Math. Program., Ser. A 85: 593–616 (1999)

Digital Object Identifier (DOI) 10.1007/s101079900092

M. Locatelli

Finiteness of conical algorithms with ω−subdivisions

Received January 24, 1996 / Revised version received December 9, 1998
Published online June 11, 1999

Abstract. In this paper the problem of finding the global optimum of a concave function over a polytope is considered. A well-known class of algorithms for this problem is the class of conical algorithms. In particular, the conical algorithm based on the so called $\omega-$subdivision strategy is considered. It is proved that, for any given accuracy $\varepsilon > 0$, this algorithm stops in a finite time by returning an $\varepsilon-$optimal solution for the problem, while it is convergent for $\varepsilon = 0$.

Key words. concave optimization – conical algorithms – $\omega-$subdivisions

1. Introduction

In this paper the problem of minimizing a concave function over a polytope is considered. Let $A = [a_1, \dots, a_m]^T \in \Re^{m\times n}$ and $b \in \Re^m$ be such that

$$P = \{x \in \Re^n \ : \ Ax \leq b\}$$

is a nonempty and full-dimensional polytope. Let f be a concave function such that its level sets

$$D_\gamma = \{x \in \Re^n \ : \ f(x) \geq \gamma\}$$

are bounded for any $\gamma \in \Re$. A point

$$x^* \in \arg\min_{x\in P} f(x) \tag{1}$$

is searched for, i.e. the aim is to detect one global minimum of the concave function f over the polytope P. Besides the global optima, many local optima may exist and the problem is a NP-hard one even when the objective function f and the feasible polytope P have some special forms, e.g. f is quadratic and P is a hypercube (see [12]). However, the concavity of f has some important consequences. In [10, page 73] it is shown that given a nonconstant concave function over a convex set, the global minimum can not be attained in the interior of the convex set. When the feasible region is a polytope we can go a little bit further and say that the global optimum value is attained in at least one vertex of the polytope.

M. Locatelli: Universitá degli studi di Firenze, Dipartimento di Sistemi e Informatica, Via di S.Marta, 3-50139 Firenze, Italy

Mathematics Subject Classification (1991): 90C26, 65K05

Different approaches for the solution of this problem have been proposed in the literature. These approaches can be subdivided into three main groups: vertex enumeration, successive approximation, and successive partitioning. For detailed reviews see [2], [5], [7], [11]. In this paper a class of algorithms belonging to the successive partitioning approach is considered, namely the class of conical algorithms, first introduced in [14]. In Sect. 2 a description of the class of conical algorithms is given and two well-known subdivision strategies, bisection and ω-subdivison, are discussed. In the theory about conical algorithms there is no proof of finiteness for the algorithm employing only ω-subdivisions. The proof will be given in this paper. In particular, in Sect. 3 the main ideas on which the proof is based are introduced, while the technical aspects of the proof are presented in the appendix.

2. Description of conical algorithms

In conical algorithms for concave minimization over polytopes the feasible polytope is subdivided through polyhedral cones, and for each cone a linear subproblem is solved, whose feasible region is the intersection of the cone with the feasible polytope. The optimal value of the linear subproblem is employed, in a way that will be made clear in what follows, both for the fathoming rule, i.e. the rule for excluding cones from further consideration, and for the selection of the cone which has to be further subdivided at a given iteration. Moreover, the optimal solution of the linear subproblem related to the selected cone can be chosen as the point with respect to which the selected cone is subdivided into subcones. In what follows conical algorithms will be described in detail.

Let $V(P)$ be the vertex set of the polytope P, and, for a given $v \in V(P)$, let $adj(v) \subset V(P)$ be the set of vertices adjacent to v in P. Given a nonsingular matrix $Q \in \Re^{n\times n}$, whose columns are the vectors $q_1, \ldots, q_n$, let

$$cone\{q_1, \ldots, q_n\} = \{x \in \Re^n : \ x = Q\lambda, \ \lambda \in \Re^n_+\} = \{x \in \Re^n : \ Q^{-1}x \geq 0\},$$

denote the polyhedral cone whose generating rays are the vectors $q_1, \ldots, q_n$. Let $e = (1, \ldots, 1) \in \Re^n$ be the row vector whose components are all equal to 1. Finally, for a given $x \neq 0$ and $\gamma \leq f(0)$, let

$$ext[x; \gamma] = \arg\max\{y \mid y = \pi x, \ \pi \geq 0, \ f(y) \geq \gamma\}. \tag{2}$$

The existence of $ext[x; \gamma]$ is guaranteed by the boundedness of D_γ.

Conical algorithms alternate between two phases.

Phase I Let $\overline{x} \in P$. Start a local search from $\overline{x}$ in order to find a vertex $\overline{v} \in V(P)$ such that

$$\forall\, v \in adj(\overline{v}), \ \ f(v) \geq f(\overline{v}).$$

Set $\tau = f(\overline{v})$ and $\gamma = \tau - \varepsilon$, where $\varepsilon \geq 0$ is fixed. Translate the origin in $\overline{v}$. Let Q_0 be a nonsingular matrix whose columns q_j^0, $j \in \{1, \ldots, n\}$, satisfy $f(q_j^0) = \gamma$, and are the generating rays of a cone $C(Q_0)$ containing the feasible polytope P, i.e.

$$C(Q_0) = cone\{q_1^0, \ldots, q_n^0\} \supset P.$$

Set $\mathcal{P} = \{C(Q_0)\}$ and $\mathcal{M} = \mathcal{P}$.

Phase II 1. For all $C(Q) \in \mathcal{P}$ solve the linear subproblem

$$\begin{aligned} \max\ & eQ^{-1}x \\ x &\in P \\ Q^{-1}x &\geq 0, \end{aligned} \tag{3}$$

whose feasible region is the intersection $C(Q) \cap P$ between the cone $C(Q)$ and the feasible polytope. The optimum value is denoted with $\mu(Q)$ and the optimal solution with $\omega(Q)$.

2. If, for some $C(\overline{Q}) \in \mathcal{P}$, it holds that $f(\omega(\overline{Q})) < \tau$, go back to Phase I with $\overline{x} = \omega(\overline{Q})$, otherwise go to the next step.
3. In $\mathcal{M}$ delete all the cones $C(Q)$ such that $\mu(Q) \leq 1$ (see (4) after the description of the algorithm).
4. Let $\mathcal{R}$ be the collection of remaining cones. If $\mathcal{R} = \emptyset$, then stop: $\overline{v}$ is a global $\varepsilon-$optimal solution, i.e.

$$f(\overline{v}) \leq \min_{x \in P} f(x) + \varepsilon.$$

5. Select a cone $C(Q^*) \in \mathcal{R}$, $C(Q^*) = cone\{q_1^*, \dots, q_n^*\}$.
6. According to a given rule, select a point $x^* \in C(Q^*)$, $x^* = \sum_{i=1}^n \eta_i^* q_i^*$. Let $q^* = ext[x^*; \gamma]$. Then define

$$\mathcal{P}^* = \{C(Q_i) \mid \eta_i^* > 0\},$$

where

$$C(Q_i) = cone\{q_1^*, \dots, q_{i-1}^*, q^*, q_{i+1}^*, \dots, q_n^*\}.$$

7. Set $\mathcal{M} = \mathcal{R} \cup \mathcal{P}^* \setminus \{C(Q^*)\}$ and $\mathcal{P} = \mathcal{P}^*$. Go back to Step 1.

First we observe that some care is needed for $\varepsilon = 0$. Indeed, while the nonsingular matrix Q_0 defined in Phase I always exists for $\varepsilon > 0$, for $\varepsilon = 0$ it may not exist. A condition under which Q_0 is guaranteed to exist is the nondegeneracy of the vertex $\overline{v}$ translated into the origin in Phase I (see also [7, page 301]).
The fathoming rule in Step 3 follows from the fact that

$$\mu(Q) \leq 1 \implies f(x) \geq \gamma, \ \forall\, x \in P \cap C(Q), \tag{4}$$

i.e. in $C(Q) \cap P$ it is not possible to find a point which improves by more than $\varepsilon = \tau - \gamma$ with respect to the current best observed value τ (see theorem V.I in [7]). Crucial steps in conical algorithms are Step 5 and 6, which define respectively the rule for the selection of the cone to subdivide and the choice of the point at which the selected cone has to be subdivided. For what concerns Step 5, a typical choice for the cone to be subdivided is a cone $C(Q^*)$ such that

$$C(Q^*) \in \arg\max\{\mu(Q),\ C(Q) \in \mathcal{R}\}. \tag{5}$$

For what concerns Step 6, in the literature there are two well-known subdivision strategies: bisection and ω-subdivison. The bisection strategy, first introduced in [4] for simplicial algorithms, corresponds to choose x^* in Step 6 equal to the midpoint of one of the longest edges of the $(n-1)$-simplex $C(Q^*) \cap \{x : eQ_0^{-1}x = 1\}$, where $C(Q^*)$ is the cone selected in Step 5. In [13] the algorithm based only on bisections was studied and proved to be finite for $\varepsilon > 0$ and convergent for $\varepsilon = 0$, but it turned out to be very slow. The ω-subdivision strategy corresponds to subdivide the selected cone $C(Q^*)$ with respect to the optimal solution $\omega(Q^*)$ of the corresponding linear subproblem, i.e. in step 6, $x^* = \omega(Q^*)$. The ω-subdivision strategy is generally considered a better choice with respect to bisection, because it better exploits the information collected by the algorithm. Such a subdivision strategy was first introduced in [14] and later also studied in [1] and [19]. Algorithms based on it have been reported to work well in practice in [3], [6], [15], [16], [19]. In particular, in [6] problems with up to 50 variables and 20 constraints have been solved, and conical algorithms have been succesfully compared with two other common algorithms for concave minimization over polytopes. Some algorithms combining bisections and ω-subdivisions have been presented in the literature and proved to be finite for $\varepsilon > 0$ and convergent for $\varepsilon = 0$ (see e.g. [17] and [18]), but none of them can essentially avoid the introduction of bisections. Until now the problem of establishing the finiteness for $\varepsilon > 0$ and convergence for $\varepsilon = 0$ of the algorithm employing only $\omega-$subdivisions has been an open one and will be solved in the following sections. It must be pointed out at this point that during the refereeing procedure of this paper, in [8] and [9] a result has been proposed proving finiteness for $\varepsilon > 0$ and convergence for $\varepsilon = 0$ for an algorithm similar to the one considered here, under the nonrestrictive assumption that the feasible polytope has at least one nondegenerate vertex. The result has been obtained through techniques which are substantially different from those employed in this paper.

Before starting the proof of finiteness for the algorithm employing only ω-subdivisions, we introduce an equivalent formulation of subproblem (3), showing that it is not necessary to compute the inverse of the matrix Q in order to solve the linear subproblem (3). After the change of variable $\lambda = Q^{-1}x$, (3) can be reformulated as follows

$$\begin{aligned} \max\ e\lambda \\ AQ\lambda \leq b \\ \lambda \geq 0. \end{aligned} \tag{6}$$

3. The main ideas of the proof

In this section we give the scheme followed in order to prove the finiteness of the algorithm for $\varepsilon > 0$ and the convergence for $\varepsilon = 0$. Only the main results and the way they connect to each other are presented, while all the proofs will be given in the appendix.

We first consider the case $\varepsilon > 0$. Let $C(Q) = cone\{q_1, \ldots, q_n\}$ be one of the cones in $\mathcal{P}$. After solving subproblem (3) related to cone $C(Q)$, three possible situations can arise.

1. $\mu(Q) \leq 1$: then we exclude the cone $C(Q)$ from further consideration in view of the fathoming rule at Step 3 of Phase II.

2. $\mu(Q) > 1$ and $f(\omega(Q)) < \tau$: then we go back to Phase I and we detect, through a local search, a new vertex of P with function value strictly smaller than τ. Note that since the number of vertices of P is finite, this can happen only a finite number of times.
3. $\mu(Q) > 1$ and $f(\omega(Q)) \geq \tau$.

We are now interested in seeing what happens if situation 3 holds. Let $q^* = ext[\omega(Q); \gamma]$. Let

$$C(Q^*) = cone\{q_1, \ldots, q_{i-1}, q^*, q_{i+1}, \ldots, q_n\},$$

be one of the cones obtained after the ω-subdivision of cone $C(Q)$ in Step 6 of Phase II. We obviously have $C(Q^*) \subset C(Q)$. Given $x \in C(Q^*) \cap P$, $x \neq 0$, then $\exists\, \alpha > 0$ such that

$$x = \alpha\Big[\sum_{j=1, j\neq i}^{n} \lambda_j q_j + \lambda^* q^*\Big], \tag{7}$$

with

$$(\lambda_1, \ldots, \lambda_{i-1}, \lambda^*, \lambda_{i+1}, \ldots, \lambda_n) \in \Lambda_n,$$

where

$$\Lambda_n = \Big\{\lambda \in \Re^n : \sum_{j=1}^{n} \lambda_j = 1,\ \lambda_j \geq 0,\ \ j \in \{1, \ldots, n\}\Big\}$$

denotes the unit simplex in $\Re^n$.

It is possible to prove the following theorem giving a relation between the value at x for the objective function of the linear subproblem related to cone $C(Q)$ and the value at x for the objective function of the linear subproblem related to cone $C(Q^*)$.

Theorem 1. *Let $x \in C(Q^*) \cap P$. If $\mu(Q) > 1$ and $f(\omega(Q)) \geq \tau$, then*

$$\forall\, \varepsilon > 0\ \ \exists\, \rho = \rho(\varepsilon, \gamma) > 0 :\ eQ^{*\,-1}x \leq \frac{eQ^{-1}x}{1 + \lambda^* \rho}.$$

Proof. See the appendix.

□

It follows from Theorem 1 that if we consider an infinite nested sequence $\{C(Q_k)\}$ of cones generated during Phase II of the algorithm, and a point $x \in \cap_{k=1}^{\infty} C(Q_k) \cap P$, then the sequence $\{eQ_k^{-1}x\}$ is nonincreasing and it strictly decreases each time x has a positive coefficient with respect to the last inserted generating ray of cone $C(Q_k)$, or, equivalently, with respect to the optimal solution $\omega(Q_{k-1})$ of the linear subproblem related to cone $C(Q_{k-1})$.

As we will see in the appendix, a consequence of Theorem 1 is the following theorem, stating that if the algorithm generates in Phase II an infinite nested sequence $\{C(Q_k)\}$ of cones, then some of the generating rays of the cones in the sequence must become fixed for k big enough.

Theorem 2. *Let $\{C(Q_k)\}$ be an infinite nested sequence of cones generated during Phase II of the algorithm. If there exists $x \in \cap_{k=1}^{\infty} C(Q_k) \cap P$ such that $\forall\, k \;\; eQ_k^{-1}x \geq 1$, then there exists a positive integer K such that*

$$\begin{aligned} \forall\, k \geq K \;\; C(Q_k) = cone\{q_1, \ldots, q_p, q_{p+1}^k, \ldots, q_n^k\} \\ q_1, \ldots, q_p, \;\; p \geq 1, \;\; \textit{fixed}, \\ q_{p+1}^k, \ldots, q_n^k \;\textit{change infinitely often}. \end{aligned} \tag{8}$$

Proof. See the appendix.

□

Now we notice that the following chain of implications holds.

Phase II of the algorithm is infinite

$$\Downarrow$$

there exists an infinite nested sequence $\{C(Q_k)\}$ of cones generated during Phase II of the algorithm such that

$$\forall\, k, \quad \mu(Q_k) > 1$$

$$\Downarrow$$

$$\exists\, x \in \cap_{k=1}^{\infty} C(Q_k) \cap P \;:\; \forall\, k, \;\; eQ_k^{-1}x \geq 1$$

$$\Downarrow$$

(see Theorem 2)

there exists a positive integer K such that (8) holds.

The following theorem proves that the last element in the chain of implications above can not be true, and, consequently, that Phase II of the algorithm is always finite.

Theorem 3. *Given a nested sequence of cones $\{C(Q_k)\}$ generated by the algorithm and $x \in \cap_{k=1}^{\infty} C(Q_k) \cap P$ such that $\forall\, k, \;\; eQ_k^{-1}x \geq 1$, it is <u>not</u> possible that (8) holds for some positive integer K.*

Proof. See the appendix.

□

Now the finiteness of the conical algorithm employing only ω-subdivisions immediately follows.

Theorem 4. *The algorithm employing only ω-subdivisons is finite for any $\varepsilon > 0$.*

Proof. Theorem 3 proves the finiteness of Phase II of the algorithm. Therefore, we only need to prove that Phase II is entered a finite number of times. But, as already remarked, each time we enter Phase II, a new vertex of P has been detected in Phase I through a local search. Therefore, it follows from the finite cardinality of $V(P)$ that Phase II is entered only a finite number of times as we wanted to prove.

□

Now let us assume that $\varepsilon = 0$. It is possible to prove the following theorem, stating the convergence of the algorithm for $\varepsilon = 0$.

Theorem 5. *If the selection rule (5) for the choice of the cone to be subdivided is employed and P only has nondegenerate vertices, then the algorithm is convergent for $\varepsilon = 0$, i.e. if $\{C(Q_k)\}$ is an infinite nested sequence generated during Phase II of the algorithm, then*

$$\mu(Q_k) \downarrow 1 \quad \textit{and} \quad \tau = f^* = \min_{x \in P} f(x).$$

Proof. See the appendix.

□

Notice that the nondegeneracy assumption for P is only introduced to guarantee the existence of the matrix Q_0 in Phase I, as commented in Sect. 2 immediately after the description of the algorithm.

Appendix

In all the proofs of Sects. A-C it is assumed that $\varepsilon > 0$. Only in Sect. D the case $\varepsilon = 0$ is considered.

Appendix A. Proof of Theorem 1

Let $d(y, x)$ denote the Euclidean distance between y and x. We first need the following observation.

Observation 1. *Let f be continuous over a compact set A, then it is uniformly continuous over A, i.e. $\forall\, \xi > 0$ there exists $\zeta = \zeta(\xi, A) > 0$ such that*

$$\forall\, x, y \in A \;:\; d(y, x) \leq \zeta, \;\mid f(x) - f(y) \mid < \xi. \tag{9}$$

Now let $A = D_\gamma$ and $\delta = \zeta(\varepsilon, D_\gamma) > 0$. Note that the continuity of f and the boundedness of D_γ imply that D_γ is a compact set. Therefore, recalling that $f(0) = \tau = \gamma + \varepsilon$, it follows from Observation 1 that

$$f(x) = \gamma \quad \Rightarrow \quad d(0, x) > \delta. \tag{10}$$

Moreover, it follows from the boundedness of D_γ that there exists $R = R(\gamma) > 0$ such that

$$\forall\, x \in D_\gamma \quad d(0, x) < R. \tag{11}$$

Now we are ready for the proof of Theorem 1.

Proof. Note that $eQ^{-1}\omega(Q) > 1$ implies $d(0, \omega(Q)) > 0$. Since, by assumption, $f(\omega(Q)) \geq \tau$, and since $f(q^*) = \gamma = \tau - \varepsilon$, it follows from Observation 1 that

$$d(\omega(Q), q^*) > \delta, \tag{12}$$

where δ is the same as in (10). Finally, it follows from (11) that

$$d(0, \omega(Q)) < R. \tag{13}$$

Note that

$$q^* = \left(1 + \frac{d(q^*, \omega(Q))}{d(0, \omega(Q))}\right)\omega(Q). \tag{14}$$

It follows from (12) and (13) that

$$\frac{d(q^*, \omega(Q))}{d(0, \omega(Q))} > \frac{\delta}{R} = \rho > 0. \tag{15}$$

Now let $x \in Q^* \cap P$ be such that $eQ^{*\,-1}x = \alpha$. Then x is given by (7). Moreover, it holds that

$$\omega(Q) = \mu(Q)\sum_{j=1}^{n} \overline{\lambda}_j q_j,$$

where $\mu(Q) > 1$ and $\overline{\lambda} \in \Lambda_n$. Let

$$\overline{t} = \mu(Q)\left(1 + \frac{d(q^*, \omega(Q))}{d(0, \omega(Q))}\right) > (1 + \rho), \tag{16}$$

where the last inequality follows from $\mu(Q) > 1$ and (15). Then, in view of (14)

$$q^* = \overline{t}\sum_{j=1}^{n} \overline{\lambda}_j q_j. \tag{17}$$

By substituting (17) in (7) it follows that

$$x = \alpha\Big[\sum_{j=1,\, j\neq i}^{n} \lambda_j q_j + \lambda^* \overline{t} \sum_{j=1}^{n} \overline{\lambda_j} q_j\Big],$$

and

$$eQ^{-1}x = \alpha[(1 - \lambda^*) + \overline{t}\lambda^*] \geq \alpha[1 + \lambda^*\rho],$$

which proves the theorem.

□

For a given cone $C(Q_k)$ let

$$E(Q_k) = conv\{q_1^k, \ldots, q_n^k\} = \left\{x \in \Re^n : x = \sum_{j=1}^{n} \lambda_j q_j^k,\ \lambda \in \Lambda_n\right\} \tag{18}$$

be the convex hull of the vectors $q_1^k, \ldots, q_n^k$. Now consider a nested sequence of cones $\{C(Q_k)\}$ generated during Phase II of the algorithm and denote the new column vector in Q_k with respect to Q_{k-1} with q_k^*. Let us consider a point $x \in \cap_{k=1}^{\infty} C(Q_k) \cap P$, $x \neq 0$, and denote with $y_{Q_k,x}$ the intersection of the line going through x and the origin, with $E(Q_k)$. Finally, let λ_k^* be the component of $y_{Q_k,x}$ in $E(Q_k)$ with respect to q_k^*. It follows from Theorem 1, with q_k^* replacing q^* and λ_k^* replacing λ^*, that

$$eQ_k^{-1}x \leq \frac{eQ_{k-1}^{-1}x}{1+\lambda_k^*\rho},$$

and, consequently

$$eQ_k^{-1}x \leq \frac{eQ_0^{-1}x}{\prod_{h=1}^{k}(1+\lambda_h^*\rho)}. \tag{19}$$

The following corollary is an immediate consequence of this fact.

Corollary 1. *If $x \in \cap_{k=1}^{\infty} C(Q_k) \cap P$, where $\{C(Q_k)\}$ is an infinite nested sequence of cones generated by the algorithm, then*

$$\forall\, k \quad eQ_k^{-1}x \geq 1 \implies \sum_{h=1}^{\infty} \lambda_h^* < \infty.$$

Proof. The proof is a simple consequence of (19) . Indeed, it must hold that

$$\prod_{h=1}^{\infty}(1+\lambda_h^*\rho) < \infty,$$

and since for any N

$$\prod_{h=1}^{N}(1+\lambda_h^*\rho) \geq 1 + \rho\sum_{h=1}^{N}\lambda_h^*$$

it follows that

$$\sum_{h=1}^{\infty}\lambda_h^* < \infty.$$

□

Appendix B. Proof of Theorem 2

Proof. Given an infinite nested sequence of cones $\{C(Q_k)\}$ generated during Phase II of the algorithm and a point $x \in \cap_{k=1}^{\infty} C(Q_k) \cap P$ such that for each k it holds that $eQ_k^{-1}x \geq 1$, we need to prove that at least one of the generating rays of the cones $C(Q_k)$ becomes fixed for k big enough. The proof will be done by contradiction. Therefore, it will be assumed that all the generating rays change infinitely often. The generating rays of each cone $C(Q_k)$ will be denoted by q_j^k, $j = 1, \ldots, n$. Now let T be a nonnegative integer. Since $x \in C(Q_T)$ and $eQ_T^{-1}x \geq 1$, it follows that

$$x = \alpha_T \sum_{i=1}^{n} \lambda_i^T q_i^T,$$

where $\alpha_T \geq 1$ and $\lambda^T \in \Lambda_n$. Let $T + S_j$, $S_j > 0$, be the first cone, after cone $C(Q_T)$, where the generating ray q_j^T is substituted by some other generating ray $q_j^{T+S_j}$. By recalling the notation of Corollary 1, it holds that $\lambda_{T+S_j}^* = \lambda_j^{T+S_j}$. Since $x \in C(Q_{T+S_j})$ and $eQ_{T+S_j}^{-1}x \geq 1$, it follows that

$$x = \alpha_{T+S_j}\Big[\sum_{i=1, i\neq j}^{n} \lambda_i^{T+S_j} q_i^{T+S_j} + \lambda_{T+S_j}^* q_j^{T+S_j}\Big], \tag{20}$$

where $\alpha_{T+S_j} \geq 1$ and $\lambda^{T+S_j} \in \Lambda_n$. Now since $q_j^{T+S_j} \in C(Q_{T+S_j-1})$, it must hold that

$$q_j^{T+S_j} = \sum_{i=1, i\neq j}^{n} \beta_i^{T+S_j} q_i^{T+S_j-1} + \beta_j^{T+S_j} q_j^T, \tag{21}$$

where $\beta^{T+S_j} \geq 0$. Note that in view of (10) it holds that $d(0, q_i^{T+S_j-1}) > \delta > 0$, $i \in \{1, \ldots, n\}$, while in view of (11) it holds that $d(0, q_j^{T+S_j}) < R$. Therefore,

$$\forall\, i \in \{1, \ldots, n\} \quad \beta_i^{T+S_j} \leq D, \tag{22}$$

for some positive constant D. By substituting (21) in (20) and by noticing that $\forall\, i \neq j$, $q_i^{T+S_j} = q_i^{T+S_j-1}$, it follows that

$$x = \alpha_{T+S_j}\left[\sum_{i=1, i\neq j}^{n} \big(\lambda_i^{T+S_j} + \beta_i^{T+S_j}\lambda_{T+S_j}^*\big) q_i^{T+S_j-1} + \lambda_{T+S_j}^* \beta_j^{T+S_j} q_j^T\right].$$

But it also holds that

$$x = \alpha_{T+S_j-1} \sum_{i=1}^{n} \lambda_i^{T+S_j-1} q_i^{T+S_j-1}, \tag{23}$$

where $\alpha_{T+S_j-1} \geq 1$ and $\lambda^{T+S_j-1} \in \Lambda_n$. Since the vectors $q_1^{T+S_j-1}, \ldots, q_n^{T+S_j-1}$ are independent, it follows that

$$\lambda_j^{T+S_j-1} = \frac{\alpha_{T+S_j}}{\alpha_{T+S_j-1}} \lambda^*_{T+S_j} \beta_j^{T+S_j} \leq D\lambda^*_{T+S_j}, \tag{24}$$

where the last inequality follows from (22) and the fact that, in view of Theorem 1, $\alpha_{T+S_j} \leq \alpha_{T+S_j-1}$.

Now let us consider (23). If $S_j > 1$, in (23) we must have a vector $q_v^{T+S_j-1},\ v \neq j$, which is the last one inserted. Since $q_v^{T+S_j-1} \in C(Q_{T+S_j-2})$, it follows that

$$q_v^{T+S_j-1} = \sum_{i=1, i\neq j}^{n} \beta_i^{T+S_j-1} q_i^{T+S_j-2} + \beta_j^{T+S_j-1} q_j^T, \tag{25}$$

where $\beta^{T+S_j-1} \geq 0$. By substituing (24) and (25) in (23), and noticing that, by definition, $\lambda^*_{T+S_j-1} = \lambda_v^{T+S_j-1}$, it follows that

$$\begin{aligned} x = \alpha_{T+S_j-1} \Bigg[& \sum_{i=1, i\neq j}^{n} \left(\lambda_i^{T+S_j-1} + \beta_i^{T+S_j-1} \lambda^*_{T+S_j-1}\right) q_i^{T+S_j-2} + \\ & + \left(\frac{\alpha_{T+S_j}}{\alpha_{T+S_j-1}} \lambda^*_{T+S_j} \beta_j^{T+S_j} + \beta_j^{T+S_j-1} \lambda^*_{T+S_j-1} \right) q_j^T \Bigg] \end{aligned} \tag{26}$$

Again, it also holds that

$$x = \alpha_{T+S_j-2} \left[\sum_{i=1, i\neq j}^{n} \lambda_i^{T+S_j-2} q_i^{T+S_j-2} + \lambda_j^{T+S_j-2} q_j^T \right], \tag{27}$$

where $\alpha_{T+S_j-2} \geq 1$ and $\lambda^{T+S_j-2} \in \Lambda_n$. Comparing (26) and (27), it follows that

$$\lambda_j^{T+S_j-2} = \frac{\alpha_{T+S_j}}{\alpha_{T+S_j-2}} \lambda^*_{T+S_j} \beta_j^{T+S_j} + \frac{\alpha_{T+S_j-1}}{\alpha_{T+S_j-2}} \lambda^*_{T+S_j-1} \beta_j^{T+S_j-1}.$$

Going back for S_j steps in this way, it holds that

$$\lambda_j^T = \sum_{h=1}^{S_j} \frac{\alpha_{T+h}}{\alpha_T} \lambda^*_{T+h} \beta_j^{T+h} \leq D \sum_{h=1}^{S_j} \lambda^*_{T+h},$$

where D is the same as in (22). Note that

$$\sum_{h=1}^{S_j} \lambda^*_{T+h} \leq \sum_{h=1}^{\infty} \lambda^*_{T+h},$$

where, in view of Corollary 1, the right side is the tail of a converging series and then tends to 0 as T increases. Therefore, by choosing T appropriately, the component of x

with respect to the vector q_j^T, can be made arbitrarily small and this can be done for each $j \in \{1, \ldots, n\}$. Then, if T is chosen big enough the condition

$$\sum_{j=1}^{n} \lambda_j^T = 1,$$

can not be fulfilled and $\lambda^T \in \Lambda_n$ is contradicted.

□

Appendix C. Proof of Theorem 3

In what follows it will be assumed by contradiction that

$$\exists\, K: \;\; \forall\, k \geq K, \;\; x \in C(Q_k) = cone\{q_1, \ldots, q_p, q_{p+1}^k, \ldots, q_n^k\}, \tag{28}$$

where $q_1, \ldots, q_p$ are fixed.

First we notice that the linear subproblem (3) can be reformulated as follows

$$\begin{aligned} \max\ \alpha \\ \alpha A Q \lambda \leq b \\ \lambda \in \Lambda_n. \end{aligned} \tag{29}$$

The new reformulation is not useful from the computational point of view, since it is not even a linear program, but it is useful for the theoretical proofs.

Given a cone $C(Q) = cone\{q_1, \ldots, q_n\}$, let $S(Q)$ be the set of rays inside cone $C(Q)$ which contain points in P different from the origin, i.e.

$$S(Q) = E(Q) \cap \{q : \; \exists\, \alpha > 0 \text{ s.t. } \alpha A q \leq b\},$$

where $E(Q)$ is defined in (18). Then we can prove the following lemma.

Lemma 1. *Let $q \in S(Q)$. Then $\exists\, \alpha' = \alpha'(q) \in (0, \mu(Q)]$ such that*

$$\alpha' A\, q \leq b \quad \textit{and} \quad \forall\, \alpha > \alpha' \;\; \alpha A q \not\leq b.$$

Proof. Let $\alpha' = \alpha'(q) = \max\{\alpha \geq 0 : \; \alpha A q \leq b\}$. It follows from the definition of $S(Q)$ that $\alpha' > 0$. Moreover, in view of formulation (29) of subproblem (3), it holds that

$$\mu(Q) = \max_{q \in S(Q)} \alpha'(q),$$

and, consequently, $\alpha' \leq \mu(Q)$.

□

It follows from Lemma 1 that there exists a constraint defining P and with index $t = t(q)$ such that

$$\alpha' a_t q = b_t \;\; \text{and} \;\; b_t > 0. \tag{30}$$

The following observation proves that, under some assumptions, the set $S(Q)$ coincides with the convex hull of the vectors $q_1, \ldots, q_n$.

Observation 2. *If P has nondegenerate vertices and at the beginning of phase II the cone $C(Q_0)$, whose generating rays coincides with the n edges adjacent to the origin, is chosen, then for each cone $C(Q)$*

$$S(Q) = E(Q).$$

Proof. Let q_i be a generating ray of cone $C(Q)$. If q_i has been obtained through a $\omega-$subdivision, then $q_i \in S(Q)$. Then, the only difficulties may be created by the generating rays of the starting cone $C(Q_0)$. But, under the assumptions above, they also belong to $S(Q)$. Therefore, for each $i \in \{1, \dots, n\}$, $q_i \in S(Q)$, so that $S(Q) = conv\{q_1, \dots, q_n\}$.

□

The following lemma states that, given a nested sequence of cones generated by the algorithm and satisfying (28), then the sum of the components of the optimal solutions $\omega(Q_k)$ with respect to the generating rays which are not fixed decreases to 0.

Lemma 2. *Let $\{C(Q_k)\}$ be an infinite nested sequence of cones and let $k \geq K$, where K is the same as in (28). Let*

$$\omega(Q_k) = \alpha_k\Big[\sum_{j=1}^{p} \lambda_j^k q_j + \sum_{j=p+1}^{n} \lambda_j^k q_j^k\Big], \tag{31}$$

where $\alpha_k = \mu(Q_k) > 1$ and $\lambda^k \in \Lambda_n$. Then

$$\forall\, k \geq K \;\; \mu(Q_k) > 1 \implies \sum_{j=p+1}^{n} \lambda_j^k \overset{k\to\infty}{\to} 0.$$

Proof. We prove the lemma by contradiction. If $\sum_{i=p+1}^{n} \lambda_i^k \not\to 0$ then there exists an infinite subsequence $S_1 = \{C(Q_{k_r})\}$, and a $\chi_1 > 0$ such that

$$\forall\, r \quad \sum_{j=p+1}^{n} \lambda_j^{k_r} \geq \chi_1,$$

and there exist $\chi_2 > 0$ and $j = j(r)$ such that for each r it holds that $\lambda_{j(r)}^{k_r} \geq \chi_2$. We can consider a further infinite subsequence S_2 of S_1 and an index $v > p$ such that

$$\forall\, r \; : \; C(Q_{k_r}) \in S_2 \;,\;\; j(r) = v.$$

Now, let $C(Q_k) \in S_2$ and $\omega(Q_k)$ be given by (31). We notice that there exists a cone $C(Q_{k-p_v})$, $p_v \geq 0$, of the complete nested sequence, such that $q_v^{k-p_v} = q_v^k$ is the last inserted vector. Since $\omega(Q_k) \in C(Q_{k-p_v})$

$$\omega(Q_k) = \alpha_{k-p_v}\Big[\sum_{j=1}^{p} \lambda_j^{k-p_v} q_j + \sum_{j=p+1, j\neq v}^{n} \lambda_j^{k-p_v} q_j^{k-p_v} + \lambda_v^{k-p_v} q_v^{k-p_v}\Big], \tag{32}$$

where $\alpha_{k-p_v} \geq \alpha_k$ in view of Theorem 1 and $\lambda^{k-p_v} \in \Lambda_n$. We also notice that $\forall\ j > p$, $j \neq v,\ q_j^k \in Q_{k-p_v}$ and then

$$q_j^k = \sum_{i=1}^{p} \phi_i^{j,k} q_i + \sum_{i=p+1}^{n} \phi_i^{j,k} q_i^{k-p_v},$$

where $\phi^{j,k} \geq 0$. By substitution in (31) for each $j > p$, $j \neq v$, it follows that

$$\omega(Q_k) = \alpha_k \Big[\sum_{j=1}^{p} \big(\lambda_j^k + \sum_{i=p+1,\ i\neq v}^{n} \lambda_i^k \phi_j^{i,k}\big) q_j + \sum_{j=p+1, j\neq v}^{n} \ \sum_{i=p+1,\ i\neq v}^{n} \lambda_i^k \phi_j^{i,k} q_j^{k-p_v} + \big(\lambda_v^k + \sum_{i=p+1,\ i\neq v}^{n} \lambda_i^k \phi_v^{i,k}\big) q_v^{k-p_v} \Big].$$

By comparing the coefficient of $q_v^{k-p_v}$ in this expression with the coefficient of the same vector in (32), and remembering that they must be equal because of the independence of the vectors $q_1, \ldots, q_p, q_{p+1}^{k-p_v}, \ldots, q_n^{k-p_v}$, it follows that

$$\alpha_{k-p_v} \lambda_v^{k-p_v} \geq \alpha_k \lambda_v^k,$$

so that

$$\lambda_v^{k-p_v} \geq \frac{\alpha_k}{\alpha_{k-p_v}} \lambda_v^k \geq c\chi_2,$$

where c is a positive constant (notice that $\alpha_k \geq 1$ and α_{k-p_v} is bounded from above by the finite optimal value $\mu(Q_0)$ of subproblem (3) over the initial cone $C(Q_0)$). In view of Theorem 1 it holds that

$$\mu(Q_k) = eQ_k^{-1}\omega(Q_k) \leq eQ_{k-p_v}^{-1}\omega(Q_k) \leq \frac{eQ_{k-p_v-1}^{-1}\omega(Q_k)}{1+\lambda_v^{k-p_v}\rho} \leq$$

$$\leq \frac{eQ_{k-p_v-1}^{-1}\omega(Q_k)}{1+c\chi_2\rho} \leq \frac{\mu(Q_{k-p_v-1})}{1+c\chi_2\rho}.$$

Since the $v-$th vector, $v > p$, is changed infinitely often in the nested sequence of cones, and $\lambda_v^k \geq \chi_2$ also holds infinitely often, it follows that $\mu(Q_k) \to 0$ which is a contradiction.

□

Let $\{C(Q_k)\}$ be the usual nested sequence of cones and let $k \geq K$, where K is the same as in (28). Let

$$q_k^* = ext[\omega(Q_k); \gamma]. \tag{33}$$

The vector q_k^* is the new column in the matrix Q_{k+1} with respect to the matrix Q_k, i.e. $q_j^{k+1} = q_k^*$ for some $j > p$, while $q_i^{k+1} = q_i^k,\ \forall\ i \neq j$. The following lemma proves that there exists a point $\overline{q}_k$ belonging to $cone\{q_1, \ldots, q_p\}$ and "close" to q_k^*.

Lemma 3. *Let q_k^* be given by (33). Then there exists $\overline{q}_k \in cone\{q_1, \ldots, q_p\}$ such that*

$$\overline{q}_k = u_k \sum_{i=1}^{p} \beta_i^k q_i,$$

and

- $\beta^k \in \Lambda_p$*;*
- $(1+\rho) \leq u_k \leq R_1$*, where $\rho > 0$ is the same as in Theorem 1 and $R_1 > 1$ is a positive constant;*
- $\overline{q}_k - q_k^* = \delta_k \to 0$.

Proof. Let $k \geq K$ and $\omega(Q_k)$ be given by (31). As already seen in the proof of Theorem 1, it holds that

$$q_k^* = \omega(Q_k)\left(1 + \frac{d(q_k^*, \omega(Q_k))}{d(0, \omega(Q_k))}\right).$$

Now let $c_k = \sum_{j=p+1}^{n} \lambda_j^k$ and

$$\overline{q}_k = \alpha_k \left(1 + \frac{d(q_k^*, \omega(Q_k))}{d(0, \omega(Q_k))}\right)\left[\sum_{j=1}^{p} \frac{\lambda_j^k}{1 - c_k} q_j\right].$$

Set

$$\beta_j^k = \frac{\lambda_j^k}{1 - c_k} \quad \forall\, j \in \{1, \ldots, p\}$$

It immediately follows that $\beta^k \in \Lambda_p$.
Then set

$$u_k = \alpha_k \left(1 + \frac{d(q_k^*, \omega(Q_k))}{d(0, \omega(Q_k))}\right) \geq 1 + \rho, \tag{34}$$

where ρ be the same as in Theorem 1 and the inequality follows from the proof of the same theorem and from $\alpha_k > 1$.
Next we observe that α_k is bounded from above by the finite optimal value $\mu(Q_0)$ of subproblem (3) over the initial cone $C(Q_0)$, while, in view of (11), $d(q_k^*, \omega(Q_k)) < R$. Moreover, we notice that there exists a positive constant ρ_2 such that

$$eQ_0^{-1}x \geq 1 \quad \Rightarrow \quad d(0, x) \geq \rho_2,$$

and since

$$1 < eQ_k^{-1}\omega(Q_k) \leq eQ_0^{-1}\omega(Q_k),$$

it follows that $d(0, \omega(Q_k)) \geq \rho_2$. Therefore, by setting

$$R_1 = \mu(Q_0)\left(1 + \frac{R}{\rho_2}\right),$$

it holds that $u_k \leq R_1$.

Finally, let

$$\delta_k = q_k^* - \overline{q}_k = u_k[-c_k \sum_{j=1}^{p} \frac{\lambda_j^k}{1-c_k} q_j + \sum_{j=p+1}^{n} \lambda_j^k q_j^k].$$

Since, in view of Lemma 2, $c_k \to 0$, it follows that $\delta_k \to 0$.

□

The following lemma is only needed to prove the finiteness result when the assumptions of Observation 2 are not fulfilled. Therefore, it can be skipped at a first reading. We recall here that given a sequence $\{s_k\}$ converging to 0, $o(s_k)$ denotes a sequence converging to 0 faster than s_k, i.e.

$$\frac{o(s_k)}{s_k} \to 0 \quad \text{as} \quad k \to \infty.$$

In what follows the notation is extended to vectors, i.e. given a vector s_k whose components all converges to 0, then $o(s_k)$ is a vector whose components all decreases to 0 faster than the corresponding components in s_k.

Lemma 4. *Consider the set $Z = conv\{z_0, \ldots, z_p\}$ where $z_0, \ldots, z_p$ are affinely independent so that Z has dimension p. Let $\{y_k\}$, $y_k \in Z$, be a sequence of points in Z such that*

$$Ay_k \le b + o(s_k),$$

where $s_k \to 0$ as $k \to \infty$. Then there exists $y_k' \in Z \cap P$ such that $d(y_k, y_k') = o(s_k)$.

Proof. During the proof of this lemma, the following notation will be used. Given an affine subspace $H \subseteq R^n$ and a set $S \subset H$, we denote with $relint_H(S)$ the relative interior of S in H. We denote with $lin(H)$ the linear subspace such that $H = v + lin(H)$ for $v \in H$. We will finally denote the set $Z \cap P$ with P'.

First we observe that $P' \neq \emptyset$. Indeed, the compactness of Z and $s_k \to 0$ imply that, possibly by passing to a subsequence

$$y_k \to \overline{y} \in Z \quad \text{and} \quad A\overline{y} \le b,$$

i.e. $\overline{y} \in P'$.

We can restrict our attention to the constraints defining P which are violated by at least one vertex z_i, $i \in \{0, \ldots, p\}$. Indeed, if $a_j z_i \le b_j$, $\forall\, i \in \{0, \ldots, p\}$, then

$$a_j x \le b_j \;\; \forall\, x \in Z.$$

Therefore, we are only interested in the inequalities whose indices are in the following set

$$J = \{j : \; a_j z_i > b_j, \text{ for some } i \in \{0, \ldots, p\}\}.$$

Let H_p be the p-dimensional affine subspace in which Z is contained. Then, only one of the two following cases is possible.

Case I

$$relint_{H_p}(P') \neq \emptyset,$$

(see Fig. 1) then we can consider a point $\overline{z} \in relint_{H_p}(P')$ and it must hold that $a_j\overline{z} < b_j$ for any $j \in J$. Indeed, if for some $j' \in J$ it holds that $a_{j'}\overline{z} = b_{j'}$, then $\exists\, \overline{\pi} > 0$ such that

$$\forall\, d \in lin(H_p),\ \|d\| = 1 \text{ and } \forall\, \pi \in [-\overline{\pi}, \overline{\pi}] \quad a_{j'}(\overline{z} + \pi d) \leq b_{j'},$$

which implies

$$a_{j'}x = b_{j'}, \quad \forall\, x \in Z,$$

thus contradicting the fact that $j' \in J$. In such a case we can consider the segment $[y_k, \overline{z}]$ and take y'_k as the intersection between this segment and the border of P. The intersection has a distance from y_k equal to $o(s_k)$, i.e we have $d(y_k, y'_k) = o(s_k)$. Indeed, let for any $j \in J$

$$\phi_j^k = \min\{\phi :\ \phi \geq 0,\ a_j[(1-\phi)y_k + \phi\overline{z}] \leq b_j\}.$$

If $a_j y_k \leq b_j$ then it holds $\phi_j^k = 0$. Otherwise, since $a_j\overline{z} < b_j$, it holds that $a_j\overline{z} = b_j - \xi$ for some $\xi > 0$. Then ϕ_j^k is given by the solution of the following equation

$$(1-\phi)a_j y_k + \phi a_j \overline{z} = b_j,$$

or, equivalently,

$$(1-\phi)(b_j + o(s_k)) + \phi(b_j - \xi) = b_j,$$

so that

$$\phi_j^k = \frac{o(s_k)}{o(s_k) + \xi} = o(s_k).$$

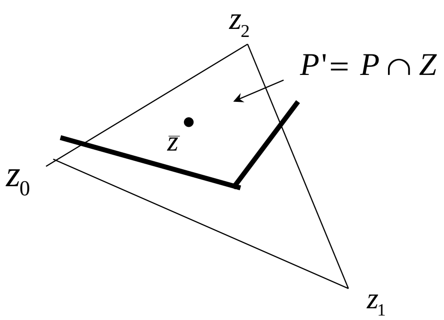

Fig. 1. Case 1

Since this is true for any $j \in J$, it must hold that

$$\overline{\phi}^k = \max\{\phi_j^k :\ j \in J\} = o(s_k),$$

and since $y'_k = (1 - \overline{\phi}^k)y_k + \overline{\phi}^k\overline{z}$, it follows that $d(y'_k, y_k) = o(s_k)$.

Case II

$$relint_{H_p}(P') = \emptyset,$$

(see Fig. 2) so that P' has dimension lower than p. In this case we first prove that $\exists\ \underline{j} \in J$ such that

$$\forall\ x \in P' :\ a_{\underline{j}}x = b_{\underline{j}}. \tag{35}$$

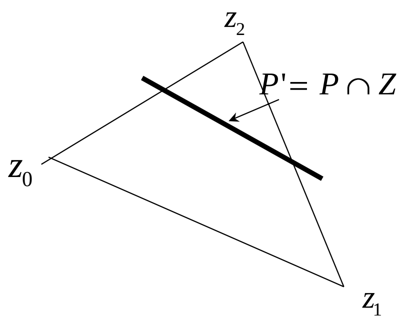

Fig. 2. Case 2

Only one of the following situations holds.

Situation I P' only consists of one point $\overline{z}$;
Situation II the points in P' generate an affine subspace $\tilde{H}$, whose dimension is at least 1 and is obviously lower than p.

In Situation I we can not have that $a_j\overline{z} < b_j$ for all $j \in J$, otherwise we would have that all points in the intersection of $conv\{z_0, \ldots, z_p\}$ with a small neighbourhood of $\overline{z}$ belong to P, thus contradicting the fact that P' only contains a single point. Therefore, (35) obviously holds. In Situation II, we can consider a point $\overline{z} \in relint_{\tilde{H}}(P')$. If we had that $a_j\overline{z} < b_j\ \forall\ j \in J$, then we would have that all points in the intersection of $conv\{z_0, \ldots, z_p\}$ with a small neighbourhood of $\overline{z}$ belong to P, thus contradicting the fact that P' has dimension lower than p. Therefore, $a_{\underline{j}}\overline{z} = b_{\underline{j}}$ for some $\underline{j} \in J$, and $\overline{z} \in relint_{\tilde{H}}(P')$ implies that (35) holds also in this situation.
Now we consider the $(p-1)$-dimensional set:

$$Z_1 = Z \cap \{x :\ a_{\underline{j}}x = b_{\underline{j}}\} \supseteq P',$$

where $\underline{j}$ is the same as in (35). Note that since $\underline{j} \in J$, then

$$\exists\ z_l,\ l \in \{0, \ldots, p\},\ \text{such that}\ a_{\underline{j}}z_l > b_{\underline{j}}.$$

We want to find a point $y_k^1 \in Z_1$ and such that $d(y_k^1, y_k) = o(s_k)$. We recall that $a_{\underline{j}}y_k = b_{\underline{j}} + o(s_k)$. If y_k is such that $a_{\underline{j}}y_k \le b_{\underline{j}}$, then we take y_k^1 as the intersection of $a_{\underline{j}}x = b_{\underline{j}}$ with the segment $[z_l, y_k]$. It holds that $d(y_k, y_k^1) = o(s_k)$. Otherwise set $r_k^0 = y_k$ and let

$$r_k^1 = z_l + \eta'(y_k - z_l),$$

where

$$\eta' = \sup\{\eta :\ a_{\underline{j}}[z_l + \eta(y_k - z_l)] > b_{\underline{j}},\ \ z_l + \eta(y_k - z_l) \in Z\}.$$

Note that $r_k^1 \in Z$ and $d(r_k^1, y_k) = o(s_k)$. If $r_k^1 \in Z_1$ (see Fig. 3) then set $y_k^1 = r_k^1$. Otherwise it holds that $a_{\underline{j}}r_k^1 > b_{\underline{j}}$ and $a_{\underline{j}}r_k^1 = b_{\underline{j}} + o(s_k)$. Moreover, r_k^1 belongs to a proper face F of Z and there must exist a vertex z_l' of this proper face such that $a_{\underline{j}}z_l' > b_{\underline{j}}$ (see Fig. 4). Then we define

$$r_k^2 = z_l' + \eta''(r_k^1 - z_l'),$$

where

$$\eta'' = \sup\{\eta :\ a_{\underline{j}}[z_l' + \eta(r_k^1 - z_l')] > b_{\underline{j}},\ \ z_l' + \eta(r_k^1 - z_l') \in F\}.$$

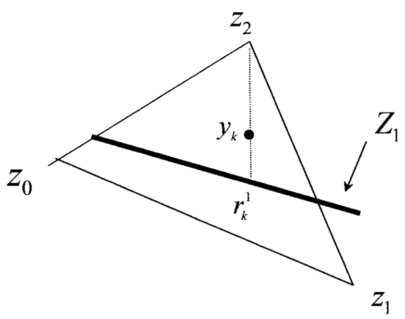

Fig. 3. $r_k^1 \in Z_1$ $(z_1 = z_2)$

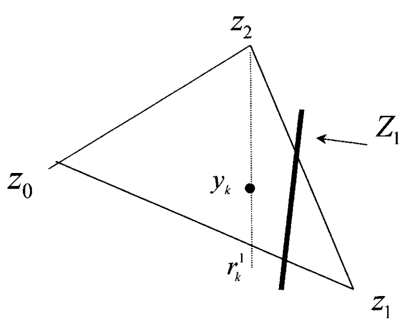

Fig. 4. $r_k^1 \notin Z_1(z_1 = z_2, z_l' = z_0)$

Again we notice that $r_k^2 \in F$ and $d(r_k^2, r_k^1) = o(s_k)$. If $r_k^2 \in Z_1$ then set $y_k^1 = r_k^2$. Otherwise it holds that $a_{\underline{j}} r_k^2 > b_{\underline{j}}$ and $a_{\underline{j}} r_k^2 = b_{\underline{j}} + o(s_k)$. Moreover, r_k^2 belongs to a proper face F' of F and there must exist a vertex z_l'' of F' such that $a_{\underline{j}} z_l'' > b_{\underline{j}}$, so that we can repeat the whole procedure.
Since the dimension of the face is reduced by at least one at each iteration, we finally stop after a finite number of iterations with $y_k^1 = r_k^j$ for some $j \in \{1, \ldots, p\}$ and since $d(r_k^{i-1}, r_k^i) = o(s_k)$ for any $i \in \{1, \ldots, j\}$, then, in view of the triangular inequality, $d(y_k^1, y_k) = o(s_k)$.
At this point we can repeat the whole reasoning with

- y_k^1 replacing y_k;
- Z_1 replacing Z;
- $$J_1 = \{j : \exists \text{ a vertex } v \text{ of } Z_1 \text{ such that } a_j v > b_j\},$$

  replacing J (notice that J_1 is strictly contained in J since the index $\underline{j}$ in (35) belongs to J but not to J_1);
- the affine subspace $H_{p-1} \supset Z_1$, with dimension $p-1$, replacing H_p.

Again two cases are possible.

Case I $relint_{H_{p-1}}(P') \neq \emptyset$.
Case II $relint_{H_{p-1}}(P') = \emptyset$.

If Case I holds a point y_k' is obtained by moving from y_k^1 to a point in the relative interior. Then y_k' satisfies $d(y_k', y_k^1) = o(s_k)$ and the triangular inequality implies $d(y_k', y_k) = o(s_k)$. If Case II holds we decrease by one the dimension and we repeat again the whole reasoning.
Obviously Case II can only hold a finite number of times and the final point y_k' will be such that $d(y_k', y_k) = o(s_k)$.

□

Finally we are ready for the proof of the theorem.

Proof. By contradiction it is assumed that (28) holds. We consider version (29) of subproblem (3) Let $\omega(Q_k)$ be given by (31). Recall that $\alpha_k = \mu(Q_k)$. Let $c_k = \sum_{j=p+1}^n \lambda_j^k$ as in the proof of Lemma 3. Finally, let $\overline{\alpha_k}$ be the optimum value of subproblem (29) with $Q = Q_k$ but with the further constraints $\lambda_i = 0 \ \forall\, i > p$. First we notice that $\forall\, k \geq K,\ c_k > 0$. Indeed

$$c_k = 0 \implies \omega(Q_k) \in cone\{q_1, \ldots, q_p\},$$

and, in view of Step 6 of the algorithm in Sect. 2, this would imply the replacement of one of the first p vectors, thus contradicting (28). But, in view of Lemma 2, it holds that $c_k \to 0$. Since, by assumption, in the nested sequence $\{C(Q_k)\}$ the i−th generating ray, $i > p$, is replaced infinitely often, then there exists $K_1 \geq K$ such that, for $k \geq K_1$

$$\exists\, \ell_i \geq 0 : \quad q^*_{k-\ell_i} = q_i^k,$$

where the notation (33) is employed. Therefore, in view of Lemma 3, it holds that

$$q_i^k = \overline{q}_i^k + \delta_i^k = w_i \sum_{j=1}^{p} \beta_j^{k-\ell_i} q_j + \delta_i^k, \tag{36}$$

where $w_i = u_{k-\ell_i}$, $\beta^{k-\ell_i} \in \Lambda_p$, and $\delta_i^k \to 0$. Then, the constraints $\alpha A Q_k \lambda \leq b$ can be written as

$$\alpha\Big[\sum_{j=1}^{p} \lambda_j A q_j + \sum_{j=p+1}^{n} \lambda_j w_j \sum_{i=1}^{p} \beta_i^{k-\ell_j} A\, q_i\Big] + \alpha\Big[\sum_{j=p+1}^{n} \lambda_j A \delta_j^k\Big] \leq b,$$

which can be further rewritten as

$$\alpha \left[\sum_{j=1}^{p} \big[\lambda_j + \sum_{i=p+1}^{n} \lambda_i w_i \beta_j^{k-\ell_i}\big] A\, q_j + \sum_{j=p+1}^{n} \lambda_j A \delta_j^k\right] \leq b. \tag{37}$$

Now we consider the optimal solution $\omega(Q_k)$ given by (31) and the quantity ν_k defined as follows

$$\nu_k = \sum_{j=1}^{p} \big[\lambda_j^k + \sum_{i=p+1}^{n} \lambda_i^k w_i \beta_j^{k-\ell_i}\big].$$

Since $\sum_{j=1}^{n} \lambda_j^k = 1$ and $\sum_{j=1}^{p} \beta_j^{k-\ell_i} = 1$, it follows that

$$\nu_k = (1 - c_k) + \sum_{i=p+1}^{n} \big[\lambda_i^k w_i \sum_{j=1}^{p} \beta_j^{k-\ell_i}\big] = (1 - c_k) + \sum_{i=p+1}^{n} \lambda_i^k w_i.$$

Since, in view of Lemma 3, $\forall\, i > p,\ w_i \geq 1 + \rho$, it holds that

$$\nu_k \geq 1 + c_k \rho. \tag{38}$$

Now we consider the point

$$\tilde{q}_k = \sum_{j=1}^{p} \tilde{\lambda}_j^k q_j,$$

where

$$\tilde{\lambda}_j^k = \frac{\lambda_j^k + \sum_{i=p+1}^{n} \lambda_i^k w_i \beta_j^{k-\ell_i}}{\nu_k}.$$

Since all the terms involved are nonnegative, it holds that $\tilde{\lambda}_j^k \geq 0$ for any $j \in \{1, \ldots, p\}$, and from the definition of ν_k, it follows that $\sum_{j=1}^{p} \tilde{\lambda}_j^k = 1$, i.e. $\tilde{q}_k \in conv\{q_1, \ldots, q_p\}$. By substituting (36) in (31) it follows that

$$\omega(Q_k) = \alpha_k \left[\sum_{j=1}^{p} \big[\lambda_j^k + \sum_{i=p+1}^{n} \lambda_i^k w_i \beta_j^{k-\ell_i}\big] q_j + \sum_{j=p+1}^{n} \lambda_j^k \delta_j^k\right].$$

Therefore, it follows from the definition of $\tilde{\lambda}_k$ that

$$\omega(Q_k) = \alpha_k \nu_k \tilde{q}_k + \alpha_k \sum_{i=p+1}^{n} \lambda_i^k \delta_i^k. \tag{39}$$

Since, in view of Lemma 3, $\delta_i^k \to 0$, it holds that

$$\omega(Q_k) = \alpha_k \nu_k \tilde{q}_k + o(c_k),$$

and the feasibility of $\omega(Q_k)$ implies that

$$\alpha_k \nu_k A \tilde{q}_k \leq b + o(c_k).$$

If the assumptions in Observation 2 are not fulfilled, then we can apply Lemma 4 with

- $y_k = \alpha_k \nu_k \tilde{q}_k$;
- $Z = conv\{z_0, \ldots, z_p\}$, with $z_0 = 0$ and $z_i = r q_i$, for $i \in \{1, \ldots, p\}$, where $r > 1$ and big enough so that

  $$y_k \in Z \text{ and } cone\{q_1, \ldots, q_p\} \cap P \subseteq Z;$$

  (such a value r exists in view of the boundedness of P and the fact that, in view of (10), $d(q_i, 0) > \delta$ for each $i \in \{1, \ldots, p\}$);
- $s_k = c_k$.

Then it follows from Lemma 4 that there exists $y_k' = \alpha' q_k' \in cone\{q_1, \ldots, q_p\} \cap P$, for some $\alpha' > 0$ and $q_k' \in conv\{q_1, \ldots, q_p\}$, such that $d(y_k', y_k) = o(c_k)$. It follows that $q_k' = \tilde{q}_k + o(c_k) \in S(Q_k)$. If the assumptions of Observation 2 are fulfilled, then $\tilde{q}_k \in S(Q_k)$, and we can directly set $q_k' = \tilde{q}_k$ without using Lemma 4.

Since $\overline{\alpha_k}$ is the optimum of subproblem (29) for $Q = Q_k$ and with the further constraints $\lambda_i = 0 \ \forall\, i > p$, it follows from the proof of Lemma 1 and its consequence (30), that there exist an index $\tilde{t}_k = \tilde{t}_k(q_k')$ and $\tilde{\alpha}_k \leq \overline{\alpha_k}$ such that

$$\tilde{\alpha}_k a_{\tilde{t}_k} q_k' = b_{\tilde{t}_k},$$

with $b_{\tilde{t}_k} > 0$. Then, obviously,

$$\overline{\alpha_k} a_{\tilde{t}_k} q_k' = \overline{\alpha_k} a_{\tilde{t}_k} (\tilde{q}_k + o(c_k)) \geq b_{\tilde{t}_k}. \tag{40}$$

Now we assume that

$$\alpha_k \nu_k \geq \overline{\alpha_k}(1 + c_k \rho). \tag{41}$$

Then from (40) and (41) it follows that

$$\alpha_k \nu_k a_{\tilde{t}_k} \tilde{q}_k \geq b_{\tilde{t}_k} + c_k \rho b_{\tilde{t}_k} + o(c_k).$$

In view of (39) and of the fact that $\omega(Q_k)$ is feasible, it holds that

$$a_{\tilde{t}_k} \omega(Q_k) = \alpha_k \nu_k a_{\tilde{t}_k} \tilde{q}_k + \alpha_k \sum_{i=p+1}^{n} \lambda_i^k a_{\tilde{t}_k} \delta_i^k \leq b_{\tilde{t}_k}.$$

Then

$$\alpha_k \sum_{i=p+1}^{n} \lambda_i^k a_{\tilde{t}_k} \delta_i^k \leq -c_k \rho b_{\tilde{t}_k} + o(c_k).$$

Therefore

$$\alpha_k \sum_{i=p+1}^{n} \lambda_i^k \parallel a_{\tilde{t}_k} \parallel\parallel \delta_i^k \parallel \geq \mid \alpha_k \sum_{i=p+1}^{n} \lambda_i^k a_{\tilde{t}_k} \delta_i^k \mid \geq c_k \rho b_{\tilde{t}_k} + o(c_k).$$

But, for a big enough k this is not possible because the term on the left is $o(c_k)$, in view of the fact that $\parallel \delta_i^k \parallel \longrightarrow 0$. Then it must hold

$$\alpha_k v_k < \overline{\alpha_k}(1 + c_k \rho),$$

But from (38) we notice that this implies $\alpha_k < \overline{\alpha_k}$, i.e. the maximum over the set $Q_k \cap P$ is lower than the maximum over a subset of $Q_k \cap P$, which is not possible.

□

Appendix D. Proof of Theorem 5

We notice that in the proof by contradiction given in the previous sections, $\varepsilon > 0$ is only used in order to prove, through Observation 1, that (10), (16) and (34) hold.

If $\varepsilon = 0$, then (10) does not hold. However, in the proofs of Sects. A-C a condition milder than (10) is actually needed. Indeed, it is enough to prove that for each $x \in C(Q_0)$, $x \neq 0$, it holds that

$$d(0, ext[x; \tau]) > \delta, \tag{42}$$

for some $\delta > 0$. In view of the nondegeneracy assumption, we can set

$$q_j^0 = ext[v_j; \tau] \quad \forall \ j \in \{1, \ldots, n\},$$

where $v_j \in adj(\overline{v})$. The matrix Q_0 whose column vectors are the vectors q_j^0 is, as required in Phase I, nonsingular and such that $C(Q_0) \supset P$. Since it holds that $d(v_j, 0) \geq \delta_1$, for some $\delta_1 > 0$, then

$$eQ_0^{-1}x \geq 1 \ \ \Rightarrow \ \ d(0, x) > \delta, \tag{43}$$

for some $\delta > 0$. Moreover, for each $x \in C(Q_0)$

$$eQ_0^{-1}x \leq 1 \ \ \Rightarrow \ \ f(x) \geq \tau. \tag{44}$$

Therefore, it follows from the definition (2) and (44) that for each $x \in C(Q_0),\ x \neq 0$

$$eQ_0^{-1} ext[x; \tau] \geq 1,$$

and it follows from (43) that (42) holds.

Next we prove that $\mu(Q_k) \downarrow 1$. Let us assume, by contradiction, that $\mu(Q_k) \downarrow (1+\theta)$, $\theta > 0$. Then (16) and (34) still hold with θ replacing ρ even for $\varepsilon = 0$. Therefore, in order to show that $\mu(Q_k) \downarrow 1$, we can employ the same proof by contradiction used in the previous sections.

Finally we assume, again by contradiction, that $\tau = f^* + \kappa_1$, for some $\kappa_1 > 0$. Let $x^* \in P$ be such that $f(x^*) = f^*$, and let $y \in [0, x^*] \subset P$ be the point such that $f(y) = f^* + \frac{\kappa_1}{2} < \tau$. It follows from Observation 1 and the boundedness of P that $x^* = (1+\kappa_2)y$, for some $\kappa_2 > 0$. Now let $C(Q)$ be a cone generated by the algorithm such that $x^* \in C(Q)$. By recalling that, for each $x \in C(Q)$, $eQ^{-1}x \leq 1$ implies $f(x) \geq \tau$ (see (4)), it follows that $eQ^{-1}y > 1$. Therefore, it holds that

$$\mu(Q) \geq eQ^{-1}x^* \geq 1 + \kappa_2 \quad \forall\, C(Q) \ni x^*. \tag{45}$$

Given an infinite nested sequence $\{C(Q_k)\}$, we proved above that $\mu(Q_k) \downarrow 1$. Therefore, there exists a positive integer K_2 such that

$$\mu(Q_{K_2}) < 1 + \kappa_2.$$

But it follows from the selection rule (5) and from (45) that $C(Q_{K_2})$ will never be selected to be further subdivided, thus contradicting the infinity of the sequence.

References

1. Bali, S. (1973): Minimization of a concave function on a bounded convex polyhedron. Ph.D. Dissertation, University of California, Los Angeles, CA
2. Benson, H.P. (1995): Concave minimization: theory, applications and algorithms. In: Horst, R., Pardalos, P., eds., Handbook of Global Optimization, pp. 43–148. Kluwer Academic Publishers, Dordrecht, The Netherlands
3. Hamami, M., Jacobsen, S.E. (1988): Exhaustive nondegenerate conical processes for concave minimization on convex polytopes. Math. Oper. Res. **13**, 479–487
4. Horst,R. (1976): An algorithm for nonconvex programming problems. Math. Program. **10**, 312–321
5. Horst, R. (1984): On the global minimization of concave functions. Introduction and survey. OR Spektrum **6**, 195–205
6. Horst, R., Thoai, N.V. (1989): Modification, implementation and comparison of three algorithms for globally solving linearly constrained concave minimization problems. Computing **42**, 271–289
7. Horst, R., Tuy, H. (1993): Global Optimization: Deterministic Approaches, 2nd edn. Springer
8. Jaumard, B., Meyer, C. (1996): On the convergence of cone splitting algorithms with $\omega-$subdivisions. Les Cahiers du GERAD G-96-36, GERAD, July 1996. Conditional acceptance in the J. Optim. Theory Appl.
9. Jaumard, B., Meyer, C. (1998): A simplified convergence proof for the cone partitioning algorithm. J. Glob. Optim. **13**, 407–416
10. Mangasarian, O.L. (1969): Nonlinear optimization. McGraw-Hill
11. Pardalos, P.M., Rosen, J.B. (1986): Methods for global concave optimization: a bibliographic survey. SIAM Rev. **26**, 367–379
12. Pardalos, P.M., Schnitger, G. (1988): Checking local optimality in constrained quadratic programming is NP-hard. Oper. Res. Lett. **7**, 33–35
13. Thoai, N.V., Tuy, H. (1980): Convergent algorithm for minimizing a concave function. Math. Oper. Res. **5**, 556–566
14. Tuy, H. (1964): Concave programming under linear constraints. Sov. Math. **5**, 1437–1440
15. Tuy, H., Khachaturov, V., Utkin, S. (1987): A class of exhaustive cone splitting procedures in conical algorithms for concave minimization. Optimization **18**, 791–808
16. Tuy, H., Horst, R. (1988): Convergence and restart in branch and bound algorithms for global optimization algorithms. Math. Program. **41**, 161–184
17. Tuy, H. (1991): Normal conical algorithm for concave minimization over polytopes. Math. Program. **51**, 229–245
18. Tuy, H. (1991): Effect of the subdivision strategy on convergence and efficiency of some global optimization algorithms. J. Glob. Optim. **1**, 23–36
19. Zwart, P.B. (1974): Global maximization of a convex function with linear inequality constraints. Oper. Res. **22**, 602–609