A Convergent Conical Algorithm with ω-Bisection for Concave Minimization

Takahito Kuno* and Tomohiro Ishihama

Graduate School of Systems and Information Engineering

University of Tsukuba, Tsukuba, Ibaraki 305-8573, Japan

April 8, 2013

Revised: April 25, 2014

Abstract

The conical algorithm is a global optimization algorithm proposed by Tuy in 1964 to solve concave minimization problems. Introducing the concept of pseudo-nonsingularity, we give an alternative proof of convergence of the algorithm with the ω-subdivision rule. We also develop a new convergent subdivision rule, named ω-bisection, and report numerical results of comparing it with the usual ω-subdivision.

Key words: Global optimization, concave minimization, conical algorithm, bisection, ω-subdivision.

1 Introduction

The concave minimization is a typical multiextremal global optimization problem, in which a locally optimal solution is not always globally optimal. From the viewpoint of computational complexity, it is known to be NP-hard [10]. To solve this intractable but valuable problem, Tuy made use of valid cuts and in 1964 proposed a first systematic solution method [12], called the *conical algorithm*, which turned out later to have no guarantee of convergence [16]. Bali [1] and Zwart [17] tried to modify the algorithm independently using the same device as in [12], i.e., ω-*subdivision*, in the early 1970s. According to ω-subdivision, cones generated in the algorithm are subdivided along a direction $\boldsymbol{\omega}$ which is given as a byproduct in the bounding process. The convergence of the algorithm with ω-subdivision, however, had remained an open question for a quarter century. During that time, a number of subdivision rules were proposed to guarantee the convergence [11, 13, 14], but none of them surpassed

*The author was partially supported by a Grant-in-Aid for Scientific Research (C 25330022) from Japan Society for the Promotion of Science. E-mail: takahito@cs.tsukuba.ac.jp

ω-subdivision in empirical efficiency. In 1998 and 99, Jaumard-Meyer [6, 7] and Locatelli [8] respectively proved the convergence of the algorithm with ω-subdivision in different ways, and finally settled the argument over the convergence.

Preceding Jaumar-Meyer and Locatelli by almost ten years, Tuy showed in [13] that the algorithm with ω-subdivision converges if a certain nonsingularity condition holds for any nested sequence of generated cones. This result is not exploited in either [6, 7] or [8], and besides it is still an open question whether the condition actually holds or not. In this paper, we introduce a similar nonsingularity condition and show that it does hold for every nested sequence of cones. Using this condition, we derive a convergence result under a more general scheme of cone subdivision, including ω-subdivision. Based on the scheme, we propose a new subdivision rule, named ω-bisection, which allows the algorithm to locate a correct solution just as ω-subdivision does.

The organization of the paper is as follows. In Section 2, we first define the d.c. feasibility problem, which needs to be solved to obtain an optimal solution of a concave minimization problem, and then illustrate how the conical algorithm solves it. In Section 3, we define the pseudo-nonsingularity, a substitute for Tuy's nonsingularity condition, and show that it holds for any nested sequence of cones generated in the conical algorithm. In Section 4, using this condition, we derive the convergence result and give an alternative proof of convergence for the conical algorithm with ω-subdivision. In Section 5, on the basis of our discussion so far, we develop the ω-bisection rule for subdividing cones, and show that the conical algorithm with ω-bisection converges to a correct solution of the d.c. feasibility problem. Lastly, in Section 6, we report the result of numerical comparison between the algorithm with ω-bisection and the usual one with ω-subdivision, and conclude the paper.

2 D.c. feasibility and the conical algorithm

Let $f : S(\subset \mathbb{R}^n) \to \mathbb{R}$ be a concave function and denote its upper level set for a real number α by

$$C(\alpha) = \{\mathbf{x} \in S \mid f(\mathbf{x}) \geq \alpha\}.$$

Also let $D \subset \mathbb{R}^n$ be a polyhedron defined as

$$D = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \leq \mathbf{b}\},$$

where $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$ ($n < m$). We assume that D has nonempty interior and is included in the interior of S, and hence f is continuous on D. Both $C(\alpha)$ and D are convex sets, but their difference $D \setminus C(\alpha)$ is not convex in general. The problem we consider is to search for a point in this *d.c. set* (difference of two convex sets) within a prescribed tolerance $\varepsilon \geq 0$, i.e.,

$$\text{(DC)}: \text{ find a point } \mathbf{x} \in D \setminus C(\alpha) \text{ if there is one, or else prove that } D \subset C(\alpha - \varepsilon),$$

which is called the *d.c. feasibility problem*. For the sake of simplicity, we assume that both $C(\alpha)$ and D are bounded sets. As a consequence, $C(\gamma)$ is also bounded for any number γ other than α, since all nonempty level sets of a concave function have the same recession cone (see e.g. Theorem 8.7 in [9]).

Associated with (DC) is the concave minimization problem

$$\left|\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{x} \in D. \end{array}\right. \tag{1}$$

It is known, e.g., [13, 14, 15], that a globally ε-optimal solution $\mathbf{x}^*$ of (1) can be computed according to the following two-phase scheme:

Let $\mathbf{z}^1 \in D$ be an initial feasible solution of (1). Also let $i \leftarrow 1$.

Phase 1 (local phase). Starting from $\mathbf{z}^i$, search the vertices of D for a local minimizer of f. Then a vertex $\mathbf{x}^i$ is obtained such that $f(\mathbf{x}^i) \leq f(\mathbf{z}^i)$ and $f(\mathbf{x}^i) \leq f(\mathbf{x})$ for every vertex $\mathbf{x}$ adjacent to $\mathbf{x}^i$.

Phase 2 (global phase). Solve (DC) for $\alpha = f(\mathbf{x}^i)$. If $D \subset C(\alpha - \varepsilon)$, then $\mathbf{x}^* \leftarrow \mathbf{x}^i$ and terminate: $\mathbf{x}^*$ is a globally ε-optimal solution of (1). Otherwise, a feasible solution $\mathbf{z} \in D$ is obtained such that $f(\mathbf{z}) < f(\mathbf{x}^i)$. Let $\mathbf{z}^{i+1} \leftarrow \mathbf{z}$, $i \leftarrow i+1$, and go to Phase 1.

Alternating between these two phases generates a sequence of vertices $\{\mathbf{x}^i \mid i = 1, 2, \ldots\}$ of D such that $f(\mathbf{x}^{i+1}) < f(\mathbf{x}^i)$. Since the number of vertices of a polyhedron is finite, it terminates after finitely many repetitions if (DC) can be solved in finite time. Our goal is therefore to solve (DC) in finite time, using the conical algorithm outlined below.

OUTLINE OF THE CONICAL ALGORITHM

Let $\gamma = \alpha - \varepsilon$, and $\mathbf{v}$ be a vertex of D such that $f(\mathbf{v}) > \gamma$. In the above two-phase scheme, such a vertex can be easily found in the process of searching for $\mathbf{x}^i$ because $\mathbf{x}^i$ is a local minimizer and $f(\mathbf{x}^i) > \gamma$ when $\varepsilon > 0$. By perturbing $\mathbf{b}$ slightly if necessary, we may assume that $\mathbf{v}$ is a nondegenerate vertex of D. The system defining D is then partitioned into

$$\mathbf{B}\mathbf{v} = \mathbf{b}_B, \quad \mathbf{N}\mathbf{v} < \mathbf{b}_N,$$

where $\mathbf{B} \in \mathbb{R}^{n \times n}$, $\mathbf{N} \in \mathbb{R}^{(m-n) \times n}$ are submatrices of $\mathbf{A}$, and $\mathbf{b}_B \in \mathbb{R}^n$, $\mathbf{b}_N \in \mathbb{R}^{(m-n)}$ are the corresponding portions of $\mathbf{b}$. Let

$$\Lambda = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{B}\mathbf{x} \leq \mathbf{b}_B\}, \quad M = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{N}\mathbf{x} \leq \mathbf{b}_N\}.$$

Then we have

$$D = M \cap \Lambda.$$

Since the vertex $\mathbf{v}$ is nondegenerate, it is an interior point of M. It should also be noted that Λ is a polyhedral cone with vertex $\mathbf{v}$ and has exactly n edges. Let $\mathbf{d}_1, \ldots, \mathbf{d}_n$ be directions of the edges of Λ. These vectors are obtained immediately from a general solution of the linear system $\mathbf{Ax} + \mathbf{w} = \mathbf{b}$, where $\mathbf{w} \in \mathbb{R}^m$ is the vector of slack variables.

To simplify the explanation, let us translate the origin $\mathbf{0}$ to $\mathbf{v}$, and again denote by $\mathbf{Nx} \leq \mathbf{b}_N$ the resulting system that defines M. We may assume in the sequel that $\mathbf{b}_N > \mathbf{0}$ because $\mathbf{0}$ has moved to the interior of M. Let $\mathbf{q}_j$ denote the *γ-extension* of $\mathbf{d}_j$, i.e.,

$$\mathbf{q}_j = \mathrm{ext}(\mathbf{d}_j) \equiv \theta_j \mathbf{d}_j, \quad j = 1, \ldots, n,$$

where

$$\theta_j = \sup\{\theta \mid f(\theta \mathbf{d}_j) \geq \gamma\}.$$

Then we have

$$\Lambda = \mathrm{con}(\mathbf{Q}) \equiv \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \sum_{j=1}^{n} \lambda_j \mathbf{q}_j,\ \boldsymbol{\lambda} \geq \mathbf{0}\},$$

where

$$\mathbf{Q} = [\mathbf{q}_1, \ldots, \mathbf{q}_n] \in \mathbb{R}^{n \times n}.$$

Note that $\mathbf{q}_j$'s are linearly independent and $\mathbf{Q}$ is invertible. Therefore, $\mathbf{q}_j$'s determine a unique hyperplane, which is the boundary of

$$G = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{eQ}^{-1}\mathbf{x} \leq 1\},$$

where $\mathbf{e} \in \mathbb{R}^n$ is the all-ones row vector. Obviously, $G \cap \Lambda$ is a simplex with $n+1$ vertices $\mathbf{q}_j$'s and $\mathbf{0}$, all belonging to $C(\gamma)$. From the convexity of $C(\gamma)$ we see that

$$G \cap \Lambda \subset C(\gamma).$$

Accordingly, if $M \cap \Lambda$ is a subset of G, we can conclude that (DC) is solved because

$$D = M \cap \Lambda \subset G \cap \Lambda \subset C(\gamma) = C(\alpha - \varepsilon).$$

The process of checking whether $M \cap \Lambda \subset G$ is usually called *bounding*. We also refer to G as the *γ-valid cut*[1] for the cone Λ.

If $M \cap \Lambda$ is not a subset of G, then either a point $\mathbf{x} \in D \setminus C(\alpha)$ is found or Λ needs to be divided into subcones for further examinations. In the latter case, an appropriate direction $\mathbf{u}$ is selected from $\Lambda \setminus \{\mathbf{0}\}$. There exists a vector $\boldsymbol{\lambda}' \geq \mathbf{0}$ such that $\mathbf{u} = \sum_{j=1}^{n} \lambda'_j \mathbf{q}$ and $\boldsymbol{\lambda}' \neq \mathbf{0}$. Let

[1] In some literature, the term "γ-valid cut" refers to the closure of the complement of G.

$J = \{j \mid \lambda'_j > 0\}$. Then Λ *is subdivided along* $\mathbf{u}$ into $|J|$ subcones:

$$\Lambda^j = \mathrm{con}(\mathbf{Q}^j), \quad j \in J,$$

where $\mathbf{Q}^j$ is referred to as a child of $\mathbf{Q}$ and defined as

$$\mathbf{Q}^j = [\mathbf{q}_1, \ldots, \mathbf{q}_{j-1}, \mathrm{ext}(\mathbf{u}), \mathbf{q}_{j+1}, \ldots, \mathbf{q}_n].$$

It is easy to see that

$$\mathrm{int}(\Lambda^i) \cap \mathrm{int}(\Lambda^j) = \emptyset \ \text{ if } \ i \neq j; \quad \Lambda = \bigcup_{j \in J} \Lambda^j.$$

In other words, the cones Λ^j's constitute a partition of Λ. This process of dividing Λ is called *branching*. After branching, the bounding process is again applied to each subcone Λ^j.

3 Pseudo-nonsingularity of the algorithm

Suppose the conical algorithm is infinite and generates a nested sequence of cones:

$$\Lambda = \Lambda_1 \supset \cdots \supset \Lambda_k \supset \Lambda_{k+1} \supset \cdots, \tag{2}$$

where Λ_{k+1} is a cone obtained by subdividing Λ_k along a direction $\mathbf{u}^k$. For each k, the cone Λ_k is spanned by an invertible matrix $\mathbf{Q}_k$, i.e.,

$$\Lambda_k = \mathrm{con}(\mathbf{Q}_k) \equiv \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \sum_{j=1}^n \lambda_j \mathbf{q}_j^k, \ \boldsymbol{\lambda} \geq \mathbf{0}\},$$

where $\mathbf{q}_j^k$ is the jth column of $\mathbf{Q}_k$ and lies on the boundary of $C(\gamma)$. Let us denote the γ-valid cut for Λ_k by

$$G_k = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{e}\mathbf{Q}_k^{-1}\mathbf{x} \leq 1\}.$$

As seen in the previous section, we have $M \cap \Lambda_k \subset C(\gamma)$ if $M \cap \Lambda_k \subset G_k$. This can be checked by solving an auxiliary problem

$$\left| \begin{array}{ll} \text{maximize} & \mathbf{e}\mathbf{Q}_k^{-1}\mathbf{x} \\ \text{subject to} & \mathbf{x} \in M \cap \Lambda_k. \end{array} \right. \tag{3}$$

Let $\boldsymbol{\omega}^k$ be an optimal solution of (3) and ζ^k the optimal value, i.e., $\zeta^k = \mathbf{e}\mathbf{Q}_k^{-1}\boldsymbol{\omega}^k$. If $f(\boldsymbol{\omega}^k) < \alpha$, then $\boldsymbol{\omega}^k$ is obviously a solution to (DC), and the conical algorithm terminates. Since the sequence (2) is infinite, that is not the case and we assume that

$$f(\boldsymbol{\omega}^k) \geq \alpha \geq \gamma, \quad k = 1, 2, \ldots. \tag{4}$$

Similarly, if $\zeta^k \leq 1$, then $M \cap \Lambda_k \subset G_k$, and we can conclude that Λ_k contains no solution to (DC). In that case, Λ_k is discarded from further consideration. However, we assume here that

$$\zeta^k > 1, \quad k = 1, 2, \ldots. \tag{5}$$

The conical algorithm is known to be convergent if the direction $\mathbf{u}^k$ of subdividing Λ_k coincides with $\boldsymbol{\omega}^k$ for every k. This subdivision rule is called ω-subdivision, and the convergence result was established independently by Jaumard-Meyer in 98 [6, 7] and by Locatelli in 99 [8]. Almost ten years earlier than those, Tuy showed that the algorithm with ω-subdivision converges if the sequence (2) is *nonsingular*[2] [13] (see also [5, 14]), i.e., there exists a subsequence $\{k_r \mid r = 1, 2, \ldots\}$ and a constant L such that

$$\|\mathbf{e}\mathbf{Q}_{k_r}^{-1}\| \leq L, \quad r = 1, 2, \ldots. \tag{6}$$

Unfortunately, it remains an open question whether (6) holds or not. In the rest of this section, we introduce another problem equivalent to (3) and show that the coefficients of its objective function satisfies a condition similar to (6). For this reason, we say that the sequence (2) is *pseudo-nonsingular*, from which we will derive a convergence result under a more general cone subdivision rule.

LINEAR PROGRAM EQUIVALENT TO (3)

The auxiliary problem (3) is a linear program of the form

$$(\mathrm{P}_k) \left| \begin{array}{ll} \text{maximize} & \mathbf{e}\mathbf{Q}_k^{-1}\mathbf{x} \\ \text{subject to} & \mathbf{N}\mathbf{x} \leq \mathbf{b}_N, \quad \mathbf{Q}_k^{-1}\mathbf{x} \geq \mathbf{0}. \end{array} \right.$$

Since the inversion of $\mathbf{Q}_k$ is not always numerically so stable, (P_k) is usually solved in the following form

$$\left| \begin{array}{ll} \text{maximize} & \mathbf{e}\boldsymbol{\lambda} \\ \text{subject to} & \mathbf{N}\mathbf{Q}_k\boldsymbol{\lambda} \leq \mathbf{b}_N, \quad \boldsymbol{\lambda} \geq \mathbf{0}. \end{array} \right. \tag{7}$$

Even if $\mathbf{Q}_k$ fails to be invertible, (7) can be defined and has an optimal solution $\boldsymbol{\lambda}^k$. The optimal solution of (P_k) is then given by $\boldsymbol{\omega}^k = \mathbf{Q}_k\boldsymbol{\lambda}^k$. The dual problem of (7) is as follows

$$\left| \begin{array}{ll} \text{minimize} & \boldsymbol{\mu}\mathbf{b}_N \\ \text{subject to} & \boldsymbol{\mu}\mathbf{N}\mathbf{Q}_k \geq \mathbf{e}, \quad \boldsymbol{\mu} \geq \mathbf{0}. \end{array} \right. \tag{8}$$

[2] Instead of "nonsingular", Tuy used the term "nondegenerate" derived from an analogous concept in [4]. However, since it is easily confused with nondegeneracy in linear programming, we use "nonsingular" in view of its relation to the invertibility of $\mathbf{Q}_k$. Also the definition here follows that in [5].

This problem also has an optimal solution $\boldsymbol{\mu}^k$, and by the assumption (5) we have

$$\mathbf{e}\boldsymbol{\lambda}^k = \boldsymbol{\mu}^k \mathbf{b}_N = \zeta^k > 1.$$

For the dual solution $\boldsymbol{\mu}^k$, let us define another linear program

$$(\mathrm{P}'_k) \left| \begin{array}{ll} \text{maximize} & \boldsymbol{\mu}^k \mathbf{N}\mathbf{x} \\ \text{subject to} & \mathbf{N}\mathbf{x} \leq \mathbf{b}_N, \quad \mathbf{Q}_k^{-1}\mathbf{x} \geq \mathbf{0}, \end{array} \right.$$

which is equivalent to (P_k) in the following sense.

Lemma 3.1. *An optimal solution of (P'_k) is $\boldsymbol{\omega}^k = \mathbf{Q}_k \boldsymbol{\lambda}^k$, and the optimal value is equal to ζ^k. Conversely, if $\mathbf{x}'$ is an optimal solution of (P'_k), then $\mathbf{x}'$ is an optimal solution of (P_k).*

Proof. Problem (P'_k) is equivalent to

$$\left| \begin{array}{ll} \text{maximize} & \boldsymbol{\mu}^k \mathbf{N}\mathbf{Q}_k \boldsymbol{\lambda} \\ \text{subject to} & \mathbf{N}\mathbf{Q}_k \boldsymbol{\lambda} \leq \mathbf{b}_N, \quad \boldsymbol{\lambda} \geq \mathbf{0}. \end{array} \right. \tag{9}$$

The dual problem is

$$\left| \begin{array}{ll} \text{minimize} & \boldsymbol{\mu} \mathbf{b}_N \\ \text{subject to} & \boldsymbol{\mu}\mathbf{N}\mathbf{Q}_k \geq \boldsymbol{\mu}^k \mathbf{N}\mathbf{Q}_k, \quad \boldsymbol{\mu} \geq \mathbf{0}. \end{array} \right. \tag{10}$$

It is obvious that $\boldsymbol{\lambda}^k$ and $\boldsymbol{\mu}^k$ are feasible for (9) and (10), respectively. From the complementary slackness between (7) and (8), we see that

$$\boldsymbol{\mu}^k (\mathbf{b}_N - \mathbf{N}\mathbf{Q}_k \boldsymbol{\lambda}^k) = 0,$$

which reduces to the duality $\boldsymbol{\mu}^k \mathbf{N}\mathbf{Q}_k \boldsymbol{\lambda}^k = \boldsymbol{\mu}^k \mathbf{b}_N$ between (9) and (10). Similarly, the converse can also be shown. □

Let us investigate the relationship between (P_k) and (P'_k) in more detail. Let

$$\Lambda_k^+ = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \sum_{j \in J_k} \mathbf{q}_j^k \lambda_j, \ \boldsymbol{\lambda} \geq \mathbf{0}\}, \quad J_k = \{j \mid \lambda_j^k > 0\}.$$

Apparently, Λ_k^+ is the minimal face of Λ_k containing the optimal solution $\boldsymbol{\omega}^k$ of (P_k) and (P'_k).

Lemma 3.2. *It holds that*

$$\boldsymbol{\mu}^k \mathbf{N}\mathbf{x} \geq \mathbf{e}\mathbf{Q}_k^{-1}\mathbf{x}, \quad \forall \mathbf{x} \in \Lambda_k. \tag{11}$$

In particular,

$$\boldsymbol{\mu}^k \mathbf{N}\mathbf{x} = \mathbf{e}\mathbf{Q}_k^{-1}\mathbf{x} \ \textit{ if } \ \mathbf{x} \in \Lambda_k^+. \tag{12}$$

Proof. If $\mathbf{x} \in \Lambda_k$, there exists a $\boldsymbol{\lambda}' \geq \mathbf{0}$ such that $\mathbf{x} = \mathbf{Q}_k \boldsymbol{\lambda}'$, and we have

$$\mathbf{e}\mathbf{Q}_k^{-1}\mathbf{x} = \mathbf{e}\boldsymbol{\lambda}' \leq \boldsymbol{\mu}^k \mathbf{N}\mathbf{Q}_k \boldsymbol{\lambda}' = \boldsymbol{\mu}^k \mathbf{N}\mathbf{x}, \tag{13}$$

by noting $\boldsymbol{\mu}^k \mathbf{N}\mathbf{Q}_k \geq \mathbf{e}$. Furthermore, if $\mathbf{x} \in \Lambda_k^+$, then $\lambda'_j = 0$ for each $j \notin J_k$, and besides we have

$$\boldsymbol{\mu}^k \mathbf{N}\mathbf{q}_j^k = 1, \quad \forall j \in J_k,$$

by the complementary slackness between (7) and (8). The equation in (12) follows from these together with (13). □

Let

$$H_k = \{\mathbf{x} \in \mathbb{R}^n \mid \boldsymbol{\mu}^k \mathbf{N}\mathbf{x} \leq 1\}.$$

Immediately from Lemma 3.2, we see the relationship between this halfspace H_k and the γ-valid cut G_k:

$$G_k \cap \Lambda_k^+ = H_k \cap \Lambda_k^+ \subset H_k \cap \Lambda_k \subset G_k \cap \Lambda_k \subset C(\gamma). \tag{14}$$

PROOF OF PSEUDO-NONSINGULARITY

Let us give here the formal definition of pseudo-nonsingularity.

Definition 3.1. The nested sequence of cones $\{\Lambda_k \mid k = 1, 2, \dots\}$ is said to be *pseudo-nonsingular* if there exists a constant L such that

$$\|\boldsymbol{\mu}^k \mathbf{N}\| \leq L, \quad k = 1, 2, \dots. \tag{15}$$

The conical algorithm is also said to be *pseudo-nonsingular* if every nested sequence of cones that it generates is pseudo-nonsingular.

Note that this definition requires the norm in (15) to be bounded from above for every k, unlike the original nonsingularity (6).

To show the pseudo-nonsingularity of the sequence (2), we only have to show the existence of a constant L satisfying (15). For this purpose, however, we need a further lemma.

Lemma 3.3. *The optimal value ζ^k of (P_k) and (P'_k) is nonincreasing in k, i.e.,*

$$\zeta^k \geq \zeta^{k+1} > 1, \quad k = 1, 2, \dots.$$

Proof. Suppose that Λ_{k+1} is spanned by $\mathbf{Q}_{k+1} = [\mathbf{q}_1^k, \dots, \mathbf{q}_{j-1}^k, \mathrm{ext}(\mathbf{u}^k), \mathbf{q}_{j+1}^k, \dots, \mathbf{q}_n^k]$. Then we have

$$\boldsymbol{\omega}^{k+1} = \sum_{i \neq j} \lambda_i^{k+1} \mathbf{q}_i^k + \lambda_j^{k+1} \mathrm{ext}(\mathbf{u}^k).$$

Let $\mathbf{u}'$ denote the intersection of the ray in direction $\mathbf{u}^k$ with the boundary of H_k. Since $\mathbf{u}' \in H_k \cap \Lambda_k$, it follows from (14) that $\mathbf{u}'$ is a point in $C(\gamma)$. On the other hand, the γ-extension $\mathrm{ext}(\mathbf{u}^k)$ lies on the boundary of $C(\gamma)$. Hence, $\mathrm{ext}(\mathbf{u}^k) = \theta \mathbf{u}'$ holds for some $\theta \geq 1$, and we have

$$\boldsymbol{\mu}^k \mathbf{N} \boldsymbol{\omega}^{k+1} = \sum_{i \neq j} \lambda_i^{k+1} \boldsymbol{\mu}^k \mathbf{N} \mathbf{q}_i^k + \lambda_j^{k+1} \theta \boldsymbol{\mu}^k \mathbf{N} \mathbf{u}' \geq \mathbf{e} \boldsymbol{\lambda}^{k+1} = \zeta^{k+1},$$

by noting that $\boldsymbol{\mu}^k \mathbf{N} \mathbf{q}_i^k \geq 1$, $\boldsymbol{\mu}^k \mathbf{N} \mathbf{u}' = 1$ and $\boldsymbol{\lambda}^{k+1} \geq \mathbf{0}$. Furthermore, we have

$$\zeta^k = \boldsymbol{\mu}^k \mathbf{N} \boldsymbol{\omega}^k \geq \boldsymbol{\mu}^k \mathbf{N} \boldsymbol{\omega}^{k+1} \geq \zeta^{k+1},$$

because $\boldsymbol{\omega}^k$ is an optimal solution of (P_k') while $\boldsymbol{\omega}^{k+1}$ is just a feasible solution. $\square$

Theorem 3.4. *The nested sequence of cones (2) is pseudo-nonsingular.*

Proof. Assume that $\|\boldsymbol{\mu}^k \mathbf{N}\| > 0$, since otherwise there is nothing to prove, and define a half-space

$$H = \{\mathbf{x} \in \mathbb{R}^n \mid \boldsymbol{\mu}^k \mathbf{N} \mathbf{x} \leq \zeta^k\}.$$

If $\mathbf{x} \in M$, then

$$\boldsymbol{\mu}^k \mathbf{N} \mathbf{x} \leq \boldsymbol{\mu}^k \mathbf{b}_N = \zeta^k,$$

which implies that M is a subset of H. Recall that the vertex $\mathbf{v}$ of Λ is a nondegenerate vertex of $D = M \cap \Lambda$ and located at $\mathbf{0}$. Therefore, $\mathbf{0}$ is an interior point of M and the distance $\delta(\mathbf{0}, \partial M)$ from $\mathbf{0}$ to the boundary of M does not vanish. Since $\delta(\mathbf{0}, \partial M)$ is a lower bound on the distance $\delta(\mathbf{0}, \partial H) = \zeta^k / \|\boldsymbol{\mu}^k \mathbf{N}\|$ from $\mathbf{0}$ to the boundary of H, we have

$$\|\boldsymbol{\mu}^k \mathbf{N}\| \leq \zeta^k / \delta(\mathbf{0}, \partial M).$$

Once $\mathbf{v}$ is selected, $\delta(\mathbf{0}, \partial M)$ stays constant. The sequence $\{\zeta^k \mid k = 1, 2, \ldots\}$ is nonincreasing, as seen in Lemma 3.3, and hence

$$\zeta^k \leq \zeta^1 = \max\{\mathbf{e} \mathbf{Q}^{-1} \mathbf{x} \mid \mathbf{x} \in D\},$$

where the right-hand-side is bounded from above because D is assumed to be bounded. $\square$

Since (2) is an arbitrary nested sequence of cones, the conical algorithm is also pseudo-nonsingular. In the next section, we will use the pseudo-nonsingularity and prove the convergence of the conical algorithm under a certain class of subdivision rules, including ω-subdivision.

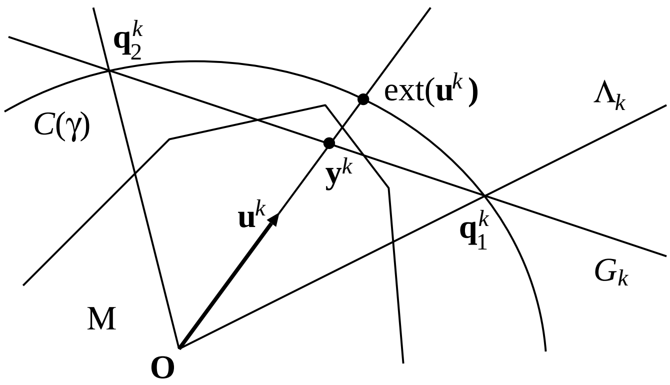

Figure 1: Geometric position between $\mathrm{ext}(\mathbf{u}^k)$ and $\mathbf{y}^k$.

4 Convergence of the algorithm with ω-subdivision

Let $\mathbf{y}^k$ denote the intersection of the ray from $\mathbf{0}$ in direction $\mathbf{u}^k$ with the boundary of G_k (see Figure 1). The main result we prove in this section is the following, which guarantees that G_k approximates $C(\gamma)$ asymptotically on Λ_k.

Theorem 4.1. *Let $\{\Lambda_k \mid k = 1,2,\dots\}$ be a nested sequence of cones such that Λ_{k+1} is obtained by subdividing Λ_k along $\mathbf{u}^k \in \Lambda_k^+$. Then,*

$$\liminf_{k\to\infty} \|\mathrm{ext}(\mathbf{u}^k) - \mathbf{y}^k\| = 0. \tag{16}$$

Before proving Theorem 4.1, we need to introduce two lemmas. Let

$$\overline{H}_k = \{\mathbf{x} \in \mathbb{R}^n \mid \boldsymbol{\mu}^k \mathbf{N}\mathbf{x} \geq 1\}.$$

Lemma 4.2. *The sequence $\{\mathbf{y}^k \mid k = 1,2,\dots\}$ is bounded and satisfies*

$$\mathbf{y}^k \in \overline{H}_\ell, \quad \ell = 1,\dots,k. \tag{17}$$

Proof. Since $\mathbf{y}^k \in G_k \cap \Lambda_k \subset C(\gamma)$ and $C(\gamma)$ is bounded, the sequence is also bounded. Let $\mathbf{q}'_j$ denote the intersection of the ray in direction $\mathbf{q}_j^k$ with ∂H_ℓ. Then $\mathbf{q}'_j \in H_\ell \cap \Lambda_\ell \subset G_\ell \cap \Lambda_\ell \subset C(\gamma)$ while $\mathbf{q}_j^k \in \partial C(\gamma)$. Hence, there is a $\theta_j \geq 1$ such that $\mathbf{q}_j^k = \theta_j \mathbf{q}'_j$. If $\mathbf{y}^k = \sum_{j=1}^n \lambda'_j \mathbf{q}_j^k$ for some $\boldsymbol{\lambda}' \geq \mathbf{0}$ with $\mathbf{e}\boldsymbol{\lambda}' = 1$, we have

$$\boldsymbol{\mu}^\ell \mathbf{N}\mathbf{y}^k = \boldsymbol{\mu}^\ell \mathbf{N} \sum_{j=1}^n \lambda'_j \mathbf{q}_j^k = \sum_{j=1}^n \theta_j \lambda'_j \boldsymbol{\mu}^\ell \mathbf{N}\mathbf{q}'_j = \sum_{j=1}^n \theta_j \lambda'_j \geq 1,$$

which implies that $\mathbf{y}^k$ is a point in $\overline{H}_\ell$. □

Lemma 4.3. *Let $\{\Lambda_k \mid k = 1,2,\dots\}$ be a nested sequence of cones such that Λ_{k+1} is obtained* *by subdividing* Λ_k *along* $\mathbf{u}^k \in \Lambda_k^+$*. Then, there exists a subsequence* $\{k_r \mid r = 1, 2, \dots\}$ *such that*

$$\mathbf{u}^{k_{2p-1}} \in \Lambda_{k_{2p-1}}^+ \cap \Lambda_{k_{2p}}^+, \quad p = 1, 2, \dots. \tag{18}$$

Proof. Let us first show that, for infinitely many k,

$$\exists \ell \geq k+1, \quad \mathbf{u}^k \in \Lambda_\ell^+. \tag{19}$$

If not, there exists a sufficiently large k such that $\mathbf{u}^k \notin \Lambda_\ell^+$ for each $\ell \geq k+1$. In other words, $\mathbf{u}^k$ remains as an edge direction in the descendants of Λ_k. Similarly, $\mathbf{u}^{k+1} \notin \Lambda_{\ell+1}^+$ for each $\ell \geq k+1$ while $\mathbf{u}^{k+1} \in \Lambda_{k+1}^+$, and so forth. Eventually, Λ_{k+n} is spanned by n vectors $\mathbf{u}^k, \mathbf{u}^{k+1}, \dots,$ and $\mathbf{u}^{k+n-1}$, at least one of which must be a point in Λ_{k+n}^+. This is a contradiction and therefore (19) holds for infinitely many k. Let us select such a k, denote it by k_1 and denote the corresponding ℓ by k_2. Then, select another $k \geq k_2 + 1$ satisfying (19), denote it by k_3 and denote the corresponding ℓ by k_4. Continuing this process yields an infinite sequence $\{k_r \mid r = 1, 2, \dots\}$, which satisfies (18). □

Now we are ready to prove theorem 4.1.

Proof of Theorem 4.1. Let $\{k_r \mid r = 1, 2, \dots\}$ be a subsequence satisfying (18), and abbreviate k_r to r. We see from Lemma 4.2 that $\{\mathbf{y}^r \mid r = 1, 2, \dots\}$ is a bounded sequence and satisfies

$$\mathbf{y}^r \in \bigcap_{\ell=1}^{r} \overline{H}_\ell.$$

According to the *bounded convergence principle* (see e.g., Lemma III.2 in [5]), we have

$$\delta(\mathbf{y}^r, \overline{H}_{r+1}) \to 0 \ \text{ as } \ r \to \infty. \tag{20}$$

Without loss of generality, suppose r is an odd number. Both $\mathbf{y}^r$ and $\mathrm{ext}(\mathbf{u}^r)$ are then points in $\Lambda_r^+ \cap \Lambda_{r+1}^+$ because $\mathbf{u}^r$ belongs to those two cones. From Lemma 3.2, it holds that

$$\boldsymbol{\mu}^{r+1} \mathbf{N} \,\mathrm{ext}(\mathbf{u}^r) = \mathbf{e} \mathbf{Q}_{r+1}^{-1} \mathrm{ext}(\mathbf{u}^r). \tag{21}$$

Since $\mathrm{ext}(\mathbf{u}^r)$ is a column, say the jth column, of $\mathbf{Q}_{r+1}$, we have $\mathrm{ext}(\mathbf{u}^r) = \mathbf{Q}_{r+1} \mathbf{e}_j$ for the jth unit vector $\mathbf{e}_j$. This, together with (21), implies that $\boldsymbol{\mu}^{r+1} \mathbf{N} \,\mathrm{ext}(\mathbf{u}^r) = 1$, or equivalently that

$$\mathrm{ext}(\mathbf{u}^r) \in \partial H_{r+1} = \partial \overline{H}_{r+1}.$$

Note that $\mathrm{ext}(\mathbf{u}^r)$ is also the γ-extension of $\mathbf{y}^r$, and $\mathrm{ext}(\mathbf{u}^r) \in \partial C(\gamma)$ while $\mathbf{y}^r \in C(\gamma)$. Therefore, $\mathbf{y}^r$ cannot be an interior point of $\overline{H}_{r+1}$, and we see that (20) can be rewritten as follows:

$$\delta(\mathbf{y}^r, \partial H_{r+1}) \to 0 \ \text{ as } \ r \to \infty. \tag{22}$$

Let $\mathbf{x}'$ and $\mathbf{y}'$ denote the points in ∂H_{r+1} closest to $\mathbf{0}$ and $\mathbf{y}^r$, respectively. Since the triangle connecting $\mathbf{x}'$, $\mathbf{0}$ and $\mathrm{ext}(\mathbf{u}^r)$ is similar to that connecting $\mathbf{y}'$, $\mathbf{y}^r$ and $\mathrm{ext}(\mathbf{u}^r)$, we have

$$\|\mathrm{ext}(\mathbf{u}^r) - \mathbf{y}^r\| = \|\mathrm{ext}(\mathbf{u}^r)\| \delta(\mathbf{y}^r, \partial H_{r+1}) \,/\, \delta(\mathbf{0}, \partial H_{r+1}). \tag{23}$$

By the pseudo-nonsingularity of $\{\Lambda_k \mid k = 1, 2, \dots\}$, we have $1 \,/\, \delta(\mathbf{0}, \partial H_{r+1}) = \|\boldsymbol{\mu}^{r+1}\mathbf{N}\| < L$ for some constant L. Also, $\|\mathrm{ext}(\mathbf{u}^r)\|$ is bounded because $\mathrm{ext}(\mathbf{u}^r)$ is a point in the bounded set $C(\gamma)$. Thus, by noting (22), we conclude that $\|\mathrm{ext}(\mathbf{u}^r) - \mathbf{y}^r\| \to 0$ as $r \to \infty$. □

The convergence result with the usual ω-subdivision can be thought of as a corollary of Theorem 4.1.

Corollary 4.4. *Let $\{\Lambda_k \mid k = 1, 2, \dots\}$ be a nested sequence of cones such that Λ_{k+1} is obtained by subdividing Λ_k along $\mathbf{u}^k = \boldsymbol{\omega}^k$. Then $\{\mathbf{y}^k \mid k = 1, 2, \dots\}$ has an accumulation point $\mathbf{y}^0 \in D$ such that $f(\mathbf{y}^0) = \gamma$.*

Proof. By Theorem 4.1, there exists a subsequence $\{k_r \mid r = 1, 2, \dots\}$ such that $\|\mathrm{ext}(\boldsymbol{\omega}^{k_r}) - \mathbf{y}^{k_r}\| \to 0$ as $r \to \infty$. We have $|f[\mathrm{ext}(\boldsymbol{\omega}^{k_r})] - f(\mathbf{y}^{k_r})| \to 0$ by the continuity of f, and hence $f(\mathbf{y}^{k_r}) \to \gamma$ by noting $f[\mathrm{ext}(\boldsymbol{\omega}^{k_r})] = \gamma$. It follows from the assumption (5) that $\mathbf{y}^{k_r}$ belongs to the segment $[\mathbf{0}, \boldsymbol{\omega}^{k_r}]$ in $M \cap \Lambda_{k_r}$, which is a subset of the compact set D. Therefore, the sequence $\{\mathbf{y}^k \mid k = 1, 2, \dots\}$ has an accumulation point $\mathbf{y}^0 \in D$ satisfying $f(\mathbf{y}^0) = \gamma$. □

Unfortunately, Theorem 4.1 does not, by itself, ensure the convergence of the algorithm to a solution of the d.c. feasibility problem (DC). It merely implies the existence of a subsequence $\{k_r \mid r = 1, 2, \dots\}$ such that the γ-extension of $\mathbf{u}^{k_r}$ approaches the γ-valid cut G_{k_r} asymptotically. To achieve the convergence to a solution of (DC), we need to further restrict the selection of the subdivision direction $\mathbf{u}^k$ for each k. One way is obviously ω-subdivision. In the next section, we will develop an alternative to ω-subdivision, named *ω-bisection*, which bisects Λ_k by splitting a two-dimensional face of Λ_k^+ into two pieces.

5 Conical algorithm based on ω-bisection

To develop the ω-bisection, we assume in the rest of the paper that f is strictly concave, i.e., if $\mathbf{x}, \mathbf{y} \in S$ and $\mathbf{x} \neq \mathbf{y}$, then we have

$$f[(1-\lambda)\mathbf{x} + \lambda\mathbf{y}] > (1-\lambda)f(\mathbf{x}) + \lambda f(\mathbf{y}), \quad \forall \lambda \in (0, 1). \tag{24}$$

Under this assumption, we can observe the following behavior in the sequence (2). As before, $\mathbf{y}^k$ denotes the intersection of the ray in direction $\mathbf{u}^k$ with the boundary of G_k.

Lemma 5.1. *Let $\{\Lambda_k \mid k = 1, 2, \dots\}$ be a nested sequence of cones such that Λ_{k+1} is obtained by subdividing Λ_k along $\mathbf{u}^k$ lying on a two-dimensional face of Λ_k^+. Then $\{\mathbf{q}_j^k \mid k = 1, 2, \dots\}$*

has an accumulation point $\mathbf{q}_j^0 \in \partial C(\gamma)$ *for each* $j = 1, \ldots, n$. *Among the* $\mathbf{q}_j^0$*'s, there exists an accumulation point* $\mathbf{y}^0$ *of* $\{\mathbf{y}^k \mid k = 1, 2, \ldots\}$.

Proof. By Theorem 4.1, we can take a subsequence $\{k_r \mid r = 1, 2, \ldots\}$ such that

$$\|\mathrm{ext}(\mathbf{u}^{k_r}) - \mathbf{y}^{k_r}\| \to 0, \quad \text{as } r \to \infty. \tag{25}$$

Since the number of possible combinations is finite, $\mathbf{u}^{k_r}$ lies on a face of Λ_{k_r} spanned by two vectors with the same pair of subscripts, say $\mathbf{q}_s^{k_r}$ and $\mathbf{q}_t^{k_r}$, for infinitely many r. By extracting a further subsequence if necessary, we have

$$\mathbf{y}^{k_r} = (1 - \lambda^{k_r})\mathbf{q}_s^{k_r} + \lambda^{k_r}\mathbf{q}_t^{k_r}, \quad \lambda^{k_r} \in (0, 1/2], \quad r = 1, 2, \ldots,$$

and besides

$$\mathbf{q}_s^{k_r} \to \mathbf{q}_s^0 \in \partial C(\gamma), \quad \mathbf{q}_t^{k_r} \to \mathbf{q}_t^0 \in \partial C(\gamma), \quad \mathbf{y}^{k_r} \to \mathbf{y}^0 \in [\mathbf{q}_s^0, \mathbf{q}_t^0], \quad \text{as } r \to \infty,$$

because $\mathbf{q}_s^k$ and $\mathbf{q}_t^k$ are generated in the compact set $\partial C(\gamma)$.

To show that $\mathbf{y}^0$ coincides with either $\mathbf{q}_s^0$ or $\mathbf{q}_t^0$, suppose on the contrary that there exists a number $\sigma > 0$ such that

$$\|\mathbf{q}_s^{k_r} - \mathbf{y}^{k_r}\| > \sigma, \quad \|\mathbf{q}_t^{k_r} - \mathbf{y}^{k_r}\| > \sigma, \quad r = 1, 2, \ldots. \tag{26}$$

For this σ, consider

$$\left| \begin{array}{ll} \text{minimize} & f[(\mathbf{x} + \mathbf{y})/2] - [f(\mathbf{x}) + f(\mathbf{y})]/2 \\ \text{subject to} & \mathbf{x}, \mathbf{y} \in \Lambda_1, \quad \|\mathbf{x} - \mathbf{y}\| \geq 2\sigma. \end{array} \right. \tag{27}$$

Since the constraints are satisfied by $(\mathbf{x}, \mathbf{y}) = (\mathbf{q}_s^{k_r}, \mathbf{q}_t^{k_r})$ and the objective function is bounded below, (27) has an optimal solution, whose value, denoted by ν, is positive by the assumption (24). Let

$$\mathbf{y}' = (1 - 2\lambda^{k_r})\mathbf{q}_s^{k_r} + 2\lambda^{k_r}\mathbf{q}_t^{k_r}.$$

By the concavity of f, we have

$$[f(\mathbf{q}_s^{k_r}) + f(\mathbf{y}')]/2 \geq (1 - \lambda^{k_r}) f(\mathbf{q}_s^{k_r}) + \lambda^{k_r} f(\mathbf{q}_t^{k_r}) = \gamma. \tag{28}$$

Note that the pair $(\mathbf{q}_s^{k_r}, \mathbf{y}')$ is a feasible solution of (27) and the objective function value $f[(\mathbf{q}_s^{k_r} + \mathbf{y}')/2] - [f(\mathbf{q}_s^{k_r}) + f(\mathbf{y}')]/2$ never falls below ν. This, together with (28), implies

$$\begin{aligned} f(\mathbf{y}^{k_r}) - f[\mathrm{ext}(\mathbf{u}^{k_r})] &= f[(\mathbf{q}_s^{k_r} + \mathbf{y}')/2] - \gamma \\ &\geq f[(\mathbf{q}_s^{k_r} + \mathbf{y}')/2] - [f(\mathbf{q}_s^{k_r}) + f(\mathbf{y}')]/2 \geq \nu > 0, \end{aligned}$$

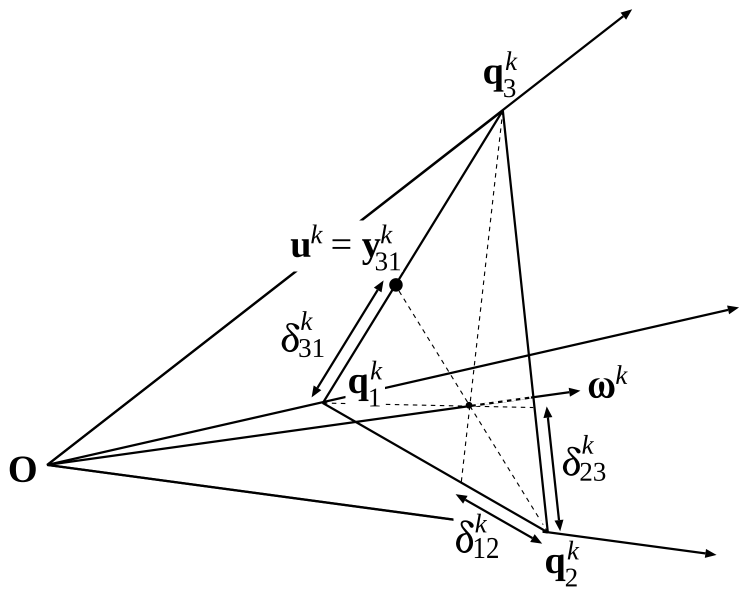

Figure 2: Process of ω-bisection when $J_k = \{1,2,3\}$.

which contradicts (25) because f is continuous. Hence, the assumption (26) is false, and we have $\mathbf{y}^0 \in \{\mathbf{q}_s^0, \mathbf{q}_t^0\}$. $\square$

ω-BISECTION RULE

On the basis of the above observation, let us now attempt to develop a systematic procedure for ω-bisection (see also Figure 2).

For each pair $\{i,j\} \subset J_k$, let

$$\mathbf{y}_{ij}^k = (\lambda_i^k \mathbf{q}_i^k + \lambda_j^k \mathbf{q}_j^k)/(\lambda_i^k + \lambda_j^k). \tag{29}$$

This point $\mathbf{y}_{ij}^k$ is the intersection of the segment $[\mathbf{q}_i^k, \mathbf{q}_j^k]$ with the hyperplane spanned by $\boldsymbol{\omega}^k$ and $n-1$ vectors $\mathbf{q}_1^k, \ldots, \mathbf{q}_{i-1}^k, \mathbf{q}_{i+1}^k, \ldots, \mathbf{q}_{j-1}^k, \mathbf{q}_{j+1}^k, \ldots, \mathbf{q}_n^k$. The segment $[\mathbf{q}_i^k, \mathbf{q}_j^k]$ is split into two pieces $[\mathbf{q}_i^k, \mathbf{y}_{ij}^k]$ and $[\mathbf{y}_{ij}^k, \mathbf{q}_j^k]$, the shorter of which has a length of

$$\delta_{ij}^k = \|\mathbf{q}_i^k - \mathbf{q}_j^k\| \min\{\lambda_i^k, \lambda_j^k\}/(\lambda_i^k + \lambda_j^k). \tag{30}$$

Among the $\mathbf{y}_{ij}^k$'s, we select as $\mathbf{u}^k$ the one with the largest δ_{ij}^k, i.e., $\mathbf{y}_{st}^k$ with

$$\{s,t\} \in \arg\max\{\delta_{ij}^k \mid \{i,j\} \subset J_k\}, \tag{31}$$

and subdivide the cone Λ_k along the direction $\mathbf{u}^k = \mathbf{y}_{st}^k$ into two subcones:

$$\Lambda_k^j = \mathrm{con}(\mathbf{Q}_k^j), \quad j = s,t, \tag{32}$$

where

$$\mathbf{Q}_k^j = [\mathbf{q}_1^k, \ldots, \mathbf{q}_{j-1}^k, \mathrm{ext}(\mathbf{y}_{st}^k), \mathbf{q}_{j+1}^k, \ldots, \mathbf{q}_n^k]. \tag{33}$$

Either Λ_k^s or Λ_k^t is adopted as Λ_{k+1} in the nested sequence of cones $\{\Lambda_k \mid k = 1, 2, \ldots\}$.

Suppose that $\{\Lambda_k \mid k = 1, 2, \ldots\}$ is generated according to the rule given by (29)–(33). Then we have the following results.

Lemma 5.2. *There exists an index set $J_0 \subset \{1, \ldots, n\}$ such that $J_k = J_0$ for infinitely many k. Moreover,*

(i) for each pair $\{i, j\} \subset J_0$, the sequence $\{\mathbf{y}_{ij}^k \mid k = 1, 2, \ldots\}$ has an accumulation point $\mathbf{y}_{ij}^0 \in \{\mathbf{q}_i^0, \mathbf{q}_j^0\}$, and

(ii) for each $j \in J_0$, the sequence $\{\lambda_j^k \mid k = 1, 2, \ldots\}$ has an accumulation point $\lambda_j^0 \geq 0$ such that $\sum_{j \in J_0} \lambda_j^0 \geq 1$.

In particular, if $\lambda_i^0, \lambda_j^0 > 0$ for $\{i, j\} \subset J_0$, then it holds that $\mathbf{y}_{ij}^0 = \mathbf{q}_i^0 = \mathbf{q}_j^0$.

Proof. Consider the same sequence $\{k_r \mid r = 1, 2, \ldots\}$ used in the proof of Lemma 5.1. The indices s and t determined in (31) belong to J_{k_r}, which is a subset of a finite set $\{1, \ldots, n\}$. This implies that there is a set J_0 such that $\{s, t\} \subset J_0 = J_{k_r}$ for infinitely many r.

(i) After extracting a subsequence, we can assume for each $\{i, j\} \subset J_0$ that $\mathbf{q}_i^{k_r} \to \mathbf{q}_i^0$, $\mathbf{q}_j^{k_r} \to \mathbf{q}_j^0$ and $\mathbf{y}_{ij}^{k_r} \to \mathbf{y}_{ij}^0 \in [\mathbf{q}_i^0, \mathbf{q}_j^0]$, as $r \to \infty$, for some $\mathbf{q}_i^0, \mathbf{q}_j^0 \in C(\gamma)$. Also, we see from Lemma 5.1 that $\mathbf{y}_{ij}^0 \in \{\mathbf{q}_i^0, \mathbf{q}_j^0\}$ because $\delta_{ij}^{k_r} \leq \delta_{st}^{k_r}$ but $\delta_{st}^{k_r} \to 0$ as $r \to \infty$.

(ii) From Lemma 3.3, we see that $\boldsymbol{\lambda}^{k_r}$ belongs to a compact set defined by $\mathbf{e}\boldsymbol{\lambda} \leq \zeta^1$ and $\boldsymbol{\lambda} \geq \mathbf{0}$. Taking a further subsequence if necessary, we have $\lambda_j^{k_r} \to \lambda_j^0$, as $r \to \infty$, for some $\lambda_j^0 \geq 0$. We also have $\sum_{j \in J_0} \lambda_j^0 \geq 1$ since $\sum_{j \in J_0} \lambda_j^{k_r} = \zeta^{k_r} > 1$.

If $\lambda_i^0, \lambda_j^0 > 0$ and $\mathbf{q}_i^0 \neq \mathbf{q}_j^0$ for some $\{i, j\} \subset J_0$, then

$$\mathbf{y}_{ij}^{k_r} = (\lambda_i^{k_r}\mathbf{q}_i^{k_r} + \lambda_j^{k_r}\mathbf{q}_j^{k_r})/(\lambda_i^{k_r} + \lambda_j^{k_r}) \to \mathbf{y}_{ij}^0 = (\lambda_i^0\mathbf{q}_i^0 + \lambda_j^0\mathbf{q}_j^0)/(\lambda_i^0 + \lambda_j^0) \notin \{\mathbf{q}_i^0, \mathbf{q}_j^0\},$$

which contradicts (i). □

Lemma 5.3. *Let $\boldsymbol{\eta}^k$ denote the intersection of the ray from $\mathbf{0}$ in direction $\boldsymbol{\omega}^k$ with ∂G_k. Then $\{\boldsymbol{\eta}^k \mid k = 1, 2, \ldots\}$ has an accumulation point $\boldsymbol{\eta}^0 \in D$ such that $f(\boldsymbol{\eta}^0) = \gamma$.*

Proof. Consider again the subsequence used in the previous lemmas. Renumbering the subscripts, we may assume that $J_0 = \{1, \ldots, p\}$ for some $p \leq n$. For $i = 3, \ldots, p$, we will show below that if

$$\mathbf{w}_{i-1}^{k_r} = \sum_{j=1}^{i-1} \lambda_j^{k_r}\mathbf{q}_j^{k_r} \to \left(\sum_{j=1}^{i-1} \lambda_j^0\right)\mathbf{q}_h^0, \quad \text{as } r \to \infty, \tag{34}$$

for some $h \in \{1,\ldots,i-1\}$, then

$$\mathbf{w}_i^{k_r} = \sum_{j=1}^{i} \lambda_j^{k_r} \mathbf{q}_j^{k_r} \to \left(\sum_{j=1}^{i} \lambda_j^0\right) \mathbf{q}_{h'}^0, \quad \text{as } r \to \infty, \tag{35}$$

for $h' \in \{h, i\}$. Note that (35) follows immediately from Lemma 5.2 if $i = 2$.

Since $\mathbf{y}_{hi}^{k_r}$ converges to either $\mathbf{q}_h^0$ or $\mathbf{q}_i^0$, there are three cases to consider:

$$\|\mathbf{q}_i^{k_r} - \mathbf{q}_h^{k_r}\| \to 0 \tag{36}$$

$$\|\mathbf{q}_i^{k_r} - \mathbf{q}_h^{k_r}\| \to \sigma, \quad \lambda_h^0 > 0, \quad \lambda_i^0 = 0 \tag{37}$$

$$\|\mathbf{q}_i^{k_r} - \mathbf{q}_h^{k_r}\| \to \sigma, \quad \lambda_h^0 = 0, \quad \lambda_i^0 > 0, \tag{38}$$

where $\sigma > 0$ is some constant. In cases (36) and (37), as $r \to \infty$, we have

$$\begin{aligned}\|\mathbf{w}_i^{k_r} - \left(\sum_{j=1}^{i} \lambda_j^{k_r}\right) \mathbf{q}_h^{k_r}\| &= \|\sum_{j=1}^{i-1} \lambda_j^{k_r} \mathbf{q}_j^{k_r} - \left(\sum_{j=1}^{i-1} \lambda_j^{k_r}\right) \mathbf{q}_h^{k_r} + \lambda_i^{k_r}(\mathbf{q}_i^{k_r} - \mathbf{q}_h^{k_r})\| \\ &\le \|\mathbf{w}_{i-1}^{k_r} - \left(\sum_{j=1}^{i-1} \lambda_j^{k_r}\right) \mathbf{q}_h^{k_r}\| + \lambda_i^{k_r}\|\mathbf{q}_i^{k_r} - \mathbf{q}_h^{k_r}\| \to 0,\end{aligned}$$

and $\mathbf{w}_i^{k_r} \to (\sum_{j=1}^{i} \lambda_j^0)\mathbf{q}_h^0$. In case (38), if $\lambda_j^0 = 0$ for $j = 1,\ldots,i-1$, then obviously

$$\|\mathbf{w}_i^{k_r} - \left(\sum_{j=1}^{i} \lambda_j^{k_r}\right) \mathbf{q}_i^{k_r}\| = \|\sum_{j=1}^{i-1} \lambda_j^{k_r}(\mathbf{q}_j^{k_r} - \mathbf{q}_i^{k_r})\| \le \sum_{j=1}^{i-1} \lambda_j^{k_r}\|\mathbf{q}_j^{k_r} - \mathbf{q}_i^{k_r}\| \to 0, \tag{39}$$

and $\mathbf{w}_i^{k_r} \to (\sum_{j=1}^{i} \lambda_j^0)\mathbf{q}_i^0$. Even if $\lambda_j^0 > 0$ for some j, we have (39) because $\mathbf{q}_i^0 = \mathbf{q}_j^0$ by Lemma 5.2. Therefore, (35) holds for all three cases. By induction, there exists an index $h \in J_0$ such that

$$\boldsymbol{\omega}^{k_r} = \mathbf{w}_p^{k_r} = \sum_{j \in J_0} \lambda_j^{k_r} \mathbf{q}_j^{k_r} \to \left(\sum_{j \in J_0} \lambda_j^0\right) \mathbf{q}_h^0, \quad \text{as } r \to \infty. \tag{40}$$

By definition, we have $\boldsymbol{\eta}^{k_r} = \boldsymbol{\omega}^{k_r} / \sum_{j \in J_0} \lambda_j^{k_r} \to \mathbf{q}_h^0$, and hence $f(\boldsymbol{\eta}^{k_r}) \to f(\mathbf{q}_h^0) = \gamma$, as $r \to \infty$. The rest of the proof is the same as that for Corollary 4.4. □

Algorithm description

Before closing this section, let us summarize the conical algorithm for solving (DC) with ω-bisection.

```
algorithm conic_ω_bisect(D, f, α, ε)
  γ ← α − ε;
```

determine a cone Λ with vertex $\mathbf{v} = \mathbf{0}$ and a polyhedron M such that $D = M \cap \Lambda$, $f(\mathbf{v}) > \gamma$, and $\mathbf{v}$ is an interior point of M;
let Λ be spanned by n vectors $\mathbf{q}_1, \ldots, \mathbf{q}_n$ with $f(\mathbf{q}_j) = \gamma$, and $\mathbf{Q} \leftarrow [\mathbf{q}_1, \ldots, \mathbf{q}_n]$;
$\mathscr{P} \leftarrow \emptyset$; $\mathscr{T} \leftarrow \{\mathbf{Q}\}$; $stop \leftarrow false$; $\mathbf{z} \leftarrow \mathbf{0}$; $k \leftarrow 1$;
while $stop = false$ do
  for each $\mathbf{Q} \in \mathscr{T}$ do
    compute an optimal solution $\boldsymbol{\lambda}(\mathbf{Q})$ of the linear program $\max\{\mathbf{e}\boldsymbol{\lambda} \mid \mathbf{Q}\lambda \in M, \boldsymbol{\lambda} \geq \mathbf{0}\}$;
    $\zeta(\mathbf{Q}) \leftarrow \mathbf{e}\boldsymbol{\lambda}(\mathbf{Q})$;
    if $\zeta(\mathbf{Q}) > 1$ then
      $\mathscr{P} \leftarrow \mathscr{P} \cup \{\mathbf{Q}\}$;
    end if
    if $f(\mathbf{Q}\boldsymbol{\lambda}(\mathbf{Q})) < \alpha$ then
      $\mathbf{z} \leftarrow \mathbf{Q}\boldsymbol{\lambda}(\mathbf{Q})$;
    end if
  end for
  if $\mathscr{P} = \emptyset$ or $f(\mathbf{z}) < \alpha$ then
    $stop \leftarrow true$;
  else
    choose $\mathbf{Q}$ with the largest $\zeta(\mathbf{Q})$ from $\mathscr{P}$, and let $\mathbf{Q}_k \leftarrow \mathbf{Q}$;
    $\boldsymbol{\lambda}^k \leftarrow \boldsymbol{\lambda}(\mathbf{Q}_k)$; $\boldsymbol{\omega}^k \leftarrow \mathbf{Q}_k\boldsymbol{\lambda}^k$;
    generate the children $\mathbf{Q}_k^s$ and $\mathbf{Q}_k^t$ of $\mathbf{Q}_k$ from $\boldsymbol{\lambda}^k$ according to (29)–(33);
    $\mathscr{P} \leftarrow \mathscr{P} \setminus \{\mathbf{Q}_k\}$; $\mathscr{T} \leftarrow \{\mathbf{Q}_k^s, \mathbf{Q}_k^t\}$; $k \leftarrow k+1$;
  end if
end while
if $\mathscr{P} \neq \emptyset$ then
  print "$\mathbf{z}$ is a point in $D \setminus C(\alpha)$.";
else
  print "D is a subset of $C(\gamma)$.";
end if
end.

Theorem 5.4. *Suppose $\varepsilon = 0$. If the algorithm* conic_ω_bisect *terminates, then it either generates a point $\mathbf{z} \in D \setminus C(\alpha)$ or proves that $D \subset C(\alpha)$. If not, the sequence $\{\boldsymbol{\omega}^k \mid k = 1, 2, \ldots\}$ has an accumulation point $\boldsymbol{\omega}^0 \in D$ such that $f(\boldsymbol{\omega}^0) = \alpha$.*

Proof. Since the claim is obvious if conic_ω_bisect terminates, it is sufficient to consider the case where it does not. In that case, the algorithm generates an infinite sequence of matrices $\{\mathbf{Q}_k \mid k = 1, 2, \ldots\}$ such that $\mathbf{Q}_{k+1}$ is a child of $\mathbf{Q}_k$ if we renumber the indices. For each k, let $\Lambda_k = \mathrm{con}(\mathbf{Q}_k)$. Then $\{\Lambda_k \mid k = 1, 2, \ldots\}$ is a nested sequence of cones generated by ω-bisection. Let $\boldsymbol{\eta}^k$ denote the intersection of the ray in direction $\boldsymbol{\omega}^k$ with ∂G_k. Since $\zeta^k =$

$\zeta(\mathbf{Q}_k) > 1$, we see that $\boldsymbol{\eta}^k$ lies on the segment $[\mathbf{0}, \boldsymbol{\omega}^k]$. By the concavity of f, we have

$$f(\boldsymbol{\eta}^k) \geq \min\{f(\mathbf{0}), f(\boldsymbol{\omega}^k)\} \geq \gamma = \alpha - \varepsilon. \tag{41}$$

However, by Lemma 5.3, there exists a subsequence $\{k_r \mid r = 1, 2, \dots\}$ such that $f(\boldsymbol{\eta}^{k_r}) \to \gamma$ as $r \to \infty$. Since $\varepsilon = 0$ and $f(\mathbf{0}) > \alpha$, this can be compatible with (41) only if $f(\boldsymbol{\omega}^{k_r}) \to \gamma = \alpha$. Moreover, $\{\boldsymbol{\omega}^k \mid k = 1, 2, \dots\}$ is generated in the compact set D, and hence it has an accumulation point $\boldsymbol{\omega}^0 \in D$ such that $f(\boldsymbol{\omega}^0) = \alpha$. □

Corollary 5.5. *If $\varepsilon > 0$, the algorithm* conic_ω_bisect *terminates in a finite number of iterations, and either generates a point $\mathbf{z} \in D \setminus C(\alpha)$ or proves that $D \subset C(\alpha - \varepsilon)$.*

Proof. As seen in the proof of Theorem 5.4, if conic_ω_bisect does not terminate, it generates a sequence $\{k_r \mid r = 1, 2, \dots\}$ such that $f(\boldsymbol{\omega}^{k_r}) \to \gamma = \alpha - \varepsilon$ as $r \to \infty$. However, if $f(\boldsymbol{\omega}^k) < \alpha$ holds, the algorithm terminates according to its stopping criterion. Therefore, conic_ω_bisect terminates in a finite number of iterations if $\varepsilon > 0$. □

6 Numerical results

In this section, we report a numerical comparison between the conical algorithm with ω-bisection and that with the usual ω-subdivision. The test problem was a concave quadratic minimization problem of the form

$$\left| \begin{array}{ll} \text{minimize} & \frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{C}\mathbf{x} + \mathbf{d}\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \\ \text{subject to} & \mathbf{A}\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \leq \mathbf{b}, \quad \mathbf{x}, \mathbf{y} \geq \mathbf{0}, \end{array} \right. \tag{42}$$

where $\mathbf{A} \in \mathbb{R}^{40\times 100}$, $\mathbf{b} \in \mathbb{R}^{40}$, $\mathbf{C} \in \mathbb{R}^{r\times r}$, $\mathbf{d} \in \mathbb{R}^{100}$, and $r \leq 100$. In order to make the feasible region bounded, $\mathbf{b}$ was set to $[1, \dots, 1, 100]^\mathsf{T}$ and each entry in the last row of $\mathbf{A}$ was set to one. Other entries of $\mathbf{A}$ were uniformly random numbers generated in the interval $[-0.5, 1.0]$. As for the objective function, $\mathbf{C}$ was a symmetric and negative-definite matrix with entries in $[-1.0, 0.0]$, and all components of $\mathbf{d}$ were random numbers in $[-300.0, 0.0]$. To solve this optimization problem, we adopted the two-phase scheme described in Section 2 and wrote two program codes, ω-bisect and ω-subdiv, in GNU Octave [3], a numerical computing environment similar to MATLAB. The only deference between ω-bisect and ω-subdiv is that the former uses ω-bisection as the cone subdivision rule for the conical algorithm in Phase 2 while the latter uses the usual ω-subdivision. These codes consist largely of solving the linear program (7) associated with (P^k), to which we applied the simplex method solver in GLPK (GNU linear programming kit) [2]. Also, to prevent the convergence from depending on the magnitude of the optimal value, we modified the tolerance in the d.c. feasibility problem and

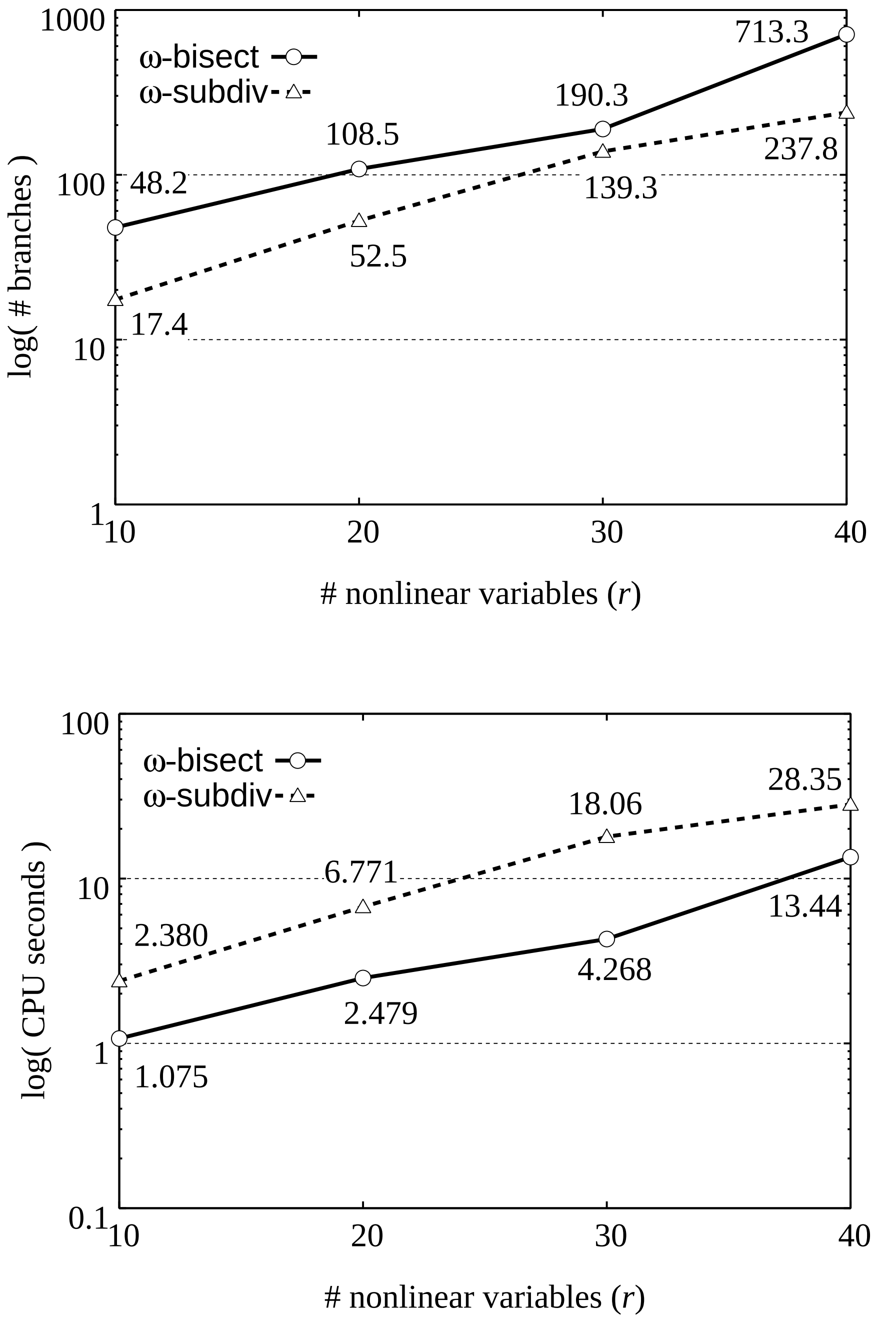

Figure 3: Comparison between ω-bisect and ω-subdiv when $\varepsilon = 10^{-5}$.

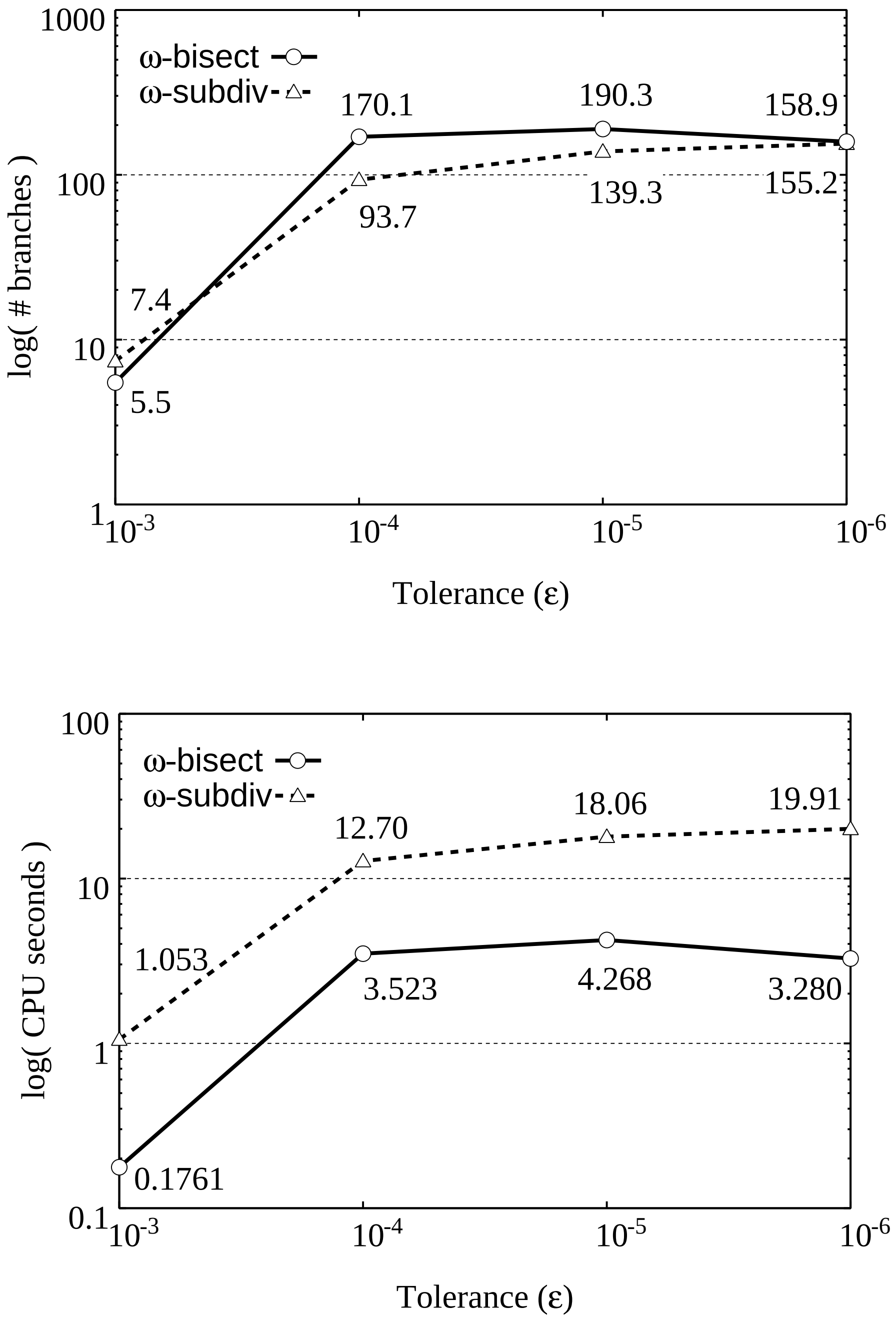

Figure 4: Comparison between ω-bisect and ω-subdiv when $r = 30$.

solved the following in Phase 2, instead of (DC) given in Section 2:

$$(\mathrm{DC}') : \text{ find a point } \mathbf{x} \in D \setminus C(\alpha) \text{ if there is one, or else prove that } D \subset C((1+\varepsilon)\alpha),$$

where we should note that $\alpha \leq 0$ for the test problem (42). We solved ten instances for each pair (r, ε) on a single core of an Intel Core i7 processor (3.33GHz). Note that, when $r < 100$, the objective function of (42) does not satisfy the condition (24); and hence ω-bisect is not guaranteed to terminate in this case. In our experiment, however, both ω-bisect and ω-subdiv solved all the instances in finite time.

Figure 3 compares the behavior of ω-bisect and ω-subdiv when the tolerance ε was fixed at 10^{-5} and the number r of nonlinear variables was changed from 10 to 40 in increments of 10. The upper plot shows the average number of total branching operations executed in each of ω-bisect and ω-subdiv, corresponding to the solid and dashed lines, respectively. The lower plot shows the average CPU time taken by each code in seconds. We see immediately from these plots that both codes actually take an exponential amount of computation time in r. What is still more remarkable is that, while ω-bisect requires more branching operations, it is much faster than ω-subdiv. This is totally due to the difference in the number of auxiliary problems generated after a single branching operation. In the worst case, ω-subdiv needs to solve $n(= 100)$ new auxiliary problems. In contrast to this, those to be solved in ω-bisect are only two. The same tendency can be observed in Figure 4, which shows the result when r was fixed at 30 and ε was changed from 10^{-3} to 10^{-6}. In this figure, the computation time for both codes seems to reach a ceiling around $\varepsilon = 10^{-5}$. It might be because unpromising cones are discarded earlier when ε is small enough. However, since this is simply a result obtained from ten instances for each ε, we cannot draw a firm conclusion.

In order to make a definite conclusion, we need to solve a wider variety of problems with different sizes. Nevertheless, it should be clear even from these limited results that ω-bisection is at least comparable with ω-subdivision in practical efficiency and promising as a new cone subdivision rule for the conical algorithm. In a subsequent paper, we will report on more detailed numerical results, together with a generalization of ω-bisection and ω-subdivision.

Acknowledgment

The authors would like to thank the anonymous referee for his/her valuable comments, which significantly improved the quality of this article.

References

[1] Bali, S., *Minimization of a Concave Function on a Bounded Convex Polyhedron*, Ph.D. Dissertation, University of California, Los Angeles (CA, 1973).

[2] GNU Linear Programming Kit, http://www.gnu.org/software/glpk/.

[3] GNU Octave, http://www.gnu.org/software/octave/.

[4] Hamami, M., and S.E. Jacobsen, "Nondegenerate conical processes for concave minimization of convex polytopes", *Mathematics of Operations Research* **13** (1988), 479–487.

[5] Horst, R. and H. Tuy, *Global Optimization: Deterministic Approaches*, 2nd ed., Springer-Verlag (Berlin, 1993).

[6] Jaumard, B., and C. Meyer, "A simplified convergence proof for the cone partitioning algorithm", *Journal of Global Optimization* **13** (1998), 407–416.

[7] Jaumard, B., and C. Meyer, "On the convergence of cone splitting algorithms with ω-subdivisions", *Journal of Optimization Theory and Applications* **110** (2001), 119–144.

[8] Locatelli, M., "Finiteness of conical algorithm with ω-subdivisions", *Mathematical Programming* **85** (1999), 593-616.

[9] Rockafellar, R.T., *Convex Analysis*, Princeton University Press (NJ, 1970).

[10] Sahni, S., "Computationally related problems", *SIAM Journal on computing* **3** (1974), 262–279.

[11] Thoai, N.V., and H. Tuy, "Convergent algorithms for minimizing a concave function", *Mathematics of Operations Research* **5** (1980), 556–566.

[12] Tuy, H., "Concave programming under linear constraints", *Soviet Mathematics* **5** (1964), 1437–1440.

[13] Tuy, H., "Normal conical algorithm for concave minimization over polytopes", *Mathematical Programming* **51** (1991), 229–245.

[14] Tuy, H., "Effect of the subdivision strategy on convergence and efficiency of some global optimization algorithms", *Journal of Global Optimization* **1** (1991), 23–36.

[15] Tuy, H., *Convex Analysis and Global Optimization*, Kluwer Academic Publishers (Dordrecht, 1998).

[16] Zwart, P.B., "Nonlinear programming: counterexamples to two global optimization algorithms", *Operations Research* **21** (1973), 1260–1266.

[17] Zwart, P.B., "Global maximization of a convex function with linear inequality constraints", *Operations Research* **22** (1974), 602–609.