \in D \right\} .$$

The $n + 1$ vertices $u_1, u_2, \dots, u_{n+1}$ of S are

$$u^1 = (\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n)$$
$$u^j = (\underline{e}_1, \dots, \underline{e}_{j-1}, \gamma - \sum_{i \neq j} \underline{e}_i, \underline{e}_{j+1}, \dots, \underline{e}_n)(j = 2, \dots, n + 1) .$$

If the bounds $\underline{e}_j$, are not explicitely given, then they can be determined by solving n convex minimization problems (with linear objective function) of the form $\underline{e}_j = \min\{x_j : x \in D\}$.

A similar construction can be applied when upper bounds $\bar{e}_j \geq x_j$ are known.

CASE 2. The set D is a polytope, i.e.,

$$D = \{x \in \mathbb{R}^n : d_i^T x + \delta_i \leq 0 \quad (i = 1, \dots, m)\}, \tag{9}$$

where $d_i \in \mathbb{R}^n$, $\delta_i \in \mathbb{R}$ $(i = 1, \dots, m)$.

Let y be a vertex of D and let $I(y) = \{i \in \{1, \dots, m\};\ d_i^T y + \delta_i = 0\}$. If the vertex y is nondegenerate, then the set $I^0(y) := I(y)$ consists of exactly n indices. In the degenerate case $(|I(y)| > n)$, one can always find a subset $I^0(y) \subset I(y)$ such that the vectors d_i $(i \in I^0(y))$ are linearly independent. Then the initial simplex S is defined by

$$S = \{x \in \mathbb{R}^n : d_i^T x + \delta_i \leq 0 (i \in I^0(y)),\ a^T x + \gamma \leq 0\}, \tag{10}$$

where

$$a = - \sum_{i \in I^0(y)} d_i,\ \gamma = -\max\{a^T x : x \in D\}$$

(for details, see Horst and Tuy (1990); for a related approach see Falk and Hoffman (1976)).

The $n + 1$ vertices of S are y and the n points where the hyperplane $\{x \in \mathbb{R}^n : a^T x + \gamma = 0\}$ intersects the edges of the cone $\{x \in \mathbb{R}^n :\ d_i^T x + \delta_i \leq 0\ (i \in I^0(y))\}$.

3.2. SUBDIVISION OF PRISMS

Let $T(U)$ be a prism generated by a simplex $U = [v^1, \dots, v^{n+1}] := \operatorname{conv}\{v^1, \dots, v^{n+1}\}$ which is defined as convex hull of its vertices $v^1, \dots,\ v^{n+1}$. Then every $r \in U$ can be represented as

$$r = \sum_{i=1}^{n+1} \lambda_i v^i,\ \sum_{i=1}^{n+1} \lambda_i = 1,\ \lambda_i \geq 0\ (i = 1, \dots, n+1).$$

Suppose that $r \neq v^i$ $(i = 1, \dots, n + 1)$. For each i satisfying $\lambda_i > 0$ let U_i be the subsimplex of U defined by

$$U_i = [v_1, \dots, v^{i-1}, r, v^{i+1}, \dots, v^{n+1}]. \tag{11}$$

Then the collection $\{U_i : \lambda_i > 0\}$ defines a partition of U, i.e., we have

$$\bigcup_{\lambda_i > 0} U_i = U,\ \operatorname{int} U_i \cap \operatorname{int} U_j = \emptyset \text{ for } i \neq j \tag{12}$$

(cf. Horst and Tuy (1990) and references therein). In a natural way the prisms $T(U_i)$ generated by the simplices U_i defined in (11) form a partition of $T(U)$ (which is defined in a similar way to (12)). In order to ensure convergence of the algorithm developed below we introduce the concept of an exhaustive subdivision process of prisms which is derived from that of simplices in a straightforward way.

DEFINITION 1. A subdivision process of prisms is called *exhaustive* if for every nested (decreasing) sequence of prisms $\{T_q\}$ generated by the subdivision process we have

$$\bigcap_{q=0}^{\infty} T_q = \tau ,$$

where τ is a line perpendicular to $\mathbb{R}^n$ (a vertical line).

Obviously, every exhaustive subdivision of simplices as defined in Horst (1986, 1988), Horst and Tuy (1987, 1990) induces an exhaustive subdivision of the corresponding prism.

A classical exhaustive subdivision process of simplices is *bisection* (cf., e.g., Horst (1976), Horst and Tuy (1990)), in which each simplex $[v^0, v^1, \ldots, v^n]$ is divided into two subsimplices by choosing in (11)

$$r = \frac{1}{2}(v^{i_1} + v^{i_2}),$$

where

$$\|v^{i_1} - v^{i_2}\| = \max\{\|v^i - v^j\| : i, j \in \{0, \ldots, n\}, i \neq j\}$$

(here $\|\cdot\|$ denotes the Euclidean norm).

Other exhaustive subdivision procedures of simplices have been investigated in Horst and Tuy (1990), Tuy (1991).

3.3. LOWER BOUNDS

Let U be an n-simplex which is derived from the initial simplex S by a subdivision process as discussed in the preceeding section, and let $T = T(U)$ be the prism generated by U. Moreover, let α denote an upper bound of $\inf\{t - g(x) : (x, t) \in \tilde{D}\}$, where $\tilde{D} = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : x \in D, f(x) \leq t\}$ is the feasible set of problem (CP). In the algorithm below, α will be the smallest value of $t - g(x)$ attained at a feasible point known so far; we assume that an initial feasible point is at hand so that we always have $\alpha < \infty$.

Furthermore, let P be a polyhedral convex set which contains $\tilde{D}$. Suppose that P is described by

$$P = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : Ax + at \leq b\} , \tag{13}$$

where A is a real $(m \times n)$-matrix and $a, b \in \mathbb{R}^m$.

We proceed to compute a lower bound $\beta(T, P, \alpha)$ of $t - g(x)$ over $T \cap \tilde{D}$ in the following way.

Let v^i $(i = 1, \ldots, n + 1)$ denote the vertices of U and define

$$I(U) = \{i \in \{1, \ldots, n + 1\} : v^i \in D\}$$

and

$$\mu = \begin{cases} \min\{\alpha, \min\{f(v^i) - g(v^i) : i \in I(U)\} , & \text{if } I(U) \neq \emptyset \\ \alpha , & \text{if } I(U) = \emptyset \end{cases} \tag{14}$$

For each $i = 1, \ldots, n+1$ consider the point (v^i, t_i) where the edge of T which passes through v^i intersects the level set $\{(x, t): t - g(x) = \mu\}$, i.e.

$$t_i = g(v^i) + \mu \quad (i = 1, \ldots, n+1).$$

Let

$$H = \{(x, t) \in \mathbb{R}^n \times \mathbb{R}: p^T x - t = \gamma\} \tag{15}$$

denote the uniquely defined hyperplane through the points (v^i, t_i), where $p \in \mathbb{R}^n$ and $\gamma \in \mathbb{R}$. Consider the two closed halfspaces

$$H_+ = \{(x, t) \in \mathbb{R}^n \times \mathbb{R}: p^T x - t \leq \gamma\} \tag{16}$$

and

$$H_- = \{(x, t) \in \mathbb{R}^n \times \mathbb{R}: p^T x - t \geq \gamma\} \tag{17}$$

generated by H. (We say H_+ (resp. H_-) is the upper (resp. lower) halfspace of H.) If $T \cap \tilde{D} \subset H_+$, then we see from the concavity of $t - g(x)$ that

$$\min\{t - g(x): (x, t) \in T \cap \tilde{D}\} \geq \min\{t - g(x): (x, t) \in T \cap H_+\} =$$
$$\min\{t - g(x): (x, t) \in \{(v^1, t_1), \ldots, (v^{n+1}, t_{n+1})\}\} = \mu .$$

If $T \cap \tilde{D} \not\subset H_+$, then we shift the hyperplane H (downward with respect to t) until it supports the set $T \cap P \cap H_-$ at a point $z = z(T)$ (z is a point in $T \cap P \cap H_-$ with the greatest distance to H). Let $\bar{H}$ denote the resulting supporting hyperplane, i.e., the hyperplane parallel to H satisfying $z \in \bar{H}$, and let $\bar{H}_+$ denote the closed upper halfspace generated by $\bar{H}$. Furthermore, for each $i = 1, \ldots, n+1$ let $z^i = (v^i, \bar{t}_i)$ be the point where the edge of T which passes through v^i intersects $\bar{H}$. Then it follows from our construction that

$$T \cap \tilde{D} \subset T \cap P \subset T \cap \bar{H}_+ ;$$

and hence

$$\min\{t - g(x): (x, t) \in T \cap \tilde{D}\} \geq \min\{t - g(x): (x, t) \in T \cap \bar{H}_+\}$$
$$= \min\{\bar{t}_i - g(v^i): i = 1, \ldots, n+1\} .$$

Let V denote the matrix with columns $v^1, \ldots, v^{n+1}$. Then the above consideration leads to the following linear program in (λ, t):

$$\max\left(\sum_{i=1}^{n+1} t_i \lambda_i - t\right) \tag{18}$$

$$\text{(LP) s.t. } AV\lambda + at \leq b \tag{19}$$

$$\sum_{i=1}^{n+1} \lambda_i = 1, \; \lambda_i \geq 0 \quad i = 1, \ldots, n+1 \tag{20}$$

where λ is a vector with components $\lambda_1, \ldots, \lambda_{n+1}$, and A, a, b are given in (13).

LEMMA 1. (*a*) *If the system* (19), (20) *has no solution, then the intersection* $T \cap \tilde{D}$ *is empty.*

(*b*) *Otherwise, let* (λ^*, t_*) *be an optimal solution of* (*LP*) *and* $c^* = \Sigma_{i=1}^{n+1} t_i \lambda_i^* - t_*$ *its optimal value, respectively. Then the following assertion holds*:

(*b*1) *If* $c^* \leq 0$, *then* $T \cap \tilde{D} \subset H_+$;

(*b*2) *if* $c^* > 0$, *then*

$$z = z(T) = (V\lambda^*, t_*),\ z^i = (v^i, \bar{t}_i) = (v^i, t_i - c^*)(i = 1, \ldots, n+1) \tag{21}$$

and

$$\bar{t}_i - g(v^i) = \mu - c^* \quad (i = 1, \ldots, n+1)\,. \tag{22}$$

Proof. Note that $(x, t) \in P$ implies $t \geq \bar{t}$ for all convex polyhedral sets P generated by the method (cf. (2)). It follows that (LP) has an optimal solution whenever its feasible set is nonempty.

Recall that the equation of H is $p^T x - t = \gamma$. It follows that determining the hyperplane $\bar{H}$ and the point z as discussed above amounts to solving the linear programming problem

$$\max\ (p^T x - t) \tag{23}$$

$$\text{s.t. } (x, t) \in T \cap P \tag{24}$$

Since every point $x \in U$ (where U is the simplex which generates T) is uniquely representable as

$$x = \sum_{i=1}^{n+1} \lambda_i v^i = V\lambda,\ \sum_{i=1}^{n+1} \lambda_i = 1,\ \lambda_i \geq 0 \quad (i = 1, \ldots, n+1)\,, \tag{25}$$

we see from (13) that the set $T \cap P$ coincides with the feasible set of problem (LP) (in other words, the systems (19), (20) describes $T \cap P$). Therefore, if the system (19), (20) has no solution, then $T \cap P = \emptyset$, and hence $T \cap \tilde{D} = \emptyset$ (because of $\tilde{D} \subset P$).

To prove part (b), we first note from (23) and (25) that

$$p^T x - t = p^T\Big(\sum_{i=1}^{n+1} \lambda_i v^i\Big) - t = \sum_{i=1}^{n+1} \lambda_i p^T v^i - t\,.$$

But, since $(v^i, t_i) \in H$, we have $p^T v^i - t_i = \gamma$ $(i = 1, \ldots, n+1)$, and hence

$$p^T x - t = \sum_{i=1}^{n+1} \lambda_i(\gamma + t_i) - t = \sum_{i=1}^{n+1} t_i \lambda_i - t + \gamma\,, \tag{26}$$

where the last equality follows from $\Sigma_{i=1}^{n+1} \lambda_i = 1$. Therefore, the linear programs (23), (24) and (LP) are equivalent, and, if γ^* denotes the optimal objective function value in (23), (24), then

$$\gamma^* = c^* + \gamma\,.$$

If $\gamma^* \leq \gamma$, then it follows from the definition of H_+ that $\bar{H}$ is obtained by a parallel shift of H in the direction H_+. Therefore, $c^* \leq 0$ implies $T \cap P \subset H_+$, and hence $T \cap \tilde{D} \subset H_+$ which proves (b1).

The equivalence of the two linear programs which we have shown above implies that the point $z = (x^*, t_*)$ with $x^* = V\lambda^*$ is an optimal solution of (23), (24).

Since $\bar{H} = \{(x, t) \in \mathbb{R}^n \times \mathbb{R}: p^T x - t = \gamma^*\}$ and $H = \{(x, t) \in \mathbb{R}^n \times \mathbb{R}: p^T x - t = \gamma\}$ we see that for each intersection point $(v^i, \bar{t}_i)$ (resp. (v^i, t_i)) of the edge of T passing through v^i with $\bar{H}$ (resp. H), we have $p^T v^i - \bar{t}_i = \gamma^*$ and $p^T v^i - t_i = \gamma$, respectively. This implies that $\bar{t}_i = t_i + \gamma - \gamma^* = t_i - c^*$, and (using $t_i = g(v^i) + \mu$) that $\bar{t}_i = g(v^i) + \mu - c^*$. □

Now recall that we have seen above that in the case (b_1) of Lemma 1 the quantity μ constitutes a lower bound of $(t - g(x))$ on $T \cap \tilde{D}$ whereas in the case (b_2) such a lower bound is given by $\min\{\bar{t}_i - g(v^i):\ i = 1, \ldots, n+1\}$. Lemma 1 thus provides the lower bound

$$\beta(T, P, \alpha) = \begin{cases} +\infty\,, & \text{if (LP) has no feasible point} \\ \mu\,, & \text{if } c^* \leq 0 \\ \mu - c^*\,, & \text{if } c^* > 0 \end{cases} \tag{27}$$

of $t - g(x)$ on $T \cap \tilde{D}$ which is calculated by solving the linear programming problem (LP). We will see in Section 3.5 that T can be deleted from further consideration when $\beta(T, P, \alpha) = \infty$ or $\beta(T, P, \alpha) = \mu$.

3.4. OUTER APPROXIMATION OF $\tilde{D}$

The polyhedral convex set $P \supset \tilde{D}$ used in the preceding section is updated in each iteration, i.e., a sequence of polyhedral convex sets $P_0, P_1, \ldots$ is constructed which satisfy

$$P_0 \supset P_1 \supset \cdots \tilde{D}\,.$$

The transition from P_k to P_{k+1} $(k = 0, 1, \ldots)$ is done in a way which is standard for pure outer approximation methods: An appropriate linear inequality $l_k(x, t) \leq 0$ is added to the constraint set which defines P_k, i.e., we set

$$P_{k+1} = P_k \cap \{(x, t) \in \mathbb{R}^n \times \mathbb{R}: l_k(x, t) \leq 0\}\,.$$

We construct the function $l_k(x, t)$ in the following way. At iteration k, we will have a lower bound β_k of $t - g(x)$ over $\tilde{D}$ which equals some $\beta(T, P_k, \alpha) < \infty$ as defined in (27) with $P = P_k$. It follows from the discussion of the preceding section that we will also have a point $(\bar{v}^k, \bar{t}_k)$, satisfying $\bar{t}_k - g(v^k) = \beta_k$. If $(\bar{v}^k, \bar{t}_k) \in \tilde{D}$, i.e., if $\bar{v}^k \in D$ and $f(\bar{v}^k) \leq \bar{t}^k$, then we are done: $(\bar{v}^k, \bar{t}_k)$ is an optimal solution of problem (CP).

Therefore, we consider the case $(\bar{v}^k, \bar{t}_k) \notin \tilde{D}$. Let $z^k = (x^{*k}, t_{*k}) \in T \cap P_k$ denote the optimal solution of the linear program (23), (24) corresponding to $P = P_k$ and to a prism T such that $\beta_k = \beta(T, P_k, \alpha)$. Recall that $x^{*k} = V^k \lambda^{*k}$, where (λ^{*k}, t_{*k}) is an optimal solution of the linear program (18), (19), (20) with $V = V^k$ and the constraints describing $T \cap P_k$.

Suppose that we have $z^k \notin \tilde{D}$, and let $\tilde{h}(x, t) := \max\{h(x), f(x) - t\}$. Clearly, $\tilde{h}$ is convex (since $h(x)$ and $f(x) - t$ are convex) and

$$\tilde{D} = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : \tilde{h}(x, t) \leq 0\} . \tag{28}$$

We set

$$l_k(x, t) = s_k^T[(x, t) - z^k] + \tilde{h}(z^k) , \tag{29}$$

where s_k is a subgradient of $\tilde{h}$ at z^k.

LEMMA 2. *The hyperplane* $\{(x, t) \in \mathbb{R}^n \times \mathbb{R}: \ l_k(x, t) = 0\}$ *strictly separates* z^k *from* $\tilde{D}$, *i.e.*

$$l_k(z^k) > 0 \tag{30}$$

and

$$l_k(x, t) \leq 0 \quad \forall (x, t) \in \tilde{D} . \tag{31}$$

Proof. Since we assume that $z^k \notin \tilde{D}$, we have $l_k(z^k) = \tilde{h}(z^k) > 0$. Relation (31) is an immediate consequence of the definition of a subgradient. □

Note that the cut (29) is standard in outer approximation methods. A great variety of other cuts can also be applied (cf. Horst, Thoai and Tuy (1987, 1989), Horst and Tuy (1990))

3.5. DELETION RULES

At each iteration of the algorithm we try to delete certain subprisms that do not contain an optimal solution. Apart from the standard branch and bound deletion rule which deletes a prism T whenever the lower bound associated with T exceeds the current best upper bound, we propose the following two deletion rules which are immediate consequences of Lemma 1.

(DR1) Delete T if (LP) has no feasible solution .

It follows from Lemma 1(a) that in this case $T \cap \tilde{D} = \emptyset$, i.e. the prism T is infeasible (cf. Horst (1988)).

(DR2) Delete T if the optimal value c^* of (LP) satisfies $c^* < 0$.

In this case we see from Lemma 1 and from the definition of μ that the current best feasible solution cannot be improved in T.

4. The Algorithm

Next we describe the algorithm for solving (P) in its transformation (CP). The notations and basic operations we use are those introduced in the preceding sections.

Initialization

Construct a simplex $S_0 \supset D$, the corresponding prism $T_0 = T(S_0)$ and a ployhedral convex set $P_0 \supset \tilde{D}$ as described in Section 3. Subdivide T_0 according to an exhaustive subdivision process (cf. Section 3.2). Let $\mathcal{S}_0$ denote the corresponding initial partition of S_0, and let

$$\mathcal{M}_0' = \{T = T(S); S \in \mathcal{S}_0\} \tag{33}$$

denote the initial prismatic partition of T_0.

Let $\bar{\alpha}_0$ be the best objective function value at the feasible points known in advance (cf. the assumption in Section 3.3), and let $V(\mathcal{S}_0)$ denote the union of the vertices of all $S \in \mathcal{S}_0$. Set

$$\alpha_0 = \min\{\bar{\alpha}_0, f(v^i) - g(v^i) : v^i \in V(\mathcal{S}_0) \cap D\}\ , \tag{34}$$

and let $(x^0, t_0) \in \tilde{D}$ satisfy $t_0 - g(x^0) = \alpha_0$.

For each $T \in \mathcal{M}_0'$ solve the linear program (LP) corresponding to α_0 and T (resp. to S when $T = T(S)$).

Delete all $T \in \mathcal{M}_0'$ which satisfy (DR1) or (DR2).

Let $\mathcal{M}_0$ denote the collection of remaining prisms in $\mathcal{M}_0$, and for each $T \in \mathcal{M}_0$ set

$$\beta(T) = \beta(T, P_0, \alpha_0) \tag{35}$$

(cf. (27)).

Let

$$\beta_0 = \min\{\beta(T) : T \in \mathcal{M}_0\}\ , \tag{36}$$

and let $(\bar{v}^0, \bar{t}_0)$ be the point satisfying $\beta_0 = \bar{t}_0 - g(\bar{v}^0)$ (cf. Lemma 1). If $(\bar{v}^0, \bar{t}_0) \in \tilde{D}$, then stop: $(\bar{v}^0, \bar{t}_0)$ is an optimal solution of problem (CP).

Iteration k $(k = 0, 1, 2, \ldots)$

At the beginning of Step k we have a polytope $P_k \supset \tilde{D}$, the best feasible point (x^k, t_k) obtained so far, and an associated upper bound $\alpha_k = t_k - g(x^k)$. Furthermore, we have a set $\mathcal{M}_k$ of prisms generated from the initial partition by deletion operation and subdivisions according to the rules stated below. Finally, for each prism $T \in \mathcal{M}_k$, a lower bound $\beta(T) \leq \inf\{t - g(x) : (x, t) \in \tilde{D} \cap T\}$ is known, and we have the bound $\beta_k \leq \text{int}\{t - g(x) : (x, t) \in \tilde{D}\}$ and a (not necessarily feasible) point $(\bar{v}^k, \bar{t}_k)$ associated with β_k, such that $\beta_k = \bar{t}_k - g(\bar{v}^k)$.

k.1. Delete all $T \in \mathcal{M}_k$ satisfying

$$\beta(T) \geq \alpha_k . \tag{37}$$

Let $\mathcal{R}_k$ be the collection of remaining prisms in $\mathcal{M}_k$. If $\mathcal{R}_k = \emptyset$, then stop: x^k is an optimal solution of problem (P) with optimal value $\beta_k = \alpha_k$.

k.2. Select a prism $T_k^* \in \mathcal{R}_k$ satisfying

$$\beta_k = \beta(T_k^*), (\bar{v}^k, \bar{t}_k) \in T_k^* ,$$

and let z^k be an optimal solution of (23), (24) corresponding to P_k and T_k^*.

If $z^k \in \tilde{D}$, then set $P_{k+1} = P_k$.

If $z^k \notin \tilde{D}$, then construct $l_k(x, t)$ according to (29), and set

$$P_{k+1} = \{(x, t) \in P_k : l_k(x, t) \leq 0\} . \tag{38}$$

k.3. Subdivide $T_k^* = T(S_k^*)$ according to the chosen exhaustive subdivision process into a finite number of subprisms $T_{k,j}(j \in J_k)$.

Let v_k^* denote the new vertex in the corresponding partition of the simplex S_k^* which generates T_k^* and set

$$\alpha'_{k+1} = \begin{cases} \alpha_k & \text{if } v_k^* \notin \tilde{D} \\ \min\{\alpha_k, f(v_k^*) - g(v_k^*)\} & \text{if } v_k^* \in \tilde{D} \end{cases} \tag{39}$$

k.4. For each $j \in J_k$ solve the linear program (LP) corresponding to $T_{k,j}$, P_{k+1} and α'_{k+1}.

Delete all $T_{k,j}(j \in J_k)$ which satisfy (DR1) or (DR2). Let $\mathcal{M}'_k$ denote the collection of remaining prisms $T_{k,j}(j \in J_k)$, and for each $T \in \mathcal{M}'_k$ set

$$\beta(T) = \max\{\beta(T_k^*), \beta(T, P_{k+1}, \alpha'_{k+1})\} \tag{40}$$

(cf. (27)).

k.5. Let F_k denote the set of new feasible points detected while evaluating (39) and solving the linear programs in Step k.4, and set

$$\alpha_{k+1} = \min\{\alpha_k, \min\{t - g(x) : (x, t) \in F_k\}\} \tag{41}$$

Let $(x^{k+1}, t_{k+1}) \in \tilde{D}$ satisfy $t_{k+1} - g(x^{k+1}) = \alpha_{k+1}$.

Set

$$\mathcal{M}_{k+1} = (\mathcal{R}_k \setminus \{T_k^*\}) \cup \mathcal{M}'_k , \tag{42}$$

and set

$$\beta_{k+1} = \min\{\beta(T) : T \in \mathcal{M}_{k+1}\} . \tag{43}$$

Let $(\bar{v}^{k+1}, \bar{t}_{k+1})$ be the point satisfying $\beta_{k+1} = \bar{t}_{k+1} - g(\bar{v}^{k+1})$ (cf. Lemma 1). If $(\bar{v}^{k+1}, \bar{t}_{k+1}) \in \tilde{D}$, then stop: $(\bar{v}^{k+1}, \bar{t}_{k+1})$ is an optimal solution of problem (CP). Otherwise, go to the next iteration.

REMARK. In practice, one would stop of course when $\alpha_k - \beta_k \leq \varepsilon$ ($\varepsilon > 0$, prescribed).

THEOREM 1. *If the algorithm is infinite, then every accumulation point of the sequence* $\{(\bar{v}^k, \bar{t}_k)\}$ *is an optimal solution of problem (CP).*

Proof. Recall that at each iteration k the point $(\bar{v}^k, \bar{t}_k) \in T_k^*$ satisfies $\beta_k = \bar{t}_k - g(\bar{v}^k)$ with $\beta_k \leq \inf\{t - g(x): (x, t) \in \tilde{D}\}$. Let $(\tilde{x}, \tilde{t})$ be an accumulation point of the sequence $\{(\bar{v}^k, \bar{t}_k)\}$, and let $\{(\bar{v}^{k_q}, \bar{t}_{k_q})\}$ be a subsequence satisfying $\lim_{q\to\infty} (\bar{v}^{k_q}, \bar{t}_{k_q}) = (\tilde{x}, \tilde{t})$. For the sake of simplicity of notation, we shall denote this subsequence by $\{(\bar{v}^q, \bar{t}_q)\}$. Since $\beta_k = \bar{t}_k - g(\bar{v}^k)$ is a lower bound of $t - g(x)$ on $\tilde{D}$, it follows from the continuity of $t - g(x)$ that $\tilde{t} - g(\tilde{x})$ is the optimal objective function value of (CP), and $(\tilde{x}, \tilde{t})$ is an optimal solution of (CP) whenever $(\tilde{x}, \tilde{t})$ is feasible. Therefore, it suffices to prove that $(\tilde{x}, \tilde{t}) \in \tilde{D}$.

A standard argument on the finiteness of the number of partition elements in each iteration which was used, e.g., in Horst (1976, 1986, 1988), Horst and Tuy (1989, 1990) show that there is a subsequence $\{(\bar{v}^{q_\nu}, \bar{t}_{q_\nu})\}$ of $\{(\bar{v}^q, \bar{t}_q)\}$ such that the sequence $\{T_{q_\nu}^*\}$ associated with $\{(v^{q_\nu}, \bar{t}_{q_\nu})\}$ (i.e., $(\bar{v}^{q_\nu}, \bar{t}_{q_\nu}) \in T_{q_\nu}^*$ and $\beta_{q_\nu} = \beta(T_{q_\nu}^*) = \bar{t}_{q_\nu} - g(\bar{v}^{q_\nu})$) is monotonically decreasing (nested). Since the subdivision process is exhaustive, and since the prisms $T_{q_\nu}^*$ are closed, we see that $T_{q_\nu}^*$ converges to a ray τ, and $(\tilde{x}, \tilde{t}) \in \tau$.

It is easy to see that $\tau \cap (\cap_\nu P_{q_\nu}) \neq \emptyset$, since $T_{q_\nu} \cap P_{q_\nu} \neq \emptyset \; \forall \nu$ (otherwise $T_{q_\nu}^*$ would have been deleted, cf. Lemma 1(a)).

Next, we consider the sequence $\{z^k\}$ of optimal solutions of the linear programs (23), (24) corresponding to T_k^* and P_k. By the same arguments which we used above for the sequence $\{(\bar{v}^k, \bar{t}_k)\}$, we see that to every accumulation point $(\tilde{x}, \tilde{t})$ of $\{(\bar{v}^k, \bar{t}_k)\}$ corresponds an accumulation point $\tilde{z}$ of $\{z^k\}$ so that $(\tilde{x}, \tilde{t})$ and $\tilde{z}$ lie on the same ray (and vice versa). Moreover, since for each q_ν of the above subsequence, we have $z^{q_\nu} \in T_{q_\nu}^*$, $(\bar{v}^{q_\nu}, \bar{t}_{q_\nu}) \in T_{q_\nu}^*$ and $z^{q_\nu} \in \bar{H}_{q_\nu}$, $(\bar{v}^{q_\nu}, \bar{t}_{q_\nu}) \in \bar{H}_{q_\nu}$ where $\bar{H}_{q_\nu}$ is a hyperplane which intersects all edges of $T_{q_\nu}^*$, we must have $\tilde{z} = (\tilde{x}, \tilde{t})$.

Finally, it follows from the general theory of outer approximation methods developed in Horst, Thoai and Tuy ((1987) and (1989)) (cf. also Horst and Tuy (1990)), that the cut (29) enforces $\tilde{z} \in \tilde{D}$. Since the cut (29) is the most straightforward of a family of cuts derived in Horst, Thoai and Tuy (1989) a direct proof is immediately available: From the construction of the polytopes P_{q_ν} it follows that

$$l_{q_\nu}(z^{q_\mu}) = s_{q_\nu}^T[(z^{q_\mu}) - (z_e^{q_\nu})] + \tilde{h}(z^{q_\nu}) \leq 0 \quad \forall \mu > \nu \,.$$

This implies in virtue of the Cauchy–Schwarz inequality that

$$\tilde{h}(z^{q_\nu}) \leq \|s_{q_\nu}\| \, \|(z^{q_\mu}) - (z^{q_\nu})\| \quad \forall \mu > \nu \,.$$

But the sequence $\{(z^{q_\nu})\}$ is convergent, and hence bounded. Therefore, the corresponding sequence of subgradients is bounded (cf. Rockafellar (1970)), and the above inequality yields (letting $\nu \to \infty$, $\mu \to \infty$ ($\mu > \nu$))

$$\tilde{h}(\tilde{z}) \leq 0, \quad \text{i.e. } \tilde{z} = (\tilde{x}, \tilde{t}) \in \tau \cap \tilde{D} . \qquad \square$$

5. Illustrative Examples

EXAMPLE 1.

$$\begin{aligned} &\min 4x_1^4 + 2x_2^2 - 4x_1^2 \\ &\text{s.t. } x_1^2 - 2x_1 - 2x_2 - 1 \leq 0 \\ &\quad -1 \leq x_1 \leq 1, -1 \leq x_2 \leq 1 \end{aligned} \tag{44}$$

Setting

$$\begin{aligned} &f(x) = 4x_1^4 + 2x_2^2 , \\ &g(x) = 4x_1^2 \\ &h(x) = \max\{h_i(x) \quad i = 1, 2, \ldots, 5\} \end{aligned}$$

where

$$\begin{aligned} &h_1(x) = x_1^2 - 2x_1 - 2x_2 - 1 \\ &h_2(x) = -x_1 - 1 \\ &h_3(x) = x_1 - 1 \\ &h_4(x) = -x_2 - 1 \\ &h_5(x) = x_2 - 1 \end{aligned}$$

problem (44) becomes

$$\begin{aligned} &\min t - g(x) \\ (\text{CP}) \;&\text{s.t. } f(x) - t \leq 0 \\ &\quad x \in D \end{aligned} \tag{45}$$

with $D = \{x \in \mathbb{R}^2 : h(x) \leq 0\}$. (46)

The origin $(0, 0)$ is in D, and we have the initial upper bound $\bar{\alpha}_0 = 0$.

Initialization

We start with a simplex $S_0 = \{(x_1 x_2) : x_1 \geq -1, x_2 \geq -1, x_1 + x_2 \leq 2\}$ with vertices $v_1 = (-1, -1)$, $v_2 = (-1, 3)$, $v^3 = (3, -1)$ and with the trivial partition $\mathcal{M}_0' = \mathcal{M}_0 = \{T(S_0)\}$, where $T_0 = T(S_0) \supset \tilde{D}$ is the prism generated by S_0. Since

$$0 \leq \min\{f(x) : x \in D\} ,$$

we set

$$P_0 = \{(x, t): -1 \leq x_1 \leq 1,\ -1 \leq x_2 \leq 1,\ t \geq 0\}\,.$$

We obtain $\alpha_0 = \bar{\alpha}_0 = 0$, $(x^0, t_0) = (0, 0, 0)$ and $\beta_0 = \beta(T_0) = \beta(T_0, P_0, \alpha_0) = -20$.

Iteration 0

We obtain $\alpha'_1 = \alpha_0 = 0$, $P_1 = P_0 \cap \{(x, t): l_0(x, t) \leq 0\}$, where

$$l_0(x, t) = 16x_1 - 4x_2 - t - 14\,.$$

Then, using bisection, we subdivide T_0 into two subprisms $T_{0,i} = T(S_{0,i})$ $(i = 1, 2)$, where $S_{0,1} = \text{conv}\{v^1, v^2, v^4\}$ and $S_{0,2} = \text{conv}[v^1, v^3, v^4]$ with $v^4 = \frac{1}{2}(v^2 + v^3) = (1, 1)$. By solving the linear programs (LP) for $T_{0,1}$ and $T_{0,2}$ we obtain the lower bounds $\beta(T_{0,1}) = -4$ and $\beta_1 = \beta(T_{0,2}) = -17$.

Iteration 1

We have $T_1 = T_{0,2}$, and we obtain $P_2 = P_1 \cap \{(x, t): l_1(x, t) \leq 0\}$, where $l_1(x, t) = 3.906x_1 - 4x_2 - t - 3.8311$. The prism T_1 is bisected into $T_{1,i} = T(S_{1,i})$ $(i = 1, 2)$, where $S_{1,1} = \text{conv}\{v^1, v^4, v^5\}$ and $S_{1,2} = \text{conv}\{v^3, v^4, v^5\}$ with $v^5 = \frac{1}{2}(v^3 + v^4) = (1, -1)$.

⋮

With a stopping criterion $\alpha_k - \beta_k \leq 0.05$ the algorithm terminates after 34 iterations at the approximate optimal solution $(x^*, t_*) = (0.7197, 0.0000, 1.0731)$ with objective function value -0.9987.

EXAMPLE 2. We applied the algorithm to the problem

$$\begin{aligned}
&\min(x_1^4 + x_2 + x_3) - (x_1 + x_2^2 - x_3)\\
&\text{s.t. } (x_1 - x_2 - 1.2)^2 + x_2 \leq 4.4\\
&x_1 + x_2 + x_3 \leq 6.5\\
&x_1 \geq 1.4\\
&x_2 \geq 1.6\\
&x_3 \geq 1.8
\end{aligned} \tag{47}$$

which was solved in Thoai (1988) by means of a cutting plane algorithm. The transformation of problem (47) in the form (CP) yields

$$\begin{aligned}
&\min t - g(x)\\
&f(x) - t \leq 0\,,\\
&h(x) \leq 0
\end{aligned}$$

where

$$f(x) = x_1^4 + x_2 + x_3 \tag{48}$$
$$g(x) = x_1 + x_2^2 - x_3$$
$$h(x) = \max\{h_i(x), i = 1, \ldots, 4\} \quad \text{with}$$
$$h_1(x) = x_1 + x_2 + x_3 - 6.5$$
$$h_2(x) = -x_1 + 1.4$$
$$h_3(x) = -x_2 + 1.6$$
$$h_4(x) = -x_3 + 1.8$$

With a tolerance $\varepsilon = 10^{-2}$, the algorithm terminated after 18 iterations at an approximate optimal solution $(x^*, t_*) = (1.400, 1.8128, 1.800, 7.454)$ (whereas the algorithm in Thoai (1988) needed 81 iterations to obtain the same accuracy).

Acknowledgement

The authors gratefully appreciate partial support of this research by the German Science Foundation (DFG).

References

Benson, H. P. (1991), Separable Concave Minimization via Partial Outer Approximation and Branch and Bound, forthcoming in *Operations Research Letters*.

Benson, H. P. and Horst, R. (1991), A Branch and Bound-Outer Approximation Algorithm for Concave Minimization over a Convex Set, forthcoming in *J. of Computers and Mathematics with Applications*.

Falk, J. E. and Hoffman, K. R. (1976), A Successive Underestimation Method for Concave Minimization Problems, *Mathematics of Operations Research* **1**, 271–280.

Giannessi, F., Jurina, L., and Maier, G. (1979), Optimal Excavation Profile for a Pipeline Freely Resting on the Sea Floor, *Engineering Structures* **1**, 81–91.

Heron, B. and Sermange, M. (1982), Nonconvex Methods for Computing Free Boundary Equilibria of Axially Symmetric Plasmas, *Applied Mathematics and Optimization* **8**, 351–382.

Hiriart-Urruty, J. B. (1985), Generalize Differentiability, Duality and Optimization for Problems Dealing with Differences of Convex Functions, *Lecture Notes in Economics and Mathematical Systems* **256**, Springer-Verlag, 37–69.

Hoffman, K. L. (1981), A Method for Globally Minmizing Concave Functions over Convex Sets, *Mathematical Programming* **20**, 22–32.

Horst, R. (1976), An Algorithm for Nonconvex Programming Problems, *Mathematical Programming* **10**, 312–321.

Horst, R. (1980), A Note on the Convergence of an Algorithm for Nonconvex Programming Problems, *Mathematical Programming* **19**, 237–238.

Horst, R. (1986), A General Class of Branch and Bound Methods in Global Optimization with Some New Approaches for Concave Minimization, *J. of Optimization Theory and Applications* **51**, 271–291.

Horst, R. (1988), Deterministic Global Optimization with Partition Sets Whose Feasibility is Not Known. Application to Concave Minimization, Reverse Convex Constraints, D.C.-Programming and Lipschitzian Optimization, *J. of Optimization Theory and Application* **58**, 11–37.

Horst, R. (1989), On Consistency of Bounding Operations in Deterministic Global Optimization, *J. of Optimization Theory and Applictions* **61**, 143–146.
Horst, R. (1990), Deterministic Global Optimization: Recent Advances and New Fields of Application, *Naval Research Logistics* **37**, 433–471.
Horst, R., Phong, T. Q. and Thoai, N. V. (1990), On Solving General Reverse Convex Programming Problems by a Sequence of Linear Programs and Line Searches, *Annals of Operations Research* **25**, 1–18.
Horst, R. and Thoai, N. V. (1989), Modification, Implementation and Comparison of Three Algorithms for Globally Solving Linearly Constrained Concave Minimization Problems, *Computing* **42**, 271–289.
Horst, R., Thoai, N. V., and Benson, H. P. (1991), Concave Minimization via Conical Partitions and Polyhedral Outer Approximation, forthcoming in *Mathematical Programming*.
Horst, R., Thoai, N. V., and Tuy, H. (1987), Outer Approximation by Polyhedral Convex Sets, *Operations Research Spektrum* **9**(3), 153–159.
Horst, R., Thoai, N. V., and Tuy, H. (1989), On an Outer Approximation Concept in Global Optimization, *Optimization* **20**, 255–264.
Horst, R., Thoai, N.V., and de Vries, J. (1987), On Finding New Vertices and Redundant Constraints in Cutting Plane Algorithms for Global Optimization, *Operations Research Letters* **7**, 85–90.
Horst, R. and Tuy, H. (1987), On the Convergence of Global Methods in Multiextremal Optimization, *J. of Optimization Theory and Applications* **54**, 253–271.
Horst, R. and Tuy, H. (1990), *Global Optimization: Deterministic Approaches*, Springer, Berlin.
Nguyen, V. H. and Strodiot, J. J. (1988), Computing a Global Solution to a Design Centering Problem, *Technical Report* **88/12**, Facultés Universitaires de Namur.
Pardalos, P. M., Glick, J. H., and Rosen, J. B. (1987), Global Minimization of Indefinite Quadratic Problems, *Computing* **39**, 281–291.
Pardalos, P. M. and Rosen, J. B. (1987), Constrained Global Optimization: Algorithms and Applications, *Lecture Notes on Computer Science* **268**, Springer XIV.
Polak, E. (1987), On the Mathematical Foundations of Nondifferentiable Optimization in Engineering Design, *SIAM Review* **29**, 21–89.
Polak, E. and Vincentelly, A. S. (1979), Theoretical and Computational Aspects of the Optimal Design Centering Tolerancing and Tuning Problem, *IEEE Transactions Circuits and Systems* **CAS-26**, 795–813.
Rockafellar, R. T. (1970), *Convex Analysis*, Princeton University Press.
Thach, P. T. (1988), The Design Centering Problem as a D.C. Program, *Mathematical Programming* **41**, 229–248.
Thieu, T. V., Tam, B. T., and Ban, V. T. (1983), An Outer Approximation Method for Globally Minimizing a Concave Function over a Compact Convex Set, *Acta Mathematica Vietnamica* **8**, 21–40.
Thoai, N. V. (1988), A Modified Version of Tuy's Method for Solving D.C. Programming Problems, *Optimization* **19**, 665–674.
Toland, J. F. (1979), A Duality Principle for Nonconvex Optimization and the Calculus of Variations, *Archive of Radional Mechanics and Analysis* **71**, 41–61.
Tuy, H. (1983), On Outer Approximation Methods for Solving Concave Minimization Problems, *Acta Mathematica Vietnamica* **8**, 3–34.
Tuy, H. (1986), A General Deterministic Approach to Global Optimization via D.C. Programming, in Hiriart-Urruty, J. B. (ed.), *Fermat Days* 1985: *Mathematics for Optimization*, North Holland, Amsterdam, 137–162.
Tuy, H. (1987), Convex Programs with an Additional Reverse Convex Constraint, *J. of Optimization Theory and Applications* **52**, 463–485.
Tuy, H. (1987a), Global Minimization of a Difference of Two Convex Functions, *Mathematical Programming Study* **30**, 150–182.
Tuy (1991), Effect of the Subdivision Strategy on Convergence and Efficiency of Some Global Optimization Algorithms, *J. of Global Optimization* **1**, 23–36.
Tuy, H. and Horst, R. (1988), Convergence and Restart in Branch and Bound Algorithms for Global

Optimization. Application to Concave Minimization and DC-Optimization Problems, *Mathematical Programming* **41**, 161–183.

Tuy, H., Thieu, T. V., and Thai, N. Q. (1985), A Conical Algorithm for Globally Minimizing a Concave Function over a Closed Convex Set, *Mathematics of Operations Research* **10**, 498–514.

Tuy, H. and Thuong, N. V. (1988), On the Global Minimization of a Convex Function under General Nonconvex Constraints, *Applied Mathematics and Optimization* **18**, 119–142.

Vidigal, L. M. and Director, S. W. (1982), A Design Centering Algorithm for Nonconvex Regions of Acceptability, *IEEE Transactions on Computer-Aided-Design of Integrated Circuits and Systems* **CAD-I,** 13–24.