

General Models in Min-Max Continuous Location: Theory and Solution Techniques^{1,2}

J. B. G. FRENK,³ J. GROMICHO,⁴ AND S. ZHANG⁵

Communicated by S. Schaible

Abstract. In this paper, a class of min-max continuous location problems is discussed. After giving a complete characterization of the stationary points, we propose a simple central and deep-cut ellipsoid algorithm to solve these problems for the quasiconvex case. Moreover, an elementary convergence proof of this algorithm and some computational results are presented.

Key Words. Location theory, min-max programming, ellipsoid algorithm, quasiconvexity.

1. Introduction

Let $\mathcal{D} := \{d_1, \dots, d_n\} \subseteq \mathbb{R}^s$, $s \geq 2$, denote a set of n different demand points, and let $x \in \mathbb{R}^s$ be the location of a facility to be chosen. If \mathcal{G}_i , $i = 1, \dots, n$, are compact convex sets with $0 \in \text{int}(\mathcal{G}_i)$ and $\text{int}(\mathcal{G}_i)$ denoting the interior of \mathcal{G}_i , we define the distance between x and d_i by $\gamma_{\mathcal{G}_i}(x - d_i)$, with $\gamma_{\mathcal{G}}$ the gauge or Minkowsky functional (Refs. 1 and 2) of the set \mathcal{G}_i , i.e.,

$$\gamma_{\mathcal{G}_i}(x) := \inf\{t > 0 : x \in t\mathcal{G}_i\}. \quad (1)$$

If \mathcal{G}_i is symmetric around 0, the corresponding gauge is called a norm. The motive to consider continuous location problems with arbitrary gauges is

¹The work of the second author was supported by JNICT (Portugal), under Contract BD/631/90-RM, during his stay at Erasmus University in Rotterdam.

²The authors would like to thank the anonymous referees for simplifying the proofs in the first part of Section 2 and for their constructive remarks improving the presentation.

³Associate Professor of Operations Research, Econometric Institute, Erasmus University, Rotterdam, Netherlands.

⁴Assistant Lecturer, DEIO, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal.

⁵Assistant Professor of Operations Research, Econometric Institute, Erasmus University, Rotterdam, Netherlands.

twofold. First of all, although most papers discuss only continuous location models with rectangular, Euclidean, or Tchebycheff norms, it is known that an l_p -norm (with p not equal to 1, 2, or ∞) or a general polyhedral or block norm (Refs. 3 and 4) sometimes fits better the empirical data (Refs. 4–7). This implies that it might be sensible to consider arbitrary norms within location models. Secondly, it is sometimes unrealistic to assume, as imposed by a norm, that the distance is symmetric (Ref. 8). Hence, to model this asymmetric behavior of distances (i.e., the distance between x and y is not the same as the distance between y and x), we need to consider arbitrary gauges. Although a gauge seems to be an intractable distance function, it turns out that the analysis of location models with gauges is relatively simple due to the convexity properties of gauges. In particular, a gauge $\gamma_{\mathcal{G}}$ of a compact convex set \mathcal{G} with $0 \in \text{int}(\mathcal{G})$ is a positively homogeneous, nonnegative finite-valued convex function (Refs. 2 and 9). It should be clear from the above observations that the selection of a distance function is central in all continuous location problems, and this topic is intensively discussed in an excellent recent survey paper by Plastria (Ref. 10). To conclude our discussion of how to measure distance, we observe that the Euclidean, rectangular, or Tchebycheff norms were used mostly to obtain easy tractable nonlinear/linear programming problems (Ref. 11). However in this paper, we present a simple deep-cut version of the ellipsoid algorithm which can handle efficiently a large subclass of the general single-facility location model with gauges. To introduce the model, let the function $\gamma: \mathbb{R}^s \rightarrow \mathbb{R}_+^n$ be given by

$$\gamma(x) := (\gamma_{\mathcal{G}_1}(x - d_1), \dots, \gamma_{\mathcal{G}_n}(x - d_n)), \quad (2)$$

and suppose that $f_i: \mathbb{R}^n \rightarrow \mathbb{R}_+$, $1 \leq i \leq n$, denote a set of nondecreasing functions on \mathbb{R}_+^n associated with the demand points. We also assume that the functions f_i are differentiable on an open set \mathcal{S} with $\mathbb{R}_+^n \subseteq \mathcal{S}$. This implies (Ref. 12) that the partial derivatives $\partial f_i / \partial z_j(z)$, $1 \leq j \leq n$, of f_i evaluated in $z \in \mathbb{R}_+^n$ exist and are nonnegative for every $1 \leq i \leq n$. The single-facility location model that we discuss in this paper is now given by

$$(P) \quad \inf_{x \in \mathbb{R}^s} \max_{1 \leq i \leq n} \varphi_i(x),$$

with $\varphi_i: \mathbb{R}^s \rightarrow \mathbb{R}$ defined by $\varphi_i(x) := f_i(\gamma(x))$. The above optimization problem is called a min-max problem and arises, for example, in the location of an emergency unit (Ref. 11). Clearly, if all the functions f_i are equal, problem (P) reduces to $\inf\{\varphi_1(x): x \in \mathbb{R}^s\}$. Most continuous unconstrained single-facility location problems are special cases of the optimization problem (P) (Ref. 11), and this problem has in general a nondifferentiable objective function (Ref. 13). Since the functions f_i are continuous on the open set \mathcal{S} and nondecreasing on \mathbb{R}_+^n , it follows that the function $z \mapsto \max_{1 \leq i \leq n} f_i(z)$

satisfies the same properties; hence, our model is a special case of a more general model discussed in Ref. 14. This implies (Refs. 10 and 14) that there exists an easily computable hyperrectangle \mathcal{R} which contains at least one optimal solution of (P). If the distance is measured by an arbitrary norm and $s = 2$, this hyperrectangle \mathcal{R} can be replaced by the smaller set $\text{conv}(\mathcal{D})$, with $\text{conv}(\mathcal{D})$ the convex hull of \mathcal{D} (Ref. 15), or by any rectangle containing $\text{conv}(\mathcal{D})$. From these remarks, it follows now that there exists an easily computable hyperrectangle \mathcal{R} such that solving (P) is equivalent to solving

$$(Q) \quad \min_{x \in \mathcal{R}} \max_{1 \leq i \leq n} \varphi_i(x).$$

From the equivalence between (Q) and (P), it is also easy to find some $a_0 \in \mathbb{R}^s$ and $r > 0$ such that an optimal solution x^* belongs to the set $a_0 + r\mathcal{B}$, with \mathcal{B} the open Euclidean unit ball. A way to construct such a sphere is given by the following procedure. Consider the center a_0 of \mathcal{R} , and take r bigger than the Euclidean distance from a_0 to one of the rectangle vertices. This sphere yields a starting ellipsoid for the ellipsoid algorithm to be proposed in Section 3. Let us now introduce

$$\varphi := \max_{1 \leq i \leq n} \varphi_i.$$

By our assumptions, it follows that the directional derivative

$$\varphi'(x; y) := \lim_{t \downarrow 0} [\varphi(x + ty) - \varphi(x)]/t$$

exists for every $x, y \in \mathbb{R}^s$, and so we can define the set

$$\Gamma_\varphi := \{x \in \mathbb{R}^s : \varphi'(x; y) \geq 0, \text{ for every } y \in \mathbb{R}^s\}.$$

Since every optimal solution of (P) belongs to Γ_φ , this means that Γ_φ is not empty. If we want to solve the nondifferentiable optimization problem (P) by some iterative procedure, we have to find out whether an iteration point is optimal or not. In general, this is a difficult computational task. It turns out to be easier to check whether a point belongs to Γ_φ ; since all optimal solutions belong to Γ_φ , this might be a good stopping strategy. So, we stop the iterative procedure if the present iteration point belongs to Γ_φ . In general, this does not mean that the point is optimal for (P). However, if the functions f_i are convex, it is well known that the set Γ_φ coincides with the set of optimal points. In Lemma 2.1, a weaker condition is presented for the equivalence between local and global optima. To derive an efficient algorithm for checking whether the present iteration point belongs to Γ_φ , we need first to give an easy characterization of this set. Among other topics, this will be discussed in the next section.

2. Properties

To give a geometrical description of elements of Γ_φ , we need to calculate the directional derivative $\varphi'(x; y)$ with $\varphi: \mathbb{R}^s \rightarrow \mathbb{R}$ given by

$$\varphi(x) := \max_{1 \leq i \leq n} \varphi_i(x) \quad \text{and} \quad \varphi_i(x) := f_i(\gamma(x)).$$

Since f_i is differentiable on the open set \mathcal{S} with $\mathbb{R}_+^n \subseteq \mathcal{S}$, it follows by the chain rule of differentiation (Ref. 12) that

$$\varphi'_i(x; y) = \sum_{j=1}^n (\partial f_i / \partial z_j)(\gamma(x)) \gamma'_{\mathcal{G}_j}(x - d_j; y), \quad (3)$$

with $\gamma'_{\mathcal{G}_j}(x - d_j; y)$ the directional derivative of the gauge $\gamma_{\mathcal{G}_j}$ at $x - d_j$ in the direction y . Since $(\partial f_i / \partial z_j)(\gamma(x)) \geq 0$, the finite-valued function $y \mapsto \varphi'_i(x; y)$ is positively homogeneous and convex, and so it is quasidifferentiable in the sense of Pshenichnyi (Ref. 16). Moreover, by relation (3), it follows that $y \mapsto \varphi'_i(x; y)$ is the support function of the compact convex set

$$\partial \varphi_i(x) = \sum_{j=1}^n (\partial f_i / \partial z_j)(\gamma(x)) \partial \gamma_{\mathcal{G}_j}(x - d_j),$$

with $\partial \gamma_{\mathcal{G}_j}(x - d_j)$ the subgradient set of $\gamma_{\mathcal{G}_j}$ at $x - d_j$. Hence by Theorem 3.4 of Ref. 16, the function $y \mapsto \varphi'(x; y)$ is quasidifferentiable in the sense of Pshenichnyi and by the same result it is the support function of the set

$$\partial \varphi(x) := \text{conv} \left(\bigcup_{i \in I(x)} \partial \varphi_i(x) \right),$$

with

$$I(x) := \{1 \leq i \leq n: \varphi(x) = \varphi_i(x)\}.$$

By these observations and $x_0 \in \Gamma_\varphi$ if and only if $0 \in \partial \varphi(x_0)$, the next result follows immediately.

Theorem 2.1. If $\mathcal{C}(x) \subseteq \mathbb{R}^s$ denotes the set

$$\text{conv} \left(\bigcup_{i \in I(x)} \left(\sum_{j=1}^n (\partial f_i / \partial z_j)(\gamma(x)) \partial \gamma_{\mathcal{G}_j}(x - d_j) \right) \right),$$

then $x_0 \in \Gamma_\varphi$ if and only if $0 \in \mathcal{C}(x_0)$.

To check the above condition for some x_0 , we need to solve a membership problem. In the next examples, this membership problem will be specified for gauges often used in location theory.

Example 2.1. Taking for each $1 \leq j \leq n$ the set \mathcal{G}_j given by

$$\mathcal{G}_j := \{x \in \mathbb{R}^s : N_{p_j}(x) \leq 1\}, \quad 1 < p_j < \infty,$$

with $N_p(x) := \|x\|_p$ the l_p -norm at x , it follows by the above result for x_0 not belonging to \mathcal{D} that $x_0 \in \Gamma_\varphi$ if and only if

$$0 \in \text{conv}(\{c_i : i \in I(x_0)\}),$$

with

$$c_i := \sum_{j=1}^n (\partial f_i / \partial z_j)(\gamma(x_0)) \nabla N_{p_j}(x_0 - d_j)$$

and $\nabla N_{p_j}(x)$ the gradient of N_{p_j} at $x \neq 0$. It is well known for $x \neq 0$ that

$$\nabla N_p(x) = \|x\|_p^{1-p} \begin{bmatrix} \delta(x_1) |x_1|^{p-1} \\ \vdots \\ \delta(x_s) |x_s|^{p-1} \end{bmatrix},$$

with δ the sign function defined as

$$\delta(x) := \begin{cases} -1, & \text{if } x < 0, \\ 0, & \text{if } x = 0, \\ 1, & \text{if } x > 0. \end{cases}$$

In this case, one has to decide upon the feasibility of the linear system

$$\sum_{i \in I(x_0)} \lambda_i c_i = 0, \quad \sum_{i \in I(x_0)} \lambda_i = 1, \quad \lambda_i \geq 0, \text{ for every } 1 \leq i \leq n.$$

A possible way to do this is by applying phase one of the simplex algorithm. Moreover, for x_0 belonging to \mathcal{D} , say $x_0 = d_k$, we have $x_0 \in \Gamma_\varphi$ if and only if

$$0 \in \text{conv}\left(\bigcup_{i \in I(x_0)} \mathcal{C}_i\right),$$

with

$$\mathcal{C}_i := \left\{x \in \mathbb{R}^s : \left\| x - \sum_{j \neq k} (\partial f_i / \partial z_j)(\gamma(x_0)) \nabla N_{p_j}(x_0 - d_j) \right\|_{q_k} \leq (\partial f_i / \partial z_k)(\gamma(x_0)) \right\}$$

and

$$(1/q_k) + (1/p_k) = 1.$$

If $I(x_0)$ consists of only one element, the above membership problem is trivial to check. To conclude this example, we observe that efficient algorithms for $s=2$ to solve both membership problems are presented in Ref. 17. In that paper, the complexity of these decision problems for $s \geq 2$ is also discussed.

Example 2.2. Consider for each $1 \leq j \leq n$ a polytope $\mathcal{G}_j \subseteq \mathbb{R}^s$, i.e.,

$$\mathcal{G}_j = \text{conv}(\{v_{1j}, \dots, v_{r_j}\}),$$

with v_{1j}, \dots, v_{r_j} its vertices and $0 \in \text{int}(\mathcal{G}_j)$. If $v_{1j}^\circ, \dots, v_{r_j}^\circ$ are the vertices of its polar \mathcal{G}_j° , it follows again by Theorem 2.1 that $x_0 \in \Gamma_\phi$ if and only if

$$0 \in \text{conv}\left(\bigcup_{i \in I(x_0)} \mathcal{C}_i\right),$$

with

$$\mathcal{C}_i := \sum_{j=1}^n (\partial f_i / \partial z_j)(\gamma(x_0)) \text{conv}(\{v_l^\circ : l \in I_j(x_0)\}),$$

and

$$I_j(x_0) := \{1_j \leq i \leq r_j^\circ : \gamma_{\mathcal{G}_j}(x_0 - d_j) = \langle x_0 - d_j, v_i^\circ \rangle\}.$$

By Lemma I.5 of Ref. 9,

$$\mathcal{C}_i := \text{conv}\left(\sum_{j=1}^n (\partial f_i / \partial z_j)(\gamma(x_0)) \{v_l^\circ : l \in I_j(x_0)\}\right),$$

and so \mathcal{C}_i is again a polytope for every $i \in I(x_0)$.

This implies that $\text{conv}(\bigcup_{i \in I(x_0)} \mathcal{C}_i)$ is also a polytope; hence, in principle, one can solve the above decision problem by linear programming.

As noticed in the introduction, the set Γ_ϕ may contain nonoptimal points. In order to ensure that Γ_ϕ coincides with the set of optimal points, we introduce the following set of functions.

Definition 2.1. See Ref. 18. Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a convex set. A function $h: \mathbb{R}^n \rightarrow \mathbb{R}$ is called quasiconvex on \mathcal{C} if the sets $\mathcal{L}_h(\alpha) \cap \mathcal{C}$ are convex for every $\alpha \in \mathbb{R}$, with $\mathcal{L}_h(\alpha) := \{x \in \mathbb{R}^n : h(x) < \alpha\}$ the so-called strict lower level set of h of level α .

Observe (Ref. 18) that $h: \mathbb{R}^n \rightarrow \mathbb{R}$ is quasiconvex on \mathcal{C} if and only if

$$h(\lambda x_1 + (1 - \lambda)x_2) \leq \max\{h(x_1), h(x_2)\},$$

for every $x_1, x_2 \in \mathcal{C}$ and $0 \leq \lambda \leq 1$.

It is now possible to prove the following result.

Lemma 2.1. If for every $1 \leq i \leq n$, the function $f_i: \mathbb{R}^n \rightarrow \mathbb{R}$ is nondecreasing on \mathbb{R}_+^n and differentiable and quasiconvex on an open convex set \mathcal{S} with $\mathbb{R}_+^n \subseteq \mathcal{S}$, and if the gradient $\nabla f_i(z)$ contains at least one positive component for every $z \in \mathcal{S}$, then the set Γ_φ of stationary points coincides with the set of optimal solutions of (P).

Proof. Clearly, every optimal solution of (P) belongs to Γ_φ . To prove the reverse, we observe the following. For every $1 \leq i \leq n$, it follows by the nonnegativity of $\nabla f_i(z)$ for every $z \in \mathcal{S}$ and the convexity of γ_{g_j} , $1 \leq j \leq n$, that

$$\begin{aligned} \varphi'_i(x; y) &= \sum_{j=1}^n (\partial f_i / \partial z_j)(\gamma(x)) \gamma'_{g_j}(x - d_j; y) \\ &\leq \sum_{j=1}^n (\partial f_i / \partial z_j)(\gamma(x)) (\gamma_{g_j}(x + y - d_j) - \gamma_{g_j}(x - d_j)) \\ &= f'_i(\gamma(x); \gamma(x + y) - \gamma(x)), \end{aligned}$$

for every $y \in \mathbb{R}^s$. Hence, for $x_0 \in \Gamma_\varphi$ and $y \in \mathbb{R}^s$ fixed, this implies that

$$\begin{aligned} 0 \leq \varphi'(x_0; y) &= \max\{\varphi'_i(x_0; y) : i \in I(x_0)\} \\ &\leq \max\{f'_i(\gamma(x_0); \gamma(x_0 + y) - \gamma(x_0)) : i \in I(x_0)\} \\ &= f'_{i(y)}(\gamma(x_0); \gamma(x_0 + y) - \gamma(x_0)), \end{aligned} \tag{4}$$

for some $i(y) \in I(x_0)$. Since we assume that $\nabla f_{i(y)}(z) \geq 0$ contains at least one positive component for every $z \in \mathcal{S}$, it follows from Theorem 2 of Ref. 19 that $f_{i(y)}$ is pseudoconvex on \mathcal{S} , and so by (4) we obtain

$$\varphi(x_0 + y) \geq f_{i(y)}(\gamma(x_0 + y)) \geq f_{i(y)}(\gamma(x_0)) = \varphi(x_0).$$

Since the above inequality holds for every $y \in \mathbb{R}^s$, it must follow that x_0 is an optimal solution of (P). □

If x_0 does not belong to Γ_φ , it follows that

$$\mathcal{F} := \{y \in \mathbb{R}^s : \varphi'(x_0; y) < 0\}$$

is nonempty, and this implies $\mathcal{F} \subseteq \mathcal{N}$, with

$$\mathcal{N} := \bigcup_{\lambda > 0} \lambda (\mathcal{L}_\varphi(\varphi(x_0)) - x_0) \quad \text{and} \quad x_0 \in \text{cl}(\mathcal{L}_\varphi(\varphi(x_0))).$$

To apply the central-cut version of the ellipsoid algorithm proposed in the next section, one needs at each iteration an element of the normal cone

$\mathcal{N}_{\mathcal{L}}(x_0)$ of $\mathcal{L}_{\varphi}(\varphi(x_0))$ at the center $x_0 \notin \Gamma_{\varphi}$ of the current ellipsoid, i.e.,

$$\mathcal{N}_{\mathcal{L}}(x_0) := \{x^* \in \mathbb{R}^s : \langle x - x_0, x^* \rangle \leq 0, \text{ for every } x \in \text{cl}(\mathcal{L}_{\varphi}(\varphi(x_0)))\}.$$

Clearly, $\mathcal{N}_{\mathcal{L}}(x_0)$ equals \mathcal{K}° , with \mathcal{K}° the polar cone of \mathcal{K} . To guarantee that $\mathcal{N}_{\mathcal{L}}(x_0)$ is nonempty, we assume in the remainder that the functions $f_i: \mathbb{R}^n \rightarrow \mathbb{R}$, $1 \leq i \leq n$, are nondecreasing on \mathbb{R}_+^n and differentiable and quasi-convex on an open convex set \mathcal{S} with $\mathbb{R}_+^n \subseteq \mathcal{S}$. These conditions imply that the objective function φ of (P) is continuous and quasiconvex on \mathbb{R}_+^n ; so, by a well-known separation result (Ref. 1), it follows that $\mathcal{N}_{\mathcal{L}}(x_0)$ is nonempty for x_0 not belonging to Γ_{φ} . For φ quasiconvex and continuous, it is shown in Proposition 1 of Ref. 20 using a more complicated proof than for the convex case (Ref. 21) that \mathcal{K} equals \mathcal{T} for \mathcal{T} nonempty. Since $y \mapsto \varphi'(x_0; y)$ is the support function of $\partial\varphi(x_0)$, one can now apply the same proof as for convex functions (see Theorem VI.1.3.5 of Ref. 21), and so the following result holds. Recall that the set $\mathcal{C}(x)$ is defined in Theorem 2.1.

Theorem 2.2. If x_0 does not belong to Γ_{φ} , then $\mathcal{N}_{\mathcal{L}}(x_0) = \bigcup_{\lambda \geq 0} \lambda \mathcal{C}(x_0)$.

As shown in the next section, it is sufficient to know an element of the normal cone $\mathcal{N}_{\mathcal{L}}(x_0)$ for the execution of the so-called central-cut version of the ellipsoid algorithm. However, if one wants to use convexlike deep cuts (Refs. 22 and 23), which have a positive influence on the convergence speed as shown in Theorem 3.1, we need to consider the following subclass of differentiable functions.

Definition 2.2. The function $h: \mathbb{R}^n \rightarrow \mathbb{R}$ differentiable on an open set \mathcal{S} , with $\mathbb{R}_+^n \subseteq \mathcal{S}$, is said to have a locally bounded gradient if, for every $z \in \mathcal{S}$, there exists some $\epsilon > 0$ and some constant M such that $\|\nabla h(y)\|_2 \leq M$ for every y belonging to $z + \epsilon \mathcal{B}$, with \mathcal{B} the open Euclidean unit ball.

A class of functions related to the above class is given in the next definition.

Definition 2.3. See Ref. 24. A function $h: \mathbb{R}^n \rightarrow \mathbb{R}$ is called Lipschitz continuous on a set $\mathcal{X} \subseteq \mathbb{R}^n$ with Lipschitz constant L if $|h(z) - h(y)| \leq L\|z - y\|_2$ for every z, y belonging to \mathcal{X} .

If the gradient of the function f_i is locally bounded, it must follow for every compact set $\mathcal{X} \subseteq \mathcal{S}$ that there exists some $M > 0$ such that

$$\|\nabla f_i(y)\|_2 \leq M, \quad \text{for every } y \in \mathcal{X}.$$

Additionally, if \mathcal{K} is also convex, this implies by Theorem 9.19 of Ref. 12 that f_i is Lipschitz continuous on \mathcal{K} . Since by Theorem 41.D of Ref. 25, every finite-valued convex function on \mathbb{R}^s is Lipschitz continuous on any compact subset, it is now easy to verify that the function $\varphi_i(x) = f_i(\gamma(x))$ is also Lipschitz continuous on any compact subset of \mathbb{R}^s . Finally, noting that Lipschitz continuity is preserved under the finite max operator, we obtain that the objective function

$$\varphi(x) = \max_{1 \leq i \leq n} \varphi_i(x)$$

of (P) is again Lipschitz continuous on any compact subset. Knowing the Lipschitz constants of the functions f_i , $1 \leq i \leq n$, and γ_{φ_j} , $1 \leq j \leq n$, one can also easily determine the Lipschitz constant of the function φ . By the above observations, we obtain that the function φ is quasiconvex on \mathbb{R}^s and Lipschitz continuous on any compact subset. A class of functions closely related to quasiconvexity and Lipschitz continuity is now given by the next definition.

Definition 2.4. See Ref. 26. A function $h: \mathbb{R}^n \rightarrow \mathbb{R}$ is called lower subdifferentiable on $\mathcal{K} \subseteq \mathbb{R}^n$ if, for every $z \in \mathcal{K}$, there exists some $z^* \in \mathbb{R}^n$ satisfying

$$h(y) \geq h(z) + \langle y - z, z^* \rangle,$$

for every y belonging to $\mathcal{L}_h(h(z)) \cap \mathcal{K}$. The set of all so-called lower subgradients z^* of the function h at z is denoted by $\partial^-h(z)$. Finally, the function $h: \mathbb{R}^n \rightarrow \mathbb{R}$ is called boundedly lower subdifferentiable on \mathcal{K} if h is lower subdifferentiable on \mathcal{K} and there exists a constant $N > 0$ such that, for every $z \in \mathcal{K}$, there exists some $z^* \in \partial^-h(z)$ with $\|z^*\|_2 \leq N$.

The following result is needed for the proof of the next theorem. Although it is known (Refs. 27 and 28), we list its proof for completeness.

Lemma 2.2. For every x , $x_0 \in r\mathcal{B}$, with $r > 0$ and \mathcal{B} the open Euclidean unit ball in \mathbb{R}^s , and for any hyperplane \mathcal{H} going through x_0 , the orthogonal projection $x_{\mathcal{H}}$ of x on \mathcal{H} belongs to $\sqrt{2}r\mathcal{B}$.

Proof. If \mathcal{H} is a hyperplane in \mathbb{R}^s going through x_0 , it follows that there exists some $x^* \in \mathbb{R}^s$ with $\|x^*\|_2 = 1$ such that

$$\mathcal{H} = \{x \in \mathbb{R}^s : \langle x - x_0, x^* \rangle = 0\}.$$

The orthogonal projection $x_{\mathcal{H}}$ of x on \mathcal{H} satisfies

$$x_{\mathcal{H}} = x + \langle x_0 - x, x^* \rangle x^*,$$

and this implies

$$\begin{aligned} \|x_{\mathcal{H}} - \langle x_0, x^* \rangle x^*\|_2^2 &= \|x - \langle x, x^* \rangle x^*\|_2^2 \\ &= \|x\|_2^2 - (\langle x, x^* \rangle)^2 \leq \|x\|_2^2. \end{aligned} \tag{5}$$

Moreover, using $x_{\mathcal{H}} \in \mathcal{H}$, we obtain by the Cauchy–Schwartz inequality that

$$\begin{aligned} \|x_{\mathcal{H}} - \langle x_0, x^* \rangle x^*\|_2^2 &= \|x_{\mathcal{H}}\|_2^2 - 2\langle x_0, x^* \rangle \langle x_{\mathcal{H}}, x^* \rangle + (\langle x_0, x^* \rangle)^2 \\ &= \|x_{\mathcal{H}}\|_2^2 - (\langle x_0, x^* \rangle)^2 \geq \|x_{\mathcal{H}}\|_2^2 - \|x_0\|_2^2. \end{aligned} \tag{6}$$

Combining Inequalities (5) and (6) and using $x, x_0 \in r\mathcal{B}$, the desired result follows. □

To prove the next result, we need to introduce the following notation. For any function $h: \mathbb{R}^s \rightarrow \mathbb{R}$, the function $h_{\mathcal{X}}$ denotes the restriction of h to $\mathcal{X} \subseteq \mathbb{R}^s$; i.e., the domain of $h_{\mathcal{X}}$ is \mathcal{X} , and $h_{\mathcal{X}}$ equals h on \mathcal{X} .

Theorem 2.3. For every $1 \leq i \leq n$, if the functions $f_i: \mathbb{R}^n \rightarrow \mathbb{R}$ are non-decreasing on \mathbb{R}_+^n and differentiable and quasiconvex on an open convex set \mathcal{S} with $\mathbb{R}_+^n \subseteq \mathcal{S}$, and if its gradients are locally bounded, then the restriction $\varphi_{\mathcal{X}}$ of the objective function $\varphi: \mathbb{R}^s \rightarrow \mathbb{R}$ of optimization problem (P) is boundedly lower subdifferentiable for every compact set $\mathcal{X} \subseteq \mathbb{R}^s$. Moreover, if x_0 does not belong to Γ_{φ} and $x_0 \in \text{int}(\mathcal{X})$, it follows that $\text{cl}(\text{cone}(\partial^{-}\varphi_{\mathcal{X}}(x_0)))$ equals $\bigcup_{\lambda \geq 0} \lambda \mathcal{C}(x_0)$ for every compact set \mathcal{X} with $\text{cl}(\text{cone}(\partial^{-}\varphi_{\mathcal{X}}(x_0)))$ denoting the smallest closed cone including $\partial^{-}\varphi_{\mathcal{X}}(x_0)$.

Proof. By the remarks after Definition 2.3, it follows that φ is quasiconvex on \mathbb{R}^s and Lipschitz continuous on any compact subset. To prove that $\varphi_{\mathcal{X}}$ is boundedly lower subdifferentiable, we will construct for each $x_0 \in \mathcal{X}$ a lower subgradient x_0^* with uniformly bounded norm. This construction is a slight modification of a similar construction in Ref. 26. Let \mathcal{X} be some compact set and $x_0 \in \mathcal{X}$, and suppose that

$$\mathcal{L}\varphi_{\mathcal{X}}(\varphi_{\mathcal{X}}(x_0)) = \mathcal{L}_{\varphi}(\varphi(x_0)) \cap \mathcal{X}$$

is nonempty. Since φ is quasiconvex on \mathbb{R}^s and Lipschitz continuous on every compact subset, it must follow that the nonempty lower level set $\mathcal{L}_{\varphi}(\varphi(x_0))$ is an open convex set. Due to x_0 not being an element of $\mathcal{L}_{\varphi}(\varphi(x_0))$, there exists by Theorem 11.3 of Ref. 1 some nonzero $u_0 \in \mathbb{R}^s$ satisfying

$$\langle x - x_0, u_0 \rangle < 0, \quad \text{for every } x \in \mathcal{L}_{\varphi}(\varphi(x_0)).$$

If x_0 does not belong to Γ_{φ} , we obtain by Theorem 2.2 that every element of the cone $\bigcup_{\lambda > 0} \lambda \mathcal{C}(x_0)$ satisfies the above property. Consider now the

hyperplane

$$\mathcal{H} := \{x \in \mathbb{R}^n : \langle x - x_0, u_0 \rangle = 0\},$$

and let $x_{\mathcal{H}}$ denote the orthogonal projection of x on \mathcal{H} . Since $\mathcal{H} \cap \mathcal{L}_{\varphi}(\varphi(x_0))$ is empty and $x_{\mathcal{H}}$ belongs to \mathcal{H} , this implies that $\varphi(x_{\mathcal{H}}) \geq \varphi(x_0)$. Since \mathcal{H} is compact, there exists some $r > 0$ satisfying $\mathcal{H} \subseteq r\mathcal{B}$, with \mathcal{B} the open Euclidean unit ball. Hence by Lemma 2.2 and the Lipschitz continuity of φ on $\sqrt{2}r \text{cl}(\mathcal{B})$, we can find some $L > 0$ satisfying

$$\varphi(x_0) - \varphi(x) \leq \varphi(x_{\mathcal{H}}) - \varphi(x) \leq L\|x_{\mathcal{H}} - x\|_2, \tag{7}$$

for every $x \in \mathcal{L}_{\varphi}(\varphi(x_0)) \cap \mathcal{H} \subseteq \mathcal{L}_{\varphi}(\varphi(x_0)) \cap r\mathcal{B}$. Moreover it follows, using $x_0 - x_{\mathcal{H}} \perp u_0$ and $x - x_{\mathcal{H}}$ parallel to u_0 , that

$$\begin{aligned} \langle x - x_0, u_0 \rangle &= \langle x - x_{\mathcal{H}}, u_0 \rangle + \langle x_{\mathcal{H}} - x_0, u_0 \rangle \\ &= \langle x - x_{\mathcal{H}}, u_0 \rangle = -\|u_0\|_2 \|x - x_{\mathcal{H}}\|_2. \end{aligned} \tag{8}$$

Hence by (7) and (8), we obtain that $Lu_0/\|u_0\|_2$ belongs to $\partial^- \varphi_{\mathcal{H}}(x_0)$, and so the first part is proved. To verify the remaining part, we observed already that any element of the cone $\bigcup_{\lambda > 0} \lambda \mathcal{C}(x_0)$ can be taken as u_0 , and this immediately implies by the above construction that

$$\bigcup_{\lambda \geq 0} \lambda \mathcal{C}(x_0) \subseteq \text{cl}(\text{cone}(\partial^- \varphi_{\mathcal{H}}(x_0))).$$

Due to the definition of $\partial^- \varphi_{\mathcal{H}}(x_0)$ and x_0 belonging to $\text{int}(\mathcal{H})$, one can easily show that $\partial^- \varphi_{\mathcal{H}}(x_0)$ must be a subset of the normal cone $\mathcal{N}_{\mathcal{L}_{\varphi}(\varphi(x_0))}(x_0)$ to $\text{cl}(\mathcal{L}_{\varphi}(\varphi(x_0)))$ at x_0 ; so, by Theorem 2.2, it follows that

$$\text{cl}(\text{cone}(\partial^- \varphi_{\mathcal{H}}(x_0))) \subseteq \bigcup_{\lambda \geq 0} \lambda \mathcal{C}(x_0).$$

Hence, the two sets are equal, and this concludes the proof. □

In the next section, we will discuss an algorithm to solve optimization problem (P).

3. Deep and Central-Cut Ellipsoid Algorithm

In this section, a so-called deep and central-cut version of the ellipsoid algorithm will be presented for the solution of our optimization problem (P). Although this method can also be applied to general unconstrained optimization problems with similar properties, we present only the method in the framework of our general location model. It is assumed that the

functions $f_i: \mathbb{R}^n \rightarrow \mathbb{R}$ are nondecreasing on \mathbb{R}_+^n and differentiable and quasi-convex on an open convex set \mathcal{S} with $\mathbb{R}_+^n \subseteq \mathcal{S}$. These conditions guarantee that φ is quasiconvex on \mathbb{R}^s . Moreover, if the gradients of the functions f_i are locally bounded, then the function φ is also Lipschitz continuous on any compact subset of its domain; by Theorem 2.3, it follows that φ is boundedly lower subdifferentiable on any compact subset. Observe for computational purposes that, in the proof of Theorem 2.3, an easy construction of a lower subgradient is given if the present iteration point does not belong to Γ_φ . However, in order to carry out this construction, we need to know the Lipschitz constant on a certain compact subset. This means that, if the function φ is not Lipschitz continuous on that set or if the Lipschitz constant is not known, we can apply only a central-cut version of the algorithm.

As verified already, the following essential assumption holds for our optimization problem. Recall that \mathcal{B} denotes the open Euclidean unit ball.

Assumption 3.1. An optimal solution x^* of the optimization problem (P) exists satisfying $x^* \in a_0 + r\mathcal{B}$, with $a_0 \in \mathbb{R}^s$ and $r > 0$ known in advance.

We will now give a description of the iterative procedure to solve (P). Each step of the algorithm tests whether the current iteration point belongs to $\Gamma_\varphi \cap (a_0 + r\mathcal{B})$. If so, the algorithm is stopped, and the present iteration point is taken as a solution. Under the additional assumption that $\nabla f_i(z)$ contains at least one positive component for every $1 \leq i \leq n$ and $z \in \mathcal{S}$, it follows by Lemma 2.1 that this iteration point is indeed optimal. Since the general framework of the ellipsoid algorithm is well known (Refs. 29 and 30), we proceed by describing each iteration step. Clearly, by Assumption 3.1 it follows that $x^* \in a_0 + r\mathcal{B}$ satisfies

$$\varphi(x^*) = \min\{\varphi(x) : x \in \mathbb{R}^s\} = \min\{\varphi(x) : x \in a_0 + r\mathcal{B}\},$$

and so the ellipsoid algorithm is started by taking $a_0 + r \text{cl}(\mathcal{B})$ as the initial ellipsoid

$$\mathcal{E}(A_0; a_0) := \{x \in \mathbb{R}^s : \langle x - a_0, A_0^{-1}(x - a_0) \rangle \leq 1\},$$

with

$$A_0 = r^2 I.$$

Suppose now the ellipsoid algorithm starts the $(m+1)$ th step, $m \geq 0$, and x^* belongs to $\mathcal{E}(A_m; a_m)$. By the stopping rule, it follows that a_k does not belong to $\Gamma_\varphi \cap (a_0 + r\mathcal{B})$ for every $k \leq m$, and so $l_m > \varphi(x^*)$, with

$$l_m := \min\{\varphi(a_k) : k \leq m, a_k \in a_0 + r\mathcal{B}\}$$

the lowest recorded function value on the set $a_0 + r\mathcal{B}$ until iteration m . Clearly, l_0 is properly defined due to $a_0 \in a_0 + r\mathcal{B}$.

In order to generate cuts we distinguish the following cases.

Case 1. a_m belongs to $a_0 + r\mathcal{B}$. Since a_m does not belong to Γ_φ , it follows by Theorem 2.2 (in case φ is not Lipschitz continuous or the Lipschitz constant is not known) that

$$\mathcal{L}_\varphi(\varphi(a_m)) \subseteq \mathcal{H}^-(\beta_m) := \{x \in \mathbb{R}^s : \langle x, a_m^* \rangle \leq \beta_m\}, \tag{9}$$

with

$$\beta_m := \langle a_m, a_m^* \rangle \quad \text{and} \quad a_m^* \in \bigcup_{\lambda > 0} \lambda \mathcal{C}(a_m).$$

Hence, the optimal solution x^* belongs to the closed halfspace $\mathcal{H}^-(\beta_m)$. Moreover, if the Lipschitz constant for the function φ on $a_0 + \sqrt{2}r \text{cl}(\mathcal{B})$ is known, there exists by Theorem 2.3 an easy computable nonzero $a_m^* \in \mathbb{R}^s$ satisfying

$$\varphi(x) \geq \varphi(a_m) + \langle x - a_m, a_m^* \rangle, \tag{10}$$

for every $x \in \mathcal{L}_\varphi(\varphi(a_m)) \cap (a_0 + r\mathcal{B})$. In order to derive a so-called deep or central cut with respect to φ , observe by (10) and using $x^* \in a_0 + r\mathcal{B}$ that

$$l_m > \varphi(x^*) \geq \varphi(a_m) + \langle x^* - a_m, a_m^* \rangle,$$

and so x^* belongs to the lower halfspace

$$\mathcal{H}^-(\beta_m) := \{x \in \mathbb{R}^s : \langle x, a_m^* \rangle \leq \beta_m\},$$

with

$$\beta_m := \langle a_m, a_m^* \rangle + l_m - \varphi(a_m).$$

It can be shown for both cases (Ref. 22) that the hyperplane $\mathcal{H}(\beta_m)$ is a valid central cut, if $l_m = \varphi(a_m)$, or a valid deep cut, if $l_m < \varphi(a_m)$, and so there exists (Ref. 29) a smaller-volume ellipsoid $\mathcal{E}(A_{m+1}; a_{m+1})$ with

$$x^* \in \mathcal{E}(A_m; a_m) \cap \mathcal{H}^-(\beta_m) \subseteq \mathcal{E}(A_{m+1}; a_{m+1}).$$

This finishes the construction of a so-called objective cut.

Case 2. a_m does not belong to $a_0 + r\mathcal{B}$. If $h(x) := \|x - a_0\|_2$, we obtain by the subgradient inequality that

$$h(x) \geq h(a_m) + \langle x - a_m, \nabla h(a_m) \rangle,$$

with

$$\nabla h(a_m) = (a_m - a_0) / \|a_m - a_0\|_2,$$

and so it follows that

$$r > h(x^*) \geq h(a_m) + \langle x^* - a_m, \nabla h(a_m) \rangle = \langle x^* - a_0, \nabla h(a_m) \rangle.$$

Hence, the optimal point x^* belongs to the lower halfspace

$$\mathcal{H}^-(\beta_m) := \{x \in \mathbb{R}^s : \langle x, \nabla h(a_m) \rangle \leq \beta_m\},$$

with

$$\beta_m := r + \langle a_0, \nabla h(a_m) \rangle.$$

Since one can also prove that the hyperplane $H(\beta_m)$ is a valid deep cut (Ref. 22), there exists again a smaller volume ellipsoid $\mathcal{E}(A_{m+1}; a_{m+1})$ satisfying

$$x^* \in \mathcal{E}(A_m; a_m) \cap H^-(\beta_m) \subseteq \mathcal{E}(A_{m+1}; a_{m+1}).$$

This finishes the construction of a so-called norm cut.

Since we have explained how to generate cuts in each step, we can now give a description of the algorithm.

Algorithm 3.1.

Step 0. Let $m := 0$ and $A_0 := r^2 I$.

Step 1. If $a_m \in \Gamma_\varphi \cap (a_0 + r\mathcal{B})$, then stop; else, go to Step 2.

Step 2. If $a_m \in a_0 + r\mathcal{B}$, then apply an objective cut; else, apply a norm cut.

Step 3. Update the ellipsoid, let $m := m + 1$, and return to Step 1.

As already observed in Section 2, the reader is referred to Ref. 17 for a discussion on the verification of the stopping rule in Step 1 by means of an easy algorithm in case of a min-max planar location problem with an l_p -norm to measure distances. This algorithm is based on the results derived in the previous section. Moreover, the linear time algorithm presented in Ref. 17 to check whether 0 belongs to the convex hull of a finite set of points in \mathbb{R}^2 can also be applied in the case of polyhedral gauges, while for \mathbb{R}^s , $s \geq 2$, this reduces to the feasibility of a linear programming problem. The above generation of simple deep cuts is known for the convex-constrained case (Refs. 22, 23 and 31). Finally, denoting the depth of a cut by

$$0 \leq \alpha_m := [\langle a_m, a_m^* \rangle - \beta_m] / \sqrt{\langle a_m^*, A_m a_m^* \rangle} < 1,$$

one can show (Ref. 29) that in Step 3 the ellipsoid $\mathcal{E}(A_{m+1}; a_{m+1})$ given by

$$A_{m+1} := \delta_m(A_m - \sigma_m b_m b_m^t),$$

$$a_{m+1} := a_m - \tau_m b_m,$$

with the updating values

$$\delta_m := s^2(1 - \alpha_m^2)/(s^2 - 1),$$

$$\sigma_m := 2(1 + s\alpha_m)/[(s+1)(1 + \alpha_m)],$$

$$\tau_m := (1 + s\alpha_m)/(s+1),$$

and

$$b_m := A_m a_m^* / \sqrt{\langle a_m^*, A_m a_m^* \rangle},$$

is the smallest volume ellipsoid containing $\mathcal{E}(A_m; a_m) \cap H^-(\beta_m)$.

In order to give an easy (contrary to the proof in Ref. 32) convergence proof of the above algorithm, we need the following lemma.

Lemma 3.1. It follows that $\mathcal{L}_\varphi(l_m) \cap (a_0 + r\mathcal{B}) \subseteq \mathcal{E}(A_m; a_m) \cap \mathcal{H}^-(\beta_m)$, for every $m \geq 0$, with $\mathcal{E}(A_m; a_m)$ the ellipsoid constructed in the m th step of the algorithm.

Proof. If $m=0$, we obtain that $a_0 \in a_0 + r\mathcal{B}$; so, for both function classes, an objective cut is performed. Hence, if the Lipschitz constant of φ is known, it follows by (10) that

$$l_0 > \varphi(x) \geq \varphi(a_0) + \langle x - a_0, a_0^* \rangle,$$

for every $x \in (a_0 + r\mathcal{B}) \cap \mathcal{L}_\varphi(\varphi(a_0))$ and this implies that

$$(a_0 + r\mathcal{B}) \cap \mathcal{L}_\varphi(\varphi(a_0)) \subseteq \mathcal{H}^-(\beta_0).$$

Moreover, if the other class of functions is considered, we obtain by (9) that again

$$(a_0 + r\mathcal{B}) \cap \mathcal{L}_\varphi(\varphi(a_0)) \subseteq \mathcal{H}^-(\beta_0);$$

and using $a_0 + r\mathcal{B} \subseteq \mathcal{E}(A_0; a_0)$, both cases yield that

$$\mathcal{L}_\varphi(l_0) \cap (a_0 + r\mathcal{B}) \subseteq \mathcal{H}^-(\beta_0) \cap \mathcal{E}(A_0; a_0).$$

Hence, the desired result holds for $m=0$; to continue the proof, we suppose that the result holds for step $m=k$. Since $l_{k+1} \leq l_k$, the induction hypothesis clearly implies that

$$\begin{aligned} \mathcal{L}_\varphi(l_{k+1}) \cap (a_0 + r\mathcal{B}) &\subseteq \mathcal{E}(A_k; a_k) \cap \mathcal{H}^-(\beta_k) \\ &\subseteq \mathcal{E}(A_{k+1}; a_{k+1}). \end{aligned} \tag{11}$$

Moreover, if $a_{k+1} \in a_0 + r\mathcal{B}$, it follows similarly as for $m=0$ that, for both cases, it must hold that

$$\mathcal{L}_\varphi(l_{k+1}) \cap (a_0 + r\mathcal{B}) \subseteq \mathcal{H}^-(\beta_{k+1}),$$

and this yields by (11) that

$$\mathcal{L}_\varphi(l_{k+1}) \cap (a_0 + r\mathcal{B}) \subseteq \mathcal{E}(A_{k+1}; a_{k+1}) \cap \mathcal{H}^-(\beta_{k+1}).$$

Finally, if $a_{k+1} \notin a_0 + r\mathcal{B}$, a norm cut will be performed; as in the previous case, using the subgradient inequality for $h(x) = \|x - a_0\|_2$, it follows that

$$a_0 + r\mathcal{B} \subseteq H^-(\beta_{k+1}).$$

Applying again (11) yields

$$\mathcal{L}_\varphi(l_{k+1}) \cap (a_0 + r\mathcal{B}) \subseteq \mathcal{E}(A_{k+1}; a_{k+1}) \cap \mathcal{H}^-(\beta_{k+1}),$$

and so the desired result is proved. □

It is now possible to prove the main result. With

$$a := (s^2 - 1)/s^2, \quad b := \sqrt[s]{(s+1)/(s-1)},$$

observe that

$$ab > 1, \quad 1/\sqrt{ab} < 1.$$

Theorem 3.1. If the ellipsoid algorithm executes an infinite number of iterations, then it follows that $l_m \downarrow \varphi(x^*)$. Moreover, if φ is Lipschitz continuous on $a_0 + \sqrt{2}r \text{cl}(\mathcal{B})$ with Lipschitz constant L , then there exists some m_0 such that

$$\begin{aligned} 0 &\leq l_m - \varphi(x^*) \\ &\leq (Lr/\sqrt{s}) (1/\sqrt{ab})^m \prod_{k=0}^{m-1} \sqrt{1 - \alpha_k^2} \sqrt[2s]{(1 - \alpha_k)/(1 + \alpha_k)}, \end{aligned}$$

for every $m \geq m_0$.

Proof. We start by evaluating $\det(A_m)$. It can be proved (Ref. 22) that

$$\det(A_m) = \det(A_0) \prod_{k=0}^{m-1} (\delta_k^s (1 - \sigma_k)).$$

Since $A_0 = r^2 I$ and A_m is positive definite for every m , we obtain after some calculations that

$$\begin{aligned} 0 &\leq \det(A_m) = r^{2s} \prod_{k=0}^{m-1} (1/ab)^s (1 - \alpha_k^2)^s [(1 - \alpha_k)/(1 + \alpha_k)] \\ &\leq r^{2s} (1/ab)^{sm}, \end{aligned}$$

and this yields $\det(A_m) \rightarrow 0$. Clearly, l_m is a nonincreasing sequence satisfying $l_m \geq \varphi(x^*)$ for every $m \geq 0$, and so it follows that $c := \lim_{m \uparrow \infty} l_m$ exists and $c \geq \varphi(x^*)$. Suppose now that $c > \varphi(x^*)$. Since $x^* \in a_0 + r\mathcal{B}$, there exists some $\epsilon > 0$ such that $x^* + \delta\mathcal{B} \subseteq a_0 + r\mathcal{B}$ for every $0 < \delta \leq \epsilon$. Moreover, since φ is continuous on $a_0 + r\mathcal{B}$ and $\varphi(x^*) < c$, there exists also some $0 < \hat{\delta} \leq \epsilon$ satisfying $x^* + \hat{\delta}\mathcal{B} \subseteq \mathcal{L}_\varphi(c)$, and so

$$x^* + \hat{\delta}\mathcal{B} \subseteq \mathcal{L}_\varphi(c) \cap (a_0 + r\mathcal{B}).$$

Hence, by Lemma 3.1 and $l_m \downarrow c$, we obtain that, for every $m \geq 0$,

$$\begin{aligned} x^* + \hat{\delta}\mathcal{B} \subseteq \mathcal{L}_\varphi(c) \cap (a_0 + r\mathcal{B}) &\subseteq \mathcal{L}_\varphi(l_m) \cap (a_0 + r\mathcal{B}) \\ &\subseteq \mathcal{E}(A_m; a_m) \cap \mathcal{H}^-(\beta_m). \end{aligned}$$

Finally, since

$$\text{vol}(\mathcal{E}(A_m; a_m)) = \sqrt{\det(A_m)} V_s, \quad \text{with } V_s := \text{vol}(\mathcal{B}),$$

it follows that

$$\begin{aligned} 0 < \hat{\delta}^s V_s = \text{vol}(x^* + \hat{\delta}\mathcal{B}) \\ &\leq \text{vol}(\mathcal{E}(A_m; a_m) \cap \mathcal{H}^-(\beta_m)) \\ &\leq (1/2) \text{vol}(\mathcal{E}(A_m; a_m)) \\ &= (1/2) \sqrt{\det(A_m)} V_s, \end{aligned}$$

for every $m \geq 0$, and this contradicts $\det(A_m) \rightarrow 0$. Hence, it must follow that $l_m \downarrow \varphi(x^*)$, and so the first part is proved. To verify the inequality, we observe that, for φ Lipschitz continuous with Lipschitz constant L and using $l_m \downarrow \varphi(x^*)$, there exists some m_0 such that

$$[l_m - \varphi(x^*)]/L \leq \epsilon,$$

for every $m \geq m_0$. Take now $x \in x^* + [(l_m - \varphi(x^*)]/L]\mathcal{B}$ with $m \geq m_0$. We obtain

$$\varphi(x) - \varphi(x^*) \leq L \|x - x^*\|_2 < l_m - \varphi(x^*),$$

and this implies $x \in L_\varphi(l_m)$ and

$$\begin{aligned} x^* + [(l_m - \varphi(x^*)]/L]\mathcal{B} &\subseteq L_\varphi(l_m) \cap (a_0 + r\mathcal{B}) \\ &\subseteq \mathcal{E}(A_m; a_m) \cap \mathcal{H}^-(\beta_m). \end{aligned}$$

Thus,

$$\begin{aligned} \text{vol}(x^* + [(l_m - \varphi(x^*)) / L] \mathcal{B}) &\leq \text{vol}(\mathcal{E}(A_m; a_m) \cap \mathcal{H}^-(\beta_m)) \\ &\leq (1/2) \text{vol}(\mathcal{E}(A_m; a_m)), \end{aligned}$$

and computing these volumes gives

$$[(l_m - \varphi(x^*)) / L]^s V_s \leq (1/2) \sqrt[r^{2s} \prod_{k=0}^{m-1} (1/ab)^s (1 - \alpha_k^2)^s [(1 - \alpha_k) / (1 + \alpha_k)]]{V_s}.$$

Dividing by V_s , raising both sides to s^{-1} , and multiplying by L yields the desired result. \square

We finish this section by pointing out that, if $\alpha_k = 0$ for every $k \geq 0$, then the result boils down to the well-known geometric rate of the unconstrained central-cut ellipsoid algorithm (Refs. 32 and 33). The convergence rate was proved to be the same in the constrained case (Ref. 34). However, in Ref. 34, only the central-cut case was analyzed. Observe that the analysis of the deep-cut case exhibits the positive influence of deep cuts in the convergence of the algorithm and can be extended to the constrained case as well (Ref. 22). However, the proof of this result for the constrained case is much more elaborate.

4. Computational Experience

In order to test the algorithm, it was coded by us in Turbo Pascal version 7.0; no commercial routines were used, except the standard functions and procedures of the language. The program includes the optimality test discussed in Ref. 17, which was applied to each center not subject to a norm cut. The program was compiled and executed on an AST Bravo 4/33, a PC/AT compatible with an Intel 80486 CPU with built-in numerical processor and clock speed of 33 MHz. The numerical precision used is the double precision (64-bit IEEE floating-point format) real numbers of Turbo Pascal. The computational experience was carried out over 300 uncorrelated planar instances of the problem taking as the disutility function of each demand point $d_i \in \mathbb{R}^2$, $1 \leq i \leq n$, the function

$$\begin{aligned} f_i(z) &= 100w_i \log(z_i + 1), & \text{for } 1 \leq i \leq \lfloor n/3 \rfloor, \\ f_i(z) &= 100w_i \arctan(z_i), & \text{for } \lfloor n/3 \rfloor + 1 \leq i \leq \lfloor 2n/3 \rfloor, \\ f_i(z) &= 5w_i z_i, & \text{for } \lfloor 2n/3 \rfloor + 1 \leq i \leq n, \end{aligned}$$

with $\lceil x \rceil$ denoting the ceiling of $x \in \mathbb{R}$. Assigning to the nonlinear (concave) disutility functions a bigger weight than to the linear ones prevents that the linear functions dominate the concave functions prohibiting φ to become convex.

It is easy to verify, due to

$$\|x\|_p \leq \|x\|_1 \leq \sqrt{s} \|x\|_2, \quad \text{for every } 1 \leq p \leq \infty \text{ and } x \in \mathbb{R}^s,$$

that for $s=2$ the constant

$$\sqrt{2} \max \left\{ \max_{1 \leq i \leq \lceil 2n/3 \rceil} \{100w_{ij}\}, \max_{\lceil 2n/3 \rceil + 1 \leq i \leq n} \{5w_{ij}\} \right\}$$

is a Lipschitz constant for the function φ .

The instances were generated randomly in the following way. We start by describing the selection of the problem parameters.

The number n of demand points belongs to $\{5, 25, 50, 100, 250\}$.

For the l_p -norm being used, we take $p_i \in \{1.1, 1.5, 1.9, 2.1, 3.0\}$ and construct two different classes of examples. For the first class, one value is assigned to all the demand points; for the second class, denoted in the tables as "mix," to each demand point a value p_i selected randomly from the set $\{1.1, 1.5, 1.9, 2.1, 3.0\}$ is assigned. The weight w_i of the demand point d_i is determined as follows. We draw uniformly numbers from the interval $[0, 1]$, say \tilde{w}_i , $1 \leq i \leq n$, and set w_i equal to

$$h_i = \tilde{w}_i / \sum_{j=1}^n \tilde{w}_j, \quad \text{for every } 1 \leq i \leq n.$$

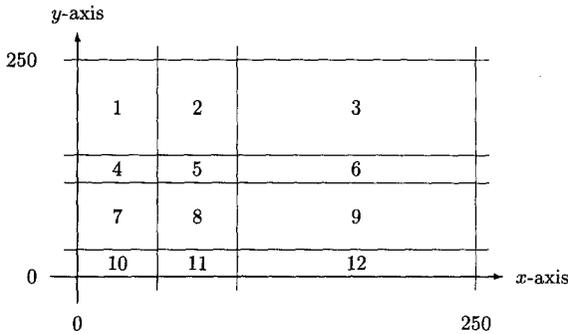


Fig. 1. Clustered problem $m_1=2, m_2=3$.

Table 1a. Results for quasiconvex functions applying central cuts.

n	p	T	It	O	N
5	1.1	0.154	91.3	91.0	0.3
5	1.5	0.166	98.2	98.1	0.1
5	1.9	0.165	97.7	97.6	0.1
5	2.1	0.160	94.4	94.4	0.0
5	3.0	0.179	98.5	98.4	0.1
5	mix	0.169	100.2	100.1	0.1
25	1.1	0.665	97.5	97.4	0.1
25	1.5	0.654	102.0	102.0	0.0
25	1.9	0.649	99.2	99.2	0.0
25	2.1	0.694	99.3	99.3	0.0
25	3.0	0.730	109.4	109.4	0.0
25	mix	0.700	106.3	106.3	0.0
50	1.1	1.308	103.4	103.3	0.1
50	1.5	1.296	102.2	102.2	0.0
50	1.9	1.402	112.1	112.1	0.0
50	2.1	1.402	108.8	108.8	0.0
50	3.0	1.256	102.0	101.8	0.2
50	mix	1.422	112.2	112.2	0.0
100	1.1	2.286	93.8	93.7	0.1
100	1.5	2.606	103.6	103.6	0.0
100	1.9	2.654	109.3	109.3	0.0
100	2.1	2.708	107.0	107.0	0.0
100	3.0	2.688	107.0	106.8	0.2
100	mix	2.503	103.7	103.7	0.0
250	1.1	6.307	102.2	102.2	0.0
250	1.5	6.572	106.7	106.7	0.0
250	1.9	6.845	111.1	111.1	0.0
250	2.1	6.979	114.1	114.1	0.0
250	3.0	6.828	111.1	111.1	0.0
250	mix	7.107	115.5	115.5	0.0

Now, we describe the procedure to generate the demand points. All the demand points are generated within the square $[0, 250] \times [0, 250]$, for which a clustered structure is created using the following procedure. First, we draw two integers m_1 and m_2 ranging from 1 to 20; then, we divide the square $[0, 250] \times [0, 250]$ into $(m_1 + 1)(m_2 + 1)$ subsquares by generating randomly m_1 x -axis coordinates and m_2 y -axis coordinates in $(0, 250)$ (cf. Fig. 1). Then, we label these subsquares from 1 to $(m_1 + 1)(m_2 + 1)$.

Subsequently, we choose randomly according to these labels some given number of subsquares. In each chosen subsquare, we draw uniformly a given number of demand points. Finally, the remaining demand points are drawn

Table 1b. Results for quasiconvex functions applying deep cuts.

<i>n</i>	<i>p</i>	<i>T</i>	<i>It</i>	<i>O</i>	$\bar{\alpha}_O$	<i>N</i>	$\bar{\alpha}_N$	<i>DO</i>	$\bar{\alpha}_{DO}$
5	1.1	0.159	90.5	90.4	0.0041	0.1	0.0035	62.7	0.0058
5	1.5	0.165	97.2	97.1	0.0033	0.1	0.0055	67.2	0.0050
5	1.9	0.165	97.3	97.2	0.0022	0.1	0.0006	71.7	0.0030
5	2.1	0.159	93.2	93.2	0.0026	0.0	0.0000	65.3	0.0037
5	3.0	0.165	97.3	97.1	0.0043	0.2	0.0181	68.5	0.0062
5	mix	0.169	99.3	99.2	0.0030	0.1	0.0045	69.7	0.0043
25	1.1	0.617	96.0	95.9	0.0033	0.1	0.0009	68.8	0.0046
25	1.5	0.669	101.1	101.1	0.0033	0.0	0.0000	75.9	0.0044
25	1.9	0.653	98.4	98.4	0.0028	0.0	0.0000	75.6	0.0037
25	2.1	0.628	97.8	97.8	0.0026	0.0	0.0000	71.1	0.0036
25	3.0	0.720	108.5	108.5	0.0027	0.0	0.0000	80.4	0.0036
25	mix	0.695	105.3	105.3	0.0030	0.0	0.0000	75.5	0.0042
50	1.1	1.318	101.0	100.9	0.0045	0.1	0.0041	77.4	0.0058
50	1.5	1.249	101.2	101.2	0.0032	0.0	0.0000	76.1	0.0043
50	1.9	1.396	110.3	110.3	0.0033	0.0	0.0000	83.4	0.0043
50	2.1	1.358	107.3	107.3	0.0034	0.0	0.0000	82.0	0.0045
50	3.0	1.338	101.2	101.0	0.0028	0.2	0.0042	76.5	0.0037
50	mix	1.360	110.4	110.4	0.0031	0.0	0.0000	80.9	0.0041
100	1.1	2.354	92.2	92.1	0.0043	0.1	0.0053	70.4	0.0057
100	1.5	2.455	101.6	101.6	0.0033	0.0	0.0000	76.1	0.0044
100	1.9	2.778	108.5	108.5	0.0035	0.0	0.0000	82.4	0.0046
100	2.1	2.549	104.8	104.8	0.0037	0.0	0.0000	78.3	0.0050
100	3.0	2.583	105.9	105.7	0.0030	0.2	0.0084	80.2	0.0039
100	mix	2.615	102.7	102.7	0.0036	0.0	0.0000	76.5	0.0049
250	1.1	5.954	100.0	100.0	0.0047	0.0	0.0000	75.0	0.0063
250	1.5	6.290	105.6	105.6	0.0040	0.0	0.0000	79.7	0.0053
250	1.9	6.530	109.6	109.6	0.0039	0.0	0.0000	83.3	0.0051
250	2.1	6.721	112.4	112.4	0.0038	0.0	0.0000	87.6	0.0049
250	3.0	6.573	110.3	110.3	0.0036	0.0	0.0000	85.7	0.0046
250	mix	6.796	114.1	114.1	0.0045	0.0	0.0000	85.9	0.0059

uniformly from the original square $[0, 250] \times [0, 250]$ and added to the already existing set of demand points, in a total of n points.

Finally, the execution is stopped with a relative error (see Ref. 22) less than 5×10^{-6} .

For each pair (n, p) , 10 uncorrelated instances of the problem were generated according to the procedure described above, and each of them was solved by the central-cut and the deep-cut versions of the algorithm. Also for each value of n , 10 uncorrelated instances were generated with mixed norms and also solved by both versions of the algorithm.

The results are summarized in Table 1a and Table 1b, where each row corresponds to averages of 10 instances.

Table 2a. Results for convex functions applying central cuts.

n	p	T	It	O	N
5	1.1	0.111	69.5	69.5	0.1
5	1.5	0.123	77.6	77.6	0.0
5	1.9	0.120	75.8	75.8	0.0
5	2.1	0.117	74.0	74.0	0.0
5	3.0	0.120	75.3	75.2	0.1
5	mix	0.134	84.5	84.5	0.0
25	1.1	0.457	77.0	77.0	0.0
25	1.5	0.537	84.8	84.8	0.0
25	1.9	0.468	76.8	76.8	0.0
25	2.1	0.522	81.2	81.2	0.0
25	3.0	0.508	80.7	80.7	0.0
25	mix	0.504	81.4	81.4	0.0
50	1.1	0.866	73.7	73.6	0.1
50	1.5	0.914	78.7	78.7	0.0
50	1.9	0.920	80.7	80.7	0.0
50	2.1	0.980	81.1	81.1	0.0
50	3.0	0.944	80.8	80.8	0.0
50	mix	0.933	81.8	81.8	0.0
100	1.1	1.766	77.1	77.1	0.0
100	1.5	1.911	82.9	82.9	0.0
100	1.9	1.846	80.3	80.3	0.0
100	2.1	1.877	81.6	81.6	0.0
100	3.0	1.934	84.1	84.1	0.0
100	mix	1.855	81.3	81.3	0.0
250	1.1	4.397	77.1	77.1	0.0
250	1.5	4.624	81.4	81.4	0.0
250	1.9	4.581	79.6	79.6	0.0
250	2.1	4.992	86.5	86.5	0.0
250	3.0	5.122	88.9	88.9	0.0
250	mix	4.676	82.0	82.0	0.0

In Table 1a, we list the average time T taken by the central-cut version of the algorithm in seconds of the AST Bravo, the average total number of iterations It , the average number of objective cuts O , and the average number of norm cuts N .

In Table 1b, we list the average time T taken by the deep-cut version of the algorithm in seconds of the AST Bravo, the average total number of iterations It , the average number of objective cuts O , the average depth of an objective cut $\bar{\alpha}_O$, the average number of norm cuts N and the corresponding average depth $\bar{\alpha}_N$, the average number of objective cuts that were deep DO ,

Table 2b. Results for convex functions applying deep cuts.

n	p	T	It	O	$\bar{\alpha}_O$	N	$\bar{\alpha}_N$	DO	$\bar{\alpha}_{DO}$
5	1.1	0.077	48.3	48.2	0.0960	0.1	0.0019	30.8	0.1498
5	1.5	0.094	59.0	59.0	0.0769	0.0	0.0000	35.4	0.1232
5	1.9	0.087	54.6	54.6	0.0892	0.0	0.0000	34.4	0.1402
5	2.1	0.083	51.8	51.8	0.0947	0.0	0.0000	33.4	0.1464
5	3.0	0.089	55.4	55.3	0.0849	0.1	0.0058	34.0	0.1367
5	mix	0.107	66.7	66.7	0.0651	0.0	0.0000	36.2	0.1074
25	1.1	0.349	51.5	51.5	0.1039	0.0	0.0000	34.7	0.1513
25	1.5	0.399	65.6	65.6	0.0685	0.0	0.0000	37.9	0.1111
25	1.9	0.334	55.6	55.6	0.0896	0.0	0.0000	33.7	0.1381
25	2.1	0.358	60.2	60.2	0.0805	0.0	0.0000	37.5	0.1267
25	3.0	0.365	60.8	60.8	0.0791	0.0	0.0000	35.6	0.1258
25	mix	0.345	54.7	54.7	0.1044	0.0	0.0000	35.7	0.1558
50	1.1	0.627	53.0	52.9	0.0830	0.1	0.0022	34.2	0.1283
50	1.5	0.692	57.4	57.4	0.0857	0.0	0.0000	35.5	0.1316
50	1.9	0.701	60.1	60.1	0.0819	0.0	0.0000	36.6	0.1275
50	2.1	0.705	60.2	60.2	0.0854	0.0	0.0000	37.9	0.1275
50	3.0	0.680	59.6	59.6	0.0822	0.0	0.0000	37.0	0.1283
50	mix	0.707	57.4	57.4	0.0931	0.0	0.0000	36.9	0.1452
100	1.1	1.251	55.1	55.0	0.0875	0.1	0.0022	34.9	0.1377
100	1.5	1.387	60.8	60.8	0.0835	0.0	0.0000	36.9	0.1308
100	1.9	1.468	63.3	63.3	0.0671	0.0	0.0000	37.7	0.1081
100	2.1	1.354	59.3	59.3	0.0884	0.0	0.0000	36.6	0.1321
100	3.0	1.558	68.0	68.0	0.0627	0.0	0.0000	39.2	0.0998
100	mix	1.377	58.8	58.7	0.0868	0.1	0.0028	35.7	0.1368
250	1.1	3.139	56.1	56.1	0.0795	0.0	0.0000	34.6	0.1277
250	1.5	3.159	56.9	56.9	0.0939	0.0	0.0000	37.4	0.1406
250	1.9	3.212	57.7	57.7	0.0892	0.0	0.0000	36.2	0.1368
250	2.1	3.702	67.1	67.1	0.0706	0.0	0.0000	38.6	0.1110
250	3.0	3.786	68.6	68.6	0.0749	0.0	0.0000	38.5	0.1176
250	mix	3.423	61.3	61.3	0.0756	0.0	0.0000	36.6	0.1224

and the corresponding average depth (average taken within the deep objective cuts only) $\bar{\alpha}_{DO}$.

Notice that the seemingly small average depth may be explained by the dependence between the value of the Lipschitz constant and the depth of the cut. For the convex case, more encouraging results are reported in Ref. 22 with an average depth of cuts ten times bigger and about 16% reduction in iterations and time.

In order to report some results on convex functions, we generated similar examples where the disutility function of every demand point is taken

linear, i.e., $f_i(z) = w_i z_i$. This is known as the Rawls problem, and the corresponding results are given in Table 2a and Table 2b. For this problem, the influence of the deep cuts is much more encouraging.

References

1. ROCKAFELLAR, R. T., *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1972.
2. MICHELOT, C., *The Mathematics of Continuous Location*, Studies in Locational Analysis, ISOLDE VI Survey Papers, Vol. 5, pp. 59–83, 1993.
3. IDRISSE, H., LORIDAN, P., and MICHELOT, C., *Approximation of Solutions for Location Problems*, Journal of Optimization Theory and Applications, Vol. 56, pp. 127–143, 1988.
4. WARD, J. E., and WENDELL, R. E., *Using Block Norms for Location Modeling*, Operations Research, Vol. 33, pp. 1074–1090, 1985.
5. LOVE, R. F., and MORRIS, J. G., *Modelling Intercity Road Distances by Mathematical Functions*, Operational Research Quarterly, Vol. 23, pp. 61–71, 1972.
6. BERENS, W., and KOERLING, F. J., *Estimated Road Distances by Mathematical Functions*, European Journal of Operational Research, Vol. 21, pp. 54–56, 1972.
7. WARD, J. E., and WENDELL, R. E., *A New Norm for Measuring Distance Which Yields Linear Location Problems*, Operations Research, Vol. 28, pp. 836–844, 1980.
8. PLASTRIA, F., *On Destination Optimality in Asymmetric Distance Fermat–Weber Problems*, Annals of Operations Research, Vol. 40, pp. 355–369, 1992.
9. MARTI, J. T., *Konvexe Analysis*, Birkhauser Verlag, Basel, Switzerland, 1977.
10. PLASTRIA, F., *Continuous Location Anno 1992: A Progress Report*, Studies in Locational Analysis, ISOLDE VI Survey Papers, Vol. 5, pp. 85–127, 1993.
11. LOVE, R. F., MORRIS, J. G., and WESOŁOWSKY, G. O., *Facility Location: Models and Methods*, Elsevier Science Publishers B.V., North Holland, Amsterdam, Netherlands, 1988.
12. RUDIN, W., *Principles of Mathematical Analysis*, 3rd Edition, McGraw-Hill, Auckland, New Zealand, 1976.
13. ROSENHOLTZ, I., *There is No Differentiable Metric on \mathbb{R}^n* , American Mathematical Monthly, Vol. 86, pp. 585–586, 1979.
14. PLASTRIA, F., *Localization in Single Facility Location*, European Journal of Operational Research, Vol. 18, pp. 215–219, 1984.
15. WENDELL, R. E., and HURTHER, A. P., JR., *Location Theory, Dominance, and Convexity*, Operations Research, Vol. 21, pp. 314–320, 1973.
16. PSHENICHNYI, B. N., *Necessary Conditions for an Extremum*, Marcel Dekker, New York, New York, 1971.
17. FRENK, J. B. G., GROMICHO, J., and ZHANG, S., *General Models in Min-Max Planar Location: Checking Optimality Conditions*, Journal of Optimization Theory and Applications, Vol. 89, pp. 65–87, 1996.

18. AVRIEL, M., DIEWERT, W. E., SCHAIBLE, S., and ZANG, I., *Generalized Convexity*, Plenum Press, New York, New York, 1988.
19. KOMLÓSI, S., *Some Properties of Nondifferentiable Pseudoconvex Functions*, *Mathematical Programming*, Vol. 26, pp. 232–237, 1983.
20. CROUZEIX, J. P., *Some Differentiability Properties of Quasiconvex Functions on \mathbb{R}^n* , *Lecture Notes in Control and Information Sciences*, Springer Verlag, New York, New York, Vol. 30, pp. 9–20, 1981.
21. HIRIART-URRUTY, J. B., and LEMARÉCHAL, C., *Convex Analysis and Minimization Algorithms, I: Fundamentals*, *Comprehensive Studies in Mathematics*, Springer Verlag, Berlin, Germany, Vol. 305, 1993.
22. FRENK, J. B. G., GROMICHO, J., and ZHANG, S., *A Deep-Cut Ellipsoid Algorithm for Convex Programming: Theory and Applications*, *Mathematical Programming*, Vol. 63, pp. 83–108, 1994.
23. SHOR, N. Z., and GERSHOVICH, V. I., *Family of Algorithms for Solving Convex Programming Problems*, *Cybernetics*, Vol. 15, pp. 502–508, 1979.
24. ORTEGA, J. M., and RHEINBOLDT, W. C., *Iterative Solutions of Nonlinear Equations in Several Variables*, Academic Press, New York, New York, 1970.
25. VARBERG, D. E., and ROBERTS, A. W., *Convex Functions*, Academic Press, New York, New York, 1973.
26. PLASTRIA, F., *Lower Subdifferentiable Functions and Their Minimization by Cutting Planes*, *Journal of Optimization Theory and Applications*, Vol. 46, pp. 37–53, 1985.
27. MARTINEZ-LEGAZ, J. E., *On Lower Subdifferentiable Functions*, *Trends in Mathematical Optimization*, Edited by K. H. Hoffman, J. B. Hiriart-Urruty, C. Lemaréchal, and J. Zowe, Birkhäuser Verlag, Basel, Switzerland, 1988.
28. FRENK, J. B. G., GROMICHO, J., PLASTRIA, F., and ZHANG, S., *A Deep-Cut Ellipsoid Algorithm and Quasiconvex Programming*, *Generalized Convexity*, Edited by S. Komlósi, T. Rapcsák, and S. Schaible, Springer Verlag, New York, New York, pp. 62–76, 1994.
29. BLAND, R. G., GOLDFARB, D., and TODD, M. J., *The Ellipsoid Method: A Survey*, *Operations Research*, Vol. 29, pp. 1039–1091, 1981.
30. GRÖTSCHEL, M., LOVÁSZ, L., and SCHRIJVER, A., *Geometric Algorithms and Combinatorial Optimization*, Springer Verlag, Berlin, Germany, 1988.
31. DZIUBAN, S. T., ECKER, J. G., and KUPFERSCHMID, M., *Using Deep Cuts in an Ellipsoid Algorithm for Nonlinear Programming*, *Mathematical Programming Study*, Vol. 25, pp. 93–107, 1985.
32. GOFFIN, J. L., *Convergence Rates of the Ellipsoid Method on General Convex Functions*, *Mathematics of Operations Research*, Vol. 8, pp. 135–150, 1983.
33. SHOR, N. Z., *Utilization of the Operation of Space Dilation in the Minimization of Convex Functions*, *Cybernetics*, Vol. 6, pp. 7–15, 1970.
34. LUTHI, H. J., *On the Solution of Variational Inequalities by the Ellipsoid Method*, *Mathematics of Operations Research*, Vol. 10, pp. 515–522, 1985.