Discussiones Mathematicae
Differential Inclusions, Control and Optimization 27 (2007) 71–93

INTERIOR PROXIMAL METHOD FOR VARIATIONAL INEQUALITIES ON NON-POLYHEDRAL SETS

Alexander Kaplan and Rainer Tichatschke

Department of Mathematics
University of Trier, 54286 Trier, Germany

Abstract

Interior proximal methods for variational inequalities are, in fact, designed to handle problems on polyhedral convex sets or balls, only. Using a slightly modified concept of Bregman functions, we suggest an interior proximal method for solving variational inequalities (with maximal monotone operators) on convex, in general non-polyhedral sets, including in particular the case in which the set is described by a system of linear as well as strictly convex constraints. The convergence analysis of the method studied admits the use of the ϵ-enlargement of the operator and an inexact solution of the subproblems.

Keywords: variational inequalities, Bregman function, proximal algorithm.

2000 Mathematics Subject Classification: 47J20, 47H05, 65J20, 65K10, 90C25.

1. Introduction

The proximal point method (PPM), originally developed by Martinet [23] to solve convex optimization problems and further investigated in a more general context (for finding zeros of a maximal monotone operator) by Rockafellar [29], has initiated a great number of new algorithms for various classes of variational inequalities and related problems.

For the variational inequality

$$\mathbf{VI}(\mathcal{Q},\mathbf{K}) \qquad find\; x \in K,\; q \in \mathcal{Q}(x): \quad \langle q, y - x\rangle \geq 0, \quad \forall\, y \in K,$$

where K is a convex, closed subset of $\mathbb{R}^n$, $\mathcal{Q} : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ is a maximal monotone operator and $\langle \cdot, \cdot \rangle$ stands for the inner product in $\mathbb{R}^n$, the application of the exact PPM can be described as follows.

Exact proximal point method:
Given $v^1 \in K$ and a sequence $\{\chi_k\}$, $0 < \chi_k \leq \bar{\chi} < \infty$. With $v^k \in K$ obtained in the previous step, define $v^{k+1} \in K$, $q^{k+1} \in \mathcal{Q}(v^{k+1})$ such that

$$\langle q^{k+1} + \chi_k \nabla_1 D(v^{k+1}, v^k), v - v^{k+1} \rangle \geq 0 \quad \forall\, v \in K.$$

Here $D : (x, y) \mapsto \|x - y\|^2$ and $\nabla_1 D$ denotes the partial gradient of D with respect to the first vector argument.

For different modifications of the PPM, also with other quadratic distance functions D, we refer to [16, 17, 21] and the references therein. In fact, the PPM can be viewed as a regularization method. According to the convergence results, the parameter χ_k does not necessarily have to tend to 0, and this ensures a much better numerical stability of regularized problems than in classical regularization methods.

In the last decade, a new branch in PPM's which deals with the use of non-quadratic distance functions has been extensively studied. The main motivation for this type of proximal methods is the following: for certain classes of problems an appropriate choice of non-quadratic distance functions permits one to preserve regularizing properties of the original version of the PPM and at the same time this choice guarantees that the iterates stay in the interior of the set K, i.e., with certain precautions, regularized problems can be treated as unconstrained ones.

Usually, a Bregman distance, an entropic φ-divergence or a logarithmic-quadratic distance are applied to construct such *interior proximal methods* (see references in [10, 21, 31] and the recent papers [2, 3, 20]). However, up to now, distance functions providing an *interior point effect* have been created only for the case where K is a polyhedral convex set or a ball.

In the present paper, using a slightly modified definition of the class of Bregman functions (see Remark 1), we extend the Bregman-function-based interior proximal methods to solve $VI(\mathcal{Q}, K)$ with

$$K := \{x \in \mathbb{R}^n : g_i(x) \leq 0, \quad i \in I\},$$

under the assumptions that g_i $(i \in I)$ are convex, continuously differentiable functions and the function

$$\max\{g_i: \ g_i \ \text{ not affine}\}$$

is strictly convex on K. The operator $\mathcal{Q}$ is supposed to be maximal monotone and paramonotone on K.

The modification in the definition of Bregman functions consists mainly in a relaxation of the standard convergence sensing condition (Stand-B4 in Remark 1) which proves to be restrictive already when K is a ball.

The convergence analysis of the method studied admits a successive approximation of the operator $\mathcal{Q}$ by means of the ϵ-enlargement concept, as well as an inexact solution of the subproblems under a criterion of the summability of the error vectors (Section 2).

In Section 3, we check the validity of the modified requirements on Bregman functions for some particular functions and discuss the Bregman-function-based methods developed earlier under the provision of modification mentioned above. In Appendix, two examples are given clarifying the relaxation of the standard convergence sensing condition.

2. Interior proximal method

The variational inequality $VI(\mathcal{Q}, K)$ is considered under the following basic assumptions:

- **A1** $K \subset I\!R^n$ *is a convex closed set,* $\mathcal{Q}: I\!R^n \to 2^{I\!R^n}$ *is a maximal monotone operator*;
- **A2** $\mathrm{dom}\mathcal{Q} \cap \mathrm{int}K \neq \emptyset$;
- **A3** *solution set* $SOL(\mathcal{Q}, K)$ *of* $VI(\mathcal{Q}, K)$ *is non-empty.*

Assumptions A1 and A2 provide the maximal monotonicity of $\mathcal{Q} + \mathcal{N}_K$, where $\mathcal{N}_K$ is the normality operator of K (cf. [28], Theorem 1).

Let f be a Bregman-type function with zone $\mathrm{int}K$. According to the terminology of Bregman functions, under a zone of f we mean an open convex set int $\mathrm{dom}f$. More precisely, we suppose:

- **B1** $\mathrm{dom}f = K$ *and* f *is continuous and strictly convex on* K;
- **B2** f *is continuously differentiable on* $\mathrm{int}K$;

B3 *With a generalized distance function*

$$D_f : (x,y) \in K \times \mathrm{int}K \mapsto f(x) - f(y) - \langle \nabla f(y), x-y \rangle,$$

for each $x \in K$ there exist constants $\alpha(x) > 0$, $c(x)$ such that

$$D_f(x,y) + c(x) \geq \alpha(x)\|x-y\|, \qquad \forall\, y \in \mathrm{int}K;$$

B4 *If $\{z^k\} \subset \mathrm{int}K$ converges to z, then at least one of the following properties is valid:*

(i) $\lim\limits_{k\to\infty} D_f(z,z^k) = 0$ *or*

(ii) $\overline{\lim}_{k\to\infty} D_f(\bar{z},z^k) = +\infty$ *if $\bar{z} \neq z$, $\bar{z} \in \mathrm{bd}K$;*

B5 (*zone coercivity*) $\nabla f(\mathrm{int}K) = I\!R^n$.

Remark 1. Assumption B4 is evidently weaker than the standard convergence sensing condition:

Stand-B4 If $\{z^k\} \subset \mathrm{int}K$ converges to z, then $\lim\limits_{k\to\infty} D_f(z,z^k) = 0$,

and it is closely related to condition (iv) in the definition of a generalized Bregman function introduced by Kiwiel, see [21], Definition 2.4. The other standard condition

Stand-B3 For any $x \in K$ and any constant α the set
$L^\alpha(x) := \{y \in \mathrm{int}K : D_f(x,y) \leq \alpha\}$ is bounded,

follows immediately from B3. Both B3 and Stand-B3 are obviously valid if K is a bounded set. $\diamond$

Referring in the sequel to *standard conditions on Bregman functions*, we mean the fulfillment of B1, B2, Stand-B3, Stand-B4, and one of the conditions B5 or

B6 (boundary coercivity)
If $\{z^k\} \subset \mathrm{int}K$, $\lim_{k\to\infty} z^k = z \in \mathrm{bd}K$, then it holds for each $x \in \mathrm{int}K$

$$\lim_{k\to\infty} \left\langle \nabla f(z^k), x - z^k \right\rangle = -\infty.$$

Remark 2. To our knowledge, among the Bregman functions (under standard conditions), considered in the literature (see [11, 12, 14] and references therein), only the function

$$\sum_{j=1}^{n} \left(x_j - x_j^{\beta}\right), \quad 0 < \beta < 1$$

on $K = I\!R^n_+$ does not satisfy B3. Assumption B3 was introduced in [20] in order to weaken the stopping criteria in generalized proximal methods. $\diamond$

In the sequel, we make use of

Lemma 1 ([30], Theorem 2.4)**.** *Let f satisfy conditions* B1 *and* B2. *If*

$$\{z^k\} \subset K, \quad \{y^k\} \subset \mathrm{int} K, \quad \lim_{k\to\infty} D_f(z^k, y^k) = 0$$

and one of these sequences converges, then the other one converges to the same limit, too.

Lemma 2. *Let conditions* B1, B2, Stand-B3 *and* B4 *be valid and* $\{v^k\} \subset \mathrm{int} K$*. Moreover, suppose that C is a non-empty subset of K, $\{D_f(x, v^k)\}$ converges for each $x \in C$, and each cluster point of $\{v^k\}$ belongs to C. Then $\{v^k\}$ converges to some $v \in C$.*

Proof. Since $\{D_f(x, v^k)\}$ converges for $x \in C$ and Stand-B3 is valid, the sequence $\{v^k\}$ is bounded. Take two subsequences $\{v^{l_k}\}$ and $\{v^{n_k}\}$ of $\{v^k\}$ with

$$\lim_{k\to\infty} v^{l_k} = v^*, \qquad \lim_{k\to\infty} v^{n_k} = \bar{v}.$$

If $\bar{v} \in \mathrm{int} K$, then $\lim_{k\to\infty} D_f(\bar{v}, v^{n_k}) = 0$ follows from B2, and since the whole sequence $\{D_f(\bar{v}, v^k)\}$ converges, one gets

$$\lim_{k\to\infty} D_f(\bar{v}, v^k) = 0.$$

In turn, applying Lemma 1 with $z^k := \bar{v}$, $y^k := v^{l_k}$, we conclude $v^* = \bar{v}$.

Now let $\bar{v} \in \mathrm{bd} K$. We make use of B4 setting $z := v^*$, $z^k := v^{l_k}$ and $\bar{z} := \bar{v}$. If B4(i) is valid, then $\lim_{k\to\infty} D_f(v^*, v^{l_k}) = 0$, hence

$$\lim_{k\to\infty} D_f(v^*, v^k) = 0.$$

Again, Lemma 1 with $z^k := v^*$, $y^k := v^{n_k}$, yields $v^* = \bar{v}$. But if B4(ii) holds, then $\bar{v} \neq v^*$ is also impossible taking into account that $\{D_f(\bar{v}, v^k)\}$ is a convergent sequence. ■

Let us remind the ϵ-enlargement of an maximal monotone operator $\mathcal{Q}$:

$$\mathcal{Q}_\epsilon(x) = \{u \in I\!R^n : \langle u - v, x - y\rangle \geq -\epsilon \quad \forall\, y \in \mathrm{dom}\mathcal{Q},\ v \in \mathcal{Q}(y)\}.$$

For properties of the ϵ-enlargement, see [8, 9].

Now we describe the method under consideration.

Interior proximal method:

Starting with an arbitrary $x^1 \in \mathrm{int}K$, the method generates two sequences $\{x^k\} \subset I\!R^n$ and $\{e^k\} \subset I\!R^n$ conforming to the recursion

$$e^{k+1} \in \mathcal{Q}^k(x^{k+1}) + \chi_k \nabla_1 D_f(x^{k+1}, x^k). \tag{1}$$

Here f is a Bregman function satisfying B1–B5 and $\mathcal{Q}^k$ is an approximation of $\mathcal{Q}$ such that $\mathcal{Q} \subset \mathcal{Q}^k \subset \mathcal{Q}_{\epsilon_k}$.

We study the convergence of the iterates x^k of this method under the conditions[1]

$$0 < \chi_k < \bar{\chi}\ (\bar{\chi} > 0 \text{ arbitrary}),\ \epsilon_k \geq 0,\ \sum_{k=1}^{\infty} \frac{\epsilon_k}{\chi_k} < \infty,\ \sum_{k=1}^{\infty} \frac{\|e^k\|}{\chi_k} < \infty. \tag{2}$$

According to [7], Lemma 1, conditions B1, B2 and B5 ensure that for each $y \in \mathrm{int}K$

$$\mathrm{dom}\ \partial_1 D_f(\cdot, y) = \mathrm{int}K, \tag{3}$$

and

$$\partial_1 D_f(x, y) = \begin{cases} \nabla f(x) - \nabla f(y) & \text{if } x \in \mathrm{int}K \\ \emptyset & \text{otherwise} \end{cases} \tag{4}$$

($\partial_1 D_f$ denotes the partial subdifferential of D_f).

[1] Considering this method with $e^k \equiv 0$ instead of the last condition in (2), assumption B3 can be replaced by Stand-B3, with minor and evident modifications in the convergence analysis below.

Moreover, the conditions A1, A2, B1, B2 and B5 imply that, for any $e \in I\!R^n$, $y \in \mathrm{int}K$ and $\chi > 0$, the inclusion

$$e \in \mathcal{Q}(x) + \chi \partial_1 D_f(x, y)$$

has a unique solution in $\mathrm{int}K$ (see [7], Theorem 1).

Thus, the assumptions mentioned guarantee that for any sequences $\{e^k\} \subset I\!R^n$ and $\{\chi_k\}$, $\chi_k > 0$, there exists a sequence $\{x^k\}$ satisfying (1), and $\{x^k\} \subset \mathrm{int}K$ is a straightforward corollary of (3).

Lemma 3. *Suppose that the sequence* $\{(x^k, e^k)\}$ *fulfills recursion* (1), *that* $\{x^k\} \subset \mathrm{int}K$ *and that conditions* A1–A3, B1–B3 *and* (2) *are valid. Then*

(i) $\{D_f(x^*, x^k)\}$ *is convergent for each* $x^* \in SOL(\mathcal{Q}, K)$;

(ii) $\{x^k\}$ *is bounded;*

(iii) $\lim\limits_{k\to\infty} D_f(x^{k+1}, x^k) = 0$.

Proof. According to (1) there exists $q^{k+1} \in \mathcal{Q}^k(x^{k+1})$ such that

$$e^{k+1} = q^{k+1} + \chi_k \nabla_1 D_f(x^{k+1}, x^k).$$

From this equality and (4) we conclude that

$$\left\langle q^{k+1} + \chi_k(\nabla f(x^{k+1}) - \nabla f(x^k)), x - x^{k+1} \right\rangle \geq -\|e^{k+1}\| \|x - x^{k+1}\| \tag{5}$$

holds for all $x \in K$. Together with the obvious identity

$$\begin{aligned} &D_f(x, x^{k+1}) - D_f(x, x^k) \\ &= -D_f(x^{k+1}, x^k) + \left\langle \nabla f(x^k) - \nabla f(x^{k+1}), x - x^{k+1} \right\rangle \end{aligned} \tag{6}$$

this yields

$$\begin{aligned} &D_f(x, x^{k+1}) - D_f(x, x^k) \leq -D_f(x^{k+1}, x^k) \\ &+ \frac{1}{\chi_k} \left\langle q^{k+1}, x - x^{k+1} \right\rangle + \frac{\|e^{k+1}\|}{\chi_k} \|x - x^{k+1}\|, \ \forall x \in K. \end{aligned} \tag{7}$$

Choose $x^* \in SOL(\mathcal{Q}, K)$ and $q^* \in \mathcal{Q}(x^*)$ satisfying

$$\langle q^*, x - x^* \rangle \geq 0, \quad \forall\, x \in K.$$

From the definition of $\mathcal{Q}_\epsilon$ and the inclusions $\mathcal{Q} \subset \mathcal{Q}^k \subset \mathcal{Q}_{\epsilon_k}$, one gets

$$\langle q^{k+1} - q^*, x^* - x^{k+1} \rangle \leq \epsilon_k. \tag{8}$$

Now, we take (7) with $x := x^*$ and insert there the following two estimates

$$\langle q^{k+1}, x^* - x^{k+1} \rangle \leq \langle q^*, x^* - x^{k+1} \rangle + \epsilon_k \leq \epsilon_k,$$

$$\|x^* - x^{k+1}\| \leq \frac{1}{\alpha(x^*)} \left[D_f(x^*, x^{k+1}) + c(x^*) \right].$$

The first one is true because of (8) and $x^{k+1} \in K$, and the second one is a consequence of B3. These insertions lead to

$$\begin{aligned} & D_f(x^*, x^{k+1}) - D_f(x^*, x^k) \\ & \leq -D_f(x^{k+1}, x^k) + \frac{\delta_k}{\alpha(x^*)} D_f(x^*, x^{k+1}) + \delta_k \left(1 + \frac{c(x^*)}{\alpha(x^*)}\right), \end{aligned} \tag{9}$$

where $\delta_k := \chi_k^{-1} \max\{\|e^{k+1}\|, \epsilon_k\}$.

Conditions (2) provide that $\frac{\delta_k}{\alpha(x^*)} < \frac{1}{2}$ for $k \geq k_0$, k_0 sufficiently large. Therefore,

$$1 \leq \left(1 - \frac{\delta_k}{\alpha(x^*)}\right)^{-1} \leq \left(1 + \frac{2\delta_k}{\alpha(x^*)}\right) < 2, \quad \forall\, k \geq k_0,$$

and (9) results in

$$\begin{aligned} & D_f(x^*, x^{k+1}) \\ & \leq \left(1 + \tfrac{2\delta_k}{\alpha(x^*)}\right) D_f(x^*, x^k) - D_f(x^{k+1}, x^k) + 2\delta_k \left(1 + \tfrac{c(x^*)}{\alpha(x^*)}\right). \end{aligned} \tag{10}$$

Taking into account that D_f is a non-negative function and $\sum_{k=1}^{\infty} \delta_k < \infty$, Lemma 2.2.2 in [26], applied to (10), guarantees that $\{D_f(x^*, x^k)\}$ converges and

$$\lim_{k \to \infty} D_f(x^{k+1}, x^k) = 0.$$

Now, condition B3 implies that the sequence $\{x^k\}$ is bounded. ■

In the sequel, we make use of the following additional assumptions on the operator $\mathcal{Q}$.

A4 *$\mathcal{Q}$ is the subdifferential of a proper convex, lower semi-continuous function J;*

A5 (a) *$\mathcal{Q}$ is paramonotone on K (for the definition and some properties, see* [15]*),*

(b) $\lim_{k\to\infty} y^k = \bar{y} \in K$, *$q^k \in \mathcal{Q}(y^k)$ implies that $\{q^k\}$ is a bounded sequence.*

If A4 is applied, we also suppose that $\mathcal{Q} \subset \mathcal{Q}^k \subset \partial_{\epsilon_k} J$, where ∂_e denotes the ϵ-subdifferential. Since the inclusion $\partial_\epsilon J \subset (\partial J)_\epsilon \equiv \mathcal{Q}_\epsilon$ is valid, no alterations in the preceding part of the paper are needed.

Assumptions A4 and A5(a) are rather standard in Bregman-function-based proximal methods. For some relaxation of A5(b) see [8, 13, 18] and [30] and Subsection 3.1 below. A Bregman-function-based proximal method for variational inequalities with non-paramonotone operators was studied in [19].

Assuming A5, the following property **(*)** of a paramonotone operator $\mathcal{A}$ on a set C is decisive (cf., [15]):

(*) *If x^* solves the variational inequality*

$$\text{(11)} \qquad find\ x \in C, q \in \mathcal{A}(x): \quad \langle q, y - x\rangle \geq 0, \quad \forall\, y \in C,$$

and for some $\bar{x} \in C$ there exists $\bar{z} \in \mathcal{A}(\bar{x})$ with

$$\langle \bar{z}, x^* - \bar{x}\rangle \geq 0,$$

then $\bar{x}$ is also a solution of (11).

Lemma 4. *Let assumptions* A1–A3, B1–B3, B5 *and one of the assumptions* A4 *or* A5 *be satisfied. Then each cluster point of the sequence $\{x^k\}$, generated by method* (1)–(2)*, belongs to $SOL(\mathcal{Q}, K)$.*

Proof. According to Lemma 3, the sequence $\{x^k\}$ is bounded, hence there exists a convergent subsequence $\{x^{j_k}\}$ with $\lim_{k\to\infty} x^{j_k} = \bar{x}$. Since K is a closed set and $x^k \in \text{int}K$ one gets $\bar{x} \in K$. Taking into account conclusion (iii) of Lemma 3, the application of Lemma 1 with $z^k := x^{j_k+1}$, $y^k := x^{j_k}$, yields $\lim_{k\to\infty} x^{j_k+1} = \bar{x}$. Moreover, using the identity (6) with $x := x^* \in SOL(\mathcal{Q}, K)$, the relation

$$\lim_{k\to\infty} \chi_{j_k}\langle \nabla f(x^{j_k+1}) - \nabla f(x^{j_k}), x^* - x^{j_k+1}\rangle = 0 \tag{12}$$

follows immediately from Lemma 3 and $\chi_k \in (0, \bar{\chi}]$ for all k.

First, let us suppose that A4 is valid. Due to the convexity of the function J and $q^{k+1} \in \mathcal{Q}^k(x^{k+1}) \subset \partial_{\epsilon_k} J(x^{k+1})$, relation (5) considered for $x := x^*$ and $k := j_k$ implies that

$$\begin{aligned} -J(x^{j_k+1}) + J(x^*) \ &+ \chi_{j_k}\langle \nabla f(x^{j_k+1}) - \nabla f(x^{j_k}), x^* - x^{j_k+1}\rangle \\ &\geq -\|e^{j_k+1}\|\|x^* - x^{j_k+1}\| - \epsilon_{j_k}. \end{aligned} \tag{13}$$

Passing to the limit in (13) for $k \to \infty$, in view of (12), (2), $\chi_k \in (0, \bar{\chi}]$ and the lower semi-continuity of J, we obtain $J(\bar{x}) \leq J(x^*)$. Together with $\bar{x} \in K$, $x^* \in SOL(\mathcal{Q}, K)$, this yields

$$0 \in \partial\left(J(\bar{x}) + \delta(\bar{x}|K)\right),$$

where $\delta(\cdot|K)$ is the indicator function of K. In view of assumption A2, Theorem 23.8 in [27] provides $\bar{x} \in SOL(\mathcal{Q}, K)$.

Now, let us turn to the case when the operator $\mathcal{Q}$ possesses property A5. Owing to the Brønsted-Rockafellar property of the ϵ-enlargement (cf., [9]) and the relation $\mathcal{Q} \subset \mathcal{Q}^k \subset \mathcal{Q}_{\epsilon_k}$, there exist $\tilde{x}^{j_k+1}$ and $q(\tilde{x}^{j_k+1}) \in \mathcal{Q}(\tilde{x}^{j_k+1})$ such that

$$\|x^{j_k+1} - \tilde{x}^{j_k+1}\| \leq \sqrt{\epsilon_{j_k}}, \qquad \|q^{j_k+1} - q(\tilde{x}^{j_k+1})\| \leq \sqrt{\epsilon_{j_k}}, \quad \forall\ k. \tag{14}$$

Hence, $\lim_{k\to\infty} \tilde{x}^{j_k+1} = \bar{x}$, and taking into account A5(b) with $y^k := \tilde{x}^{j_k+1}$, both sequences $\{q(\tilde{x}^{j_k+1})\}$ and $\{q^{j_k+1}\}$ are bounded. Together with the second inequality in (14), this allows us to conclude, without loss of generality, that

$$\lim_{k\to\infty} q^{j_k+1} = \bar{q} \quad \text{and} \quad \lim_{k\to\infty} q(\tilde{x}^{j_k+1}) = \bar{q}.$$

In turn, the maximal monotonicity of $\mathcal{Q}$ ensures $\bar{q} \in \mathcal{Q}(\bar{x})$.

Thus, inserting $x := x^* \in SOL(\mathcal{Q}, K)$ into (5) and passing to the limit for $k := j_k$, $k \to \infty$, we infer from (12) and (2) that

$$\langle \bar{q}, x^* - \bar{x} \rangle \geq 0.$$

Finally, property (*) of the paramonotone $\mathcal{Q}$ provides $\bar{x} \in SOL(\mathcal{Q}, K)$. ■

Now, from $\{x^k\} \subset \mathrm{int} K$ and Lemmata 2–4, the main convergence result follows immediately.

Theorem 1. *Let assumptions* A1–A3, B1–B5 *as well as one of the assumptions* A4 *or* A5 *be valid. Then the sequence* $\{x^k\}$*, generated by method* (1)–(2)*, belongs to* $\mathrm{int} K$ *and converges to a solution of* $VI(\mathcal{Q}, K)$*.*

3. Bregman functions with non-polyhedral zones

Bregman functions with zone $\mathrm{int} K$ are of main interest in applications. Exactly in this case we deal with interior proximal methods. However, as it was already mentioned, up to now such Bregman functions have been constructed only for linearly constrained sets K or when K is a ball. The corresponding convergence results are not applicable, for instance, if K is the intersection of a half-space and a ball. The reason is that neither condition B4(i) nor B4(ii) considered separately[2] is guaranteed in this case (see Example 2 in Appendix).

In this section, the application of methods (1)–(2) is studied for the case

$$K := \{x \in \mathbb{R}^n : g_i(x) \leq 0,\ i \in I\}, \quad I = I_1 \cup I_2, \tag{15}$$

where I is a finite index set, g_i $(i \in I_1)$ are affine functions, g_i $(i \in I_2)$ are convex continuously differentiable functions and $\max_{i \in I_2} g_i$ is supposed to be strictly convex on K. Further, Slater's condition

$$\exists\ \tilde{x} : g_i(\tilde{x}) < 0\ \forall\ i \in I \tag{16}$$

is supposed to be valid.

The following statement clarifies the choice of Bregman functions with zone $\mathrm{int} K$.

[2]Usually, for linearly constrained K condition Stand-B4 (i.e., B4(i)) is supposed. In the case when K is a ball, instead of Stand-B4 condition B4(ii) was introduced in [11].

Theorem 2. *Let φ be a strictly convex, continuous and increasing function with* $\mathrm{dom}\varphi = (-\infty, 0]$, *and φ be continuously differentiable on* $(-\infty, 0)$. *Moreover, let*

$$\lim_{t \uparrow 0} \varphi'(t)t = 0, \tag{17}$$

$$\lim_{t \uparrow 0} \varphi'(t) = +\infty. \tag{18}$$

Then the function[3]

$$h(x) := \sum_{i \in I} \varphi(g_i(x)) + \theta \sum_{j=1}^{n} |x_j|^\gamma, \quad \gamma > 1 \textit{ is fixed,} \tag{19}$$

with $\theta := 0$ if K is certainly bounded and $\theta := 1$ if the boundedness of K is unknown, satisfies the conditions B1–B5.

Proof. Step by step we check whether the assumptions B1–B5 are satisfied for the function h to be a Bregman function.

• Since φ is a convex increasing function, the convexity of the composition $\varphi \circ g_i$ (and hence, of h) on the set K is guaranteed. Thus, if $\theta = 1$, the strict convexity of h is evident.

But, if $\theta = 0$ (i.e., K is bounded), we have to consider two cases.

(a) $I_2 \neq \emptyset$.
Assume that $\sum_{i \in I_2} \varphi \circ g_i$ is not strictly convex on K. Then one can choose points $x^1, x^2 \in K$, $x^1 \neq x^2$, and $\lambda \in (0, 1)$ such that

$$\sum_{i \in I_2} \varphi(g_i(\lambda x^1 + (1-\lambda)x^2)) = \lambda \sum_{i \in I_2} \varphi(g_i(x^1)) + (1-\lambda) \sum_{i \in I_2} \varphi(g_i(x^2)).$$

Since each function $\varphi \circ g_i$ is convex, the last equality means that for each $i \in I_2$,

$$\varphi(g_i(\lambda x^1 + (1-\lambda)x^2)) = \lambda \varphi(g_i(x^1)) + (1-\lambda)\varphi(g_i(x^2)) \tag{20}$$

[3] As it will be clear from the proof of this theorem, instead of $\sum_{i=1}^{n} |x_j|^\gamma$ in (19) one can take any strictly convex and continuously differentiable function guaranteeing the fulfilment of B3 for h.

is valid. Since g_i is convex, φ is increasing and $x^1, x^2 \in K$, we obtain

$$(21) \quad \varphi(\lambda g_i(x^1) + (1-\lambda)g_i(x^2)) \geq \lambda\varphi(g_i(x^1)) + (1-\lambda)\varphi(g_i(x^2)), \quad i \in I_2.$$

The strict convexity of φ implies, that only the equality is possible in (21), and $g_i(x^1) = g_i(x^2)$ holds for each $i \in I_2$. Finally, in view of (20) and the increase of φ one gets for $i \in I_2$

$$g_i(\lambda x^1 + (1-\lambda)x^2) = \lambda g_i(x^1) + (1-\lambda)g_i(x^2),$$

which contradicts the strict convexity of $\max_{i\in I_2} g_i$.

(b) $I_2 = \emptyset$, i.e., all functions g_i are affine:

$$g_i(x) := \langle a^i, x\rangle - b_i, \quad i \in I_1 = I.$$

Then, due to Slater's condition and the boundedness of K, the rank of the matrix $A = \{a^i\}_{i\in I}$ equals n. Using this fact, the strict convexity of the function h can be easily established following [24].

Thus, taking into account also the differentiability properties of g_i and φ, assumptions B1 and B2 for the function h are always valid.

• Of course, assumption B3 is guaranteed if K is a bounded set. In the case of an unbounded K, B3 is evident if $\gamma = 2$ in (19); for arbitrary $\gamma > 1$ condition B3 follows from Proposition 2 in [20] and

$$D_h(x, y) \geq h_0(x) - h_0(y) - \langle\nabla h_0(y), x - y\rangle,$$

(where $h_0(x) := \sum_{i=1}^n |x_i|^\gamma$).

• Now, we check the validity of assumption B4. Denote

$$I_<(y) := \{i : g_i(y) < 0\}, \quad I_=(y) := \{i : g_i(y) = 0\},$$

and let $\{z^k\} \subset \mathrm{int}K$, $\lim_{k\to\infty} z^k = z$.

(a) At first, suppose that $I_<(z) \supset I_2$. For $i \in I_<(z)$, one gets

$$\lim_{k\to\infty} g_i(z^k) = g_i(z) < 0 \quad \text{and} \quad \lim_{k\to\infty} \langle\nabla g_i(z^k), z - z^k\rangle = 0,$$

whereas for $i \in I_=(z)$ we have $\lim_{k\to\infty} g_i(z^k) = g_i(z) = 0$, $b_i = \langle a^i, z\rangle$. Thus, for $i \in I_=(z)$,

$$\lim_{k\to\infty} \varphi'(\langle a^i, z^k\rangle - b_i)\langle a^i, z - z^k\rangle = \lim_{k\to\infty} \varphi'(\langle a^i, z^k - z\rangle)\langle a^i, z - z^k\rangle = 0$$

follows from $\langle a^i, z^k - z\rangle \to 0$ and (17). Using these relations and the identity

$$\begin{aligned} D_h(x,y) \;=\; & h(x) - h(y) - \sum_{i\in I_=(y)} \varphi'(\langle a^i, y\rangle - b_i)\langle a^i, x - y\rangle \\ & - \sum_{i\in I_<(y)} \varphi'(g_i(y))\langle \nabla g_i(y), x - y\rangle - \theta\langle \nabla h_0(y), x - y\rangle, \end{aligned}$$

one can easily conclude that $\lim_{k\to\infty} D_h(z, z^k) = 0$.

(b) Now, let $g_{i_0}(z) = 0$ be valid for some $i_0 \in I_2$. Denote

$$I_2(y) := \{i \in I_2 : g_i(y) = \max_{j\in I_2} g_j(y)\}$$

and take $\bar z \in K$, $\bar z \neq z$. In view of the convexity of the functions $\varphi \circ g_i$ and h_0 the following holds

$$D_h(\bar z, z^k) \geq \varphi(g_{i_0}(\bar z)) - \varphi(g_{i_0}(z^k)) - \varphi'(g_{i_0}(z^k))\langle \nabla g_{i_0}(z^k), \bar z - z^k\rangle. \tag{22}$$

Obviously, relation (18) implies

$$\lim_{k\to\infty} \varphi'(g_{i_0}(z^k)) = +\infty, \tag{23}$$

whereas

$$\lim_{k\to\infty} \langle \nabla g_{i_0}(z^k), \bar z - z^k\rangle = \langle \nabla g_{i_0}(z), \bar z - z\rangle. \tag{24}$$

But, using the structure of the subdifferential of a max-function and the strict convexity of $\max_{i\in I_2} g_i$, we obtain

$$\max_{i\in I_2} g_i(\bar z) - \max_{i\in I_2} g_i(z) > \langle \nabla g_j(z), \bar z - z\rangle, \quad \forall\, j \in I_2(z).$$

Since $\bar z, z \in K$ and $g_{i_0}(z) = 0$, the last inequality yields

$$\langle \nabla g_j(z), \bar z - z\rangle < 0, \quad j \in I_2(z). \tag{25}$$

Relations (22)–(25) and the continuity of $\varphi \circ g_{i_0}$ ensure that

$$\lim_{k\to\infty} D_h(\bar z, z^k) = +\infty. \tag{26}$$

In fact, we have proved a stronger property than B4(ii): relation (26) holds for each $\bar z \in K$, $\bar z \neq z$, whereas B4(ii) supposes

$$\overline{\lim}_{k\to\infty} D_h(\bar z, z^k) = +\infty \quad \text{if} \quad \bar z \in \mathrm{bd}K,\ \bar z \neq z.$$

• To prove the fulfillment of assumption B5, we use Theorem 4.5 in [5], which states, in particular, that

> *For a function f with properties* B1 *and* B2, *the boundary coercivity implies zone coercivity if K is bounded or the super-coercivity condition* $\lim_{x\in K,\ \|x\|\to\infty} \frac{f(x)}{\|x\|} = +\infty$ *is valid.*

Let $z \in \mathrm{bd}K$, $\{z^k\} \in \mathrm{int}K$, $\lim_{k\to\infty} z^k = z$ and $x \in \mathrm{int}K$. Then

$$\lim_{k\to\infty} \langle \nabla g_i(z^k), x - z^k\rangle = \langle \nabla g_i(z), x - z\rangle,$$

and for $i \in I_=(z)$ the relations

$$\langle \nabla g_i(z), x - z\rangle < 0, \qquad \lim_{k\to\infty} \varphi'(g_i(z^k)) = +\infty$$

are obvious. Hence, if $i \in I_=(z)$, then

$$\lim_{k\to\infty} \varphi'(g_i(z^k))\langle \nabla g_i(z^k), x - z^k\rangle = -\infty, \tag{27}$$

whereas for $i \in J_<(z)$ it holds

$$\lim_{k\to\infty} \varphi'(g_i(z^k))\langle \nabla g_i(z^k), x - z^k\rangle = \varphi'(g_i(z))\langle \nabla g_i(z), x - z\rangle. \tag{28}$$

From (27) and (28) and the continuous differentiability of the function h_0, we immediately conclude that the function h is boundary coercive, hence according to Theorem 4.5 in [5], h satisfies assumption B5. ■

Particular functions satisfying the conditions of Theorem 2 are

$$\varphi(t) = -(-t)^p, \quad p \in (0,1) \text{ arbitrarily chosen,} \tag{29}$$

$$
(30) \qquad \varphi(t) = \begin{cases} -t\ln(-t) + t & \text{if } -\dfrac{1}{2} \le t \le 0 \\ -\ln 2 \ln\Big(-t + \dfrac{1}{2}\Big) - \dfrac{1}{2}\ln 2 - \dfrac{1}{2} & \text{if } t < -\dfrac{1}{2} \end{cases},
$$

where by convention $\varphi(0) = 0$.

The principal idea for constructing function (30) consists in the following: Constraints, which turn out to be active at the limit point, are handled mainly by the first term in φ, whereas those which become inactive at the limit point, are observed by both terms of φ.

Combining Theorems 1 and 2 we immediately obtain the following result.

Theorem 3. *Let the set K be described by* (15), (16) *and the function h be defined as in Theorem* 2 (*in particular, φ of form* (29) *or* (30) *can be chosen*). *Moreover, suppose that $VI(\mathcal{Q}, K)$ satisfies assumptions* A1–A3 *and one of the assumptions* A4 *or* A5.

Then, the sequence $\{x^k\}$, generated by method (1)–(2) *with the use of function h* (*in place of f*), *belongs to* $\mathrm{int} K$ *and converges to $x \in SOL(\mathcal{Q}, K)$.*

Corollary 1. *Suppose that the hypotheses of Theorem* 3 *are fulfilled. If $VI(\mathcal{Q}, K)$ has more than one solution, then the sequence $\{x^k\}$, generated by method* (1)–(2) *with Bregman function* (19), *converges to a solution x such that $g_i(x) < 0$, $i \in I_2$.*

Indeed, the opposite assumption that $g_i(x) = 0$ holds for some $i \in I_2$ leads immediately to a contradiction between the statement (i) of Lemma 3 and relation (26) given with $z^k := x^k$ and $\bar{z} \in SOL(\mathcal{Q}, K)$, $\bar{z} \neq x$.

Remark 3. If functions g_i in (15) satisfy

$$
g_i(x) \ge -1 \quad \forall\, x \in K, \ \forall\, i \in I_2,
$$

one can quite similarly prove that the function $h_l : K \to I\!R$, defined by

$$
(31) \qquad h_l(x) := \sum_{i \in I} [(-g_i(x)) \ln(-g_i(x)) + g_i(x)] + \theta \sum_{j=1}^{n} |x_j|^{\gamma}
$$

(with θ, γ as in (19)), possesses properties B1–B5, too. $\diamond$

3.1. Embedding of the original Bregman-function-based methods

Let us analyze the original proof technique of Bregman-function-based proximal methods destined to variational inequalities on *polyhedral* sets. Denote by D a distance function and by v^ℓ an iterate of such a method. To our knowledge, original convergence results for these methods establish convergence of the sequence $\{D(x, v^\ell)\}$ for each solution x *without* the use of assumption Stand-B4. Hence, for interior proximal methods involving *standard* requirements on Bregman functions (see, in particular, [8, 13, 30]), the original convergence analysis can be preserved (with minor alterations in the final stage only) if we replace Stand-B4 by B4 with $\bar{z} \in K$ instead of $\bar{z} \in \mathrm{bd}K$ in B4(ii)[4].

Indeed, let a subsequence $\{v^{\ell_k}\}$ of $\{v^\ell\}$ with $\lim_{k\to\infty} v^{\ell_k} = v$ be chosen such that there exists $x \in SOL(\mathcal{Q}, K)$, $x \neq v$ (if this is impossible, clearly $\lim_{\ell\to\infty} v^\ell = v$ and $SOL(\mathcal{Q}, K) = \{v\}$). The convergence of $\{D(x, v^\ell)\}$ implies that condition B4(ii) with $\bar{z} := x$, $z := v$, $z^k := v^{l_k}$ is violated. Hence, B4(i) is valid, i.e., $\lim_{k\to\infty} D(v, v^{\ell_k}) = 0$, and B4 turns into Stand-B4. Referring now to the original convergence results, we immediately conclude that $v \in SOL(\mathcal{Q}, K)$. Then, convergence of $\{D(v, v^\ell)\}$ implies $\lim_{\ell\to\infty} D(v, v^\ell) = 0$ and Theorem 2.4 in [30] yields $\lim_{\ell\to\infty} v^\ell = v$.

On this way, inserting the function (19) (for instance, with φ defined by (29) or (30)) in methods where Bregman functions are described exactly by standard conditions (see the note after Remark 1), one can extend these methods to solve $VI(\mathcal{Q}, K)$ on *non-polyhedral* sets K given by (15), (16).

Return to the convergence analysis in Section 2. For method (1)–(2) without an approximation of the operator $\mathcal{Q}$, i.e. with $\mathcal{Q}^k \equiv \mathcal{Q}$, the same arguments as given just above permit us to weaken assumption A5, replacing A5(b) by the condition that $\mathcal{Q}$ is a pseudo-monotone operator in the sense of Brézis-Lions[5]. With this alteration, the proof of Lemma 4 can be performed quite similarly to the proof of the case D3 in Lemma 3 of [18].

Moreover, for the exact version of method (1)–(2), with $\mathcal{Q}^k \equiv \mathcal{Q}$, $e^k \equiv 0$, assumption A5(b) can be omitted at all according to the analysis of Solodov and Svaiter [30].

---

[4]Let us recall that, for K given by (15), (16), the conditions of Theorem 2 ensure the fulfillment of the so modified assumption B4.

[5]There are several notions of pseudo-monotonicity. This notion was introduced in [6] and [22] for single-valued operators, for multi-valued operators see for instance [25].

In Appendix, we give examples showing that

- ∘ although assumption B4 is valid for the function h defined by (19), (29), each of the conditions B4(i) and B4(ii) itself can be violated (of course, it means, in particular, that Stand-B4 is not satisfied, in general);
- ∘ assumption B4 can be violated for this function h if the function $\max_{i\in I_2} g_i$ is not strictly convex.

3.2. On entropy-like and logarithmic-quadratic proximal methods

We have tried to extend in a similar manner entropy-like and logarithmic-quadratic proximal methods (see [1, 3, 4, 31]). Applications of these methods to the dual of linearly constrained programs or to variational inequalities on polyhedral convex sets provide attractive properties of subproblems (for instance, C^∞ multiplier methods with bounded Hessians are constructed in this way in [3]).

However, the basic requirements on the kernel functions in these methods seem to be not appropriate for such an extension. Indeed, in case $K := I\!R^m_+$ the corresponding distance functions have the form

$$d(u,v) = \sum_{i=1}^{m} v_i^\alpha \left[\varphi\left(\frac{u_i}{v_i}\right) + \beta \left(\frac{u_i}{v_i} - 1\right)^2 \right] \tag{32}$$

($\alpha := 1,\ \beta := 0$ in entropy-like methods and $\alpha := 2,\ \beta > 0$ in logarithmic-quadratic methods). If

$$K := \{x \in I\!R^n : g_i(x) := \langle a_i, x\rangle - b_i \le 0,\ i = 1, \dots, m\},$$

the distance D is given by

$$D(x,y) := d(-g(x), -g(y)), \quad g = (g_1, \dots, g_m). \tag{33}$$

As to the kernel function φ, it is supposed, in particular, that $\mathrm{dom}\varphi \subseteq [0, +\infty)$, φ is twice continuously differentiable on $(0, +\infty)$ and strictly convex on its domain, and

$$\varphi(1) = \varphi'(1) = 0,\ \varphi''(1) > 0. \tag{34}$$

Already these conditions enforce that the function $D(\cdot, y)$ defined by (33) is, in general, nonconvex for some $y \in \mathrm{int}K$, if at least one function g_i is not affine.

Example 1. Let the kernel function φ with the properties mentioned above be fixed, and $K := \{x \in I\!R^1 : g(x) \leq 0\}$, where the choice of the convex and sufficiently smooth function $g : I\!R^1 \to I\!R^1$ will be specified below. According to (32), (33)

$$D(x,y) = (-g(y))^{\alpha} \left[\varphi\left(\frac{g(x)}{g(y)}\right) + \frac{\beta}{2}\left(\frac{g(x)}{g(y)} - 1\right)^2 \right].$$

Assume first that $\varphi(\frac{1}{2}) \leq \varphi(\frac{3}{2})$ and take g with values

$$g(x^1) = -\frac{1}{2}, \quad g(1) = -\frac{3}{2}, \quad g(x^2) = -1,$$

where x^1, x^2 are given points, such that $(x^1 - 1)(x^2 - 1) < 0$. Then, for $y = x^2$ one gets

$$D(x^1, y) = \varphi\left(\frac{1}{2}\right) + \frac{\beta}{8}, \quad D(1,y) = \varphi\left(\frac{3}{2}\right) + \frac{\beta}{8}, \quad D(x^2, y) = 0.$$

Because of $\varphi(\frac{1}{2}) \leq \varphi(\frac{3}{2})$ and $\varphi(t) \geq 0 \; \forall\, t \in \mathrm{dom}\varphi$ (the last inequality follows from the convexity of φ and (34)), we conclude that the functions $D(\cdot, y)$ and $\varphi \circ (-g)$ are not convex.

But, if $\varphi(\frac{1}{2}) > \varphi(\frac{3}{2})$, the same conclusion holds true with $y = x^1$ and

$$g: \; g(x^1) = -1, \quad g(1) = -\frac{1}{2}, \quad g(x^2) = -\frac{3}{2}.$$

4. APPENDIX

The following example shows that, although assumption B4 is valid for the function h defined by (19), (29), each of the conditions B4(i) and B4(ii) itself can be violated.

Example 2. Let $K := \{x \in I\!R^2 : g_i(x) \leq 0, \; i = 1, 2\}$, where

$$g_1(x) := x_1^2 + x_2^2 - 1, \quad g_2(x) := -x_1 + x_2 - 1.$$

According to (19), (29), let $h(x) = h_1(x) + h_2(x)$, with

$$h_1(x) := -(1 - x_1^2 - x_2^2)^p, \quad h_2(x) := -(1 + x_1 - x_2)^p, \quad p \in (0,1).$$

Choose the sequence $\{z^k\}$, $z^k = (z_1^k, z_2^k)$, with

$$z_1^k := \left(1 - \sigma_k^{\frac{1}{1-p}} - (1 - \sigma_k^{\frac{1}{1-p}} - \sigma_k)^2\right)^{1/2},$$
$$z_2^k := 1 - \sigma_k^{\frac{1}{1-p}} - \sigma_k, \text{ where } \sigma_k > 0, \lim_{k\to\infty} \sigma_k = 0.$$

Obviously, $z^k \in \mathrm{int}K$ for a large k ($k \geq k_0$), and $\lim_{k\to\infty} z^k = z = (0,1) \in \mathrm{bd}K$. Taking into account the convexity of h_1 and h_2 and $\lim_{k\to\infty}(h_1(z) - h_1(z^k)) = 0$, condition B4(i) should imply

$$\lim_{k\to\infty} (-g_1(z^k))^{p-1}\langle \nabla g_1(z^k), z - z^k\rangle = 0.$$

But, a straightforward calculation yields

$$(-g_1(z^k))^{p-1}\langle \nabla g_1(z^k), z - z^k\rangle = -2, \quad \forall\ k \geq k_0.$$

Now, choose $\bar{z} = (0,1)$ and $z^k = (z_1^k, z_2^k)$ with

$$z_1^k = -\frac{1}{2} + \sigma_k, \qquad z_2^k = \frac{1}{2}, \quad \text{where, as before, } \sigma_k > 0,\ \lim_{k\to\infty} \sigma_k = 0.$$

Then, because of $g_1(z) < 0$ for $z = \lim_{k\to\infty} z^k = (-\frac{1}{2}, \frac{1}{2})$, the function

$$d_1 :\ v \mapsto h_1(\bar{z}) - h_1(v) - \langle \nabla h_1(v), \bar{z} - v\rangle$$

is continuous in a neighborhood of z, hence $\lim_{k\to\infty} d_1(z^k) = d_1(z) = (\frac{1}{2})^p < \infty$. But, with

$$d_2 :\ v \mapsto h_2(\bar{z}) - h_2(v) - \langle \nabla h_2(v), \bar{z} - v\rangle$$

we obtain

$$d_2(z^k) = (1-p)\left(1 + z_1^k - z_2^k\right)^p = (1-p)\sigma_k^p \to 0 \quad \text{as } k \to \infty.$$

Finally,

$$\lim_{k\to\infty} D_h(\bar{z}, z^k) = \lim_{k\to\infty}\left(d_1(z^k) + d_2(z^k)\right) = (\frac{1}{2})^p < \infty,$$

i.e., condition B4(ii) is not valid, too. $\diamond$

The next example indicates that assumption B4 may be violated for function (19) if the function $\max_{i \in I_2} g_i$ is not strictly convex.

Example 3. Let $K := \{x \in I\!R^4 : x_1^2 + x_2^2 + x_3 + x_4 \le 1\}$.
With $g(x) := x_1^2 + x_2^2 + x_3 + x_4 - 1$, take

$$h(x) := -(-g(x))^p + \sum_{j=1}^{4} x_j^2, \quad p \in (0,1),$$

and $\{z^k\}$ such that

$$\begin{aligned} &\text{if } k = 2l-1: && z_1^k = \left(1 - \sigma_k^{\frac{1}{1-p}} - (1 - \sigma_k^{\frac{1}{1-p}} - \sigma_k)^2\right)^{1/2}, \\ & && z_2^k = 1 - \sigma_k^{\frac{1}{1-p}} - \sigma_k, \ z_3^k = z_4^k = 0; \\ &\text{if } k = 2l: && z_1^k = 0, \ z_2^k = 1, \ z_3^k = -\sigma_k, \ z_4^k = 0, \quad \sigma_k > 0, \ \lim_{k\to\infty} \sigma_k = 0. \end{aligned}$$

Then $z = \lim_{k\to\infty} z^k = (0,1,0,0)$ and for a large l

$$(-g(z^k))^{p-1} \langle \nabla g(z^k), z - z^k \rangle = -2, \quad \text{if } k = 2l-1.$$

But, choosing $\bar{z} := (0,1,-1,1)$, one gets

$$(-g(z^k))^{p-1} \langle \nabla g(z^k), \bar{z} - z^k \rangle = \sigma_k^p, \quad \text{if } k = 2l.$$

Thus, neither B4(i) nor B4(ii) is valid. ◇

REFERENCES

[1] A. Auslender and M. Haddou, *An interior-proximal method for convex linearly constrained problems and its extension to variational inequalities*, Math. Programming **71** (1995), 77–100.

[2] A. Auslender and M. Teboulle, *Entropic proximal decomposition methods for convex programs and variational inequalities*, Math. Programming Ser. A **91** (2001), 33–47.

[3] A. Auslender, M. Teboulle and S. Ben-Tiba, *Interior proximal and multiplier methods based on second order homogeneous kernels*, Mathematics of Oper. Res. **24** (1999), 645–668.

[4] A. Auslender, M. Teboulle and S. Ben-Tiba, *A logarithmic-quadratic proximal method for variational inequalities*, Computational Optimization and Applications **12** (1999), 31–40.

[5] H. Bauschke and J. Borwein, *Legendre functions and the method of random Bregman projections*, J. Convex Analysis **4** (1997), 27–67.

[6] H. Brézis, *Équations et inéquations non linéaires dans les espaces vectoriels en dualité*, Ann. Inst. Fourier **18** (1968), 115–175.

[7] R. Burachik and A. Iusem, *A generalized proximal point algorithm for the variational inequality problem in Hilbert space*, SIAM J. Optim. **8** (1998), 197–216.

[8] R. Burachik, A. Iusem and B. Svaiter, *Enlargements of maximal monotone operators with application to variational inequalities*, Set-Valued Analysis **5** (1997), 159–180.

[9] R. Burachik and B. Svaiter, *ϵ-enlargement of maximal monotone operators in Banach spaces*, Set-Valued Analysis **7** (1999), 117–132.

[10] R. Burachik and B. Svaiter, *A relative error tolerance for a family of generalized proximal point methods*, Math. of Oper. Res. **26** (2001), 816–831.

[11] Y. Censor, A. Iusem and S.A. Zenios, *An interior point method with Bregman functions for the variational inequality problem with paramonotone operators*, Math. Programming **81** (1998), 373–400.

[12] Y. Censor and S.A. Zenios, *Proximal minimization algorithm with d-functions*, J. Optim. Theory Appl. **73** (1992), 451–464.

[13] J. Eckstein, *Approximate iterations in Bregman-function-based proximal algorithms*, Math. Programming **83** (1998), 113–123.

[14] A. Iusem, *On some properties of generalized proximal point methods for quadratic and linear programming*, JOTA **85** (1995), 593–612.

[15] A. Iusem, *On some properties of paramonotone operators*, J. of Conv. Analysis **5** (1998), 269–278.

[16] A. Kaplan and R. Tichatschke, Stable Methods for Ill-Posed Variational Problems - Prox-Regularization of Elliptic Variational Inequalities and Semi-Infinite Optimization Problems, Akademie Verlag, Berlin 1994.

[17] A. Kaplan and R. Tichatschke, *Proximal point approach and approximation of variational inequalities*, SIAM J. Control Optim. **39** (2000), 1136–1159.

[18] A. Kaplan and R. Tichatschke, *Convergence analysis of non-quadratic proximal methods for variational inequalities in Hilbert spaces*, J. of Global Optimization **22** (2002), 119–136.

[19] A. Kaplan and R. Tichatschke, *Interior proximal method for variational inequalities: Case of non-paramonotone operators*, Journal of Set-Valued Analysis **12** (2004), 357–382.

[20] A. Kaplan and R. Tichatschke, *On inexact generalized proximal methods with a weakened error tolerance criterion*, Optimization **53** (2004), 3–17.

[21] K. Kiwiel, *Proximal minimization methods with generalized Bregman functions*, SIAM J. Control Optim. **35** (1997), 1142–1168.

[22] J.L. Lions, Quelques Méthodes de Résolution de Problèmes Nonlinéaires, Dunod, Paris 1969.

[23] B. Martinet, *Régularisation d'inéquations variationelles par approximations successives*, RIRO **4** (1970), 154–159.

[24] N. Megiddo, *Pathways to the optimal set in linear programming*, In Progress in Mathematical Programming, Interior Point and Related Methods (1989), N. Megiddo, Ed., Springer, New York, pp. 131–158.

[25] D. Pascali and S. Sburlan, Nonlinear Mappings of Monotone Type, Editura Academiei, Bucharest 1978.

[26] B.T. Polyak, Introduction to Optimization, Optimization Software, Inc. Publ. Division, New York 1987.

[27] R.T. Rockafellar, Convex Analysis, Princeton University Press, Princeton 1970.

[28] R.T. Rockafellar, *On the maximality of sums of nonlinear monotone operators*, Trans. Amer. Math. Soc. **149** (1970), 75–88.

[29] R.T. Rockafellar, *Monotone operators and the proximal point algorithm*, SIAM J. Control Optim. **14** (1976), 877–898.

[30] M. Solodov and B. Svaiter, *An inexact hybrid generalized proximal point algorithm and some new results on the theory of Bregman functions*, Math. Oper. Res. **25** (2000), 214–230.

[31] M. Teboulle, *Convergence of proximal-like algorithms*, SIAM J. Optim. **7** (1997), 1069–1083.

Received 16 February 2006