Journal of Global Optimization 9: 65–93, 1996.

Subdivision of Simplices Relative to a Cutting Plane and Finite Concave Minimization

MICHAEL NAST
University of Trier, Fachbereich IV, Department of Mathematics, Trier, Germany
e-mail: nast@uni-trier.de

Abstract. In this article, our primary concern is the classical problem of minimizing globally a concave function over a compact polyhedron (Problem (P)). We present a new simplicial branch and bound approach, which combines triangulations of intersections of simplices with halfspaces and ideas from outer approximation in such a way, that a class of finite algorithms for solving (P) results. For arbitrary compact convex feasible sets one obtains a not necessarily finite but convergent algorithm. Theoretical investigations include determination of the number of simplices in each applied triangulation step and bounds on the number of iterations in the resulting algorithms. Preliminary numerical results are given, and additional applications are sketched.

Key words: Concave minimization, branch and bound, simplicial subdivisions, triangulation, global optimization

1. Introduction

One of the most interesting and fundamental global optimization problems is of the form

$$\min f(x), \text{ subject to } x \in X, \qquad (P)$$

where X is a nonempty, compact polyhedral set (a polytope) in $\mathbb{R}^n$, and $f : D \to \mathbb{R}$ is a concave function on some open set D suitably containing the feasible domain X. Problem (P) is frequently called the concave minimization problem, and it is well–known that f may possess (possibly very many) local minima on X that are different from the global minimum we want to find. On the other hand, it is also well–known that there exists an extreme point of X which solves problem (P), which is one of the key observations when it comes to the task to develop solution methods for (P). Surveys on concave minimization including discussions of its various applications, related problems and existing solution methods are given, for instance, in Ref. [HoTu93], [HoPaTh95] and, in particular, in the very comprehensive recent survey of Benson in [Be95].

According to Benson ([Be95]), existing solution methods for solving problem (P) or generalizations of it can be classified by looking at their relationship to or their usage of three fundamental algorithmic approaches: enumeration, successive approximation, and successive partitioning (branch and bound). Due to extreme point optimality, enumerative methods and (outer) approximation schemes yield finite algorithms for problem (P) that are able to find exact extreme point optimal solutions. To guarantee this property, in worst case instances of (P), all extreme points of the feasible domain X are possibly investigated.

On the other hand, the vast majority of proposed algorithms use branch and bound schemes in various forms. Although practically often quite efficient, almost all the branch and bound methods cannot guarantee finiteness (see, e.g. [HoTu93], [Be95]). To our knowledge, only comparatively few finite partitioning methods are known from the literature; the following enumeration is given without any claim for completeness.

When f is separable, then the algorithm of Falk and Soland and a later variant by Soland are proved to be finite, using *rectangular* partitioning sets combined with lower bounds obtained via linear programming ([FaSo69], [So74]).

To generate finitely convergent *conical* methods, Hamami and Jacobsen introduced a subclass of exhaustive cone splitting processes, the so called exhaustive nondegenerate (or END, for short) cone subdivision ([HaJa88]). This process has been used both in [HaJa88] and [HoTu93] (Chapter VII.1) to give conical partitioning schemes that find exact, extreme point optimal solutions for problem (P) under certain additional (mild) assumptions.

At least partially inspired by ([So74]), Benson ([Be85]) proposed a *simplicial* branch and bound method using LP bounds and radial subdivisions, which does not require the separability of the objective function. As a follow–up, Benson and Sayin developed a variant of Bensons approach by incorporating a process called neighbor generation, which ensures both implementability and finiteness ([BeSa94]). Finally, the so called exact simplicial algorithm due to Ban (cf. [TaBa85], [HoTu93], Chapter VII.3) achieves finite convergence by introducing the concept of so called "trivial" simplices and a special bisection method for "nontrivial" simplices.

For a survey including almost all of these methods, we refer to [Be95]. In this article, we introduce a new subdivision procedure for simplicial partition sets: given an n–simplex S and a linear inequality, we construct a simplicial partition of the part of S which satisfies the given inequality. The application of this subdivision strategy in the branching step leads to a new finite partitioning method, which is closely related to the Ban–Algorithm. Moreover, it will allow us to investigate some interesting theoretical questions regarding bounds on the number of generated partition sets and bounds on the number of iterations for branch and bound schemes using this method.

The plan for this article is as follows: In Section 2, we give an algorithm for the subdivision step and hints for efficient implementations. After analyzing the possible number of new partition sets generated by this procedure, we will show, in Section 3, that no other method can improve on the proposed one in this respect. In Section 4, we will derive a class of finite algorithms for problem (P) which uses the new subdivision procedure, discuss the relation between these algorithms and Ban's method, and sketch some modifications for solving generalizations of problem (P). The article closes with a discussion of numerical examples in Section 5.

2. Subdivision of Simplices Relative to a Cutting Plane

Let $S = [v^0, \ldots, v^n]$ be an n–simplex with vertex set $V(S) = \{v^0, \ldots, v^n\}$, $v^i \in \mathbb{R}^n$, $0 \le i \le n$, and, for $a \in \mathbb{R}^n, b \in \mathbb{R}$, let $\{H = x \in \mathbb{R}^n : ax - b = 0\}$ be a hyperplane generating the half–spaces

$$H^{\leq} = \{x \in \mathbb{R}^n : ax - b \le 0\} \quad \text{and} \quad H^{\geq} = \{x \in \mathbb{R}^n : ax - b \ge 0\}.$$

Define the corresponding open half–spaces

$$H^- := H^{\leq} \setminus H, \quad H^+ := H^{\geq} \setminus H,$$

the vertex sets

$$\begin{aligned} V^-(S) &:= V(S) \cap H^-, \\ V^+(S) &:= V(S) \cap H^+, \\ V^=(S) &:= V(S) \cap H, \end{aligned}$$

and their cardinalities

$$n^-(S) := |V^-(S)|, \quad n^+(S) := |V^+(S)|, \quad n^=(S) := |V^=(S)|.$$

DEFINITION 1. The hyperplane H is called irredundant for S, iff

$$S \cap H^{\leq} \neq S \neq S \cap H^{\geq}.$$

Otherwise H is called redundant for S.

LEMMA 2. *H is irredundant for S* $\Longleftrightarrow \min\{n^+(S), n^-(S)\} \ge 1$.

Proof. Let $\Lambda^n := \{\lambda \in \mathbb{R}^{n+1} : \lambda \ge 0, \sum_{j=0}^n \lambda_j = 1.\}$. Then H is irredundant for S, iff we can find two points $s^1, s^2 \in S$ satisfying $s^1 \in H^-$, $s^2 \in H^+$. The latter condition is equivalent to finding two vertices $v^1, v^2 \in V(S)$ with $v^1 \in V^-(S)$ and $v^2 \in V^+(S)$, since $s^i = \sum_{j=0}^n \lambda_j^i v^j$ for some $\lambda^i \in \Lambda^n$, $i = 1, 2$ □

DEFINITION 3. Let $P \subset \mathbb{R}^n$ be a polytope satisfying int$P \neq \emptyset$, and let I be a finite set of indices. A family $\mathcal{S} := \{S_i : i \in I\}$ of n–simplices S_i is said to be a simplicial partition of P, if

$$P = \bigcup_{i \in I} S_i \qquad \text{and} \qquad \mathrm{int} S_i \cap \mathrm{int} S_j = \emptyset \ \ \forall i, j \in I, \ i \neq j.$$

If $V(S_i) \subseteq V(P)$ $\forall i \in I$, then $\mathcal{S}$ is called a triangulation of P.

Now let S be an n–simplex and H a hyperplane irredundant for S. Assume in the following, without loss of generality, that $n \geq 1$. The main purpose of this section is to construct simplicial partitions of the n–polytopes

$$P^{\leq} := S \cap H^{\leq} \quad \text{and} \quad P^{\geq} := S \cap H^{\geq}, \tag{1}$$

respectively (the trivial case when H is redundant for S will also be considered).

Let conv(X) denote the convex hull of a set X. Since $V(S) = V^-(S) \cup V^+(S) \cup V^=(S)$, we have $S = \text{conv}(V^-(S) \cup V^+(S) \cup V^=(S))$, where $V^-(S) \neq \emptyset \neq V^+(S)$. For any pair of vertices $u \in V^-(S)$, $v \in V^+(S)$, there exists a unique intersection point $h = e \cap H$ of the edge $e = [u, v]$ with the hyperplane H, given by

$$h = h(u, v, H) = \lambda u + (1 - \lambda)v,$$

where $\lambda = (av - b) / (av - au) \in (0, 1)$.

Let $V(S, H) := \{h(u, v, H) : u \in V^-(S),\ v \in V^+(S)\}$. It is well-known that

$$V(P^{\leq}) = V^-(S) \cup V^=(S) \cup V(S, H) \quad \text{and} \tag{2a}$$
$$V(P^{\geq}) = V^+(S) \cup V^=(S) \cup V(S, H), \tag{2b}$$

for a proof see Ref. [HoTu93], Lemma II.1.

For any $h := h(u, v, H) \in V(S, H)$, the radial subdivision (cf., e.g., [HoTu93], [Tuy91]) of the simplex S at the point h yields the two n–simplices

$$S_1 = \text{conv}\,(\ (V(S) \setminus \{u\}) \cup \{h\}\ )$$

and

$$S_2 = \text{conv}\,(\ (V(S) \setminus \{v\}) \cup \{h\}\ ),$$

satisfying $S_1 \cup S_2 = S$, int$S_1 \cap$ int$S_2 = \emptyset$. By construction, we have

$$n^-(S_1) = n^-(S) - 1\ ,\ n^+(S_1) = n^+(S)$$
$$n^-(S_2) = n^-(S) \qquad ,\ n^+(S_2) = n^+(S) - 1$$

and $n^=(S_1) = n^=(S_2) = n^=(S) + 1$, since $h \in H$.

DEFINITION 4. The radial subdivision of S at h is called a bisection of S with respect to u, v and H, or short a bisection with respect to h.

The following **Algorithm 1 (A1)** applies bisections on a given n–simplex S, until a given hyperplane H redundant or irredundant for S is redundant for *every* generated subsimplex:

Iteration 0: Set $\mathcal{L} \leftarrow \{S\}$, $\mathcal{M}_S^- \leftarrow \emptyset$, $\mathcal{M}_S^+ \leftarrow \emptyset$, $k \leftarrow 1$.

Iteration k:

$k.1$: If $\mathcal{L} = \emptyset$, then stop.

$k.2$: Choose $S_k \in \mathcal{L}$ and set $\mathcal{L} \leftarrow \mathcal{L} \setminus \{S_k\}$.

$k.3$: If $n^+(S_k) = 0$, set $\mathcal{M}_S^- \leftarrow \mathcal{M}_S^- \cup \{S_k\}$ and go to Step $k.6$.

$k.4$: If $n^-(S_k) = 0$, set $\mathcal{M}_S^+ \leftarrow \mathcal{M}_S^+ \cup \{S_k\}$ and go to Step $k.6$.

$k.5$: Choose $u \in V^-(S)$, $v \in V^+(S)$ and bisect S_k with respect to $h(u, v, H)$, generating two n–simplices S_{k_1}, S_{k_2}.
Set $\mathcal{L} \leftarrow \mathcal{L} \cup \{S_{k_1}, S_{k_2}\}$.

$k.6$: Set $k \leftarrow k + 1$ and go to Iteration k.

PROPOSITION 5. *Let S be a n–simplex, H be a hyperplane, $n^+ := n^+(S)$, $n^- := n^-(S)$, $n^= := n^=(S)$ be the cardinalities of the sets $V^+(S)$, $V^-(S)$, $V^=(S)$ as defined above (with respect to H). Then we have:*

(i) Algorithm 1 terminates after a finite number, i_S, of iterations. If K denotes the total number of n–simplices generated in $\mathcal{M}_S^- \cup \mathcal{M}_S^+$, then $i_S = 2K + 1$.

(ii) $\mathcal{M}_S := \mathcal{M}_S^- \cup \mathcal{M}_S^+$ forms a simplicial partition of S.
If $n^+ > 0$, then $\mathcal{M}_S^+$ forms a triangulation of $P^{\geq}$.
If $n^- > 0$, then $\mathcal{M}_S^-$ forms a triangulation of $P^{\leq}$.
If $n^+ = 0$, then $\mathcal{M}_S^+ = \emptyset$ and $\mathcal{M}_S = \mathcal{M}_S^- = \{S\}$.
If $n^- = 0$, then $\mathcal{M}_S^- = \emptyset$ and $\mathcal{M}_S = \mathcal{M}_S^+ = \{S\}$.

(iii) $K = |\mathcal{M}_S| = \binom{n^+ + n^-}{n^+}$, $K^+ := \left|\mathcal{M}_S^+\right| = \binom{n^+ + n^- - 1}{n^-}$ and $K^- := \left|\mathcal{M}_S^-\right| = \binom{n^+ + n^- - 1}{n^+}$, where we understand $\binom{n}{k} = 0$ for $k > n$.

Proof. To show part (i), first note that (A1) changes the size $\ell(\mathcal{L}) := |\mathcal{L}|$ of the set $\mathcal{L}$ in every iteration step $k < i_S \leq \infty$ in exactly one of the following two ways:

$\ell(\mathcal{L}) \leftarrow \ell(\mathcal{L}) + 1$, if $k = 0$ and in all iteration steps $k \geq 1$ where Step $k.5$ is executed, i.e., for all k satisfying $n^+(S_k) \geq 1$ and $n^-(S_k) \geq 1$ ('a size increasing step').

$\ell(\mathcal{L}) \leftarrow \ell(\mathcal{L}) - 1$ in all iteration steps $k \geq 1$ where Step $k.5$ is not executed, i.e. for all k satisfying $n^+(S_k) = 0$ or $n^-(S_k) = 0$ ('a size decreasing step').

Since the bisection operation used increases $n^=(S)$ by one in every iteration for the newly generated simplices S_{k_1}, S_{k_2}, every nested sequence of partition sets must be finite, ending up in a finite set of simplices $\{\widetilde{S}_i\}$ with either $n^+(\widetilde{S}_i) = 0$ or $n^-(\widetilde{S}_i) = 0$. We conclude that (A1) stops with $\mathcal{L} = \emptyset$ after a finite number of iterations, i.e. $i_S < \infty$. Since K is the number of produced output n–simplices collected in $\mathcal{M}_S$, and hence equals the total number of size decreasing steps, we must have $i_S = 2K + 1$ (for every size decreasing step there is exactly one size increasing step; the final iteration stating $\mathcal{L} = \emptyset$ adds one).

If $n^- = 0$ or $n^+ = 0$, the statements in **(ii)** are trivial. Otherwise, part (ii) is an obvious consequence of (i), Lemma 2 and the fact that every bisection process in Step $k.5$ generates a simplicial partition $\{S_{k_1}, S_{k_2}\}$ of S_k. It follows by induction that $\mathcal{L} \cup \mathcal{M}_S$ forms a simplicial partition of S in every iteration step k. In Step $i_S.1$ we have $\mathcal{L} = \emptyset$, and (ii) holds because of Lemma 2 and steps $k.3$, $k.4$ of (A1). Note that, by Step $k.5$ and (2a), (2b), every simplex $\tilde{S}$ in $\mathcal{M}_S^+$ or $\mathcal{M}_S^-$ satisfies $V(\tilde{S}) \subseteq V(P^{\geq})$ or $V(\tilde{S}) \subseteq V(P^{\leq})$, respectively, so that the simplicial partitions constructed in $\mathcal{M}_S^+$ and $\mathcal{M}_S^-$ are in fact triangulations of $P^{\geq}$ and $P^{\leq}$.

We show part **(iii)** using induction on $n^+ + n^-$:

(1) For $n^+ + n^- = 1$, assume (α): $n^+ = 1,\ n^- = 0$. We have $S \subseteq H^{\geq}$, and H is a supporting hyperplane for S. (A1) executes Step 1.4, setting $\mathcal{M}_S^+ = \{S\}$, and stops in Step 2.1 with $\mathcal{L} = \emptyset = \mathcal{M}_S^-$ after three iterations, with $k = 2$. We obtain
$$K = 1 = \binom{1+0}{1}, \quad K^+ = 1 = \binom{1+0-1}{0}, \quad K^- = 0 = \binom{1+0-1}{1}.$$
For case (β): $n^+ = 0,\ n^- = 1$ one concludes the result by analogous arguments.

(2) Now let us assume that (iii) holds for $n^+ + n^- = k,\ k \geq 1$. We have to show that (iii) holds for every simplex S with $n^+ + n^- = k + 1$. Assume (α): $n^+ = k + 1,\ n^- = 0$: We stop after three iterations with (iii), see (1)(α). Assume (β): $n^- = k + 1,\ n^+ = 0$: We stop after three iterations with (iii), see (1)(β).
Now assume (γ): $n^+ \geq 1,\ n^- \geq 1$: Step 1.5 of (A1) bisects $S_1 = S$ into two subsimplices S_{1_1}, S_{1_2} satisfying $n^+(S_{1_j}) + n^-(S_{1_j}) = k,\ j \in \{1, 2\}$, and
$$\begin{aligned} n^+(S_{1_1}) &= n^+ & \qquad n^+(S_{1_2}) &= n^+ - 1 \\ n^-(S_{1_1}) &= n^- - 1 & \qquad n^-(S_{1_2}) &= n^- \end{aligned}$$
After this initial step, (A1) executes exactly the steps as if one would apply it separately on S_{1_1} and S_{1_2}, but in an order dependent on the selections made in Step $k.2$, $k \geq 2$. Using the induction hypothesis, we see that (A1) produces
$$\binom{k-1}{n^- - 1} \text{ subsimplices generated from } S_{1_1} \text{ in } \mathcal{M}_S^+, \quad \binom{k-1}{n^+} \text{ in } \mathcal{M}_S^-,$$
$$\binom{k-1}{n^-} \text{ subsimplices generated from } S_{1_2} \text{ in } \mathcal{M}_S^+, \quad \binom{k-1}{n^+ - 1} \text{ in } \mathcal{M}_S^-,$$
and summing up using $k = n^+ + n^- - 1$ we see that
$$\begin{aligned} K^+ &= \binom{k-1}{n^- - 1} + \binom{k-1}{n^-} = \binom{k}{n^-} = \binom{n^+ + n^- - 1}{n^-} \\ K^- &= \binom{k-1}{n^+ - 1} + \binom{k-1}{n^+} = \binom{k}{n^+} = \binom{n^+ + n^- - 1}{n^+} \\ K = K^+ + K^- &= \binom{n^+ + n^- - 1}{n^+ - 1} + \binom{n^+ + n^- - 1}{n^+} = \binom{n^+ + n^-}{n^+} \end{aligned}$$

This completes the proof of Proposition 5. □

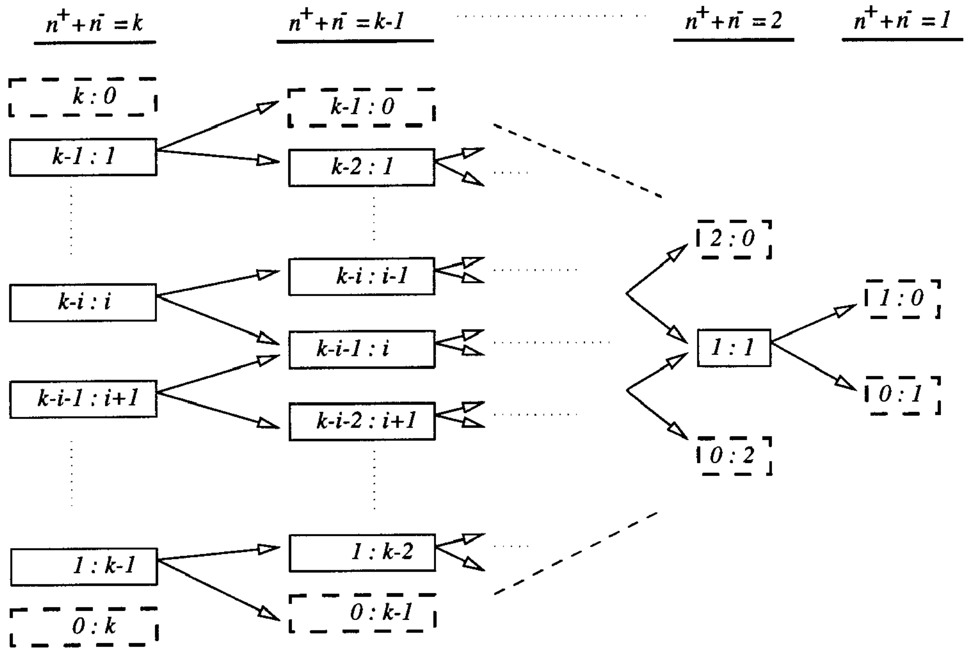

Figure 1. Graph G of subdivision

REMARKS:

(a) The subdivision process of Algorithm 1 can be visualized by a directed graph G as shown in Figure 1. A column in G consists of nodes $\boxed{i : j}$ representing the $k+1$ possible configurations of $i \geq 0$ vertices lying in V^+ and $j \geq 0$ vertices lying in V^-, if the sum $n^+ + n^-$ is fixed to $k \geq 1$. The leaves in G are nodes with either $i = 0$ or $j = 0$, representing a simplex S with $S \subset H^{\leq}$ or $S \subset H^{\geq}$, respectively. From every node $N = \boxed{i : j} \in G$ that is not a leaf, there are exactly two directed edges emanating from N, representing the bisection process. Starting from such a node, one finds the simplices generated by Algorithm 1 applied to a corresponding simplex S with $n^+(S) = i$, $n^-(S) = j$ by visiting all leaves in G reachable from N.

(b) We can prescribe any order of selection in Step k.2 of Algorithm 1, for example by choosing a specific data structure for the implementation of the set $\mathcal{L}$. Organizing $\mathcal{L}$ as a stack, i.e. retrieving elements of $\mathcal{L}$ in a last–in, first–out order, corresponds to a Depth First Search (DFS) leaf–visiting process in G.

(c) For our applications, it seems worthwhile to note that Algorithm 1 is only conceptual. Actual implementations can be greatly enhanced regarding storage requirements and runtime efficiency, for example by avoiding the (explicit) usage of the sets $\mathcal{L}, \mathcal{M}_S^+, \mathcal{M}_S^-$ and by using specialized storage schemes in Step k.5.

Let us clarify the last idea by giving an outline of a recursive variant of Algorithm 1 developed for the following purpose: given a simplex S and a hyperplane H irredundant for S, construct a simplicial partition of $S \cap H^{\leq}$. We will represent simplices by certain associated integral index sets. To be more specific, let

$$S = \left[u^1, \dots, u^{n^-(S)}, v^1, \dots, v^{n^+(S)}, w^1, \dots w^{n^=(S)}\right],$$

where $u^i \in V^-(S)$, $v^j \in V^+(S)$, $w^k \in V^=(S)$. We encode this initial simplex S by the triple $t_S := (k, m, I)$ where $k = n^-(S)$, $m = n^+(S)$, $I = \emptyset$. In general, the initially empty set

$$I = \left\{ (i_1, j_1), \dots, (i_{|I|}, j_{|I|}) \right\}$$

consists of index pairs referencing elements of the set

$$H_S := \left\{ h_{ij} := h(u^i, v^j, H) : 1 \leq i \leq n^-(S),\ 1 \leq j \leq n^+(S) \right\}$$

of intersection points h_{ij}. Bisecting any given simplex S represented by a triple (k, m, I) with respect to h_{km} into S_1, S_2 leads to the triple representations

$$S_1 = (k-1, m, I_1) \quad , \quad S_2 = (k, m-1, I_2)$$

where $I_1 = I_2 = I \cup \{(k, m)\}$. On the other hand, given a triple representation $t_{\widetilde{S}} := (i_1, i_2, J)$ one can reconstruct the associated simplex $\widetilde{S} = \widetilde{S}\,(i_1, i_2, J)$ as

$$\widetilde{S} = \left[u^1, \dots, u^{i_1}, v^1, \dots, v^{i_2}, w^1, \dots w^{n^=(S)}, h_{j_1 j_2} : (j_1, j_2) \in J\right].$$

procedure `Subdivide(`k, m, I, $\mathcal{M}_S^-$`)`
begin
  if $m = 0$
    $\mathcal{M}_S^- \leftarrow \mathcal{M}_S^- \cup S\,(k, m, I)$;
  else
    `Subdivide(`k, $m-1$, $I \cup \{(k, m)\}$, $\mathcal{M}_S^-$`)`;
    if $k > 1$
      `Subdivide(`$k-1$, m, $I \cup \{(k, m)\}$, $\mathcal{M}_S^-$`)`;
end

Figure 2. The Procedure `Subdivide(·)`

Using this notation, the recursive procedure `Subdivide(·)` shown in Figure 2 can be used to produce the set $\mathcal{M}_S^-$ by a top level call with initial parameters $k = n^-(S)$, $m = n^+(S)$, $I = \emptyset$ and $\mathcal{M}_S^- = \emptyset$. Used in such a manner, `Subdivide(·)` visits all leaves of the corresponding graph G representing simplices $\widetilde{S} \in \mathcal{M}_S^-$ in a DFS-order; by further usage of the index encoding scheme, it is relatively easy to derive a subdivision algorithm which widely avoids a multiple storage of the vertex coordinates.

Table I. Calling sequence for $n^-(S) = 2 = n^+(S)$

i	ℓ_i	k_i	m_i	I_i	S_i	$\mathcal{M}_S^-$
1	1	2	2	$\emptyset$	—	$\emptyset$
2	2	2	1	$\{(2,2)\}$	—	$\emptyset$
3	3	2	0	$\{(2,2),(2,1)\}$	$[u^1,u^2,h_{22},h_{21}]$	$\{S_3\}$
4	3	1	1	$\{(2,2),(2,1)\}$	—	$\{S_3\}$
5	4	1	0	$\{(2,2),(2,1),(1,1)\}$	$[u^1,h_{22},h_{21},h_{11}]$	$\{S_3,S_5\}$
6	2	1	2	$\{(2,2)\}$	—	$\{S_3,S_5\}$
7	3	1	1	$\{(2,2),(1,2)\}$	—	$\{S_3,S_5\}$
8	4	1	0	$\{(2,2),(1,2),(1,1)\}$	$[u^1,h_{22},h_{12},h_{11}]$	$\{S_3,S_5,S_8\}$

EXAMPLE 6. Let us illustrate the manner of working of `Subdivide`$(\cdot)$. Assume that S is a 3–simplex, and assume that H divides the vertex set of S into $V^-(S) = \{u^1,u^2\}$ and $V^+(S) = \{v^1,v^2\}$. Table I shows the calling sequence for the procedure `Subdivide`$(\cdot)$. Its first column gives the number i of the current call. In each row (call) i, ℓ_i denotes the current call level, which equals the current depth of the associated call tree, and k_i, m_i, I_i are the actual parameters. The S_i column gives the simplex $S_i := S(k_i, m_i, I_i)$, if one is generated in the i–th call, and the last column lists the state of the output set $\mathcal{M}_S^-$ just before call number $i+1$. Note at this point that k_i, m_i, and I_i are given by value to `Subdivide`$(\cdot)$, while $\mathcal{M}_S^-$ is given by reference.

The calling sequence in Table I applies for any 3–simplex S and hyperplane H satisfying $n^-(S) = 2 = n^+(S)$, and it applies for any higher dimensional example satisfying this condition: just add to every generated simplex in $\{S_3, S_5, S_8\}$ the vertex set $V^=(S)$. Therefore, at least in principle, it is even possible to precompute a database containing the corresponding index sets for each possible pair $(n^-(S), n^+(S))$ once and use it to generate from these indices the partition sets in $\mathcal{M}_S^-$.

3. Subdivisions with Minimal Cardinality

In this section, we give an answer to the natural question of what is the minimal number of simplices in a simplicial partition or triangulation of the polytope $P^\leq = S \cap H^\leq$ introduced in (1). It is obvious that we can restrict our considerations to the (full–dimensional) case when $n^-(S) \geq 1$: otherwise, $P^\leq = \text{conv}(V^=(S))$ holds, so that $P^\leq$ is a simplicial face of S if $V^=(S) \neq \emptyset$, and $P^\leq = \emptyset$, if $V^=(S) = \emptyset$. Also, the case of $P^\geq$ is completely symmetrical, so that there is no need to discuss it seperately.

We will prove that the procedures developed in Section 2 generate minimal partitions (and triangulations) in the sense described above. For this purpose, we will transform $P^\leq$ into a "standard"–polytope $E \subseteq \mathbb{R}^n$, which is *projectively* *equivalent* (see, e.g. [Gr67]) to $P^{\leq}$. For the sake of completeness, let us recall the definition and put together some basic properties of a projective transformation, tailored to our need (and each with our own direct verification, given in Appendix A, starting on page 26).

DEFINITION 7. Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a transformation defined by

$$f(x) := \frac{Ax + a}{c^T x + \gamma}, \tag{3}$$

where $A \in \mathbb{R}^{n \times n}$, $a, c \in \mathbb{R}^n$ and $\gamma \in \mathbb{R}$, with at least one of c and γ being different from 0. Then f is called a projective transformation from $\mathbb{R}^n$ into $\mathbb{R}^n$, defined on $\mathbb{R}^n \setminus N(f)$, where $N(f) := \{x \in \mathbb{R}^n : c^T x + \gamma = 0\}$. For $X \subseteq \mathbb{R}^n$, f is said to be permissible for X if $X \cap N(f) = \emptyset$. If the matrix

$$M_f := \begin{pmatrix} A & a \\ c^T & \gamma \end{pmatrix} \tag{4}$$

is regular, then f is called nonsingular.

LEMMA 8. *Let $X \subseteq \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}^n$ be a nonsingular projective transformation permissible for X. Let $Y := f(X) := \{y \in \mathbb{R}^n : y = f(x), x \in X\}$ be the image of X under f. Then f is invertible on X with inverse $g : Y \to X$, which is a nonsingular projective transformation permissible for Y. Moreover, if*

$$f^{-1}(y) = g(y) := \frac{By + b}{d^T y + \delta},$$

where f is given by (3), and $\eta = f(\xi)$ for some $\xi \in X$, then

$$d^T \eta + \delta = \left(c^T \xi + \gamma\right)^{-1}.$$

LEMMA 9. *Let $P \subset \mathbb{R}^n$ be a polytope with vertex set*

$$V(P) \subseteq X := \{x^1, \dots, x^k\} \subset P,$$

where $x^i \neq x^j$ for $i \neq j$. Let $f : \mathbb{R}^n \to \mathbb{R}^n$, $f(x) := \frac{Ax+a}{c^T x+\gamma}$ be a projective transformation permissible for X satisfying $c^T x^i + \gamma > 0$, $1 \leq i \leq k$, and define $Y := f(X) := \{f(x^1), \dots, f(x^k)\} =: \{y^1, \dots, y^k\}$. Then the image

$$P' := f(P) := \{y \in \mathbb{R}^n : y = f(x), x \in P\}$$

of P under f satisfies $P' = \mathrm{conv}(Y)$.

LEMMA 10. *In the situation of Lemma 9, assume that f is nonsingular. Then we have:*
(a) If $V(P) = X$, then $V(P') = Y$, and, if $V(P') = Y$, then $V(P) = X$.

(b) *If S_d is a d–simplex contained in P, then $f(S_d)$ is a d–simplex contained in P'. Conversely, if S'_d is a d–simplex contained in P', then $f^{-1}(S'_d)$ is a d–simplex contained in P.*

(c) *If $\mathcal{D} := \{S_i : i \in I\}$ is a simplicial partition of P, then $\mathcal{D}' := f(\mathcal{D}) := \{f(S) : S \in \mathcal{D}\}$ is a simplicial partition of P' with the same cardinality. Conversely, if $\mathcal{D}' := \{S'_i : i \in I'\}$ is a simplicial partition of P', then $\mathcal{D} := f^{-1}(\mathcal{D}') := \{f^{-1}(S') : S' \in \mathcal{D}'\}$ is a simplicial partition of P with the same cardinality.*

Using these results, we are now able to projectively transform $P^{\leq}$ into an n–polytope $E \subseteq \mathbb{R}^n$, whose vertex set is suitably contained in $V(C_n)$, where C_n denotes the unit cube in $\mathbb{R}^n$. More precisely, using the notation of Section 2, we obtain:

PROPOSITION 11. *Consider a given pair $(S, H^{\leq})$, and let $P^{\leq} := S \cap H^{\leq}$. Let*

$$k := n^-(S) - 1 \geq 0, \quad l := n^+(S) \geq 0 \ \text{ and } n^=(S) \geq 0. \tag{5}$$

Then there exists a nonsingular projective transformation $f : \mathbb{R}^n \to \mathbb{R}^n$ permissible for $P^{\leq}$ with

$$f(P^{\leq}) = \operatorname{conv}(\{0, e_1, \ldots, e_n\} \cup \{e_i + e_j : i \in I,\ j \in J\}) =: E_{k,l}\,, \tag{6}$$

where $e_1, \ldots, e_n$ denote the unit vectors in $\mathbb{R}^n$,

$$I := \{i : 1 \leq i \leq k\}, \quad J := \{j : k+1 \leq j \leq k+l\}$$

and $\{e_i + e_j : i \in I,\ j \in J\} = \emptyset$ if $I = \emptyset$ (i.e. $k = 0$) or $J = \emptyset$ (i.e. $l = 0$). Moreover, this representation of $E_{k,l}$ is unshrinkable, i.e.,

$$V(E_{k,l}) = \{0, e_1, \ldots, e_n\} \cup \{e_i + e_j : i \in I,\ j \in J\}. \tag{7}$$

Proof. Let $S = [v^0, \ldots, v^n]$ and $H^{\leq} = \{x \in \mathbb{R}^n : h(x) = a^T x - b \leq 0\}$. Assume an ordering of $V(S)$ satisfying

$$\begin{aligned} V^-(S) &= \{v^0, \ldots, v^k\}, \\ V^+(S) &= \{v^{k+1}, \ldots, v^{k+l}\}, \\ V^=(S) &= \{v^{k+l+1}, \ldots, v^n\}\,. \end{aligned}$$

Here we understand $V^+(S) = \emptyset$, if $l = 0$, $V^=(S) = \emptyset$, if $k + l = n$, i.e. $n^=(S) = 0$. Note that $V^-(S) \neq \emptyset$ by (5). Then the vertex set of $P^{\leq}$ is given by

$$V(P^{\leq}) = V^-(S) \cup V^=(S) \cup \{h_{ij} : i \in I \cup \{0\},\ j \in J\}\ ,$$

where $h_{ij} := v^i + \lambda_{ij}(v^j - v^i)$ for some $\lambda_{ij} \in (0,1)$, $i \in I \cup \{0\}$, $j \in J$, if $J \neq \emptyset$. Let $h_j := h(v^j)$, $0 \leq j \leq n$. Since $h(h_{ij}) = 0$, the scalars λ_{ij} are given by the equation $h_i + \lambda_{ij}(h_j - h_i) = 0$, i.e.,

$$\lambda_{ij} = \frac{h_i}{h_i - h_j} \quad \text{for } \ i \in I \cup \{0\},\ j \in J, \tag{8}$$

satisfying $\lambda_{ij} \in (0,1)$, since $h_i < 0$ and $h_j > 0$. Define

$$V := \left(v^1 - v^0, \dots, v^n - v^0\right) \in \mathrm{I\!R}^{n \times n}, \tag{9}$$

$$D := \operatorname{diag}\left(\tfrac{h_1}{h_0}, \dots, \tfrac{h_k}{h_0}, \quad \tfrac{-h_{k+1}}{h_0}, \dots, \tfrac{-h_{k+l}}{h_0}, \quad 1, \dots, 1\right) \in \mathrm{I\!R}^{n \times n}, \tag{10}$$

$$b^T := \left(\tfrac{h_1 - h_0}{h_0}, \dots, \tfrac{h_k - h_0}{h_0}, \quad -1, \dots, -1, \quad 0, \dots, 0\right) \in \mathrm{I\!R}^{k+l+m}, \tag{11}$$

where, in (11), $m := n^{=}(S)$, i.e. $k + l + m = n$. In (10) and (11), we understand that the first, second or third group of columns vanishes if $k = 0$, $l = 0$ or $m = 0$, respectively. Consider the projective transformation $f : \mathrm{I\!R}^n \to \mathrm{I\!R}^n$ given by

$$f(x) := \frac{DV^{-1}(x - v^0)}{b^T V^{-1}(x - v^0) + 1}. \tag{12}$$

(a) f is nonsingular, since the corresponding matrix M_f can be written as the product of two nonsingular matrices:

$$M_f = \begin{pmatrix} DV^{-1} & -DV^{-1}v^0 \\ b^T V^{-1} & 1 - b^T V^{-1} v^0 \end{pmatrix} = \begin{pmatrix} D & 0 \\ b^T & 1 \end{pmatrix} \cdot \begin{pmatrix} V^{-1} & -V^{-1}v^0 \\ 0^T & 1 \end{pmatrix}.$$

(b) f is permissible for $P^{\leq}$. To see this, let $\ell(x) := b^T V^{-1}(x - v^0) + 1$, $b^T = (b_1, \dots, b_n)$ and $b_0 := 0$. Then, for the vertices of $P^{\leq}$, we have, by (9), (11), $h_i < 0\, \forall i \in \{0, \dots, k\}$ and $\lambda_{ij} \in (0,1)$:

$$\ell\left(v^i\right) = b_i + 1 = \begin{cases} 1 > 0 & \text{for } k + l + 1 \leq i \leq n \\ h_i / h_0 > 0 & \text{if } i \in \{0, \dots, k\} \end{cases} \tag{13a}$$

$$\begin{aligned} \ell(h_{ij}) &= \ell\left(v^0 + (1 - \lambda_{ij})(v^i - v^0) + \lambda_{ij}(v^j - v^0)\right) \\ &= (1 - \lambda_{ij})\, b_i + \lambda_{ij} b_j + 1 && (13b) \\ &= (1 - \lambda_{ij})(b_i + 1) && (13c) \\ &= (1 - \lambda_{ij}) \frac{h_i}{h_0} && (13d) \\ &> 0, && \text{for } (i,j) \in (I \cup \{0\}) \times J, \text{ if } J \neq \emptyset \end{aligned}$$

Therefore, $\ell(x) > 0\ \forall x \in P^{\leq}$. (Note that (13c) follows from (13b) since $b_j = -1$ for $j \in J$ by (11)).

By (a) and (b), the requirements of Lemmas 9 and 10 are fulfilled, so that $f\left(P^{\leq}\right)$ is the polytope with vertex set

$$V(f(P^{\leq})) = f(V(P^{\leq})).$$

Using (8)–(12), (13a), (13d), and the notation $D = (d_1, \dots, d_n)$, $d_0 := 0 \in \mathrm{I\!R}^n$, we see that $f(V(P^{\leq}))$ is given by

$$f\left(v^i\right) = \frac{d_i}{\ell\left(v^i\right)} = \begin{cases} \frac{1}{\ell(v^0)} 0 = 0 & \text{if } i = 0 \\ \frac{h_i/h_0}{h_i/h_0} e_i = e_i & \text{if } i \in I, I \neq \emptyset \end{cases}$$

$$f\left(v^i\right) = \frac{d_i}{\ell\left(v^i\right)} = \frac{e_i}{1} = e_i\ , \quad \text{for } k+l+1 \le i \le n$$

$$\begin{aligned} f\left(h_{ij}\right) &= f\left(v^0 + (1-\lambda_{ij})\left(v^i - v^0\right) + \lambda_{ij}(v^j - v^0)\right) \\ &= \frac{(1-\lambda_{ij})\, d_i + \lambda_{ij} d_j}{(1-\lambda_{ij})\,\frac{h_i}{h_0}} & \text{(14a)} \\ &= \frac{h_0}{h_i} d_i + \frac{h_i}{h_i - h_j}\cdot\frac{h_i - h_j}{-h_j}\cdot\frac{h_0}{h_i}\cdot\frac{-h_j}{h_0} e_j & \text{(14b)} \\ &= \begin{cases} 0 + e_j = \ e_j & \text{if } i = 0,\ j \in J,\, J \neq \emptyset \\ \frac{h_0}{h_i}\frac{h_i}{h_0} e_i + e_j = \ e_i + e_j & \text{if } i \in I,\ j \in J, I \neq \emptyset \neq J \end{cases} \end{aligned}$$

Here, (14a) follows from (12), (13d), while (14b) follows from (8), (10). This completes the proof of Proposition 11. □

To show our main result concerning the minimality of the triangulation of $P^{\leq}$ developed in Section 2, let us now consider simplicial partitions of $E_{k,l}$ (see also Ref. [Ha91], Lemma 2):

PROPOSITION 12. *Every simplicial partition* $\mathcal{D} := \{S_i : i \in I\}$ *of* $E_{k,l}$ *satisfying*

$$V\left(S_i\right) \subseteq V\left(E_{k,l}\right)\ \forall i \in I \tag{15}$$

has the cardinality $c := n! \cdot \operatorname{vol}\left(E_{k,l}\right)$*, where* $\operatorname{vol}(K)$ *denotes the volume of an* n*-dimensional polytope* $K \subseteq \mathbb{R}^n$*. Moreover, if* $\mathcal{D}'$ *is a simplicial partition of* $E_{k,l}$ *for which (15) does not hold, then*

$$\left|\mathcal{D}'\right| \geq c. \tag{16}$$

Proof. To show the first assertion, let S be an n–simplex contained in $E_{k,l}$ satisfying $V(S) \subseteq V\left(E_{k,l}\right)$. If $S = \left[v^0, \ldots, v^n\right]$, $V := \left(v^0, \ldots, v^n\right)$ and $\mathbf{1}_k := (1, \ldots, 1) \in \mathbb{R}^k$, then it is well–known that

$$n! \cdot \operatorname{vol}(S) = \left|\det\begin{pmatrix} V \\ \mathbf{1}_{n+1}\end{pmatrix}\right| .$$

We aim to show that, for every possible choice of S, we have

$$\left|\det\begin{pmatrix} V \\ \mathbf{1}_{n+1}\end{pmatrix}\right| = 1\,. \tag{17}$$

For this, consider first the case $n = 2$: we have $V(S) \subseteq \{0, e_1, e_2, e_1 + e_2\} = V\left(C_2\right)$, where C_2 is the unit square in $\mathbb{R}^2$. Since every triangle having its vertices in $V\left(C_2\right)$ has area $\frac{1}{2}$, the assertion follows.

For $n \geq 3$, assume first that $0 \in V(S)$. Then

$$\left|\det\begin{pmatrix} V \\ \mathbf{1}_{n+1} \end{pmatrix}\right| = \left|\det\begin{pmatrix} 0 & V' \\ 1 & \mathbf{1}_n \end{pmatrix}\right| = \left|\det(V')\right| = 1 ,$$

since V' is a regular submatrix of V, and V is totally unimodular (cf., e.g., Ref. [PaSt82], Theorem 13.3). If $0 \notin V(S)$, we must have $e_i \in V(S)$ for at least one $i \in \{1, \ldots, n\}$. To see this, assume the contrary. Then

$$V(S) \subseteq \{e_i + e_j : i \in I,\ j \in J\} \subseteq \{x \in \mathbb{R}^n : \mathbf{1}_n^T x = 2\},$$

and hence the vertices of S would be affinely dependent, contradicting the assumption $\dim(S) = n$. So assume without loss of generality $v^0 = e_i$. Let $V' := (v^1, \ldots, v^n)$, so that

$$\left|\det\begin{pmatrix} V \\ \mathbf{1}_{n+1} \end{pmatrix}\right| = \left|\det\begin{pmatrix} e_i & V' \\ 1 & \mathbf{1}_n \end{pmatrix}\right| ,$$

and denote by $V'_r := \begin{pmatrix} V'_{r1} \\ V'_{r2} \end{pmatrix}$ the matrix obtained by deleting the r–th row of V'. We consider two cases:

(a) If $v_i^j = 0$ for all j satisfying $1 \leq j \leq n$, then, with $\mathbf{0}_n := (0, \ldots, 0) \in \mathbb{R}^n$:

$$\left|\det\begin{pmatrix} e_i & V' \\ 1 & \mathbf{1}_n \end{pmatrix}\right| = \left|\det\begin{pmatrix} & V'_{i1} \\ e_i & \mathbf{0}_n \\ & V'_{i2} \\ 1 & \mathbf{1}_n \end{pmatrix}\right| = \left|\det\begin{pmatrix} V'_i \\ \mathbf{1}_n \end{pmatrix}\right| .$$

(b) If $v_i^k = 1$ for some $k \in \{1, \ldots, n\}$, then $v^k = e_i + e_j$ for some $j \in J$, since otherwise $v^k = v^0$. Let v'_j denote the j–th row of V'. Subtracting column k from column 0 gives

$$\left|\det\begin{pmatrix} e_i & V' \\ 1 & \mathbf{1}_n \end{pmatrix}\right| = \left|\det\begin{pmatrix} & V'_{j1} \\ -e_j & v'_j \\ & V'_{j2} \\ 0 & \mathbf{1}_n \end{pmatrix}\right| = \left|\det\begin{pmatrix} V'_j \\ \mathbf{1}_n \end{pmatrix}\right| .$$

In both cases, we see that (17) follows by induction, so that $\mathrm{vol}(S) = \frac{1}{n!}$.

Now let $\mathcal{D} := \{S_i : i \in I\}$ be a simplicial partition of $E_{k,l}$ satisfying (15). Then $\sum_{i \in I} \mathrm{vol}(S_i) = \mathrm{vol}(E_{k,l})$, and from $\mathrm{vol}(S_i) = \frac{1}{n!}\ \forall i \in I$ we conclude that $|I| = n! \cdot \mathrm{vol}(E_{k,l}) = c$, i.e., the first assertion of Proposition 12.

Regarding the second assertion, consider an n–simplex $S = [v^0, \ldots, v^n]$ with $V(S) \subseteq E_{k,l}$, and assume that $\rho(S) := \left|\{v^i : v^i \notin V(E_{k,l})\}\right| > 0$. Assume without loss of generality $v^0 \notin V(E_{k,l})$. Then the function $\Psi_S : \mathbb{R}^n \to \mathbb{R}$ defined by

$$\Psi_S(x) := |\psi_S(x)| := \left|\det\begin{pmatrix} x & v^1 & \cdots & v^n \\ 1 & 1 & \cdots & 1 \end{pmatrix}\right|$$

is convex, since $\psi_S(x)$ is affine. Therefore $\max\{\Psi_S(x) : x \in E_{k,l}\} = \Psi_S(x^0)$ for some $x^0 \in V(E_{k,l})$, so that $\mathrm{vol}(S) \leq \mathrm{vol}(S')$, where $S' := [x^0, v^1, \dots, v^n]$. Hence for any simplex $S \subseteq E_{k,l}$ with $\rho(S) > 0$ we can find a simplex $S' \subseteq E_{k,l}$ satisfying $\rho(S') = \rho(S) - 1$ and $\mathrm{vol}(S) \leq \mathrm{vol}(S')$. Using induction we see that there is a simplex S'' with $V(S'') \subseteq V(E_{k,l})$ and $\mathrm{vol}(S) \leq \mathrm{vol}(S'') = \frac{1}{n!}$. Therefore, for any simplicial partition $\mathcal{D}'$ of $E_{k,l}$ we have $\mathrm{vol}(E_{k,l}) = \sum_{S \in \mathcal{D}'} \mathrm{vol}(S) \leq |\mathcal{D}'| \cdot \frac{1}{n!}$, which is equivalent to (16). □

Now it becomes easy to state the main result of this section:

THEOREM 13. *Let $\mathcal{D} = \{S_i : i \in I\}$ be a simplicial partition of $P^{\leq} = S \cap H^{\leq}$. If $V(S_i) \subseteq V(P^{\leq})$ $\forall i \in I$, then*

$$|I| = \binom{n^+(S) + n^-(S) - 1}{n^+(S)} =: d\left(S, H^{\leq}\right) =: d,$$

and d is the minimal cardinality of any arbitrary simplicial partition of $P^{\leq}$. In particular, the triangulation generated by Algorithm 1 (or the procedure `Subdivide`*) is minimal in this sense.*

Proof. The partition $\mathcal{P}$ of $P^{\leq}$ developed in Section 2 uses d n–simplices, and satisfies $V(S) \subseteq V(P^{\leq})$ $\forall S \in \mathcal{P}$. The Theorem follows immediately by applying Lemma 10 (c), Proposition 11 and Proposition 12, taking notice of the fact that the projective transformation used in Proposition 11 maps vertices of $P^{\leq}$ onto vertices of $E_{k,l}$. □

4. Application to Concave Minimization

In this section, based on the subdivision procedure of Section 2, a finite simplicial branch and bound algorithm is derived for the problem

$$\min_{x \in P} f(x), \tag{18}$$

where $P \subseteq \mathbb{R}^n$ is a full–dimensional polytope and $f : D \to \mathbb{R}$ is concave on a suitable set $D \subseteq \mathbb{R}^n$ containing P.

4.1. Basic Operations

4.1.1. *Methods to Compute an Initial Simplex*

The algorithm starts with a simplicial approximation $S \supseteq P$ of the polytope P. Assume that P is given in inequality representation

$$P = \{x \in \mathbb{R}^n : a_i^T x \leq b_i, \ i \in I\}, \tag{19}$$

where I is a finite index set, and $a_i \in \mathbb{R}^n,\ b_i \in \mathbb{R},\ i \in I$. Let u be a vertex of P. Then there exists an index set $I(u) = \{i \in I : a_i^T u = b_i\}$ such that $|I(u)| = n$, and the vectors $a_i,\ i \in I$, are linearly independent. We propose to choose

$$S = \{x : a_i^T x \le b_i,\ i \in I(u),\ d^T x \le \alpha\}, \tag{20}$$

where $d = -\frac{1}{n} \sum\limits_{i \in I(u)} a_i$ and $\alpha = \max\{d^T x : x \in P\}$ (cf., e.g., [HoTu93] and [HoTh88]). If the vertex u is nondegenerate, then the set $I(u)$ is unique; a detailed discussion of the degenerate case is given in [HoTu93], Chapter 3. Geometrically, the simplex S in (20) is the intersection of a polyhedral cone K with vertex u and n edges with the hyperplane $\{x : d^T x = \alpha\}$. The vertex set $V(S)$ of S is easily determined.

If $P \subset \mathbb{R}^n_+$, a simpler but usually much less tight initial simplex is given by

$$S = \{x \in \mathbb{R}^n_+ : e^T x \le \beta\}, \tag{21}$$

where $e^T = (1, \dots, 1) \in \mathbb{R}^n$, and $\beta = \max\{e^T x : x \in P\}$.

Finally, if the inequality representation of P includes upper and lower bounds $l \le x \le u$ (i.e. we have a rectangle R containing P), then it is always possible to triangulate R into $n!$ simplices in the standard way (cf. [Tod76]). In fact, $n!$ is the maximum number of simplices in any triangulation of R, and more efficient triangulation methods are known in the literature (cf., e.g. [Ha91] and references therein), especially for small dimensions $n \le 6$ (cf. [Hu93], [HuAn93]). Starting with a triangulation of R might be useful when n is small or if $V(R)$ contains many feasible vertices, see the following discussion of the bounding procedures. Notice that each simplicial branch–and–bound method can be modified in an obvious way to start with a finite union $\bigcup_{i=1}^r S_i$ of simplices satisfying $P \subseteq \bigcup_{i=1}^r S_i$, $\operatorname{int} S_k \cap \operatorname{int} S_j = \emptyset \quad \forall k, j \in \{1, \dots, r\},\ k \ne j$.

4.1.2. *Lower Bounds*

Let $S = [v^0, \dots, v^n]$ be an n–simplex, P a polytope given in the form (19) and let $f : S \longrightarrow \mathbb{R}$ be concave on S. Then to calculate a lower bound for $f^\star(S) := \min\{f(x) : x \in S \cap P\}$, we propose to compute either

$$\mu_1(S) := \min_{0 \le j \le n} f(v^j) \tag{22}$$

or, at the expense of solving a linear program,

$$\begin{aligned} \mu_2(S) \ := \ & \min_{\lambda \in \mathbb{R}^{n+1}} \ \textstyle\sum_{j=0}^n \lambda_j f(v^j) \\ & \text{s.t.} \quad \lambda \ge 0,\ e^T\lambda = 1,\ a_i^T V \lambda \le b_i,\ i \in I_S \end{aligned} \tag{23}$$

where $e := (1, \dots, 1)^T \in \mathbb{R}^{n+1}$, $V := (v^0, \dots, v^n) \in \mathbb{R}^{n \times (n+1)}$,

$$I_S := \left\{ i \in I : \exists j \in \{0, \dots, n\} : a_i^T v^j > b_i \right\}, \tag{24}$$

and $\mu_2(S) = \infty$ if (23) has no feasible solution (cf. also [FaHo76], [Ho76]).

PROPOSITION 14. *Both* $\mu_1(S)$ *and* $\mu_2(S)$ *as defined above are valid lower bounds for* $f^\star(S)$. *If* $\tilde{S}$ *is an* n*–simplex containing* S, *then* $\mu_j(S) \geq \mu_j(\tilde{S})$, $j = 1, 2$. *Moreover,* $\mu_2(S) \geq \mu_1(S)$ *and, if* $I_S = \emptyset$, *then* $\mu_1(S) = \mu_2(S) = f^\star(S)$.

Proof. From the theory of convex envelopes (cf., e.g., [HoTu93] and references there), we know that the best convex function $\varphi_S(x)$ underestimating $f(x)$ on S is given by the unique affine function satisfying

$$\varphi_S(v^j) = f(v^j), \;\; 0 \leq j \leq n \; .$$

Using barycentric coordinates $\lambda \in \mathbb{R}^{n+1}$, i.e.

$$x \in S \Longleftrightarrow x = V\lambda, \; \lambda \geq 0, \; e^T\lambda = 1,$$

we see that, since $I = I_S + \left\{ i \in I : a_i^T x \leq b_i \; \forall x \in S \right\}$:

$$\begin{aligned} f^\star(S) &\geq \min\left\{\varphi_S(x) : x \in S \cap P\right\} && \text{(25a)} \\ &= \min_{\lambda \in \mathbb{R}^{n+1}} \left\{ \textstyle\sum_{j=0}^{n} \lambda_j f(v^j) : \lambda \geq 0, \; e^T\lambda = 1, \; a_i^T V\lambda \leq b_i, \; i \in I \right\} \\ &= \mu_2(S) \\ &\geq \min_{\lambda \in \mathbb{R}^{n+1}} \left\{ \textstyle\sum_{j=0}^{n} \lambda_j f(v^j) : \lambda \geq 0, \; e^T\lambda = 1 \right\} && \text{(25b)} \\ &= \min_{0 \leq j \leq n} f(v^j) \\ &= \mu_1(S) \end{aligned}$$

Now let $S \subseteq \tilde{S}$. Then, by concavity of $f(x)$, we have

$$\mu_1(S) = \min f(V(S)) = \min f(S) \geq \min f(\tilde{S}) = \min f(V(\tilde{S})) = \mu_1(\tilde{S}),$$

and

$$\mu_2(S) = \min\{\varphi_S(x) : x \in S \cap P\} \geq \min\{\varphi_{\tilde{S}}(x) : x \in \tilde{S} \cap P\} = \mu_2(\tilde{S}),$$

since $\varphi_S(x) \geq \varphi_{\tilde{S}}(x) \; \forall x \in S$ by definition of the convex envelope and $S \cap P \subseteq \tilde{S} \cap P$. Finally, if $I_S = \emptyset$, then $S \subseteq P$ and equality holds in (25a), (25b) because of $\min f(S) = \min \varphi_S(S)$. □

4.1.3. *Upper Bounds*

For any simplex S generated in the course of solving problem (18), let $Q(S) := V(S) \cap P$ be the set of feasible vertices of S. If (23) is used for lower bounding, then add to $Q(S)$ the feasible optimal solution obtained when calculating $\mu_2(S) < \infty$. Obviously, the (possibly infinite) number $\gamma(S) := \min\{f(x) : x \in Q(S)\}$ yields an upper bound for $f^\star(S)$, and

$$\gamma := \min_{S \in \mathcal{S}} \gamma(S), \tag{26}$$

taken over the set $\mathcal{S}$ of all generated simplices, is an upper bound for $\min f(P)$. Note that, if $V(S) \subseteq P$, then $f^\star(S) = \gamma(S)$ by concavity of $f(x)$.

DEFINITION 15. Used in conjunction with $\gamma(S)$ as defined above, a lower bounding rule $\mu(S)$ is called exact (for feasible simplices), iff $\mu(S) = f^\star(S) = \gamma(S)$ holds whenever $V(S) \subseteq P$.

It is obvious from Proposition 14, that both $\mu_1(S)$ and $\mu_2(S)$ are exact in this sense. Moreover, every reasonable lower bounding rule $\mu(S)$ for problem (18) will be exact, at least if one combines it for example with $\mu_1(S)$ by using $\hat{\mu}(S) := \max\{\mu_1(S), \mu(S)\}$ instead of $\mu(S)$. Therefore, *we will assume in the following that* $\mu(S)$ *denotes an exact lower bounding rule.*

4.1.4. *Subdivision of Simplices, Deletion by Infeasibility*

Let $S = [v^0, \ldots, v^n]$ be an n–simplex with lower bound $\mu(S) < \gamma$, where γ is the upper bound defined in (26). Then $\mu(S) < \gamma(S)$ and $I_S \neq \emptyset$, since $I_S = \emptyset$ would imply $S \subseteq P$ and therefore $\mu(S) = f^\star(S) = \gamma(S) \geq \gamma$. Choose $i \in I_S$ and subdivide S with respect to the cutting plane $H_i := \{x \in \mathbb{R}^n : a_i^T x - b_i = 0\}$ into the n–simplices contained in the set $\mathcal{M}_S^- = \mathcal{M}_S^-(H_i)$, using, for example, the procedure `Subdivide()` developed in Section 2. Let $\mathcal{M}_S^- = \{S_1, \ldots, S_\ell\}$, where $\ell = K^-$ is given by Proposition (5)(iii). Then we propose to replace S by $\{S_1, \ldots, S_\ell\}$. Note that, if for some $i_0 \in I_S$, one has $V(S) \subseteq H_{i_0}^{\geq}$, the corresponding set $\mathcal{M}_S^-$ is empty. In this case, by Proposition (5)(iii), we have $\ell = 0$, and we propose to eliminate S from the set of simplices under consideration without any further subdivision. Applying this (implicit) deletion rule, one eliminates partition sets S with $S \cap P \subseteq \partial P$, i.e. one cuts off at most boundary points of P.

4.2. THE ALGORITHM

Here is the simplicial algorithm for solving Problem (18). It uses the basic operations defined in Section 4.1:

Algorithm 2 (A2):

Iteration 0:

Determine an initial n–simplex $S_0 \supseteq P$, the lower bound $\mu(S_0)$ and the set $Q(S_0)$. Set $Q_0 \leftarrow Q(S_0)$, $\gamma_0 \leftarrow \min\{f(x) : x \in Q_0\}$ and choose $y_0 \in Q_0$ satisfying $f(y_0) = \gamma_0$, if $Q_0 \neq \emptyset$. Set $\mathcal{P}_0 \leftarrow \{S_0\}$, $\mu_0 \leftarrow \mu(S_0)$, $k \leftarrow 1$.

Iteration k:

$k.1$: If $\gamma_{k-1} = \mu_{k-1}$, then stop. (y_{k-1} is an optimal solution to Problem (18) with optimal function value γ_{k-1})

$k.2$: Select $S_k \in \mathcal{P}_{k-1}$ satisfying $\mu(S_k) = \mu_{k-1}$.

$k.3$: We have $I_{S_k} \neq \emptyset$, since $\mu(S_k) < \gamma_{k-1}$. Choose $i_k \in I_{S_k}$ and compute the set $\mathcal{M}_{S_k}^-$ with respect to the cutting plane H_{i_k} (as described in Section 4.1.4). Let $\ell := |\mathcal{M}_{S_k}^-|$, $\mathcal{M}_{S_k}^- = \{S_{k_1}, \ldots, S_{k_\ell}\}$. Compute the lower bounds $\mu(S_{k_j})$ for $1 \leq j \leq \ell$.

$k.4$: Set

$$
\begin{aligned}
Q_k &\leftarrow Q_{k-1} \cup \textstyle\bigcup_{j=1}^{\ell} Q(S_{k_j}),\\
\mathcal{P}_k &\leftarrow \mathcal{P}_{k-1} \setminus \{S_k\} \cup \textstyle\bigcup_{j=1}^{\ell} \{S_{k_j}\},\\
\gamma_k &\leftarrow \min\{f(x) : x \in Q_k\},\\
\mu_k &\leftarrow \min\{\mu(S) : S \in \mathcal{P}_k\}.
\end{aligned}
$$

If $\gamma_k < \infty$, choose $y_k \in Q_k$ satisfying $f(y_k) = \gamma_k$.

$k.5$: Set $\mathcal{P}_k \leftarrow \mathcal{P}_k \setminus \{S \in \mathcal{P}_k : \mu(S) \geq \gamma_k\}$. If $\mathcal{P}_k = \emptyset$, set $\mu_k \leftarrow \gamma_k$.

$k.6$: Set $k \leftarrow k + 1$ and go to Iteration k.

4.3. Convergence

The following Lemma shows that the implicit deletion rule described in Section 4.1.4 does not cut off any feasible point from the current partition $\mathcal{P}_k$ which is better than the current best known solution y_k.

LEMMA 16. *Let, in Problem (18), f be continuous on P. Then, in any iteration $K \geq 0$ of Algorithm 2, we have*

$$\{x \in P : f(x) < \gamma_K\} \subseteq \mathcal{P}_K \ .$$

Proof. For $K = 0$ we have $P \subseteq S_0 \in \mathcal{P}_0$. Assume to the contrary, that $K \geq 1$, $x \in P$, $x \notin S \;\forall S \in \mathcal{P}_K$ and $f(x) < \gamma_K$. By continuity of f and $\dim P = n$, we can find a sequence $\{x_\nu\}_{\nu=1}^{\infty}$ with $x_\nu \in \operatorname{int} P$, $f(x_\nu) < \gamma_K$ and $\lim_{\nu\to\infty} x_\nu = x$. We conclude that, for every ν, no simplex S with $x_\nu \in S \in \mathcal{P}_k$ is deleted in any iteration $k \leq K$. Deletion of such a simplex cannot occur neither in Step $k.3$, since the rule described in Section 4.1.4 cuts off at most boundary points of P, nor in Step $k.5$, since $\mu(S) \leq f(x_\nu) < \gamma_K \leq \gamma_k$ (note that the sequence $\{\gamma_k\}$ is nonincreasing by construction). Therefore, for every x_ν we can find a simplex $S_\nu \in \mathcal{P}_K$ containing x_ν. Since $|\mathcal{P}_K| < \infty$, there is an $S \in \mathcal{P}_K$ with $\{x_q\} \subseteq S$, where $\{x_q\}$ is a suitable infinite subsequence of $\{x_\nu\}$. It follows that $x = \lim_{q\to\infty} x_q \in S \in \mathcal{P}_K$, since S is closed. □

In the following, let, for $x \in \mathbb{R}$, $\lfloor x \rfloor$ denote the greatest integer less than or equal to x.

THEOREM 17. *In Problem (18), let f be continuous on P, where the polytope P is given in the form (19) with $|I| = m$. Then Algorithm 2 stops at iteration $K+1$ yielding an exact solution y_K with $\mu_K = \gamma_K = f(y_K)$. An upper bound for the number of iterations is given by*

$$K \leq \frac{M^m - 1}{M-1}, \tag{27}$$

the total number N of simplices generated satisfies

$$N \leq \frac{M^{m+1}-1}{M-1}, \tag{28}$$

and the maximal size of any partition $\mathcal{P}_k$ is bounded by

$$|\mathcal{P}_k| \leq M^{m-1} + M - 1, \tag{29}$$

where $M = \binom{n}{\lfloor \frac{n}{2} \rfloor}$.

Proof. A directed graph G can be associated with Algorithm 2 in a natural way. The nodes of G consist of S_0 and all partition elements generated by the procedure. Two nodes S_i, S_j are connected by an arc (i,j) if and only if S_j is obtained by a direct partition of S_i in some Step k.3, i.e., $S_j \in \mathcal{M}^-_{S_i}(H_{i_k})$. Obviously, in terms of graph theory, G is a rooted tree with root S_0. A path in G corresponds to a decreasing sequence $\{S_{k_q}\}$ of successively refined partition sets. For any such path in G, the corresponding sequence $\{I_{k_q}\} := \{I_{S_{k_q}}\}$ satisfies

$$I_{k_{q+1}} \subseteq I_{k_q} \quad \text{and} \quad \left|I_{k_{q+1}}\right| \leq \left|I_{k_q}\right| - 1 , \tag{30}$$

since $S_{k_{q+1}} \subseteq \mathcal{M}^-_{S_{k_q}}(H_{i_{k_q}})$ by the rules proposed in Section 4.1.4.

Assume that there is an infinite sequence $\{S_{k_q}\}$, and let p denote the path in G associated with it. Since, with $S_{k_0} = S_0$ we have $|I_{k_0}| \leq |I| = m$, it follows from (30) that $I_{k_v} = \emptyset$ for some $v \leq m$. By (26) and since μ is exact, we have $\mu(S_{k_v}) = f^\star(S_{k_v}) = \gamma(S_{k_v}) \geq \gamma_k$, where k is the iteration where S_{k_v} is generated. Therefore, S_{k_v} gets deleted in Step k.5 of Algorithm 2, and p has a length $\ell(p) \leq m$, a contradiction. Hence every path p in G satisfies $\ell(p) \leq m$.

From Proposition (5)(iii), we know that, for any node $S \in G$, the number of immediate descendants is bounded by

$$K^-(S) = \binom{n^+(S)+n^-(S)-1}{n^+(S)} \leq \binom{n}{\lfloor \frac{n}{2} \rfloor} = M; \tag{31}$$

hence the total number $d(S_0)$ of descendants from S_0 satisfies $d(S_0) \leq \sum_{i=1}^m M^i$, which demonstrates (28).

On the other hand, there is a one to one correspondence between the iteration steps in Algorithm 2 and the nodes in G having a direct descendant, and therefore after $K \leq \sum_{i=1}^{m-1} M^i$ iteration steps we have $\mathcal{P}_K = \emptyset$ in Step K.5, i.e. we stop in Step $(K+1)$.1 with $\gamma_K = \mu_K$.

The maximal possible partition size is attained if G has a maximal number of M^i nodes in every layer i, for $0 \leq i \leq m$. Since every node in the last layer gets deleted immediately after its generation in Step $k.5$, (29) follows for the case when the selection in Step $k.2$ is done in a Breadth First Search manner in G and a worst–case problem is encountered.

Finally, from Lemma 16 we conclude that $\{x \in P : f(x) < \gamma_K\} \subseteq \mathcal{P}_K = \emptyset$, i.e. y_K solves Problem (18). □

REMARKS:

(a) Following the arguments in the proof of Theorem 17, we see that we are totally free in the choice of $i_k \in I_{S_k}$ in Step $k.3$ of Algorithm 2. For example, one can choose a cutting hyperplane H_{i_k} which generates a minimal number of new simplices $\ell = \ell(S_k, H_{i_k}) = \left|\mathcal{M}^-_{S_k}(H_{i_k})\right|$, to avoid an excessive growth of partition sets in the current Step k. Note that ℓ can be computed in advance based on the knowledge of $n^+(S)$, $n^-(S)$ by evaluating the left equation in (31). Another possible objective would be to choose a hyperplane which cuts off as many infeasible vertices $v^j \in V(S_k)$ as possible, in particular, vertices satisfying $f(v^j) < \gamma_k$. Clearly, there are many imaginable combinations of these and other selection strategies. However, in any case, in order to prune branches of G as soon as possible, it seems to be reasonable to apply the test for deletion of simplices described in Section 4.1.4.

(b) It can be expected that, for a given practical problem, Algorithm 2 terminates considerably faster than indicated by the bounds (27), (28). It is even not clear wether we can construct any reasonable problem of the form (18) where these bounds are attained. In such a problem,

  (1) every hyperplane chosen in Step $k.3$ must subdivide the vertex set $V(S_k)$ in two parts $V^-(S_k)$, $V^+(S_k)$ of (nearly, if n odd) equal size, and $V^=(S_k) = \emptyset$ for every generated simplex (note at this point, that degeneration in the sense of $V^=(S_k) \neq \emptyset$ makes the subdivision procedure *easier*, since the number of new partition sets becomes smaller);

  (2) every subdivision Step $k.3$ must generate a maximal number of simplices $\{S_{k_1}, \ldots, S_{k_\ell}\}$ with $\left|I_{S_{k_j}}\right| = |I_{S_k}| - 1$ for $1 \leq j \leq \ell$, i.e. every hyperplane H_i irredundant for S_k must remain irredundant for all S_{k_j}, except H_{i_k};

  (3) no generated simplex $S \not\subseteq P$ is deleted in a Step $k.5$.

(c) The idea of Algorithm 2 is closely related to the exact simplicial (ES) algorithm and variants of it as presented, for example, in Chapter VII, 3 of [HoTu93], which is based on an idea due to Ban (cf. also [TaBa85]). Here are the main differences:

  (1) The ES–type algorithms perform only one bisection per iteration, so that the chosen cutting plane is likely to remain irredundant for the new simplices for many further subdivisions. It is clear, on the other hand, that this strategy might become an advantage if some of the generated partition sets

can be deleted in an early step, because the growth of $|\mathcal{P}_k|$ per iteration step is bounded by one.

(2) To prove convergence for the ES–type algorithms, one prescribes a fixed selection scheme for the choice of the cutting planes (see [HoTu93]). If, after the application of a bisection to a simplex S with cutting plane H, this cutting plane remains irredundant for an unfathomable descendent $\tilde{S}$ of S, then H will be chosen again to subdivide $\tilde{S}$. Therefore, in the worst case, i.e. if one has to apply subdivisions until H is redundant for every generated subsimplex of S, the ES method computes as many bisections as Algorithm 1 would do, if one applies it to S. It is easily seen that the total number of generated and bounded subsimplices in this process adds up to $2 \cdot (K^-(S) + K^+(S) - 1)$. In this case, one call of `Subdivide(·)` can be much more efficient, since it generates only $K^-(S)$ subsimplices, and H is redundant for all of them.

(3) The so called modified ES algorithms (see [HoTu93]) build partitions $\mathcal{P}_k$ containing simplices of different dimensions. A simplex S is replaced by a face of it, if S contains only boundary points of the feasible set P. If P is full dimensional, then, with arguments similar to those in Lemma 16, it is possible to introduce deletion rules as in Section 4.1.4 to overcome this.

(4) ES–type algorithms are proved to be finite, but no bounds on the number of iterations or generated partition sets are given, whereas regarding Algorithm 2 we have Theorem 17, and, moreover, the choice of cutting planes can be based on the precomputable number of new partition sets in every step k. Using the arguments introduced in (c), (2), it is now possible to give worst case bounds for ES–type algorithms.

(d) In [HoTu93], an extension of the exact simplicial algorithm to the case of unbounded feasible sets is given by the description of a conical variant. It should be possible to generalize Algorithm 2 for unbounded feasible sets along similar lines.

(e) If we incorporate our ideas into the well–known pure outer approximation schemes (cf., e.g. [HoTu93], [HoPaTh95] or [Be95]), it is easily seen that we obtain a convergent algorithm for the case when the feasible set X is compact and convex. A simple idea for doing this is to use Algorithm 2 for solving the linearly constrained concave subproblems (P_k) arising in the outer approximation process. Note that the feasible set of (P_{k+1}) originates from that of its predecessor (P_k) only by adding a single cutting plane, so that Algorithm 2 can be used to construct a new feasible partition by updating the simplices contained in the optimal partition of (P_k), which violate the new inequality.

(f) Finally, note that if we replace the subdivision process in Step k.3 by the well known exhaustive bisection (combined with suitable tests for the deletion of infeasible partition sets in case of using the lower bound μ_1), we are led to classical (convergent, but in general infinite) simplicial branch and bound procedures for problem (18) introduced by Horst in [Ho76]. Some of these variants will be used in Section 5 for numerical comparisons.

5. Numerical Examples

In view of Proposition 5 (iii), it is immediately clear that the number $K^-(S)$ of new simplices possibly generated by the subdivision procedure introduced in Section 2 will set some natural bounds for numerical applications. In fact, in the worst case, the number $K^-(S)$ exponentially grows in the dimension of the problem. On the other hand, it is analytically completely unknown how this drawback compares to the real world performance of existing convergent branch and bound methods, which rely, in the end, on exhaustiveness of the subdivision process used and continuity arguments. To give an impression of a possible outcome comparing the new method with a convergent one, let us start with two low dimensional examples of very regular and simple structure. As mentioned before, we check our method (abbreviated as M2 in the following) against the classical one of Horst ([Ho76], (M1)), which combines bisection of a longest edge with the bound μ_2.

Table II. Results in $\mathbb{R}^2$ and $\mathbb{R}^3$

Method	μ	Acc. ε	Iter (P_1)	Mps (P_1)	Iter (P_2)	Mps (P_2)
M1:	μ_2	$1.E$-02	17	2	205	30
	μ_2	$1.E$-06	41	2	325	30
	μ_2	$1.E$-10	69	2	441	30
	μ_2	$1.E$-14	84	2	534	30
M2:	μ_1	$1.E$-14	3	2	53	19
	μ_2	$1.E$-14	2	1	42	21

EXAMPLE 18. The first example problem, stated in $\mathbb{R}^2$,

$$\min\left\{-\|x\|^2 : 1 \le x_1 + x_2 \le 3,\ -1 \le x_1 - x_2 \le 1\right\} \qquad (P_1)$$

asks us to find a point in a square with maximal Euclidean norm. (P_1) has two global solutions, given by $\hat{x}^1 := (1,2)^T$ and $\hat{x}^2 := (2,1)^T$. A starting simplex is given by $S = \{(x_1, x_2) \ge 0 : x_1 + x_2 \le 3\}$. The numerical results are shown in the (P_1)–columns of Table II. If we apply the simple bound μ_1 in M2, then the method finds an exact (machine precision) solution in 3 iterations, while using μ_2 we need 2 iterations. M1 finds an approximate solution after 17 iterations when we require an absolute accuracy ε set to $1.E$-02, i.e., the global lower and upper bounds are allowed to differ from each other at most by ε. This increases up to 84 iterations if we prescribe machine precision as in M2. Looking at the generated partitions, we see that both M1 and M2 use a maximal number of 2 remaining unfathomable partition sets (Mps) in each iteration. M1 keeps exactly 2 simplices $S_{k,1}$ and $S_{k,2}$ in each iteration step k, with $\hat{x}^i \in S_{k,i}$, $i = 1, 2$. The procedure has to shrink these simplices using bisection until the diameters are small enough to compute lower bounds with the prescribed precision. It is impossible for M1 to compute exact lower bounds since each $S_{k,i}$ contains at least one infeasible vertex having a smaller objective function value than $f^\star(P)$.

EXAMPLE 19. A similar behavior can be observed in $\mathbb{R}^3$ when applying the methods to

$$\min\left\{-\|x-e\|^2 : a \le Ax \le b\right\} \qquad (P_2)$$

where $e = (1,1,1)^T$ and the feasible set P is given by the following data:

$$\begin{bmatrix} a^T \\ A^T \\ b^T \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & -3 & -3 & -3 & 3 \\ -1 & 2 & 2 & 5 & -1 & -1 & 1 \\ 2 & -1 & 2 & -1 & 5 & -1 & 1 \\ 2 & 2 & -1 & -1 & -1 & 5 & 1 \\ 12 & 12 & 12 & 15 & 15 & 15 & 9 \end{bmatrix}.$$

The polytope P is given by 14 facet defining inequalities, it possesses 24 vertices. Problem (P_2) has 6 global solutions $\hat{x}^i$, which are given by the 6 possible permutations of the coordinates in $\hat{x} := (2,3,4)^T$, with optimal objective function value $f^*(P) = -14$. The numerical results are shown in the (P_2)–columns of Table II. The interpretation is much the same as for (P_1). Again, if the iteration counter exceeds 200, Horst's method M1 bisects 6 small simplices containing the optimal solutions $\hat{x}^i$ until the prescribed precision ε is reached.

Note at this point, that although the superior method M2 solves the proposed 2 problems much faster than one could suspect given the worst case bound (27), it computes many more vertices than the feasible sets possess. This seems to be a general drawback of (simplicial) branch and bound approaches for problem (18).

A second point worth mentioning is the strong dependence of the efficiency of M2 on the selection rule for the new cutting plane in step k.3 of (A2). We tested some of the variants sketched in the Remarks to Theorem 17, and found that the following heuristics lead to a satisfying overall performance of (A2), with respect to our test examples:

- ▷ Choose a vertex $v \in \arg\min f(V(S))$. From the set of possible cuts referenced by I_S, choose a cutting plane H which cuts off v in such a way, that either $n^-(S)$ or $K^-(S)$ (or both, if possible) are minimized.

Clearly, because this part of the algorithm seems to be very important from a practical point of view, this topic has to be investigated further.

Finally, it is obvious that (P_1) and (P_2) are far from being representative. The fact that both examples have multiple optimal solutions is somewhat disadvantageous for M1 compared to M2, and at least the special structure of the feasible set of (P_1) shows favor towards M2. The intent of both examples is to demonstrate what *can* happen and why it happens if the situation is favorable for M2. If different feasible sets and/or objective functions are used, the performance of M1 often counterbalances that of M2. This will become apparent in the next section.

Table III. Objective functions used

Funct. no. i	Functional Form $f_i(x)$	Source
1	$-\left\lvert x_1 + \sum_{j=2}^{n} \frac{j-1}{j} x_j \right\rvert^{\frac{3}{2}}$	[BeSa94], [HoTh88]
2	$-\left(1 + \left(\sum_{j=1}^{n} j x_j\right)^2\right)^{\frac{1}{2}}$	[BeSa94], [HoTh88]
3	$-\left\lvert \sum_{j=1}^{n} \frac{1}{j} x_j \right\rvert \cdot \ln\left(1 + \left\lvert \sum_{j=1}^{n} \frac{1}{j} x_j \right\rvert\right)$	[BeSa94], [HoTh88]
4	$-3 \cdot \sum_{j=1}^{n} x_j^2 + 2\left(\sum_{j=1}^{n-1} x_j x_{j+1}\right)$	[BeSa94]
5	$-\left(\sum_{j=1}^{n} x_j^2\right) \cdot \ln\left(1 + \sum_{j=1}^{n} x_j^2\right)$	[BeSa94], [HoTh88]

5.1. Experimental Results from Randomly Generated Problems

Considering higher dimensional examples than before, we compared M1 and M2 using randomly created instances $(P_{n,m}^i)$ of problem (18) for dimensions $4 \leq n \leq 10$. The construction of the examples was done in a fashion similar to that used in [BeSa94]: the feasible set $P_{n,m}$ of $(P_{n,m}^i)$ is an n–dimensional polytope given by $P_{n,m} := \{x \in \mathbb{R}^n : x \geq 0, Bx \leq b\}$, where $B \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ define $1 \leq m \leq 9$ additional inequalities, generated randomly with the method described in [HoTh88]. Finally, using the objective functions given in Table III, each problem instance $(P_{n,m}^i)$ is defined by

$$\min \{f_i(x) : x \in P_{n,m}\} \qquad (P_{n,m}^i)$$

where $1 \leq i \leq 5$, $4 \leq n \leq 10$, $1 \leq m \leq 9$, yielding a total number of 315 examples. Table IV compares the overall performance of M1 and M2 in the following way: let $\tau_1(i,m,n)$, $\tau_2(i,m,n)$ be the CPU times needed to solve $(P_{n,m}^i)$ using M1 and M2, respectively. Then the quotient

$$s_{m,n} := \frac{\sum_{i=1}^{5} \tau_1(i,m,n)}{\sum_{i=1}^{5} \tau_2(i,m,n)}$$

is used as a measure for the relative performance of M2 compared to M1, taken over five instances $(P_{n,m}^i)$ asking for the minimization of five different objective

Table IV. Relative speedup (slowdown) $s_{m,n}$ using (M2) instead of (M1)

$n\backslash m$	1	2	3	4	5	6	7	8	9
4	4.6	3.4	3.6	2.7	1.3	2.1	4.8	2.1	2.7
5	23.2	17.9	22.9	18.1	7.2	2.9	1.4	1.3	0.6
6	198.7	85.0	9.9	18.5	2.1	0.1	0.4	0.04	0.2
7	1.1	0.9	1.0	172.3	2.4	0.2	0.9	0.4	0.3
8	1.3	1.6	1.0	0.9	1.0	23.9	9.6	23.7	0.3
9	1.0	29.5	322.0	270.5	449.8	*45.6	*241.9	*66.0	*87.5
10	0.9	*8033.9	198.5	165.6	86.2	0.1	*2.5	*0.2	--

functions on a fixed feasible polytope $P_{n,m}$ of dimension n with up to $m+n$ facets. Table IV shows the values of $s_{m,n}$ for the different values of m and n.
Values marked by a star in Table IV give a lower bound for the true speedup $s_{m,n}$, since the method M1 failed at least for one objective function in the corresponding group of examples, due to exhausted computing resources. The last example is a special case; since both methods were unable to compute a solution in at least one problem instance, $s_{9,10}$ can not be given. Both methods used μ_2 for lower bounding. For all examples, we required an absolute accuracy $\varepsilon = 0.1$ for the final bounds, which turns out to be a strong precision goal for the majority of the examples. All computations were done on a Sun Sparcstation 20 using C++. To solve the linear programs, we used a code from the NAG Mark 16 FORTRAN library.

Appendix

A. Proofs

A.1. Proof of Lemma 8

Let $M_g := \begin{pmatrix} B & b \\ d^T & \delta \end{pmatrix}$ be the inverse matrix of M_f, with $B \in \mathbb{R}^{n\times n}$, b, $d \in \mathbb{R}^n$ and $\delta \in \mathbb{R}$. Then by definition of the inverse matrix

$$M_g \cdot M_f = \begin{pmatrix} B & b \\ d^T & \delta \end{pmatrix} \cdot \begin{pmatrix} A & a \\ c^T & \gamma \end{pmatrix} = \begin{pmatrix} E & 0 \\ 0^T & 1 \end{pmatrix}, \tag{32}$$

where E denotes the $n\times n$ identity matrix. Hence for $y \in Y, x \in X$ with $f(x) = y$ we see that

$$d^T y + \delta = d^T \frac{Ax + a}{c^T x + \gamma} + \delta = \frac{(d^T A + \delta c^T)x + (d^T a + \delta\gamma)}{c^T x + \gamma} = \frac{1}{c^T x + \gamma},$$

so that $g : Y \to \mathbb{R}^n$ with $g(y) := \frac{By+b}{d^T y+\delta}$ is a projective transformation permissible for Y. From (32) we conclude that g is nonsingular and that, for all $x \in X$, we have

$$g \circ f(x) = (c^T x + \gamma)(Bf(x) + b) = (BA + bc^T)x + (Ba + \gamma b) = x$$

This yields the invertibility of f on X with inverse $f^{-1} = g$. □

A.2. Proof of Lemma 9

Let $y \in P'$. Then there exists $x \in P$ with $x = \sum_{i=1}^k \lambda_i x^i$, $f(x) = y$, $\sum_{i=1}^k \lambda_i = 1$, $\lambda_i \geq 0$, $1 \leq i \leq k$, and therefore

$$s := \textstyle\sum_{i=1}^k \lambda_i(c^T x^i + \gamma) > 0.$$

Then

$$y = \frac{A(\sum_{i=1}^k \lambda_i x^i) + a(\sum_{i=1}^k \lambda_i)}{c^T(\sum_{i=1}^k \lambda_i x^i) + \gamma(\sum_{i=1}^k \lambda_i)} = \sum_{i=1}^k \frac{\lambda_i(c^T x^i + \gamma)}{s} f(x^i) = \sum_{i=1}^k \mu_i y^i,$$

where $\mu_i := \lambda_i \left(c^T x^i + \gamma\right) s^{-1} \geq 0$, $1 \leq i \leq k$, and $\sum_{i=1}^k \mu_i = 1$, i.e., $y \in \operatorname{conv}(Y)$.

Now let $y \in \operatorname{conv}(Y)$, i.e., $y = \sum_{i=1}^k \mu_i y^i$ with $\sum_{i=1}^k \mu_i = 1$ and $\mu_i \geq 0$, $1 \leq i \leq k$. To show that $y \in P'$, define for $1 \leq i \leq k$

$$l_i := c^T x^i + \gamma \quad \text{and} \quad \lambda_i := \tfrac{\mu_i}{\sigma l_i}, \quad \text{where} \quad \sigma := \textstyle\sum_{j=1}^k \frac{\mu_j}{l_j}.$$

Then, clearly, $\sigma > 0$, $\sum_{i=1}^k \lambda_i = 1$ and $\lambda_i \geq 0$ for $1 \leq i \leq k$. Moreover, for $x := \sum_{i=1}^k \lambda_i x^i \in P$, we see that

$$f(x) = \sum_{i=1}^k \lambda_i \frac{Ax^i + a}{\sum_{j=1}^k \lambda_j l_j} = \sum_{i=1}^k \frac{\mu_i}{\sigma l_i} \frac{Ax^i + a}{\sum_{j=1}^k \frac{\mu_j}{\sigma l_j} l_j} = \sum_{i=1}^k \mu_i \frac{Ax^i + a}{l_i} = y,$$

and therefore $y \in P'$. □

To show Lemma 10, we need the following result:

LEMMA 20. *Let $X_k := \{x^1, \ldots, x^k\} \subset \mathbb{R}^n$ be a finite set of affinely dependent vectors, and $f : \mathbb{R}^n \to \mathbb{R}^n$ be a projective transformation permissible for X_k. Then $f(X_k) := \{f(x^1), \ldots, f(x^k)\}$ is an affinely dependent set of vectors.*

Proof. Let $\lambda_i \in \mathbb{R}$, $1 \leq i \leq k$, not all equal zero satisfy $\sum_{i=1}^k \lambda_i x^i = 0$ and $\sum_{i=1}^k \lambda_i = 0$. Then, for $f(x) = \frac{Ax+a}{c^T x+\gamma}$ and $l_i := c^T x^i + \gamma \neq 0$:

$$\sum_{i=1}^k (\lambda_i l_i) \cdot f(x^i) = \sum_{i=1}^k \lambda_i (Ax^i + a) = A(\sum_{i=1}^k \lambda_i x^i) + a(\sum_{i=1}^k \lambda_i) = 0,$$

and $\sum_{i=1}^{k} \lambda_i l_i = c^T(\sum_{i=1}^{k} \lambda_i x^i) + \gamma(\sum_{i=1}^{k} \lambda_i) = 0$, where $\lambda_i l_i \neq 0$ iff $\lambda_i \neq 0$. □

A.3. Proof of Lemma 10

Lemma 9 tells us that $P' \subseteq \mathbb{R}^n$ is the polytope $P' = \operatorname{conv}(Y)$, thus $V(P') \subseteq Y = \{y^1, \ldots, y^k\} \subseteq P'$. By Lemma 8, f is invertible, and since $x^i \neq x^j$ for $i \neq j$, it follows that $y^i = f(x^i) \neq f(x^j) = y^j$ for $i \neq j$. Moreover, again by Lemma 8, the inverse $g : P' \to P$ of f given by $g(y) = \frac{By+b}{d^T y+\delta}$ is a nonsingular projective transformation permissible for Y satisfying $d^T y + \delta > 0 \; \forall y \in Y$. Finally, by Lemma 9, the image of Y under g is $g(Y) = X$. These considerations show that, for nonsingular f, if the requirements of Lemma 9 are satisfied by P, X and f, they are also satisfied for P', Y and f^{-1}, respectively. Therefore, it is sufficient to show the respective first assertions of (a), (b) and (c), since application of analogous reasonings to the symmetric situation for P', Y and f^{-1} will give the second.

(a) To show that $V(P) = X \Rightarrow V(P') = Y$, assume $V(P) = X$ but $V(P') \neq Y$. For $1 \leq j \leq k$, define $Y_j := Y \setminus \{f(x^j)\}$ and $X_j := X \setminus \{x^j\}$. Since $V(P') \subseteq Y$, we must have $V(P') \subseteq Y_i$ for at least one $i \in \{1, \ldots, k\}$. Let $g : P' \to P$ be the inverse of f on P'. Applying Lemma 9 to P' and g, we see from $V(P') \subseteq Y$ that

$$P'' = g(P') = \operatorname{conv}(g(Y)) = \operatorname{conv}(X) = P,$$

whereas from $V(P') \subseteq Y_i$ one concludes, again by Lemma 9, that

$$P'' = g(P') = \operatorname{conv}(g(Y_i)) = \operatorname{conv}(X_i),$$

which leads to the contradiction $P = \operatorname{conv}(X_i)$.

(b) Now let $S_d = [v^0, \ldots, v^d] \subseteq P$ be a d–simplex. Then $c^T v^i + \gamma > 0$, $0 \leq i \leq d$, since $c^T x + \gamma > 0 \; \forall x \in P = \operatorname{conv}(X)$. The restriction $\bar{f} : S_d \to f(S_d) = S'_d$ of f to S_d is nonsingular, and therefore

$$S'_d = [\bar{f}(v^0), \ldots, \bar{f}(v^d)] \quad \text{with} \quad V(S'_d) = \{\bar{f}(v^0), \ldots, \bar{f}(v^d)\},$$

by (a) applied to S_d and $\bar{f}$. Assume that the set $V(S'_d)$ is affinely dependent. Then, by Lemma 20, the image of $V(S'_d)$ under the (projective) inverse $\bar{g}$ of $\bar{f}$ is an affinely dependent set, which contradicts the assumption that $\dim S_d = d$.

(c) To prove the first assertion in (c), note first that, by (b), every member S' of $\mathcal{D}'$ is an n–simplex. We have to show that (i) $P' = \mathcal{D}'$ and (ii) that two simplices in $\mathcal{D}'$ do not share any interior points. Regarding (i), the inclusion $\mathcal{D}' \subseteq P'$ is obvious. For every $y \in P'$, the corresponding $x \in P$ with $f(x) = y$ lies in some n–simplex $S \in \mathcal{D}$. By Lemma 9, we know that $y \in f(S) \in \mathcal{D}'$, so that $P' \subseteq \mathcal{D}'$. To show (ii), assume that, for S'_i and S'_j in $\mathcal{D}'$ we have $y \in \operatorname{int} S'_i \cap \operatorname{int} S'_j$. Thus one can find $\varepsilon > 0$ so that the open ball $B_\varepsilon(y)$ satisfies $B_\varepsilon(y) \subseteq S'_i \cap S'_j$, and therefore $f^{-1}(B_\varepsilon(y))$ is an open set contained in $S_i \cap S_j$. Since $\mathcal{D}$ is a simplicial partition of P, we conclude that $i = j$. □

References

Be85 Benson, H.P. (1985), A Finite Algorithm for Concave Minimization over a Polyhedron, *Naval Research Logistics Quaterly,* 32, pp. 165–177

Be95 Benson, H.P. (1995), Concave Minimization: Theory, Applications and Algorithms, in R. Horst and P.M. Pardalos (eds.), *Handbook of Global Optimization,* Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 43–148

BeSa94 Benson, H.P. and Sayin, S. (1994), A Finite Concave Minimization Algorithm Using Branch and Bound and Neighbor Generation, *Journal of Global Optimization,* 5, pp. 1–14

FaHo76 Falk, J.E. and Hoffman, K.L. (1976), A Successive Underestimation Method for Concave Minimization Problems, *Mathematics of Operations Research,* 1, pp. 251-259

FaSo69 Falk, J.E. and Soland, R.M. (1969), An Algorithm for Separable Nonconvex Programming Problems, *Management Science,* 15, pp. 550-569

Gr67 Grünbaum, B. (1967), *Convex Polytopes,* Wiley–Interscience, London

Ha91 Haiman, M. (1991), A Simple and Relatively Efficient Triangulation of the n–Cube, *Discrete Computational Geometry,* Vol. 6, No. 4, pp. 287–289

HaJa88 Hamami, M. and Jacobsen, S.E. (1988), Exhaustive Nondegenerate Conical Process for Concave Minimization on Convex Polytopes, *Mathematics of Operations Research,* 13, pp. 479–487

Ho76 Horst, R. (1976), An Algorithm for Nonconvex Programming Problems, *Mathematical Programming,* Vol. 10, pp. 312–321

HoPaTh95 Horst,R., Pardalos, P.M. and Thoai, N.V. (1995), *Introduction to Global Optimization,* Kluwer Academic Publishers, Dordrecht, The Netherlands

HoTh88 Horst,R. and Thoai, N.V. (1988), Modification, Implementation and Comparison of Three Algorithms for Globally solving Concave Minimization Problems, *Computing,* Vol. 42, pp. 271–289

HoTu93 Horst,R. and Tuy, H. (1993), *Global Optimization (Deterministic Approaches),* 2nd Edition, Springer, Berlin

Hu93 Hughes, R.B. (1993), Minimum–cardinality triangulations of the d–cube for $d = 5$ and $d = 6$, *Discrete Mathematics,* 118, pp. 75–118

HuAn93 Hughes, R.B. and Anderson, M.R. (1993), A triangulation of the 6–cube with 308 simplices, *Discrete Mathematics,* 117, pp. 253–256

PaSt82 Papadimitriou, C.H. and Steiglitz, K. (1982), *Combinatorial Optimization: Algorithms and Complexity,* Prentice–Hall, Englewood Cliffs, New Jersey

So74 Soland, R.M. (1974), Optimal Facility Location with Concave Costs, *Operations Research,* 22, pp. 373–382

TaBa85 Tam, B.T. and Ban, V.T. (1985), Minimization of a Concave Function Under Linear Constraints. *Ekonomika i Matematicheskie Metody,* 21, 709–714, (in Russian)

Tod76 Todd, M.J. (1976), The Computation of Fixed Points and Applications, *Lecture Notes in Economics and Mathematical Systems,* Vol. 124, Springer, Berlin

Tuy91 Tuy, H. (1991), Effect of the Subdivision Strategy on Convergence and Efficiency of Some Global Optimization Algorithms, *Journal of Global Optimization,* Vol. 1, Nr. 1, pp. 23–36

Acknowledgements

The author is indebted to Reiner Horst, Nguyen Van Thoai and to an anonymous referee for their helpful suggestions and comments regarding this article.

Address for correspondence: nast@uni–trier.de