Algorithms for the Quasiconvex Feasibility Problem

Yair Censor (yair@math.haifa.ac.il)
and
Alexander Segal (asegal@math.haifa.ac.il)

Department of Mathematics, University of Haifa
Mt. Carmel, Haifa 31905, Israel

July 25, 2004. Revised: January 9, 2005. Final Revision: January 27, 2005.

Abstract

We study the behavior of subgradient projections algorithms for the quasiconvex feasibility problem of finding a point $x^* \in R^n$ that satisfies the inequalities $f_1(x^*) \leq 0, f_2(x^*) \leq 0, \ldots, f_m(x^*) \leq 0$, where all functions are continuous and quasiconvex. We consider the consistent case when the solution set is nonempty. Since the Fenchel-Moreau subdifferential might be empty we look at different notions of the subdifferential and determine their suitability for our problem. We also determine conditions on the functions, that are needed for convergence of our algorithms. The quasiconvex functions on the left-hand side of the inequalities need not be differentiable but have to satisfy a Lipschitz or a Hölder condition.

1 Introduction

In this paper we study the behavior of iterative subgradient projections algorithms for solving systems of inequalities with continuous quasiconvex functions on the left-hand side. This problem, called the *quasiconvex feasibility* *problem* (QFP), is defined as follows. Let R^n be the n-dimensional Euclidean space, and let $f_1(x), f_2(x), \cdots, f_m(x)$ be continuous quasiconvex functions defined on R^n. The quasiconvex feasibility problem is to find a point x^*, such that $f_1(x^*) \leq 0, f_2(x^*) \leq 0, \cdots, f_m(x^*) \leq 0$. We consider the consistent case, i.e., the case when a solution exists. The notion *quasiconvex feasibility problem* was introduced by Goffin, Luo and Ye in [18], where they used cutting planes algorithms and only the differentiable case was considered there.

The *convex feasibility problem* (CFP), which is a special case of the quasiconvex feasibility problem, was well-studied in the last decades. This fundamental problem has many applications in and outside mathematics in fields such as: optimization theory (see, e.g., Polyak [33], Eremin [15], Censor and Lent [4] and Chinneck [11]), approximation theory (see, e.g., von Neumann [39], Halperin [20] and Deutsch [14]), image reconstruction from projections and computerized tomography (see, e.g., Herman [21], [22], Censor [6] – [10]) and other areas.

The algorithmic approach to solving the CFP was comprehensively investigated, see, e.g., Bauschke and Borwein [2], Censor [5], for general overviews of algorithms and, e.g., Crombez [12] and [13] for some recent results. In this study we investigate the possibilities of modifying and adapting some of these algorithmic schemes so that they become applicable to the QFP. In particular, we look at the *cyclic subgradient projections* (CSP) (Censor and Lent [4]), *parallel subgradient projections* (PSP) (dos Santos [35], [36]) and *Eremin's algorithmic scheme* [16]. The common idea of all these algorithms is to employ projections of different types, with respect to the individual level-sets of the functions, to generate a sequence of points that converges to a solution. When the functions on the left-hand side of the inequalities are quasiconvex the situation is much more complicated because such functions lack separation properties that convex functions have. Straightforward generalizations of the aforementioned algorithms are not possible because the subdifferential of Fenchel-Moreau might be empty at some points, thus, inapplicable to quasiconvex functions.

Using different notions for subdifferentials, we develop algorithms for the QFP for functions that are not necessarily differentiable, but have to satisfy a Lipschitz or a Hölder condition. In Section 2 we present preliminary material and discuss several notions of subdifferentials. In Section 3 we present and study our algorithms for solving quasiconvex feasibility problems and clarify the relation between them and existing methods for subgradient minimization. In Section 4 we present additional algorithms for the QFP, based on a class of algorithms of Eremin.

2 Background and Preliminaries

We use the books of Rockafellar [34], Hiriart-Urruty and Lemaréchal [23], as our desk-references for convex analysis. We work in the n-dimensional Euclidean space R^n where $\langle x, y\rangle$ and $\|x\|$ are the Euclidean inner product and norm, respectively. A function $f : X \to R \cup \{+\infty\}$ is a *proper function* if $\operatorname{dom}(f) := \{x \in R^n \mid f(x) < +\infty\}$ is nonempty. For any $a \in R$ the *level (*respectively*, strict level) set of* f*, corresponding to* a, is the set

$$lev_f(a) = \{x \in R^n \mid f(x) \le a\}, \tag{1}$$

respectively,

$$lev_f^{<}(a) = \{x \in R^n \mid f(x) < a\}. \tag{2}$$

Given a set $C \subseteq R^n$, we denote by $\operatorname{int} C$, $\operatorname{ri} C$, $\operatorname{cl} C$ and $\operatorname{bd} C$ its interior, relative interior, closure and boundary, respectively.

Definition 1 ***(Normal cone)*** *A* `normal cone to a set` $C \subseteq R^n$ `at a point` $z \in R^n$ *is denoted and defined by*

$$N_C(z) := \{q \in R^n \mid \langle q, y - z\rangle \le 0, \textit{ for all } y \in C\}. \tag{3}$$

Observe that this definition does not require that $z \in \operatorname{cl} C$, see, e.g., Gromicho [19, p. 15].

Definition 2 ***(Orthogonal projection)*** *Given a set* $C \subseteq R^n$ *and a point* $z \in R^n$*, an* `orthogonal projection of` z `onto` C*, denoted* $P_C(z)$*, is a point* $P_C(z) \in C$*, such that*

$$\|z - P_C(z)\| = \inf\{\|z - y\| \mid y \in C\}. \tag{4}$$

If C is nonempty, closed and convex then the projection exists and is unique, see, e.g., [23, p. 46]. The following notion of subdifferential plays an important role in convex analysis and in algorithms for solving the CFP.

Definition 3 ***(The Fenchel-Moreau subdifferential)*** *Given a function f and a point z, the* `Fenchel-Moreau (FM) subdifferential of f at z` *is defined by*

$$\partial^{FM} f(z) = \{t \in R^n \mid \langle t, x - z\rangle \leq f(x) - f(z), \text{ for all } x \in R^n\}. \quad (5)$$

Definition 4 ***(Quasiconvex function)*** *Let $f : C \to R$, where C is a nonempty convex set in R^n. The function f is said to be* `quasiconvex` *if, for all $x, y \in C$, the following inequality holds*

$$f(\theta x + (1-\theta)y) \leq \max\{f(x), f(y)\}, \text{ for all } \theta \in (0,1). \quad (6)$$

Quasiconvexity has a geometrical interpretation, indeed f is quasiconvex if and only if its level-sets $lev_f(a)$ are convex for all $a \in R$ which, in turn, is true if and only if its strict level-sets $lev_f^<(a)$ are convex for all $a \in R$. Convex functions have convex level sets (see, e.g., [34, Theorem 4.6]), and, therefore, are quasiconvex, but the converse is not true (e.g., the function $\log x$ on $(0, +\infty)$). Applications of quasiconvex functions which are not convex can be found in approximation theory (fractional programming), see, e.g., Bajona-Xandri and Martinez-Legaz [1], Boncompte and Martinez-Legaz [3], Stancu-Minasian [38], location theory, see, e.g., Gromicho [19], microeconomic theory (utility functions), see, e.g., Mas-Colell, Whinston and Green [29].

Using (2) and Definition 1 we introduce the notation for the cone

$$N_{lev_f^<}(z) := \{q \in R^n \mid \langle q, y - z\rangle \leq 0, \text{ for all } y \in lev_f^<(f(z))\} \quad (7)$$

where $f : R^n \to R$ is a quasiconvex function. This cone is never empty because it contains the origin and it follows directly from a separation argument [34, Theorem 11.3] that $N_{lev_f^<}(z)$ never reduces to the origin alone. Techniques for computing elements of the normal cone to a level-set can be found, e.g., in the recent book by Gromicho [19].

Definition 5 ***(Hölder condition)*** *A function $f : R^n \to R$ is said to satisfy the* `Hölder condition with degree` β `at a point` z `on a set` $C \subseteq R^n$ *if there exists a number $L < \infty$ and a $\beta \in (0,1]$ such that*

$$|f(y) - f(z)| \leq L \|y - z\|^{\beta}, \quad \text{for all } y \in C. \quad (8)$$

A Hölder condition can be verified by estimating the growth behavior of a function. Note that if a function satisfies a Hölder condition then it is uniformly continuous and, therefore, continuous. The Hölder condition with degree 1 is called the *Lipschitz condition.*

2.1 Various subdifferentials and their connections

For generalization of gradient methods to nondifferentiable quasiconvex functions we need to use a broader notion than the FM-subdifferential because the FM-subdifferential is usually empty even for a differentiable nonconvex function on R^n, e.g., the real-valued single variable function $y = x^3$ at $x = 0$. For functions that are not convex, concave or saddle and are not differentiable several notions of subdifferentials have been proposed in the literature. In the last thirty years there have been several attempts to define an appropriate notion of subdifferential for quasiconvex functions. The oldest one is the Greenberg-Pierskalla (GP) subdifferential [17].

Definition 6 ***(Greenberg-Pierskalla subdifferential)*** *Given a function f and a point z, the* `GP-subdifferential of` f `at` z*, is defined by*

$$\partial^{GP} f(z) = \{t \in R^n \mid \langle t, x - z\rangle \geq 0 \Longrightarrow f(x) \geq f(z)\}. \tag{9}$$

Independently from Greenberg and Pierskalla, this same notion has been introduced by Zabotin, Korablev and Khabibullin [40] under the name *generalized support*. The GP-subdifferential is often called *quasi-subdifferential*. A variation of the GP-subdifferential is the *star-subdifferential*.

Definition 7 ***(Star-subdifferential)*** *Given a function f and a point z, the* `star-subdifferential of` f `at` z*, is defined by*

$$\partial^{\star} f(z) := \begin{cases} \{t \in R^n \backslash \{0\} \mid \langle t, x - z\rangle > 0 \Longrightarrow f(x) \geq f(z)\}, & z \notin \Gamma, \\ R^n, & z \in \Gamma, \end{cases} \tag{10}$$

where Γ is the set of minimizers of f.

Obviously, $\partial^{GP} f(z)\backslash\{0\} \subseteq \partial^{\star} f(z)$. If f is quasiconvex on R^n and finite at z, then $\partial^{\star} f(z) \neq \emptyset$, see, e.g., the review paper of Penot [30, Proposition 22]. If f is continuous, then $\partial^{\star} f(z) = \partial^{GP} f(z)$, [30, Proposition 8]. Note that (10) is equivalent to

$$\partial^{\star} f(z) = \{t \in R^n \backslash \{0\} \mid f(x) < f(z) \Longrightarrow \langle t, x - z\rangle \leq 0\}. \tag{11}$$

Therefore, if f is a quasiconvex, continuous function on R^n and z is not a minimizer of f, then

$$\partial^{GP} f(z) = \partial^{\star} f(z) = N_{lev_f^{<}}(z)\backslash\{0\} \neq \emptyset. \tag{12}$$

Denoting by $S(0,1) := \{z \in R^n \mid \|z\| = 1\}$ the unit sphere, (12) guarantees that

$$S(0,1) \cap \partial^\star f(z) \neq \emptyset. \tag{13}$$

Plastria introduced and explored, in [31], properties of his lower subdifferential.

Definition 8 ***(Plastria's lower subdifferential)*** *Given a function f and a point z, the* `Plastria (P) lower subdifferential of` f `at` z *(denoted in [31] as $\partial^- f$), is defined by*

$$\partial^P f(z) = \{t \in R^n \mid f(x) < f(z) \Longrightarrow \langle x - z, t\rangle \leq f(x) - f(z)\}. \tag{14}$$

A function f is called *lower subdifferentiable* (*lsd*) on $K \subseteq R^n$ if it admits at least one P-lower subgradient at each point. It is clear that every convex function is *lsd*, since $\partial^{FM} f(z) \subseteq \partial f^P(z)$, but not conversely, as the real-valued single variable function $f(x) = |x|^{1/2}$ shows. Moreover, Plastria shows in [31] that every Lipschitzian quasiconvex function on R^n has $\partial^P f(z) \neq \emptyset$, for every $z \in R^n$.

Theorem 9 *[31] For any function f and point $z \in R^n$, $\partial^P f(z)$ is a closed convex set, and $0 \in \partial^P f(z)$ if and only if z is a global minimizer of f, in which case $\partial^P f(z) = R^n$.*

Lower subdifferentiability was investigated by Plastria and Martinez-Legaz, see, for example, [31], [32], [27], [28]. For applications of lower subdifferentiability in the field of fractional programming, see, e.g., [1], [3].

2.2 Konnov's result

In his recent work [26] Konnov considers a normalized subgradient method for minimization of quasiconvex functions which employs the stepsize rule based on a priori knowledge of the optimal value of the cost function. Konnov's algorithm is a modification of the well-known algorithm developed by Polyak [33] for convex functions. Suppose that the function f is continuous and quasiconvex. Assume that it attains its global minimum f^* on R^n and let $D^* = \text{argmin}\{f(x) \mid x \in R^n\}$. Then D^* is nonempty, closed and convex. We further assume that f satisfies the Hölder condition with constant L and degree β at a point $x^* \in D^*$. Konnov proved a somewhat extended version of the following proposition.

Proposition 10 *[26, Proposition 2.1] Suppose that the function f satisfies the Hölder condition with degree $\beta > 0$ at a point $x^* \in D^*$ on the set* $\mathrm{cl}\, lev_f^<(f(z))$ *for some point $z \in R^n \backslash D^*$. Then we have*

$$f(z) - f^* \leq L \langle t, z - x^* \rangle^\beta \text{, for all } t \in S(0,1) \cap N_{lev_f^<}(z). \tag{15}$$

3 Algorithms for the quasiconvex feasibility problem

Consider a family of sets

$$D_i = \{x \in R^n \mid f_i(x) \leq 0\} \text{ for } i = 1, 2, \ldots, m, \tag{16}$$

where all f_i are continuous and quasiconvex and let

$$D = \cap_{i=1}^m D_i \tag{17}$$

represent a *quasiconvex feasibility problem*. Our algorithms deal with quasiconvex functions satisfying a Hölder condition. Later on we use also the following property.

Definition 11 *Given a set $Q \subseteq R^n$, a sequence $\{x^k\}_{k=0}^\infty$ is* ***Fejér-monotone with respect to*** *Q if for every $x \in Q$,*

$$\left\|x^{k+1} - x\right\| \leq \left\|x^k - x\right\| \text{, for all } k \geq 0. \tag{18}$$

Some of the methods studied below use a specific control sequence. A *control sequence* $\{i(k)\}_{k=0}^\infty$ is a sequence of indices according to which individual sets D_i may be chosen for the execution of an iterative step of the algorithm.

Definition 12 ***(Control sequences)***

1. ***Almost cyclic control.*** *A control sequence $\{i(k)\}_{k=0}^\infty$ is* **almost cyclic** *on $\{1, 2, \ldots, m\}$ if $1 \leq i(k) \leq m$, for all $k \geq 0$, and there exists an integer $\sigma \geq m$ (called the* **almost cyclicality constant***) such that, for all $k \geq 0$, $\{1, 2, \ldots, m\} \subseteq \{i(k+1), i(k+2), \ldots, i(k+\sigma)\}$. An almost cyclic control with $\sigma = m$ is called* **cyclic***.*

2. ***Most violated constraint control.*** *This control sequence* $\{i(k)\}_{k=0}^{\infty}$ *is obtained by determining which constraint is most violated by the iterate* x^k. *If* $D_i = \{x \in R^n \mid f_i(x) \leq 0\}$, *are the sets in the feasibility problem then* $i(k)$ *is the* `most violated constraint control index` *if* $f_{i(k)}(x^k) > 0$ *and*

$$f_{i(k)}(x^k) = \max\{f_i(x^k) \mid i = 1, 2, \ldots, m\}. \tag{19}$$

Next we present an iterative algorithm with the most violated constraint control for solving the QFP. We denote by $g^+(x)$ the positive part $g^+(x) := \max\{0, g(x)\}$.

Algorithm 13

Initialization: $x^0 \in R^n$ *is arbitrary.*

Iterative step: *Given the current iterate* x^k, *calculate the next iterate* x^{k+1} *by*

$$x^{k+1} = x^k - \lambda_k \left(\frac{f^+_{i(k)}(x^k)}{L_{i(k)}} \right)^{1/\beta_{i(k)}} t^k, \tag{20}$$

where $t^k \in S(0,1) \cap \partial^\star f_{i(k)}(x^k)$ *and* $\beta_{i(k)}$ *and* $L_{i(k)}$ *are the Hölder constant and degree, respectively, of* $f_{i(k)}$.

Relaxation parameters: $\{\lambda_k\}_{k=0}^{\infty}$ *are confined to the interval* $\varepsilon_1 \leq \lambda_k \leq 2 - \varepsilon_2$, *for all* $k \geq 0$, *with some arbitrarily small* $\varepsilon_1, \varepsilon_2 > 0$.

Control: *Most violated constraint control.*

The convergence of this algorithm can be secured by our following theorem.

Theorem 14 *Let the following assumptions hold: (i) the functions* $f_i(x)$ *are quasiconvex on* R^n, *(ii) the problem (17) is consistent, i.e.,* $D \neq \emptyset$, *and (iii) the functions* f_i *satisfy, for every* i, *Hölder conditions with constants* L_i *and degrees* β_i, *for all* $x \in D$, *respectively, on* R^n.

Under these assumptions any sequence $\{x^k\}_{k=0}^{\infty}$, *generated by Algorithm 13, converges to a solution of the problem (17).*

Proof. Our proof consists of the following three steps:

Step 1: $\left\{x^k\right\}_{k=0}^{\infty}$ is Fejér-monotone with respect to D.

Step 2: $\lim_{k\to\infty} f_i^+(x^k) = 0$, for every *fixed* $i \in \{1, 2, \ldots, m\}$.

Step 3: $\lim_{k\to\infty} x^k = x^* \in D$.

We now proceed with the proof of each step.

Step 1: If $x^k \in D$ for some $k \geq 0$ then the iterates remain at this point and the problem is solved. Therefore, lets assume that $x^k \notin D$ for all k. Take some $x \in D$. From (20) and using the fact that $\left\|t^k\right\| = 1$ we have

$$\begin{aligned}\left\|x^{k+1} - x\right\|^2 &= \|x^k - \lambda_k \left(\frac{f_{i(k)}^+(x^k)}{L_{i(k)}}\right)^{1/\beta_{i(k)}} t^k - x\,\|^2 \\ &= \left\|x^k - x\right\|^2 - 2\lambda_k \left(\frac{f_{i(k)}^+(x^k)}{L_{i(k)}}\right)^{1/\beta_{i(k)}} \left\langle t^k, x^k - x\right\rangle \\ &+ \lambda_k^2 \left(\frac{f_{i(k)}^+(x^k)}{L_{i(k)}}\right)^{2/\beta_{i(k)}}. \end{aligned} \tag{21}$$

From assumption (iii) of the theorem follows

$$\left|f_{i(k)}(x^k) - f_{i(k)}(x)\right| \leq L_{i(k)} \left\|x^k - x\right\|^{\beta_{i(k)}}. \tag{22}$$

Then, $x^k \notin D$ and $x \in D$ imply that $f_{i(k)}^+(x^k) = f_{i(k)}(x^k)$ and $f_{i(k)}(x) \leq f_{i(k)}^+(x) = 0$, thus, we obtain $f_{i(k)}^+(x^k) \leq L_{i(k)} \left\|x^k - x\right\|^{\beta_{i(k)}}$, i.e., $f_{i(k)}^+$ also satisfies a Hölder condition at the point x with constant $L_{i(k)}$ and degree $\beta_{i(k)}$ on the level-set which is defined by $f(x^k)$. Since x is a minimizer of $f_{i(k)}^+$ we use (13) to deduce that there exists a t^k

$$t^k \in S(0,1) \cap \partial^\star f_{i(k)}(x^k), \tag{23}$$

and (15) to get

$$f_{i(k)}^+(x^k) \leq L_{i(k)} \left\langle t^k, x^k - x\right\rangle^{\beta_{i(k)}}. \tag{24}$$

Therefore, from (21) and (24), we have

$$\begin{aligned}\left\|x^{k+1} - x\right\|^2 &\leq \left\|x^k - x\right\|^2 - 2\lambda_k \left(\frac{f_{i(k)}^+(x^k)}{L_{i(k)}}\right)^{2/\beta_{i(k)}} + \lambda_k^2 \left(\frac{f_{i(k)}^+(x^k)}{L_{i(k)}}\right)^{2/\beta_{i(k)}} \\ &\leq \left\|x^k - x\right\|^2 - \lambda_k(2 - \lambda_k) \left(\frac{f_{i(k)}^+(x^k)}{L_{i(k)}}\right)^{2/\beta_{i(k)}}. \end{aligned} \tag{25}$$

The fact that $\varepsilon_1 \leq \lambda_k \leq 2 - \varepsilon_2$, for all $k > 0$, yields

$$\left\| x^{k+1} - x \right\|^2 \leq \left\| x^k - x \right\|^2 - \varepsilon_1 \varepsilon_2 \left(\frac{f_{i(k)}^+(x^k)}{L_{i(k)}} \right)^{2/\beta_{i(k)}}, \tag{26}$$

from which Fejér-monotonicity follows.

Step 2: For $x \in D$ the sequence $\left\{ \left\| x^k - x \right\| \right\}_{k=0}^{\infty}$ is monotonically decreasing and bounded below, therefore, there exist the limit $\lim_{k\to\infty} \left\| x^k - x \right\| = d$. This implies, via (26), that

$$\lim_{k\to\infty} \left(\frac{f_{i(k)}^+(x^k)}{L_{i(k)}} \right)^{2/\beta_{i(k)}} = 0, \tag{27}$$

thus,

$$\lim_{k\to\infty} f_{i(k)}^+(x^k) = 0. \tag{28}$$

Then the most violated constraint control implies that

$$\lim_{k\to\infty} f_i^+(x^k) = 0, \text{ for every } \textit{fixed } i \in \{1, 2, \ldots, m\}. \tag{29}$$

Step 3: Fejér-monotonicity of $\left\{ x^k \right\}_{k=0}^{\infty}$ with respect to D, proven in Step 1, implies boundedness. Therefore, $\left\{ x^k \right\}_{k=0}^{\infty}$ must have a convergent subsequence,

$$\lim_{s\to\infty} x^{k_s} = \widetilde{x}. \tag{30}$$

From (29) and the continuity of f_i^+ we know that $\widetilde{x} \in D$. In Step 2 we showed that $\lim_{k\to\infty} \left\| x^k - \widetilde{x} \right\| = d$, but now $\lim_{s\to\infty} \left\| x^{k_s} - \widetilde{x} \right\| = 0$, thus, $\lim_{k\to\infty} \left\| x^k - \widetilde{x} \right\| = 0$ and the proof is complete. ■

An appropriate modification allows us to formulate and prove convergence for an almost cyclically controlled version of Algorithm 13.

Algorithm 15

$\left.\begin{array}{l}\textbf{\textit{Initialization:}}\\ \textbf{\textit{Iterative step:}}\\ \textbf{\textit{Relaxation parameters:}}\end{array}\right\}$ *Same as in Algorithm 13.*

Control: *The sequence* $\{i(k)\}_{k=0}^{\infty}$ *is almost cyclic on* $\{1, 2, \ldots, m\}$.

Theorem 16 *Under the assumptions of Theorem 14, any sequence* $\{x^k\}_{k=0}^{\infty}$*, generated by Algorithm 15, converges to a solution of the problem (17).*

Proof. The proof consists of the following five steps:
Step 1: $\left\{x^k\right\}_{k=0}^{\infty}$ is Fejér-monotone with respect to D.
Step 2: $\lim_{k\to\infty} f_{i(k)}^{+}(x^k) = 0$.
Step 3: $\lim_{k\to\infty} \left\|x^{k+1} - x^k\right\| = 0$.
Step 4: $\lim_{k\to\infty} f_i^{+}(x^k) = 0$, for every *fixed* $i \in \{1, 2, \ldots, m\}$.
Step 5: $\lim_{k\to\infty} x^k = x^* \in D$.
We now proceed with the proof of each step.

Step1: The proof of this step is identical with the proof of **Step1** in Theorem 14.

Step 2: For $x \in D$ the sequence $\left\{\left\|x^k - x\right\|\right\}_{k=0}^{\infty}$ monotonically decreases and is bounded from below, thus, there exists the limit $\lim_{k\to\infty} \|x^k - x\| = d$. This implies, via (26), that

$$\lim_{k\to\infty} \left(\frac{f_{i(k)}^{+}(x^k)}{L_{i(k)}}\right)^{2/\beta_{i(k)}} = 0, \tag{31}$$

and

$$\lim_{k\to\infty} f_{i(k)}^{+}(x^k) = 0. \tag{32}$$

Step 3: By substitution from Algorithm 15 and from the fact that $\left\|t^k\right\| = 1$, for all $k \geq 0$, we get

$$\left\|x^{k+1} - x^k\right\|^2 = \lambda_k^2 \left(\frac{f_{i(k)}^{+}(x^k)}{L_{i(k)}}\right)^{2/\beta_{i(k)}}, \tag{33}$$

and the right-hand side of this equation tends to zero as $k \to \infty$, see (32). Note that this implies, by the triangle inequality, also that

$$\lim_{k\to\infty} \left\|x^{k+j} - x^k\right\| = 0 \tag{34}$$

for every integer j.

Step 4: Let $i \in \{1, 2, \ldots, m\}$ be a *fixed* index. Then, for any l,

$$\left|f_i^{+}(x^k)\right| \leq \left|f_i^{+}(x^k) - f_i^{+}(x^l)\right| + \left|f_i^{+}(x^l)\right|. \tag{35}$$

Choose now l to be the integer larger than but closest to k such that $i = i(l)$ (its existence is guaranteed by the almost cyclic control). Let us denote

$$\Gamma_{\widehat{x}} := \left\{ x \in R^n \mid \|x - \widehat{x}\| \leq \left\|x^0 - \widehat{x}\right\| \right\}.$$

The set $\Gamma_{\widehat{x}}$ is compact, therefore, $f_i^+(x)$ is uniformly continuous on it. Thus, (34) implies that $\lim_{k\to\infty} \left|f_i^+(x^k) - f_i^+(x^l)\right| = 0$. Since $i = i(l)$, (32) implies $\lim_{k\to\infty} \left|f_i^+(x^l)\right| = 0$. Thus, (35) gives the required result that

$$\lim_{k\to\infty} f_i^+(x^k) = 0 \text{ for every fixed } i \in I. \tag{36}$$

Step 5: Fejér-monotonicity of $\left\{x^k\right\}_{k=0}^{\infty}$, proven in Step 1, implies boundedness. Therefore, $\left\{x^k\right\}_{k=0}^{\infty}$ must have a convergent subsequence, i.e.,

$$\lim_{s\to\infty} x^{k_s} = \widetilde{x}. \tag{37}$$

From (36) and the continuity of f_i^+ we know that $\widetilde{x} \in D$. In Step 2 we showed that $\lim_{k\to\infty} \left\|x^k - \widetilde{x}\right\| = d$, but now we have the additional information that $\lim_{s\to\infty} \left\|x^{k_s} - \widetilde{x}\right\| = 0$; Thus, $\lim_{k\to\infty} \left\|x^k - \widetilde{x}\right\| = 0$ and the proof is complete. ■

Now we present a parallel algorithm for solving the QFP.

Algorithm 17

Initialization: *$x^0 \in R^n$ is arbitrary.*

Iterative step: *Given the current iterate x^k, calculate the next iterate x^{k+1} by*

$$x^{k+1} = x^k - \lambda_k \sum_{i=1}^{m} \alpha_i \left(\frac{f_i^+(x^k)}{L_i} \right)^{1/\beta_i} t^{i,k}, \tag{38}$$

where $t^{i,k} \in S(0,1) \cap \partial^\star f_i(x^k)$, and $0 < \alpha_i < 1$, for all i, and $\sum_{i=1}^{m} \alpha_i = 1$. The β_i and L_i are the Hölder constants and degrees, respectively, of f_i.

Relaxation parameters: *$\{\lambda_k\}_{k=0}^{\infty}$ are confined to the interval $\varepsilon_1 \leq \lambda_k \leq 2 - \varepsilon_2$, for all $k \geq 0$ with some arbitrary small $\varepsilon_1, \varepsilon_2 > 0$.*

Theorem 18 *Under the assumptions of Theorem 14, any sequence $\{x^k\}_{k=0}^{\infty}$, generated by Algorithm 17, converges to a solution of the problem (17).*

Proof. The proof consists of the following three steps:
Step 1: $\left\{x^k\right\}_{k=0}^{\infty}$ is Fejér-monotone with respect to D.
Step 2: $\lim_{k\to\infty} f_i^+(x^k) = 0$ for all $i \in \{1, 2, \ldots, m\}$.
Step 3: $\lim_{k\to\infty} x^k = x^* \in D$.
We proceed with the proof of each step.

Step 1: Take some $x \in D$. From (38) and using the fact that $\|t^{i,k}\| = 1$, we have

$$\begin{aligned}\left\|x^{k+1} - x\right\|^2 &= \|x^k - \lambda_k \sum_{i=1}^m \alpha_i \left(\frac{f_i^+(x^k)}{L_i}\right)^{1/\beta_i} t^{i,k} - x\|^2 \\ &= \left\|x^k - x\right\|^2 + \lambda_k^2 \|\sum_{i=1}^m \alpha_i \left(\frac{f_i^+(x^k)}{L_i}\right)^{1/\beta_i} t^{i,k}\|^2 \\ &- 2\lambda_k \left\langle \sum_{i=1}^m \alpha_i \left(\frac{f_i^+(x^k)}{L_i}\right)^{1/\beta_i} t^{i,k} \ , \ x^k - x \right\rangle. \end{aligned} \tag{39}$$

By an argument similar to the one given in the discussion of (24) in the proof of Theorem 14 we obtain

$$f_i^+(x^k) \le L_i \left\langle t^{i,k}, x^k - x\right\rangle^{\beta_i}, \tag{40}$$

hence,

$$\begin{aligned}\left\|x^{k+1} - x\right\|^2 &\le \left\|x^k - x\right\|^2 + \lambda_k^2 \|\sum_{i=1}^m \alpha_i \left(\frac{f_i^+(x^k)}{L_i}\right)^{1/\beta_i} t^{i,k}\|^2 \\ &- 2\lambda_k \sum_{i=1}^m \alpha_i \left(\frac{f_i^+(x^k)}{L_i}\right)^{2/\beta_i}. \end{aligned} \tag{41}$$

Due to the convexity of $\|\cdot\|^2$ we have

$$\begin{aligned}\left\|x^{k+1} - x\right\|^2 &\le \left\|x^k - x\right\|^2 + \lambda_k^2 \sum_{i=1}^m \alpha_i \left(\frac{f_i^+(x^k)}{L_i}\right)^{2/\beta_i} \\ &- 2\lambda_k \sum_{i=1}^m \alpha_i \left(\frac{f_i^+(x^k)}{L_i}\right)^{2/\beta_i} \end{aligned} \tag{42}$$

$$\le \left\|x^k - x\right\|^2 + (\lambda_k^2 - 2\lambda_k) \sum_{i=1}^m \alpha_i \left(\frac{f_i^+(x^k)}{L_i}\right)^{2/\beta_i}. \tag{43}$$

From $\varepsilon_1 \leq \lambda_k \leq 2 - \varepsilon_2$, for all $k > 0$, we get

$$\left\|x^{k+1} - x\right\|^2 \leq \left\|x^k - x\right\|^2 - \varepsilon_1\varepsilon_2 \sum_{i=1}^{m} \alpha_i \left(\frac{f_i^+(x^k)}{L_i}\right)^{2/\beta_i}, \tag{44}$$

and the sum in the right-hand side is positive, thus, Fejér-monotonicity follows.

Step 2: For $x \in D$ the sequence $\left\{\left\|x^k - x\right\|\right\}_{k=0}^{\infty}$ is monotonically decreasing and bounded below, therefore, there exists the limit $\lim_{k\to\infty} \left\|x^k - x\right\| = d$. This implies, via (44), that

$$\lim_{k\to\infty} \left(\frac{f_i^+(x^k)}{L_i}\right)^{2/\beta_i} = 0, \tag{45}$$

and

$$\lim_{k\to\infty} f_i^+(x^k) = 0 \text{ for all } i \in \{1, 2, \ldots, m\}. \tag{46}$$

Step 3: Similar to the proof of Step 3 in Theorem 16. ■

It is interesting to note the relation between our algorithms for the QFP and existing results. In [24], Kiwiel studies methods for subgradient minimization of quasiconvex functions that employ a variety of subdifferentials. His subdifferentials include those that we use although the notations are different. How do those results relate to the work presented here? A basic tool for deriving feasibility algorithms from minimization methods is (see, e.g., Shor [37, p. 39]) to define $\Psi(x) := \max\{f_i(x) \mid i = 1, 2, \ldots, m\}$ and apply a subgradient minimization method to the function $\Psi^+(x)$. Doing so with Kiwiel's subgradient minimization methods of [24] generates subgradient algorithms for the QFP which differ in a fundamental way from the algorithms that we present here. In our algorithms there appear parameters $\{\lambda_k\}_{k=0}^{\infty}$, called relaxation parameters, that are confined to the interval $\varepsilon_1 \leq \lambda_k \leq 2-\varepsilon_2$, for all $k \geq 0$ with some arbitrary small $\varepsilon_1, \varepsilon_2 > 0$. Except for this restriction to the interval, these parameters are free and commonly user-chosen. Their practical significance in experimental work with algorithms for convex feasibility problems cannot be exaggerated, see, e.g., Censor and Herman [9, Section 6]. In the subgradient minimization algorithms of Kiwiel in [24], there appear instead of relaxation parameters, other quantities, called there "standard divergent-series stepsizes" which must fullfil the conditions that $\lambda_k > 0$, for all $k \geq 0$, $\lim_{k\to\infty} \lambda_k = 0$ and $\sum_{k=0}^{\infty} \lambda_k = +\infty$. Therefore, such parameters will also appear in any algorithm for the QFP derived from Kiwiel's minimization algorithm of [24].

4 Algorithms for solving systems of inequalities with quasiconvex Lipschitz continuous functions on the left-hand side

In this section we extend the validity of the class of Eremin's algorithms to the QFP with quasiconvex Lipschitz continuous functions on the left-hand side. We present first a number of useful facts about Plastria's P-lower subdifferential, which we employ here. Penot showed, in [30, Proposition 12], that if f is Lipschitz continuous on R^n then

$$\begin{aligned} \partial^{GP} f(z) &= \cup_{\lambda \geq 0} \lambda \partial^P f(z) \\ &= \cup_{\alpha \in [0,1]} \alpha \partial^P f(z). \end{aligned} \tag{47}$$

One says, see Plastria [31], that f is *boundedly lower subdifferentiable (blsd)* on a set $\Theta \subseteq R^n$, if at each point of Θ there exists a lower subgradient of f of norm not exceeding a constant L. The constant L is called a *blsd-bound* of f. The following theorem and its proof are useful for our further discussion.

Theorem 19 *[31, Theorem 2.3] Every quasiconvex function f on R^n that satisfies a Lipschitz condition with constant L is blsd on R^n with blsd-bound L.*

This theorem guarantees the nonemptiness of P-lower subdifferentials. The following theorem relates the P-lower subdifferential of a quasiconvex function which satisfies a Lipschitz condition to the normal cone of level sets, and, therefore, serves as our tool for calculating P-lower subgradients. Its proof is inspired by the proof of [31, Theorem 2.3].

Theorem 20 *Let f be a quasiconvex function on R^n that satisfies a Lipschitz condition with constant L and suppose that $z \in R^n$. Then, for all $v \in S(0,1) \cap N_{lev_f^<}(z)$, the vector*

$$\tilde{v} = Lv \tag{48}$$

belongs to $\partial^P f(z)$*. Additionally, for all* $u \in \partial^P f(z) \backslash \{0\}$*, the vector*

$$\tilde{u} = L \frac{u}{\|u\|} \tag{49}$$

belongs to $\partial^P f(z)$*.*

Proof. Let $z \in R^n$. Since f is Lipschitz, it is continuous, and $lev_f^<(z)$ is an open convex set not containing z. Then there exists a vector $v \in S(0,1) \cap N_{lev_f^<}(z)$. Set $\tilde{v} = Lv$. For any $y \in lev_f^<(z)$, let $P(y)$ be the orthogonal projection of y on the hyperplane

$$\{x \in R^n \mid \langle x - z, \tilde{v} \rangle = 0\} \tag{50}$$

that passes through z and is perpendicular to $\tilde{v}$. Then $P(y) \notin lev_f^<(z)$ thus, $f(P(y)) \geq f(z)$. Furthermore, since $P(y) - y$ is co-linear with $\tilde{v}$ and the latter has length L, we have

$$\langle z - y, \tilde{v} \rangle = \|P(y) - y\| \cdot \|\tilde{v}\| = L \|P(y) - y\|. \tag{51}$$

Thus, from Lipschitzity and (51), we have

$$f(z) - f(y) \leq f(P(y)) - f(y) \leq L \|P(y) - y\| = \langle z - y, \tilde{v} \rangle, \tag{52}$$

showing that $\tilde{v} \in \partial^P f(z)$. The additional assertion follows immediately from (47). See Figure 1 for a geometrical illustration. ■

Now we are ready to study Eremin's algorithms for solving systems of inequalities with quasiconvex Lipschitz continuous functions $\{f_i\}_{i=1}^m$ on the left-hand side. Assume that $\{K_i\}_{i=1}^m$ is a set of real positive numbers and let $I(x) = \{j \mid \max\{f_i(x) \mid i = 1, 2, \ldots, m\} = f_j(x)\}$ and $s(x) = \{i \mid f_i(x) > 0\}$. The following definition was given by Eremin [16].

Definition 21 *Let* $D \subseteq R^n$ *be a closed convex set, let* $d(x)$ *be a continuous real-valued function, defined on* R^n*, that satisfies* $\{x \mid d(x) \leq 0\} = D$*. Let* $e(x)$ *be a vector-valued function that is defined and nowhere equal to zero on* $R^n \backslash D$*. Assume also that* $e(x)$ *is bounded on any bounded set. Such a pair of functions* $d(x)$ *and* $e(x)$ *is said to have the* **`d-e property`** *if for arbitrary* $z \notin D$ *the half-space*

$$\Omega = \{x \in R^n \mid \langle e(z), x - z \rangle + d(z) \leq 0\} \tag{53}$$

contains D*.*

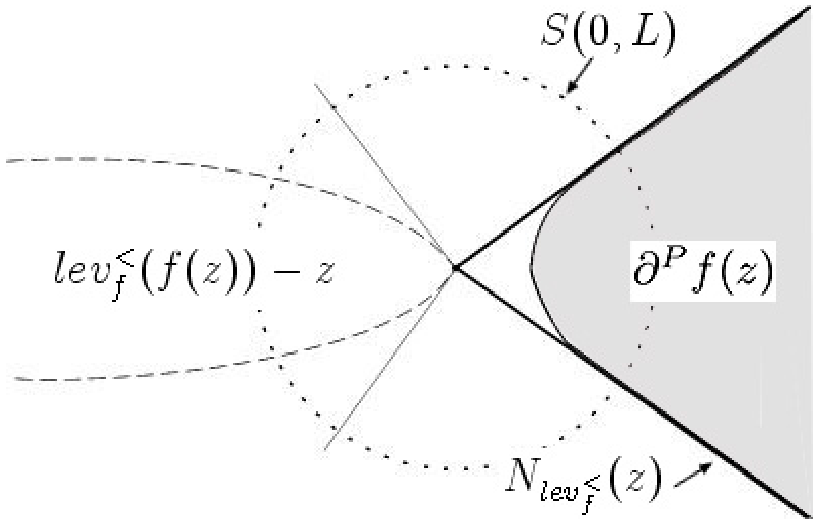

Figure 1: Geometric illustration of the theorem.

Algorithm 22 ***(Eremin's algorithmic scheme)***
Initialization: *$x^0 \in R^n$ is arbitrary.*
Iterative step: *Given x^k, calculate the next iterate x^{k+1} from*

$$x^{k+1} = \begin{cases} x^k - \lambda_k \dfrac{d(x^k)}{\parallel e(x^k) \parallel^2} e(x^k), & \text{if } d(x^k) > 0, \\ x^k, & \text{if } d(x^k) \leq 0, \end{cases} \tag{54}$$

where the pair $d(x)$ and $e(x)$ are user-chosen functions that have the d-e property.

Relaxation parameters: *$\{\lambda_k\}_{k=0}^{\infty}$ are confined to the interval $\varepsilon_1 \leq \lambda_k \leq 2 - \varepsilon_2$, for all $k \geq 0$ with some arbitrary small $\varepsilon_1, \varepsilon_2 > 0$.*

While Eremin discussed this algorithmic scheme only for convex and differentiable functions we are able to extend the scope of convergence, as the following theorem shows.

Theorem 23 *Let the following assumptions hold*
(i) the functions $f_i(x)$ are quasiconvex and Lipschitz continuous with Lipschitz constants L_i on R^n, for all $i \in \{1, 2, \dots, m\}$,
(ii) the problem (17) is consistent, i.e., $D \neq \emptyset$,

Then any sequence $\{x^k\}_{k=0}^{\infty}$, generated by Algorithm 22, converges to a point $x^* \in D$, if the pairs of functions $d(x)$ and $e(x)$ are chosen by one of the following methods.

Method 1:

$$d(x) = f_j(x) \quad \text{and} \quad e(x) = L_j \frac{t^j}{\| t^j \|}, \tag{55}$$

where $t^j \in \partial^P f_j(x)$ and j is any index from $I(x)$.

Method 2:

$$d(x) = \begin{cases} \sum_{i \in s(x)} K_i f_i(x), & \text{if } s(x) \neq \emptyset, \\ 0, & \text{if } s(x) = \emptyset, \end{cases} \tag{56}$$

and

$$e(x) = \sum_{i \in s(x)} K_i L_i \frac{t^i}{\| t^i \|}, \tag{57}$$

where $t^i \in \partial^P f_i(x)$ for all $i \in \{1, 2, \dots, m\}$.

Method 3:

$$d(x) = \begin{cases} \sum_{i \in s(x)} f_i^2(x), & \text{if } s(x) \neq \emptyset, \\ 0, & \text{if } s(x) = \emptyset, \end{cases} \tag{58}$$

and

$$e(x) = \sum_{i \in s(x)} L_i f_i(x) \frac{t^i}{\| t^i \|}, \tag{59}$$

where $t^i \in \partial^P f_i(x)$ for all $i \in \{1, 2, \dots, m\}$.

Proof. The above should have been phrased in the language of multivalued functions and selectors. Recall that a *multivalued function* from R^n to itself is a function $\mathcal{F} : R^n \to \mathcal{P}$, where $\mathcal{P}$ is the power set of R^n (consisting of all subsets of R^n) and a continuous function $t(x)$ from R^n to itself is a *selector* for $\mathcal{F}$ if $t(x) \in \mathcal{F}(x)$, for all $x \in R^n$. Using this language, in Method 2, for example, $\Phi(x) = \partial^P f_i(x)$ is a multivalued function and $t(x)$ is a selector of $\Phi(x)$. However, for simplicity we do not use the language of selectors here because there is no ambiguity. To prove the theorem we show that all $d(x)$ and

$e(x)$ pairs have the d-e property for $D = \cap_{i\in I} D_i$. Then the required result will follow from Eremin's theorem which states that if $d(x)$ and $e(x)$ have the d-e property then convergence of his algorithm is achieved, see [16, Lemma 2]. In all three methods $d(x)$ is a continuous real-valued function defined on R^n and satisfying $\{x \mid d(x) \le 0\} = D$. From Theorem 19 it follows that $e(x)$ is well-defined, from Theorem 9 we know that it is nowhere equal to zero on $R^n \backslash D$ and, by its construction, we know that it is bounded on any bounded set. Suppose that $z \notin D$ and $\Omega = \{x \in R^n \mid \langle e(z), x - z\rangle + d(z) \le 0\}$. We must verify the inclusion $D \subseteq \Omega$ for each of the three methods of choosing $d(x)$ and $e(x)$. From $z \notin D$ and $y \in D$ we obtain that $f(y) < f(z)$, where f stands for f_j, with $j \in I(z)$ in the case of Method 1, and for f_i, with $i \in s(z)$, for all other cases. Therefore, we can make use of Definition 8 and Theorem 20 in considering all three methods. Indeed, in Method 1

$$\begin{aligned}\langle e(z), y - z\rangle + d(z) &= \left\langle L_j \frac{t^j}{\| t^j \|}, y - z \right\rangle + f_j(z) \\ &\le f_j(y) \le 0, \text{ where } j \in I(z).\end{aligned} \tag{60}$$

In Method 2

$$\begin{aligned}\langle e(z), y - z\rangle + d(z) &= \left\langle \sum_{i\in s(z)} K_i L_i \frac{t^i}{\| t^i \|}, y - z \right\rangle + \sum_{i\in s(z)} K_i f_i(z) \\ &= \sum_{i\in s(z)} K_i \left(\left\langle L_i \frac{t^i}{\| t^i \|}, y - z \right\rangle + f_i(z) \right) \\ &\le \sum_{i\in s(z)} K_i f_i(y) \le 0.\end{aligned} \tag{61}$$

In Method 3

$$\begin{aligned}\langle e(z), y - z\rangle + d(z) &= \left\langle \sum_{i\in s(z)} L_i f_i(z) \frac{t^i}{\| t^i \|}, y - z \right\rangle + \sum_{i\in s(z)} f_i^2(z) \\ &= \sum_{i\in s(z)} f_i(z) \left(\left\langle L_i \frac{t^i}{\| t^i \|}, y - z \right\rangle + f_i(z) \right) \\ &\le \sum_{i\in s(z)} f_i(z) f_i(y) \le 0,\end{aligned} \tag{62}$$

and the proof is complete. ■

Acknowledgments. We are grateful to Prof. Adi Ben-Israel from Rutgers Center for Operations Research, Rutgers, The State University of New Jersey, and Prof. Dan Butnariu from the Department of Mathematics at the University of Haifa for their useful comments. The insightful reports of two anonymous referees helped us to improve the first version of this paper. This research was partially supported by grant No. 522/04 of the Israel Science Foundation (ISF), by NIH grant No. HL70472, and by grant No. 2003275 of the United States-Israel Binational Science Foundation (BSF) and partially done at the Center for Computational Mathematics and Scientific Computation (CCMSC) at the University of Haifa.

References

[1] C. Bajona-Xandri and J.E. Martinez-Legaz, Lower subdifferentiability in minimax fractional programming, *Optimization*, **45** (1999), 1–12.

[2] H.H. Bauschke, J.M. Borwein, On projection algorithms for solving convex feasibility problems, *SIAM Review,* **38** (1996), 367–426.

[3] M. Boncompte and J.E. Martinez-Legaz, Fractional programming by lower subdifferentiability techniques, *Journal of Optimization Theory and Applications*, **68** (1991), 95–116.

[4] Y. Censor and A. Lent, Cyclic subgradient projections, *Mathematical Programming*, **24** (1982), 233–235.

[5] Y. Censor, Iterative methods for the convex feasibility problem, in: M. Rosenfeld and J. Zaks (Editors), *Convexity and Graph Theory,* Elsevier Science Publishers, Amsterdam, The Netherlands, 1984, pp. 83–91.

[6] Y. Censor, Row-action methods for huge and sparse systems and their applications, *SIAM Review*, **23** (1981), 444–466.

[7] Y. Censor, Parallel application of block iterative methods in medical imaging and radiation therapy, *Mathematical Programming*, **42** (1988), 307–325.

[8] Y. Censor, On variable block algebraic reconstruction techniques, in: G.T. Herman, A.K. Louis and F. Natterer (Editors), *Mathematical Methods in Tomography*, Lecture Notes in Mathematics, Vol. **1497,** Springer-Verlag, Berlin, Heidelberg, Germany, 1991, pp. 133–140.

[9] Y. Censor and G.T. Herman, Block-iterative algorithms with underrelaxed Bregman projections, *SIAM Journal on Optimizations*, **13**, (2002), 283–297.

[10] Y. Censor and S. A. Zenios , *Parallel Optimization: Theory, Algorithms, and Applications*, Oxford University Press, New York, NY, USA, 1997.

[11] J.W. Chinneck, The constraint consensus method for finding approximately feasible points in nonlinear programs, *INFORMS Journal on Computing,* **16** (2004), 255–265.

[12] G. Crombez, Non-monotoneous parallel iteration for solving convex feasibility problems, *Kybernetika*, **39** (2003), 547–560.

[13] G. Crombez, Finding common fixed points of a class of paracontractions, *Acta Mathematica Hungarica*, **103** (2004), 233–241.

[14] F. Deutsch, The method of alternating orthogonal projections, in: S.P. Singh (Editor), *Approximation Theory, Spline Functions and Applications*, Kluwer Academic publishers, Dordrecht, The Netherlands, 1992, pp. 105–121.

[15] I.I. Eremin, Fejér mappings and convex programming, *Siberian Mathematical Journal*, **10** (1969), 762–772.

[16] I.I. Eremin, On systems of inequalities with convex functions in the left sides, *American Mathematical Society Translations*, **88** (1970), 67–83.

[17] H.P. Greenberg and W.P. Pierskalla, Quasi-conjugate functions and surrogate duality, *Cahiers Centre Études Recherche Opér.*, **15** (1973), 437–448.

[18] J.-L. Goffin, Z.Q. Luo and Y. Ye, On the complexity of a column generation algorithm for convex or quasiconvex feasibility problems, in: *Large Scale Optimization: State of the Art*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1994, pp. 182–191.

[19] J. Gromicho, *Quasiconvex Optimization and Location Theory*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

[20] I. Halperin, The product of projection operators, *Acta Scientiarum Mathematicarum (Szeged)*, **23** (1962), 96–99.

[21] G.T. Herman, *Image Reconstruction From Projections: The Fundamentals of Computerized Tomography*, Academic Press, New York, NY, USA, 1980.

[22] G.T. Herman and L.B. Meyer, Algebraic reconstruction techniques can be made computationally efficient, *IEEE Transactions on Medical Imaging*, **MI-12** (1993), 600–609.

[23] J.-B. Hiriart-Urruty and C. Lemaréchal, *Fundamentals of Convex Analysis*, Springer-Verlag, Berlin, Heidelberg, Germany, 1993.

[24] K.C. Kiwiel, Convergence and efficiency of subgradient methods for quasiconvex minimization, *Mathematical Programming, Series A*, **90** (2001), 1–25.

[25] I. Konnov, Properties of support and quasisupport vectors, *Issledovania Prikladnoi Matematiki,* **17** (1990), 50–57 (in Russian).

[26] I. Konnov, On convergent properties of a subgradient method, *Optimization Methods and Software*, **17** (2003), 53–62

[27] J.E. Martinez-Legaz, On lower subdifferentiable functions, in: K.H. Hoffman, J.-B. Hiriart-Urruty, C. Lemaréchal and J. Zowe (Editors), *Trends in Mathematical Optimization,* Birkhäuser-Verlag, Basel, Switzerland, 1988, pp. 197–232.

[28] J.E. Martinez-Legaz, Lower subdifferentiability of quadratic functions, *Mathematical Programming*, **60** (1991), 93–113.

[29] A. Mas-Colell, M.D. Whinston and J.R. Green, *Microeconomic Theory,* Oxford University Press, New York, NY, USA, 1995.

[30] J.-P. Penot, Are generalized derivatives useful for generalized convex functions? in: J.-P. Crouzeix, J.E. Martinez-Legaz and M. Volle (Editors), *Generalized convexity, Generalized Monotonicity: Recent Results*,

Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998, pp. 3–59.

[31] F. Plastria, Lower subdifferentiable functions and their minimization by cutting planes, *Journal of Optimization Theory and Applications*, **46** (1985), 37–53.

[32] F. Plastria, The minimization of lower subdifferentiable functions under nonlinear constraints: An all feasible cutting plane algorithm, *Journal of Optimization Theory and Applications*, **57** (1988), 463–484.

[33] B.T. Polyak, Minimization of unsmooth functionals, *USSR Computational Mathematics and Mathematical Physics*, **9** (1969), 14–29.

[34] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, USA, 1970.

[35] L.T. Santos, *Metodos de Projecao do Subgradiente para o Problema de Factibilidade Convexa*, Tese de Mestrado, IMECC-UNICAMP, State University of Campinas, Campinas, SP, Brazil, (1985).

[36] L.T. Santos, A parallel subgradient projections method for the convex feasibility problem, *Journal of Computational and Applied Mathematics*, **18** (1987), 307–320.

[37] N.Z. Shor, *Minimization Methods for Non-Differentiable Functions*, Springer-Verlag, Berlin, Heidelberg, Germany, 1985.

[38] I.M. Stancu-Minasian, *Fractional Programming: Theory, Methods, and Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1997.

[39] J. von Neumann, *Functional Operators, Vol. 2: The Geometry of Orthogonal Spaces*, Princeton University Press, Princeton, NJ, USA, 1950.

[40] Y.I. Zabotin, A.I. Korablev and R.F. Khabibullin, On optimization of quasiconvex functional, *Izvestia VUZov: Matematika*, **10** (1972), 23–27 (in Russian).