Journal of Global Optimization **5:** 333–348, 1994.

Constraint Decomposition Algorithms in Global Optimization

REINER HORST and NGUYEN VAN THOAI
Fachbereich IV-Mathematik, Universitaet Trier, D-54286 Trier, Germany.

(Received: 9 July 1993; accepted: 24 February 1994)

Abstract. Many global optimization problems can be formulated in the form

$$\min\{c(x,y) : x \in X, y \in Y, (x,y) \in Z, y \in G\}$$

where X, Y are polytopes in $\mathbb{R}^p$, $\mathbb{R}^n$, respectively, Z is a closed convex set in $\mathbb{R}^{p+n}$, while G is the complement of an open convex set in $\mathbb{R}^n$. The function c:$\mathbb{R}^{p+n} \rightarrow \mathbb{R}$ is assumed to be linear.

Using the fact that the nonconvex constraints depend only upon the y-variables, we modify and combine basic global optimization techniques such that some new decomposition methods result which involve global optimization procedures only in $\mathbb{R}^n$. Computational experiments show that the resulting algorithms work well for problems with small n.

Key words: Global optimization, decomposition, canonical d.c. program, conical branch and bound algorithms, outer approximation, cutting plane algorithms.

1. Introduction

Let X, Y be polytopes (bounded polyhedral sets) in $\mathbb{R}^p_+$, $\mathbb{R}^n_+$, respectively, and let G be the complement of an open convex set in $\mathbb{R}^n$. Further, let Z be a closed convex set in $\mathbb{R}^{p+n}$ and c: $\mathbb{R}^{p+n} \rightarrow \mathbb{R}$ a linear function. We consider the global optimazation problem

$$(P) \quad \min\{c(x,y) : x \in X, y \in Y, (x,y) \in Z, y \in G\}.$$

Problem (P) includes several important classes of global optimization problems. As examples let us consider the problems

$$(P_1) \quad \min\{f(u) - g(v) : u \in U, v \in V, (u,v) \in W\},$$

where U, V are rectangles in $\mathbb{R}^{n_1}$, $\mathbb{R}^{n_2}$, respectively, W is a convex set in $\mathbb{R}^{n_1+n_2}$, and f, g are convex functions defined on $\mathbb{R}^{n_1}$, $\mathbb{R}^{n_2}$, respectively, and

$$(P_2) \quad \min\{f(x) : x \in U, \prod_{i=1}^{n} g_i(x) \leq 1\},$$

where U is a convex set in $\mathbb{R}^p$, f is a linear function and $g_i (i = 1, \cdots, n)$ are positive convex functions defined on U.

In Problem (P_1) the objective function is a separated *d.c. function* (difference of two convex functions). Frequently, (P_1) is called *d.c. programming problem* (cf., e.g., Horst and Tuy, 1993). An interesting special case arises when f is a linear function and n_1 is much larger than n_2. Problem (P_2) belongs to a class of programs dealing with a so-called *multiplicative terms*, which have recently attracted the attention of several authors because of its wide range of applications (cf., e.g., Muu, 1993; Thach, Burkard, and Oettli, 1991; Thoai, 1993; Tuy, 1992).

By introducing two additional variables, t_1, t_1 in (P_1) and n additional variables, $y_1, \cdots, y_n$ in (P_2), we can transform these problems into the form (P) as follows:

$$(P_1)' \quad \min\{t_1 - t_2 : u \in U, v \in V, (u,v) \in W, f(u) - t_1 \leq 0, g(v) - t_2 \leq 0\},$$

and

$$(P_2)' \quad \min\{f(x) : x \in U, g_i(x) - y_i \leq 0, (i = 1, \cdots, n), \prod_{i=1}^{n} y_i \leq 1\}.$$

Since in $(P_1)'$ the function $g(v) - t_2$ is convex, we have $G = \{(v, t_2) : g(v) - t_2 \geq 0\}$, and since in $(P_2)'$ the function $\prod_{i=1}^{n} y_i$ is quasi-concave we can define $G = \{(y_1, \cdots, y_n) : \prod_{i=1}^{n} y_i \leq 1\}$.

Problem (P) belongs to a class of nonconvex programming problems called *canonical d.c. programs*, and the constraint $y \in G$ is often called *reverse convex constraint*. A collection of results on canonical d.c. programs can be found in the book of Horst and Tuy (1993). However, it is a matter that up to now the existing methods for solving canonical programs can only work successfully on problems with small size ($p+n \leq 10$). In view of the special structure of Problem (P) that the reverse convex constraint depends only upon the y-variables, we intend to modify and combine basic global optimization techniques such that new decomposition algorithms result which use global optimization procedures only in the y-space. This gives hope to handle effectively the often occuring cases where the whole size of a problem under consideration may be fairly large while the number of nonconvex variables is small.

In this article we present two realizations of the above idea. In Section 2 we propose conical branch and bound techniques and polyhedral outer approximation, whereas the second approach, which we develop in Section 3, relies on cutting plane techniques. Throughout this article we need the following assumptions.

$$\begin{aligned} &(i) \quad \operatorname{int} \Omega \neq \emptyset, \text{ where } \Omega = \{(x,y) : x \in X, y \in Y, (x,y) \in Z\}, \\ &(ii) \quad \text{a point } y^o \in \mathbb{R}^n \backslash G \text{ is available.} \end{aligned} \tag{1.1}$$

Assumption (*i*) is usually needed for polyhedral outer approximations of a convex set. It is worth noting that a point y^0 satisfying (*ii*) can be determined independently of the structure of the set Ω. For establishing the cutting plane algorithm in Section 3, Assumtion (*ii*) must be replaced by a more strict condition.

Numerical examples to illustrate these algorithms are given in Section 4.

2. A Branch and Bound Algorithm

As usual, every algorithm of branch and bound type consists of two basic operations: branching and bounding. In our algorithms we need an additional basic operation for the successive outer approximation of the compact convex set Ω by a sequence of polytopes. We begin to establish our algorithm for solving Problem (P) with these basic operations.

2.1. Polyhedral Partitions

Let y^0 be a point in $\mathbb{R}^n$. A collection $\{C_1, \cdots, C_r\}$ of subsects of $\mathbb{R}^n$ is called a *conical partition* of $\mathbb{R}^n$, if each set $C_j\,(j = 1, \cdots, r)$ is a convex polyheral cone of dimension n, having exactly n edges emanating from y^0, such that $\bigcup_{j=1}^r C_j = \mathbb{R}^n$ and int $C_j \cap$ int $C_i = \emptyset$ for $j \neq i$. Throughout this article, by a *cone* or *conical partition set* in $\mathbb{R}^n$ we always mean a convex polyhedral cone having the above structure. The conical partition $\{C_1, \cdots, C_r\}$ of a cone C is defined similarly (cf. e.g., Horst and Tuy, 1993; Horst, Thoai, Benson, 1991; Horst and Thoai, 1992).

A collection $\{F_1 \cdots, F_r\}$ of subsets of $\mathbb{R}^{p+n}$ is called a *C-partition* of $\mathbb{R}^{p+n}$, if

$$F_j = \mathbb{R}^p \times C_j (j = 1, \cdots, r), \tag{2.1}$$

and $\{C_i, \cdots, C_r\}$ forms a conical partition in $\mathbb{R}^n$. The sets F_j are called *C-partition sets*. A C-partition of an element of a C-partition is defined similarly by using a conical partition of the corresponding cone C, i.e., we say that the collection $\{F_1, \cdots, F_r\}$, forms a C-partition of $F = \mathbb{R}^p \times C$, where C is a cone, if $F_j = \mathbb{R}^p \times C_j, (j = 1, \cdots, r)$, and $\{C_1, \cdots, C_r\}$, forms a conical partition of C.

In the context of conical branch and bound algorithms in $\mathbb{R}^n$, at the beginning of an algorithm, the space $\mathbb{R}^n$ is usually divided into $n+1$ cones, and a first collection of $r \leq n+1$ of these cones is chosen such that its union contains the set Y. Thereafter, at each iteration a cone is divided into finitely many subcones using certain standard partition rules. For more details on various conical partition rules we refer, e.g., to Horst and Tuy (1993), Horst, Thoai, and Benson (1991), Horst and Thoai (1992), see also the numerical example in Section 4.

For convergence proofs of conical algorithms, the most useful characterization of a partition process is the concept of *exhaustiveness*. A nested subsequence $\{C_q\}, C_q \supset C_{q+1} \forall q$, is called exhaustive if the intersection $\cap_{q=1}^{\infty} C_q$ is a ray (a halfline emanating from a point y^0). A conical partition process is called exhaustive if every nested subsequence of cones generated throughout the algorithm is exhaustive. A typical example for exhaustive partition processes is the well-known *bisection process*. Other classes of exhaustive partitions are discussed in Tuy, Khachaturov, and Utkin (1987), Horst and Tuy (1993), Horst, Thoai, and de Vries (1992 and 1992a).

2.2. Lower bounds

Let C be a conical partition set in $\mathbb{R}^n$, i.e., C is a polyhedral convex cone of dimension n having n edges emanating from a point $y^0 \in \mathbb{R}^n$, and let P be a polytope in $\mathbb{R}_+^{p+n}$ which contains the convex set Ω defined in (1.1). We propose here a procedure for computing a lower bound μ of the linear function $c(x, y)$ over the set $\{(x, y) \in \mathbb{R}^{p+n} : (x, y) \in \Omega \cap F, y \in G\}$, where $F = C \times \mathbb{R}^p$. Recall that in Problem (P) the set G is the complement of an open convex set in $\mathbb{R}^n$.

Assume that the polytope P is described as the solution set of the system

$$Ax + By \leq d \tag{2.2}$$

where A, B and d are matrices and vector, respectively, of appropriate sizes.

Note that, by Assumption (ii), the point y^0 satisfies $y^0 \notin G$. For each $i = 1, \cdots, n$, let $v^i \neq y^0$ be a point on the $i - th$ edge of C, and let z^i be a point on this edge determined by

$$z^i = y^0 + \theta(v^i - y^0) \text{ with } \theta = \min\{\theta_1 : \sup\{\lambda : y^0 + \lambda(v^i - y^0) \notin G\}\}, \tag{2.3}$$

where θ_1 is a given positive (usually large) number. Then every point $(x, y) \in F$ is uniquely representables as

$$\begin{aligned}(x, y) \;=\; & (0^p, y^0) + \sum_{i=1}^{p} \alpha_i e^i + \sum_{i=1}^{n} \lambda_i((0^p, z^i) \\ & -(0^p, y^0)), \lambda_i \geq 0 (i = 1, \ldots, n)\end{aligned} \tag{2.4}$$

where 0^p is the origin of $\mathbb{R}^p$, and e^i is the $i - th$ unit vector of $\mathbb{R}^{p+n}$ $(i = 1, \cdots, p)$.

Denoting by U the matrix with n columns $(z^1 - y^0), \cdots, (z^n - y^0)$ we can write (2.4) in the form

$$(x, y) = (0^p, y^0) + (\alpha, U\lambda), \lambda \geq 0 \tag{2.5}$$

where $\alpha = (\alpha_1, \cdots, \alpha_p)$ and $\lambda = (\lambda_1, \cdots, \lambda_n)$. The computation of the lower bound μ is based on the following

THEOREM 2.1. *Let the linear function $c(x, y)$ be defined by $c(x, y) = c^p x + c^n y$, where $c^n \in \mathbb{R}^n$ and $c^p \in \mathbb{R}^p$. Then the number*

$$\begin{aligned}\mu \;&=\; \mu(C, P) \\ &=\; c^n y^0 + \min\{c^p\alpha + c^n U\lambda : A\alpha + BU\lambda \\ &\quad \leq d - By^0, \sum_{i=1}^{n} \lambda_i \geq 1, \alpha \in \mathbb{R}_+^p, \lambda \in \mathbb{R}_+^n\}\end{aligned} \tag{2.6}$$

is a lower bound of c(x, y) over the set $\{(x, y) : (x, y) \in \Omega \cap F, y \in G\}$. (We understand that $\mu = \infty$, if the linear program in (2.6) has no feasible solution.)

Proof. Let H^n be the hyperplane in $\mathbb{R}^n$ containing the n point $z^1, \cdots, z^n$ and let $H = H^n \times \mathbb{R}^p$. Let H_+ be the halfspace generated by H which does not contain the point $(0^p, y^0)$. First, we assert that

$$\{(x,y) : (x,y) \in F, y \in G\} \subseteq \{(x,y) : (x,y) \in F \cap H_+\}.$$

Indeed, by definition, the sets H and H_+ are described as

$$H = \{(x,y) : (x,y) = (0^p, y^0) + (\alpha, U\lambda), \alpha \in \mathbb{R}^p, \lambda \in \mathbb{R}^n_+, \sum_{i=1}^{n} \lambda_i = 1]\}$$

and

$$H_+ = \{(x,y) : (x,y) = (0^p, y^0) + (\alpha, U\lambda), \alpha \in \mathbb{R}^p, \lambda \in \mathbb{R}^n_+, \sum_{i=1}^{n} \lambda_i \geq 1.$$

Let (x,y) be any point of F such that $y \in G$. Since $y^0 \notin G$ and the complement, $\bar{G}$, of G is convex it follows that the line segment $[y, y^0]$ intersects the boundary ∂G of G at an unique point $\hat{y}$ and meets H at an unique point $\tilde{y}$ and we have $\tilde{y} \in [\hat{y}, y^0]$. This implies that $(x, \hat{y}) \in H_+$, and hence $(x,y) \in F \cap H^+$.

From the above assertion we see that

$$\begin{aligned} \mu &= \min\{c(x,y) : (x,y) \in P \cap F \cap H_+\} \\ &\leq \min\{c(x,y) : (x,y) \in P \cap F, y \in G\} \\ &\leq \min\{c(x,y) : (x,y) \in \Omega \cap F, y \in G\} \end{aligned} \tag{2.7}$$

Note that $P \subset \mathbb{R}^{p+n}_+$, and

$$\begin{aligned} \mathbb{R}^{p+n}_+ \cap H_+ \cap F &= \{(x,y) : (x,y) = (0^p, y^0) + (\alpha, U\lambda), \\ &\qquad \sum_{i=1}^{n} \lambda_i \geq 1, \alpha \in \mathbb{R}^p_+, \lambda \in \mathbb{R}^n_+\}. \end{aligned} \tag{2.8}$$

Therefore, it follows from (2.2) that

$$\begin{aligned} \mu &= \min\{c(x,y) : (x,y) \in P \cap F \cap H_+\} \\ &= \min\{c^n y^0 + c^p \alpha + c^n U\lambda : A\alpha + BU\lambda \leq d - By^0, \\ &\qquad \sum_{i=1}^{n} \lambda_i \geq 1, \alpha \in \mathbb{R}^p_+, \lambda \in \mathbb{R}^n_+\} \\ &= c^n y^0 + \min\{c^p \alpha + c^n U\lambda : A\alpha + BU\lambda \leq d - By^0, \\ &\qquad \sum_{i=1}^{n} \lambda_i \geq 1, \alpha \in \mathbb{R}^p_+, \lambda \in \mathbb{R}^n_+\}. \end{aligned}$$

□

REMARK. If C' is a conical partition set such that $C' \subseteq C$ and P' is a polytope such that $\Omega \subseteq P' \subseteq P$, then we obviously have $\mu(C', P') \geq \mu(C, P)$. If an optional solution $(\bar{\alpha}, \bar{\lambda})$ of the linear program in (2.6) satisfies

$$(\bar{x}(C), \bar{y}(C)) = (O^p, y^0) + (\bar{\alpha}, U\bar{\lambda}) \in \Omega \cap G, \tag{2.9}$$

then $\mu(C, P) = c(\bar{x}(C), \bar{y}(C))$ is also an upper bound of the optimal value of Problem (P). Therefore, in the context of a branch and bound algorithm, the partition set C can be immediately removed from further consideration.

If (2.9) does not hold, we will be interested in the point $(\bar{x}(C), \bar{y}(C))$ defined by

$$(\bar{x}(C), \bar{y}(C)) = (\bar{x}(C), y^0) + \theta((\bar{x}(C), \bar{y}(C)) - (\bar{x}(C), y^0)) \tag{2.10}$$

where

$$\theta = \min\{\theta_1; sup\{t : y^0 + t(\bar{y}(C) - y^0) \notin G\}\} \tag{2.11}$$

with θ_1 being a given positive (usually large) number (cf. (2.3)). Geometrically, $\hat{y}(C)$ is the unique intersection point of the ray emanating from y^0 through $\bar{y}(C)$ with the boundary ∂G of G, if this intersection point exists; otherwise, it is a point on this ray that stands far enough from y^0. So, we see that the computation of the point $(\bar{x}(C), \hat{y}(C))$ requires only operations in $\mathbb{R}^n$. If

$$(\bar{x}(C), \hat{y}(C)) \in \Omega, \tag{2.12}$$

then $c(\bar{x}(C), \hat{y}(C))$ can be used to improve the current upper bound of the optimal value c^* of Problem (P).

2.3. Approximation by Polytopes

From the description of Problem (P) we see that the set $P_1 = X \times Y$ is the polytope in $\mathbb{R}^{p+n}$ which contains the bounded convex set Ω. Within our algorithm a sequence $\{P_k\}, k = 1, 2, \cdots$ will be iteratively constructed satisfying $P_1 \supset P_2 \supset \cdots \supset \Omega$. This approximation process needs the following basic operation:
Given a polytope P in $\mathbb{R}^{p+n}$ satisfying $P \supset \Omega$ and $P \backslash \Omega \neq \theta$, construct a polytope $\bar{P}$ satisfying $P \supset \bar{P} \supset \Omega$.

To our purpose we assume that the set Ω is defined by

$$\Omega = \{(x, y) : \phi(x, y) \leq 0\} \tag{2.13}$$

where $\phi(x, y)$ is a convex function usually defined as the maximum of a finite family of convex functions on the set $P_1 = X \times Y$.

Let C be a conical partition set as described in Section 2.1. Assume that in the lower bound estimation procedure of Section 2.2 we obtained a point $(\bar{x}, \bar{y})$ which does not satisfy (2.9), (recall that whenever (2.9) is fulfilled, the partition set C is removed from further consideration). The polytope $\bar{P}$ can be constructed by using $(\bar{x}, \bar{y})$ as a so-called *approximation point* as discussed in the following two cases.
(a) If $(\bar{x}, \bar{y}) \in \Omega$, then set $\bar{P} \leftarrow P$.
(b) If $(\bar{x}, \bar{y}) \notin \Omega$, then compute a subgradient s of ϕ at $(\bar{x}, \bar{y})$. Set

$$\bar{P} \leftarrow P \cap \{(x, y) : l(x, y) \leq 0\}, \tag{2.14}$$

where

$$l(x,y) = ((x,y) - (\bar{x},\bar{y}))s + \phi(\bar{x},\bar{y}). \tag{2.15}$$

The affine function $l(x,y)$ in (2.15) is often called *cutting function*. If a point $(u,v) \in \text{int}\Omega$ is available, this function can alternatively be defined as

$$l(x,y) = ((x,y) - (\tilde{x},\tilde{y}))s, \tag{2.16}$$

where $(\tilde{x},\tilde{y})$ is the unique intersection point of the line segment $[(\bar{x},\bar{y},(u,v)]$ with the boundary $\partial\Omega$ of Ω, and s is a subgradient of ϕ at $(\tilde{x},\tilde{y})$.
Cuts of type (2.16) are usually "deeper" than cuts of type (2.15).

2.4. THE ALGORITHM

Using notions and the three basic operations above we establish a branch and bound algorithm for solving Problem (P).

ALGORITHM 1.

Initialization

Construct a polytope P satisfying $P \supseteq X \times Y$;
Construct a conical partition $C_1, \cdots, C_{n+1}$ of $\mathbb{R}^n$ (with common vertex y^0, cf. Assumption (*ii*) and Section 2.1);
Compute lower bounds $\mu(C_i) = \mu(C_i, P)(i = 1, \cdots, n+1)$;
Set $\mathcal{C} \leftarrow \{C_1, \cdots, C_{n+1}\}$;
Compute an upper bound γ by evaluating the function c at all points $(\bar{x}(C), \bar{y}(C)), C \in \mathcal{C}$ and at all points $(\bar{x}(C), \hat{y}(C)), C \in \mathcal{C}$ satisfying (2.9) and (2.12), respectively.
If no point of $\Omega \cap G$ is found, then set $\gamma = \infty$;
Set $\mathcal{R} \leftarrow \{C \in \mathcal{C} : \mu(C) < \gamma\}$;
Set $\mu \leftarrow \min\{\mu(C) : C \in \mathcal{R}\}$;
Choose $C \in \mathcal{R}$ satisfying $\mu(C) = \mu$;
Set $stop \leftarrow false, k \leftarrow 1$.

While *stop* = *false* **do**

if $\mathcal{R} = \emptyset$ **then**

stop $\leftarrow$ *true*; (the point (x,y) with $c(x,y) = \gamma < \infty$ is optimal solution, or Problem (P) is infeasible if $\gamma = \infty$).

else

Perform a conical partition of C obtaining $C_1, \cdots, C_r$;
Construct a polytope $\bar{P}$ satisfying $P \supseteq \bar{P} \supseteq \Omega$ as described in (2.14);
Compute lower bounds $\mu(C_i) = \mu(C_i, \bar{P}), (i = 1, \cdots, r)$;
Update γ by using all feasible points obtained while computing the lower bounds for the cones $C_i, (i = 1, \cdots, r)$;
Set $\mathcal{C} \leftarrow \mathcal{C} \backslash \{C\} \bigcup \{C_1, \cdots, r\}$,
Set $\mathcal{R} \leftarrow \{C \in \mathcal{C} : \mu(C) < \gamma;\}$,

$\mu \leftarrow \{\mu(C) : C \in \mathcal{R}\}$;
Choose $C \in \mathcal{R}$ satisfying $\mu(C) = \mu$.
endif
Set $P \leftarrow \bar{P}, k \leftarrow k + 1$;
endwhile

To examine the convergence of the above algorithm, let us assign the index k to every set or quantity determined at iteration k.

If the algorithm does not terminate after finitely many iterations, either yielding an optimal solution of Problem (P) or showing that the problem is infeasible, it must generate an infinite sequence $\{(\bar{x}, \bar{y})^k\} \subset P_1 \backslash \Omega$, where $c((\bar{x}, \bar{y})^k) = \mu_k \forall k$. The convergence of Algorithm 1 is stated by means of the following results.

THEOREM 2.2. *If the conical partition process is exhaustive, then every cluster point of the sequence $\{(\bar{x}, \bar{y})^k\}$ is an optimal solution of Problem (P).*

Proof. Let (x^*, y^*) be a cluster point of $\{(\bar{x}, \bar{y})^k\}$, and let $\{(\bar{x}, \bar{y})^q\}$ be a subsequence converging to (x^*, y^*). By passing to a subsequence if necessary, we can assume that $(\bar{x}, \hat{y})^q \to (x^*, \hat{y}^*)$, (where for each k, $(\bar{x}, \hat{y})^k$ is the point defined by (2.10)–(2.11) corresponding to the cone C^k). Again, by passing to a suitable subsequence if necessary, we can assume that the corresponing sequence $\{C^q\}$ is decreasing, i.e., $C^q \supset C^{q+1} \forall q$.

From the well-known theory of outer approximation algorithms (cf. e.g. Horst, Thoai, and Tuy 1987; Horst and Tuy, 1993 and references given there), it follows that $(x^*, y^*) \in \Omega$.

On the other hand, by exhaustiveness of the conical partition process, $\{C^q\}$ converges to a ray. This implies that $\hat{y}^*$ is the intersection point of ∂G with the ray passing through y^*. Moreover, since $(\bar{x}, \bar{y})^k \in H_+^k \cap P_k$ for every k we can conclude that $\hat{y}^*$ belongs to the line segment $[y^0, y^*]$. Therefore, from the convexity of the complement $\bar{G}$ of G and the convexity of Ω it follows that $(x^*, y^*) \cap \Omega \cap G$ which imples that (x^*, y^*) is feasible to Problem (P), and hence $\lim_{q \to \infty} \mu_q = \mu^* = c(x^*, y^*) \geq c^*$, where c^* is the optimal value of Problem (P), i.e. (x^*, y^*) is an optimal solution. □

In fact, to guarantee the convergence of Algorithm 1, the exhaustiveness of the conical partition process can be weakened as follows. At each iteration $k \geq 1$ we apply the following conical partition rule:
(i) If $\hat{y}^k \in [y^0, \bar{y}^k]$, then divide C^k by an arbitrary partition rule,
(ii) otherwise, divide C^k by an exhaustive partition rule

THEOREM 2.3. *If throughout Algorithm 1 the conical partition process is performed by the above rule, then every cluster point of the sequence $\{(\bar{x}, \bar{y})^k\}$ is an optimal solution of Problem (P).*

Proof. Let (x^*, y^*) be any cluster point of $\{(\bar{x}, \bar{y})^k\}$ and let $\{C^q\}$ be a decreasing subsequence of cones such that $(\bar{x}, \bar{y})^q \to (x^*, y^*)$. If $\{C^q\}$ contains an infinite subsequence generated by an exhaustive partition process, then, as shown in Theorem 2.2, (x^*, y^*) is an optimal solution. Otherwise, by construction, there exists an infinite decreasing subsequence satisfying $\hat{y}^q \in [y^0, \hat{y}^q]$ for all q, and therefore, $\hat{y}^* \in [y^0, y^*]$, i.e. $y^* \in G$ (because we have $y^0 \in \bar{G}$ and $\hat{y}^* \in \partial G$). Since $(x^*, y^*) \in \Omega$ this implies again that (x^*, y^*) is feasible to Problem (P), and hence it is an optimal solution. □

3. A Cutting Plane Algorithm

3.1. The Algorithm

As mentioned in the introduction, in order to establish a cutting plane algorithm for solving Problem (P) we replace Assumption *(ii)* by the following assumption.

(ii)' There is a point $(x^0, y^0) \in \text{int}\Omega$ such that $y \notin G$ and $c(x^0, y^0) < c^*$.

(Here, as above, c^* denotes the optimal solution of Problem (P)). In fact, this assumption means that the reverse convex costraint $y \in G$ is essential. A point (x^0, y^0) can be found by a perturbation of an optimal solution of the convex program $\min\{c(x, y) : (x, y) \in \Omega\}$. When Ω is a polytope, and there exists a vertex $(x^0, y^0) \in \Omega$ satisfying $c(x^0, y^0) < c^0$, then Assumption *(ii)'* can be omitted.

Before stating our algorithm we establish an optimality criterion for Problem (P). To this purpose, let us assume that the set Ω is given by (2.13) and the set G is explicitely given by

$$G = \{y \in \mathbb{R}^n : g(y) \leq 0\}, \tag{3.1}$$

where g is a continous quasi-concave function on $\mathbb{R}^n_+$. Further, define

$$D = \{(x, y) \in \mathbb{R}^{p+n} : (x, y) \in \Omega, y \in G\}. \tag{3.2}$$

Let P be a polytope satisfying $P \supseteq D$. For each point $w = (u, v) \in \Omega$, let us denote by $Y(P, w)$ the projection of the set $\{(x, y) : x, y \in P, c(x, y) \leq c(u, v)\}$ on the space of y-variables, i.e.

$$Y(P, w) = \{y : (\exists x \in X)(x, y) \in P, c(x, y) \leq c(u, v)\}. \tag{3.3}$$

The following optimality criterion for Problem (P) is closely related to an optimality criterion for canonical d.c. programs, cf. Tuy (1986), Horst and Tuy (1993).

THEOREM 3.1. *Assume that the feasible set D of Problem (P) is robust (in the sense that $cl(\text{int}D) = D$, where cl(A) is the closure of a set A). Let P be a polytope satisfying $P \supseteq D$. Then a feasible point $w^* = (u^*, v^*)$ of Problem (P) is an optimal solution if*

$$\min\{g(y) : y \in Y(P, w^*)\} = 0. \tag{3.4}$$

Proof. Assume that there exists a point $(u, v) \in D$ satifying $c(u, v) < c(u^*, v^*)$. Let B be a ball around (u, v) such that $c(x, y) < c(u^*, v^*)$ for all $(x, y) \in B$. Since the set D is robust there exists a sequence $(x^q, y^q) \subset \text{int}D$ converging to (u, v). This implies that there exists an index q_0 such that $(x^{q_0}, y^{q_0}) \in \text{int}D \cap B$, i.e., $g(y^{q_0}) < 0$ and $c(x^{q_0}, y^{q_0}) < c(u^*, v^*)$. Thus, since $P \supset D$ it follows that $\min\{g(y) : (x, y) \in Y(P, w^*)\} = \min\{g(y) : (x, y) \in P, c(x, y) \leq c(u^*, y^*)\} \leq \min\{g(y) : (x, y) \in D, c(x, y) \leq c(u^*, y^*)\} \leq g(y^{q_0}) < 0$ which is a contradition to (3.4). □

Actually, in Theorem 3.1 the polytope P can be replaced by any closed set containing D. However, in order to establish an implementable algorithm for solving Problem (P) which is based on the optimality criterion (3.4), we shall construct iteratively a polytope P containing the feasible set D.

The following decomposition algorithm for solving Problem (P) is established based on the main idea of the algorithm for solving a canonical d.c. program developed by Tuy (1986).

ALGORITHM 2.

Initialization

Construct a polytope P satisfying $P \supseteq X \times Y$;
Compute $(x^0, y^0) \in \mathbb{R}^{n+p}$ satisfying Assumption $(ii)'$
(Note that we have $y^0 \notin G$, i.e. $g(y^0) > 0$).
Set $w \leftarrow (\infty)$; $c(w) \leftarrow +\infty$, $k \leftarrow 1$; $stop \leftarrow false$;

while $stop = false$ **do**

Solve the subproblem

$$\min\{g(y) : y \in Y(P, w)\} \tag{3.5}$$

obtaining an optimal solution s and the optimal value θ. (See Section 3.2.1. below).

If $\theta > 0$, $stop \leftarrow true$ (the feasible set D is empty, cf. the Corollary of Theorem 3.3 below).

If $\theta = 0$ and $w \neq (\infty)$ **then**

$stop \leftarrow true$ (w is an optimal solution of Problem P, by Theorem 3.1.)

else

Compute the intersection point v of the set $\{y : g(y) = 0\}$ with the line segment $[s, y^0]$;
Compute a point u satisfying $(u, v) \in \{(x, y) \in P : c(x, y) \leq c(w)\}$.
(cf. Section 3.2.2.);

if $(u, v) \in \Omega$ **then**

$w \leftarrow (u, v)$

else

Construct a polytope $\bar{P} \subset P$ by using (u, v) as an approximation point, as outlined in Section 2.3

Set $P \leftarrow \bar{P}$;
endif
endif
$k \leftarrow k+1$
endwhile

REMARK. Throughout Algorithm 2 two kinds of cutting planes are applied. The first one is used for the polyhedral approximation of the convex set Ω, the second one defines level sets of the linear objective function c. Therefore, we call Algorithm 2 a cutting plane algorithm.

3.2. Implementation and Convergence

In this section we give some details of an implementation of Algorithm 2 and state its convergence properties.

3.2.1. *Solving Subproblem (3.5)*

The most expensive operation in Algorithm 2 is solving Subproblem (3.5) at each iteration. Below we show the way to solve these subproblems within an unique outer approximation in $\mathbb{R}^n$. Assume that the set $\{(x,y) \in P : c(x,y) \leq c(w)\}$ is described by a system of linear inequalities,

$$Ax + By \leq b, x \geq 0, y \geq 0, \tag{3.6}$$

where A is an $(m \times p)$-matrix, B is an $(m \times n)$-matrix and $b \in \mathbb{R}^m$. The algorithm which we propose for solving Problem (3.5) is based on the following result.

THEOREM 3.2. *Let* $C = \{z \in \mathbb{R}^m : A^T z \geq 0, z \geq 0\}$, *where* A^T *is the transpose of* A, *and let* $E(C)$ *denote the set of all extreme rays of* C. *Then the set* $Y(P,w)$ *is defined by*

$$Y(P,w) = \{y \in \mathbb{R}^n_+ : z(b - By) \geq 0, z \in E(C)\}.$$

Proof. This is a classical result on projections of polyhedral sets. We present here a simple proof. We have $Y(P,w) = \{y \in \mathbb{R}^n_+ : (\exists x \in \mathbb{R}^p_+) By \leq b - Ax\}$. By the well-known Farkas Lemma it holds, for each $y \in \mathbb{R}^n_+$, that $\{x \in \mathbb{R}^p_+ : Ax \leq b - By\} \neq \theta$ if only if $z(b - By) \geq 0$ for all $z \in C$, i.e., $z(b - By) \geq 0$ for all $z \in E(C)$. □

The following result is an immediate consequence of Theorem 3.2.

COROLLARY (cf. Tuy, 1985; Thoai 1991). **a.** *The set* $Y(P,w)$ *is a polytope consisting of all points* $y \geq 0$ *satisfying*

$$\max\{(By - b)z : -A^T z \leq 0, ez \leq 1, z \geq 0\} = 0 \tag{3.7}$$

where $e \in \mathbb{R}^m$ is a vector with all components 1.

b. *For each point $\bar{y} \in \mathbb{R}^n \setminus Y(P,w)$ there exists a vertex $\bar{z}$ of the polytope $\{z : A^T x \leq 0, ez \leq 1, z \geq 0\}$ such that the affine function $l(y) = \bar{z}By - \bar{z}b$ satisfies $l(\bar{y}) > 0$ and $l(y) \leq 0$ for $y \in Y(P,w)$.*

ALGORITHM 3 (For solving Subproblem (3.5)).

Initialization

Construct a polytope S satisfying $S \supseteq Y(P,w)$ and
the vertex set $V(S)$ (See Remark (b) below);
Set $q \leftarrow 1$; $stop \leftarrow false$

while *stop = false* **do**

Select a point s satisfying $g(s) = \min\{g(y) : y \in V(S)\}$;
Compute, by the simplex method,
$\xi = \max\{(Bs - b)z : -A^T z \leq 0, ez \leq 1, z \geq 0\}$;

if $\xi = 0$ **then**

stop $\leftarrow$ *true* (s is an optimal solution of Problem (3.5)).

else

Select a basic feasible point $\bar{z}$ with $(Bs - b)\bar{z} > 0$;
Construct the affine function $l(y) = \bar{z}By - \bar{z}b$;
Set $S \leftarrow S \cap \{y : l(y) \leq 0\}$ and complete the vertex set $V(S)$ (cf. Remark (b) below);
Set $q \leftarrow q + 1$;

endif

endwhile

REMARK. **(a)** As shown in Thoai (1991), Algorithm 3 terminates after finitely many iterations yielding an optional solution s of Problem (3.5).

(b) At each iteration k of Algorithm 2 we have to solve a problem of the form (3.5). However, throughout Algorithm 2 these subproblems are solved successively within an outer approximation process. More precisely, the last polytope generated while solving the subproblem at iteration $k (k \geq 1)$ is used as the first polytope for solving the subproblem at iteration $k + 1$.

For calculating the vertex set of a polytope defined as intersection of a polytope with a halfspace, the methods discussed in Horst, Thoai, and de Vries (1988) and in Chen, Hansen, and Jaumard (1993) can be used, (see also Horst and Tuy, 1993).

3.2.2. *Computation of the Points u*

Let s be an optimal solution of Problem (3.5) with $g(s) < 0$, and let v be the intersection point of the set $\{y : g(y) = 0\}$ and the line segment $[s, y^0)]$. (Since $g(y^0) > 0$ and the set $\{y : g(y) \geq 0\}$ is convex, the point v is uniquely determined). Let

$$v = s + \varrho(y^0 - s) \quad (1 \geq \varrho \geq 0). \tag{3.8}$$

A point u satisfying $(u, v) \in \{(x, y) \in P : c(x, y) \leq c(w)\}$ can then easily be determined as follows.

Each optimal solution s of Problem (3.5) satisfies

$$\max\{(Bs - b)z : -A^T z \leq 0, ez \leq 1, z \geq 0\} = 0.$$

Considering the dual of this linear program we see that

$$\min\{\lambda : -Ax + \lambda e \geq Bs - b, x \geq 0, \lambda \geq 0\} = 0.$$

Let (r, λ^*) be an optimal solution of the dual program. Then it is easy to verify that the point (r, s) satisfies the system (3.6), i.e, $(r, s) \in \{(x, y) \in P : c(x, y) \leq c(w)\}$. (Note that, from the duality theory in linear programming, the point r is simply determined by the simplex tableau corresponding to an optimal solution of the primal program. For more details, see Thoai 1991.

Finally, a point u satisfying $(u, v) \in \{(x, y) \in P : c(x, y) \leq c(w)\}$ is computed by

$$u = r + \varrho(x^0 - r), \tag{3.9}$$

where the number ϱ is defined in (3.8).

3.2.3. *Convergence*

The following convergence properties of Algorithm 2 can be derived from the results on solving canonical d.c. programming problems presented in Horst and Tuy (1993), (see also Thoai, 1993).

THEOREM 3.3. *Assume that the feasible set D of Problem (P) is robust. Then, if Algorithm 2 does not terminate after finitely many iterations, it generates an infinite sequence $\{(u^k, v^k)\}$, every cluster point of which is an optimal solution to Problem (P).*

COROLLARY. *Problem (P) is infeasible (i.e. the set D is empty) if and only if Algorithm 2 terminates after a finite number of iternations because $\theta > 0$*

4. Illustrative Examples

To illustrate the algorithms established in previous sections we take the following problem.

$$\min c^p x + c^n y \tag{4.1}$$

$$s.t. Ax + By \leq d \tag{4.2}$$

$$\psi_i(x) = y_i \leq 0, (i = 1, \cdots, 3) \tag{4.3}$$

$$0 \leq x_j \leq 100, (j = 1 \cdots, p) \tag{4.4}$$

$$0 \leq y_j \leq 100, (j = 1 \cdots, n) \tag{4.5}$$

$$g(y) \leq 0 \tag{4.6}$$

where $p = 4, n = 3$,

$$A = \begin{pmatrix} 0.488509 & 0.063565 & 0.945686 & 0.210704 \\ -0.324014 & -0.501754 & -0.719204 & 0.099562 \\ 0.445225 & -0.346896 & 0.637939 & -0.257623 \\ -0.202821 & 0.647361 & 0.920135 & -0.983091 \end{pmatrix}$$

$$B = \begin{pmatrix} 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 \end{pmatrix}, d = \begin{pmatrix} 3.583525 \\ 0.337279 \\ 1.714630 \\ 0.479719 \end{pmatrix},$$

$$\begin{aligned} c^p &= (-0.022089, -0.761110, 0.861208, -0.348973), \\ c^n &= (0.0, 0.0, 0.0), \\ \psi_1(x) &= 0.2(0.5 = \sum_{j=1}^{4} jx_j)^2, \\ \psi_2(x) &= 0.1 \exp(0.15 + 0.1 \sum_{j=1}^{4} x_j/j), \\ \psi_3(x) &= 0.2 + \sum_{j=1}^{4} x_j/j, \\ g(y) &= \sum_{j=1}^{3} y_j - 1 \end{aligned}$$

4.1. Application of Algorithm 1

For solving this test problem Algorithm 1 is modified as follows. At each iteration, instead of setting $\mathcal{R} \leftarrow \{C \in \mathcal{C} : \mu(C) < \gamma\}$ we set $\mathcal{R} \leftarrow \{C \in \mathcal{C} : \mu(C) < \gamma - \varepsilon\}$, where ε is taken as 1% of γ. As a result, we obtain an ε-optimal solution (x^*, y^*) in the sense that $c^p x^* + c^n y^* \leq c^p x + c^n y - \varepsilon$ for all feasible points (x, y) of Problem (P). A first polytope P^1 is define by $P^1 = \{(x, y) : (x, y)$ satisfies (4.2), (4.4), (4.5)$\}$. A point $(u, v) \in \mathrm{int}\Omega$ is (1., 1., 1., 1., 100., 100., 100.) We choose $y^0 = (2., 2., 2.)$. Using the n-simplex $[v^1, v^2, v^3, v^4] = [e^1, e^2, e^3, -\frac{1}{3}(e^1 + e^2 + e^3)]$, where e^i is the $i - th$ unit vector $(i = 1, 2, 3)$, we construct a conical partition $\{C_1, C_2, C_3, C_4\}$, of $\mathbb{R}^3$, where for each $i \in \{1, \cdots, 4\}$ C_i, is the cone generated by 3 rays emanating from y^0 and having the directions $v^i, j \neq i$, respectively.

Iteration 1:
Lower bound for cone $C_1 : \mu(C_1, P^1) = -17.885049$,
Lower bound for cone $C_2 : \mu(C_2, P^1) = -4.935237$,
Lower bound for cone $C_3 : \mu(C_2, P^1) = -1.233849$,
Lower bound for cone $C_4 : \mu(C_4, P^1) = +\infty$, since the linear subproblem according to C_4 is infeasible.

At this iteration no feasible point is found. While computing lower bound $\mu(C_1, P^1)$ we obtain the point $(\bar{x}(C_1), \bar{y}(C_1)) = (0.0, 18.221018, 0.0, 11.510489, 1.0, 1.0, 1.0)$ which is infeasible. Computing the intersection point of $\partial\Omega$ with the line segment $[(\bar{x}(C_1), \bar{y}(C_1)), (u, v)]$ we construct a cutting function

$$\ell_1(x, y) = 8.295x_1 + 16.591x_2 + x_3 + 24.886 + 33.181x_4 - 1.000y_1$$
$$-81.868$$

The cone C_1 is divided into two subcones by a conical bisection. A first feasible point is found at iteration 10. This problem is solved after 17 iterations. ε-optimal solution is $x^* = (0.0, 1.590179, 0.0, 0.559156), y^* = (7.019073, 0.125630, 1.131579)$, with $c^p x^* + c^n y^* = -1.405432$

4.2. APPLICATION OF ALGORITHM 2

For computing an ε-optimal solution as mentioned above, Algorithm 2 is modified as follows. Whenever a feasible point $w = (u, v)$ is found, the set $\{(x, y) \in P : c(c, y) \le c(w)\}$ is replaced by $\{(x, y) \in P : c(x, y) \le c(w) - \varepsilon\}$ where ε is taken as 1% of $c(w)$. A first polytope P^1 is defined by

$$P^1 = \{(x, y) : (x, y) \text{ satisfies } (4.2), (4.4), y \in Y^1\},$$

where Y^1 is a polytope in $\mathbb{R}^3$ defined by

$$Y^1 = \{y \in \mathbb{R}^3 : x_1 + x_2 + x_3 \le 10^5, x_1 \ge 10^{-5}\}.$$

A point (x^0, y^0) satisfying Condition $(ii)'$ is $x^0 = (0.0, 18.220982, 0.0, 11.510461)$, $y^0 = (1377.263, 0.385, 12.188)$ with $c(x^0, y^0) = -17.885017$.

Iteration 1:
An optimal solution of Subproblem (3.5) is $s^1 = (0.000010, 0.000010, 0.000010)$ with $\theta^1 = \prod_{i=1}^3 s_i^1 - 1 < 0$.
Intersection point $(u^1, v^1) = (0.0, 0.977819, 0.0, 0.617703, 73.910081, 0.020686, 0.654078)$ is infeasible ($\psi_3(u^1) - v_3^1 = 0.189258 > 0$). Therefore, a cutting function is constructed:
$\ell_1(x, y) = x_1 + 0.5x_2 + 0.333333x_3 + 0.25x_4 - y_3 - 0.20$

Iteration 2:
An optimal solution of Subproblem (3.5) is $s^2 = (0.00010, 0.000010, 0.2)$ with $\theta^2 = \prod_{i=1}^3 s_i^2 - 1 < 0$.
Intersection point $(u^2, v^2) = (0.0, 0.891621, 0.0, 0.563250, 67.394623, 0.018863, 0.786623)$ is infeasible ($\psi_2(u^2) - v_2^2 = 0.104340 > 0$). Cutting function:
$\ell_2(x, y) = 0.012320x_1 + 0.006160x_2 + 0.004107x_3 + 0.003080x_4 - y_2 - 0.115976$

Iteration 3:
An optimal solution of Subproblem (3.5) is $s^3 = (0.000010, 0.115976, 0.2)$ with $\theta^3 = \prod_i^3 -1 < 0$.

Intersection point $(u^3, v^3) = (0.0, 0.284381, 0.0, 0.179648, 21.495438, 0.120179, 0.387103)$ is feasible. Therefore, set $w = (u^3, v^3)$ with $c(w) = -0.280930$.

...

This problem is solved after 12 iterations.
ε-optimal solution: $x^* = (0.0, 1.582390, 0.0, 0.554026), y^* = (6.936546, 0.127426, 1.130901)$, with $c(x^*, y^*) = -1.397713$.

Computational experiments on several types of test problems and a comparison of these two algorithms will be reported on an other occasion.

References

Chen, P. C., Hansen, P., and Jaumard B. (1991), On-Line and Off-Line Vertex Enumeration by Adjacency Lists, *Operations Research Letters* **10**, 403–409.

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. and Thoai, N. V. (1992), Conical Algorithm for the Global Minimisation of Linearly Constrained Decomposable Concave Minimization Problems, *J. Opt. Theory a. Appl.* **74**, 469–486.

Horst, R., Tuy, H. (1993), *Global Optimization: Deterministic Approaches*, Springer-Verlag.

Horst, R., Phong, T. Q. and Thaoi, N. V. (1990), On Solving General Reserve Convex Programming Problems by a Sequence of Linear Programs and Line Searches, *Annals of Oper. Res.* **25**, 1–18.

Horst, R., Thoai, N. V., Benson, H. P. (1991) Concave Minimization via Conical Partitions and Polyhedral Outer Approximation, *Mathematical Programming* **50**, 259–274.

Horst, R., Thoai, N. V., and Tuy, H. (1987), Outer Approximation by Polyhedral Convex Sets, *Oper. Res. Spektrum* **9**, 153–159.

Horst, R., Thoai, N. V., de Vries, J. (1992), A New Simplical Cover Technique in Constrained Global Optimization, *Journal Global Optimization* **2**, 1–19.

Horst, R., Thoai, N. V., de Vries, J. (1992a), On Geometry and Convergence of a Class of Simplicial Covers, *Optimization* **25**, 53–64.

Muu, L. D. (1993), An Algorithm for Solving Convex Programs with an Additional Convex-Concave Constraint, *Mathematical Programming* **61**, 75–87.

Thach, P. T., Burkard, R. E. and Oettli, W. (1991), Mathematical Programs with a Two-dimensional Reverse Convex Constraint, *J. Global Optimization* **1**, 145–154.

Thoai, N. V. (1991), A Global Optimization Approach for Solving the Convex Multiplicative Programming Problem, *J. Global Optimization* **1**, 341–357.

Thoai, N. V. (1993), Canonical D.C. Programming Techniques for Solving a Convex Program with an Additional Constraint of Multiplication Type, *Computing* **50**, 241–253.

Thoai, N. V., Tuy, H. (1980), Convergent Algorithms for Minimizing a Concave Function, *Mathematics of Oper. Res.* **5**, 556–566.

Tuy, H (1985), Concave Minimization under Linear Constraints with Special Structure, *Optimization* **16**, 335–352.

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*, Elsevier, Amsterdam, 137–162.

Tuy, H., Khachaturov, V., Utkin, S. (1987), A Class of Exhaustive Cone Splitting Procedures in Conical Algorithms for Concave Minimization, *Optimization* **18**, 791–807.

Tuy, H. (1992), The Complementary Convex Structure in Global Optimization. *J. Global Optimization* **2**, 21–40.