Journal of Global Optimization **22:** 119–136, 2002.

Convergence analysis of non-quadratic proximal methods for variational inequalities in Hilbert spaces

ALEXANDER KAPLAN and RAINER TICHATSCHKE*
Department of Mathematics, University of Trier, 54286 Trier, Germany
**Corresponding author (E-mail: tichat@uni-trier.de)*

Abstract. We consider a general approach for the convergence analysis of proximal-like methods for solving variational inequalities with maximal monotone operators in a Hilbert space. It proves to be that the conditions on the choice of a non-quadratic distance functional depend on the geometrical properties of the operator in the variational inequality, and — in particular — a standard assumption on the strict convexity of the kernel of the distance functional can be weakened if this operator possesses a certain 'reserve of monotonicity'. A successive approximation of the 'feasible set' is performed, and the arising auxiliary problems are solved approximately. Weak convergence of the proximal iterates to a solution of the original problem is proved.

Key words: Variational inequalities, Monotone operators, Proximal point methods, Regularization

1. Introduction

Let $(X, \|\cdot\|)$ be a Hilbert space with the topological dual X' and the duality pairing $\langle\cdot,\cdot\rangle$ between X and X'. We consider the variational inequality

$$(P) \quad \text{find } x^* \in K \text{ such that } \exists q \in \mathcal{Q}(x^*) : \langle q, x - x^* \rangle \geqslant 0 \quad \forall x \in K,$$

where $K \subset X$ is a convex closed set and $\mathcal{Q} : X \to 2^{X'}$ is a maximal monotone operator.

We generally suppose that (P) is solvable and denote by X^* its solution set.

The proximal point method (PPM), originally introduced by Martinet [22] to solve convex variational problems and later on investigated in a more general setting by Rockafellar [27], has initiated a lot of new algorithms for solving various classes of variational inequalities and related problems.

The exact proximal point method, applied to the variational inequality (P), can be described as follows:
$x^0 \in K$ and a sequence $\{\chi_k\}$, $0 < \chi_k \leqslant \bar{\chi} < \infty$, are given;
$x^{k+1} \in K$ is defined such that

$$\exists\, q(x^{k+1}) \in \mathcal{Q}(x^{k+1}) : \langle q(x^{k+1}) + \chi_k \nabla_1 D(x^{k+1}, x^k), y - x^{k+1} \rangle \geqslant 0 \quad \forall y \in K,$$

where $D(x, y) = \frac{1}{2}\|x - y\|^2$ and ∇_1 is the partial gradient w.r.t. x.

For different modifications of the PPM, also with other quadratic functionals D, we address the reader to [17, 19, 20], where numerous references can be found.

In the last decade, a new direction in the PPM's has been extensively developed, which is based on using *non*-quadratic 'distance functionals' D. The main motivation for such proximal methods is the following:

– the use of a non-quadratic proximal term permits, for certain classes of problems, to preserve the main merits of the classical PPM (good stability of the auxiliary problems and convergence of the whole sequence of iterates to a solution of the original problem) and, at the same time, to guarantee that the iterates stay in the interior of the set K;
– the application of non-quadratic proximal techniques (as in [3, 28, 29]) to the dual of a smooth convex program leads to multiplier methods with twice or higher differentiable augmented Lagrangian functionals. Moreover, in [3] the Hessians of these functionals are bounded.

More motivation of non-quadratic proximal methods can be found in [2, 4, 11, 24]. For infinite-dimensional convex optimization problems such methods have been studied in [1, 8, 7], and for variational inequalities in Hilbert spaces see [6].

The purpose of the present paper is a uniform approach for analyzing convergence of proximal-like methods for solving variational inequalities in Hilbert spaces. In comparison with preceding publications dealing with non-quadratic methods,

– the standard requirement of the strict monotonicity of the operator $\nabla_1 D(\cdot, y)$ (usually formulated as the strict convexity of Bregman's or an other function generating D) is weakened. This leads to an analogy of methods with weak regularization and regularization on a subspace (developed on the basis of the classical PPM in [18, 19]);
– a successive approximation of the set K is included.

Comparing with [6], here the class of operators $\mathcal{Q}$ is also extended (see the case D1 in Lemma 3 and Theorem 2) and the auxiliary problems are supposed to be solved approximately.

The scheme studied here and called generalized proximal point method (GPPM) can be described as follows: Taking a linear monotone operator $\mathcal{B} : X \to X'$ such that $\mathcal{Q} - \mathcal{B}$ is still monotone, we choose a convex continuous functional $h : \bar{S} \to \mathbb{R}$ so that

$$x \to \frac{1}{2}\langle \mathcal{B}x, x\rangle + h(x)$$

possesses properties like usually required for Bregman functions (with a zone S).

At the k-th step of the GPPM, with $x^k \in K^{k-1} \cap S$ obtained at the previous iteration, the iterate $x^{k+1} \in K^k \cap \bar{S}$ is defined such that

$$\begin{aligned} &\exists q(x^{k+1}) \in \mathcal{Q}(x^{k+1}) : \langle q(x^{k+1}) + \chi_k(\nabla h(x^{k+1}) - \nabla h(x^k)), x - x^{k+1}\rangle \\ &\qquad \geqslant -\delta_k \sqrt{\Gamma_1(x, x^{k+1})} \quad \forall x \in K^k \cap \bar{S}. \end{aligned} \tag{1}$$

Here, $\{K^k\}$ is a sequence of convex sets approaching K, $\{\chi_k\}$ as above, $\{\delta_k\}$ is a given non-negative sequence and

$$\Gamma_1(x, y) = \min\{\alpha\|x - y\|^2, \Gamma(x, y) + 1\}, \alpha > 0 - \text{const.}, \tag{2}$$

with

$$\Gamma(x, y) = \frac{1}{2}\langle \mathcal{B}(x - y), x - y\rangle + h(x) - h(y) - \langle \nabla h(y), x - y\rangle \tag{3}$$

considered on $\text{dom}\Gamma = \bar{S} \times D(\nabla h)$ and used below as a Lyapunov function. The choice $\delta_k \equiv 0$ corresponds to the exact variant of the GPPM.

This scheme and the required conditions on h in Section 2 do not exclude the use of quadratic functionals h, in particular, the choice $h(x) = \frac{1}{2}\|x\|^2$ leads to a perturbed version of the classical proximal point method (for this version with more general assumptions w.r.t. data approximation see [15]). Therefore, in the sequel the notion 'non-quadratic' (methods) means, as well as in a series of preceding papers, not only quadratic (methods) and indicates the predominant aspect of investigations.

The paper is organized as follows: In Section 2 conditions w.r.t. the successive approximation of Problem (P) and the regularizing functional are formulated as well as discussed, and the criterion of inexact iterations (1) is analyzed. In Section 3 solvability of the auxiliary problems is studied, and finally in Section 4 convergence of the GPPM is proved.

2. Generalized proximal point method

We make use of the following notations: $S \subset X$ is an open convex set, its closure is denoted by $\bar{S}$; $\{K^k\} \subset X$, $K^k \supset K$, is a family of convex closed sets approximating K;

$$\mathcal{N}_K : y \to \begin{cases} \{z \in X' : \langle z, y - x\rangle \geqslant 0 \quad \forall x \in K\} & \text{if } y \in K \\ \emptyset & \text{otherwise} \end{cases}$$

is the normality operator for K. The symbol $\rightharpoonup$ indicates weak convergence in X. With $\mathcal{B}$ and $h : \bar{S} \to \mathbb{R}$ as introduced, we define the functional

$$\eta(x) = \begin{cases} \frac{1}{2}\langle \mathcal{B}x, x\rangle + h(x) & \text{if } x \in \bar{S} \\ +\infty & \text{otherwise.} \end{cases}$$

Now the following basic assumptions are considered.

ASSUMPTION 1. *(On the successive approximation of Problem (P) and the choice of the controlling parameters)*
(A1) For each k, the operator $\mathcal{Q} + \mathcal{N}_{K^k}$ is maximal monotone;
(A2) $S \cap D(\mathcal{Q}) \cap K^k \neq \emptyset \quad \forall k$;

(A3) *For each k it holds*

$$\langle q(x) - q(y), x - y\rangle \geqslant \langle \mathcal{B}(x-y), x-y\rangle \\ \forall x, y \in D(\mathcal{Q}) \cap K^k, \quad \forall q(\cdot) \in \mathcal{Q}(\cdot),$$

where $\mathcal{B} : X \to X'$ is a given linear continuous and monotone operator with the symmetry property $\langle \mathcal{B}x, y\rangle = \langle \mathcal{B}y, x\rangle$;

(A4) *Any weak limit point of an arbitrary sequence $\{v^k\}$, $v^k \in S \cap D(\mathcal{Q}) \cap K^k$, belongs to $K \cap D(\mathcal{Q})$;*

(A5) *The non-negative sequences $\{\varphi_k\}$ (accuracy of approximation), $\{\chi_k\}$ (regularization parameter) and $\{\delta_k\}$ (exactness for solving the auxiliary problems) satisfy*

$$0 < \chi_k \leqslant 1, \quad \sum_{k=1}^{\infty} \frac{\varphi_k}{\chi_k} < \infty, \quad \sum_{k=1}^{\infty} \frac{\delta_k}{\chi_k} < \infty;$$

(A6) *For some $x^* \in X^* \cap \bar{S}$ and $q^*(x^*) \in \mathcal{Q}(x^*)$ obeying*

$$\langle q^*(x^*), y - x^*\rangle \geqslant 0 \quad \forall y \in K,$$

and for an arbitrary sequence $\{v^k\}$, $v^k \in S \cap D(\mathcal{Q}) \cap K^k$, there exists a sequence $\{w^k(v^k)\} \subset K \cap S$ such that

$$\langle q^*(x^*), w^k(v^k) - v^k\rangle \leqslant c\left(\Gamma(x^*, v^k) + 1\right)\varphi_k \ \ (c \geqslant 0 - const.). \tag{4}$$

Condition A6 seems to be rather artificial, especially, due to the unknown element $q^*(x^*)$. However, for a series of variational inequalities in mechanics and physics, we have a helpful a priori information about $q^*(x^*)$. So, for the problem on a steady movement of a fluid in a domain Ω bounded by a semi-permeable membrane (see [13], Sect. 1) $q^*(x^*) = 0$ has to be.

In the general situation, one can replace (4) by

$$\|w^k(v^k) - v^k\| \leqslant c_1\varphi_k.$$

Because in (4) c is an arbitrary (non-negative) constant, this causes no alterations in the analysis below.

REMARK 1. In [19] and [15], using the functional h such that

$$\exists m > 0 : \ \Gamma(x, y) \geqslant m\|x - y\|^2 \quad \forall x, y,$$

more general approximations have been considered ($K^k \supset K$ is not supposed), mainly inspired by finite element methods in mathematical physics. Here we renounce it in order to avoid too much technicalities.

ASSUMPTION 2. *(Defining the regularizing functional h)*

(B1) $h : \bar{S} \to \mathbb{R}$ *is a convex and continuous functional;*
(B2) h *is Gâteaux-differentiable on* S*;*
(B3) *The functional* η *is strictly convex on* $\bar{S}$*;*
(B4) $X^* \cap \bar{S} \neq \emptyset$*;*
(B5) *The set* $L_1(x, \delta) = \{y \in S : \Gamma(x, y) \leqslant \delta\}$ *is bounded for each* $x \in \bar{S}$ *and each* δ*;*
(B6) *If the sequences* $\{v^k\} \subset S$, $\{y^k\} \subset S$ *converge weakly to* v *and* $\lim_{k\to\infty} \Gamma(v^k, y^k) = 0$, *then*

$$\lim_{k\to\infty} \left[\Gamma(v, v^k) - \Gamma(v, y^k)\right] = 0;$$

(B7) *If* $\{v^k\} \subset S$ *is bounded,* $\{y^k\} \subset S$, $y^k \rightharpoonup \bar{y}$ *and* $\lim_{k\to\infty} \Gamma(v^k, y^k) = 0$, *then* $\lim_{k\to\infty} \|v^k - y^k\| = 0$;
(B8) *If* $\{v^k\} \subset S$, $\{y^k\} \subset S$, $v^k \rightharpoonup v$, $y^k \rightharpoonup y$ *and* $v \neq y$, *then*

$$\underline{\lim}_{k\to\infty} \left|\langle \nabla h(v^k) + \mathcal{B}v^k - \nabla h(y^k) - \mathcal{B}y^k, v - y\rangle\right| > 0;$$

(B9) $\quad \forall z \in X' \; \exists x \in S : \; \nabla h(x) + \mathcal{B}x = z.$

In [18] (Sect. 5), for two problems in elasticity theory the chosen regularizing functionals satisfy the Assumptions 2 and A3, but they are not strictly convex (see, there regularization on the kernel).

As it can be concluded from [7] (Sect. 2.1.2), condition B7 is equivalent to the following sequential consistency property for the functional η on S: for any non-empty bounded subset $E \subset S$ and any $t > 0$

$$\inf_{x\in E} \inf\{\eta(y) - \eta(x) - \langle \nabla\eta(x), y - x\rangle : y \in \bar{S}, \|y - x\| = t\} > 0.$$

For the case $\mathcal{B} = \mathbf{0}$, the totality of conditions B1–B9 is similar to the system of hypothesizes for Bregman functions in [6], only B7 is stronger than the corresponding assumption in the paper mentioned, where $v^k \rightharpoonup \bar{y}$ stands in place of $\lim_{k\to\infty} \|v^k - y^k\| = 0$. In the cases D2 and D3 (see Lemma 2 and Theorem 2 below), this assumption from [6] suffices if $\mathcal{B}$ is a compact operator. At the same time, the use of B7 permits us, in particular, to extend the class of operators $\mathcal{Q}$ by including the case D1.

If $\mathcal{B} = \mathbf{0}$, $X = \mathbb{R}^n$, the conditions B1–B9 can be derived from the standard hypothesizes for Bregman functions (see the analysis in [6], Sect. 7).

The conditions B2 and B3 ensure that $\Gamma(x, y) > 0, \Gamma_1(x, y) > 0$ hold for $x \neq y$, and obviously $\Gamma(x, x) = 0, \Gamma_1(x, x) = 0$.

The consideration of an approximation of K by $\{K^k\}$ addresses, in particular, the situation when K is given in the form $K = K_1 \cap K_2$ and we choose h by taking into account the set K_1 only. In this case $K^k = K_1 \cap K_2^k$ is natural.

Let us give a simple example illustrating the choice of the functional h. Let $X = \mathbb{R}^n$,

$$K = \left\{ x \in \mathbb{R}^n : x_j \geqslant 0,\ j = 1, \ldots, n_1;\ \sum_{j=n_1+1}^{n_2} j|x_j| \leqslant 1 \right\}$$

with $0 < n_1 < n_2 < n$,

$$\mathcal{Q} : x \to (\mathcal{A}(x_1, \ldots, x_{n_1}), x_{n_1+1} - 1, \ldots, x_n - 1),$$

where $\mathcal{A} : \mathbb{R}^{n_1} \to \mathbb{R}^{n_1}$ is an arbitrary continuous and monotone operator such that the corresponding Problem (P) is solvable. Then, considering the approximation

$$K^k = \left\{ x \in \mathbb{R}^n :\ x_j \geqslant 0,\ j = 1, \ldots, n_1,\ \sum_{j=n_1+1}^{n_2} j\sqrt{x_j^2 + \tau_k} \leqslant 1 + \sqrt{\tau_k} \sum_{j=n_1+1}^{n_2} j \right\},$$

where $\tau_k \to +0$, take

$$\mathcal{B} : x \to (0, \ldots, 0, x_{n_1+1}, \ldots, x_n).$$

Since

$$K^k \subset \left\{ x \in \mathbb{R}^n : x_j \geqslant 0,\ \ j = 1, \ldots, n_1;\ \sum_{j=n_1+1}^{n_2} j|x_j| \leqslant 1 + \sqrt{\tau_k} \sum_{j=n_1+1}^{n_2} j \right\},$$

the estimate

$$\min_{z \in K} \|z - v^k\| \leqslant c_1\sqrt{\tau_k} \quad \forall \{v^k\},\ v^k \in K^k,$$

is evident.

Now, it is easy to verify that the choice of $\{K^k\}$,

$$h(x) = \sum_{j=1}^{n_1} x_j \ln x_j - x_j \text{ (with } 0 \times \ln 0 = 0 \text{ by convention)}$$

and

$$S = \{x \in \mathbb{R}^n : x_j > 0,\ j = 1, \ldots, n_1\}$$

satisfies the conditions A1–A4, B1–B9, and A6 is fulfilled with $\varphi_k = \sqrt{\tau_k}$. The second condition in A5 forces $\sum_{k=1}^{\infty} \frac{\sqrt{\tau_k}}{\chi_k} < \infty$.

Now, let us recall the method under consideration.

Generalized proximal point method (GPPM):

Let $x^1 \in S$ be arbitrarily chosen and at the $(k-1)$-th step let $x^k \in K^{k-1} \cap S$ be defined. In the k-th step solve

$$(P^k_{\delta_k}) \qquad \text{find } x^{k+1} \in K^k \cap \bar{S} \text{ such that } \exists\, q(x^{k+1}) \in \mathcal{Q}(x^{k+1}) : \\ \langle q(x^{k+1}) + \chi_k(\nabla h(x^{k+1}) - \nabla h(x^k)), x - x^{k+1}\rangle \\ \geqslant -\delta_k \sqrt{\Gamma_1(x, x^{k+1})} \quad \forall x \in K^k \cap \bar{S}. \tag{5}$$

By (P^k_0), $\bar{x}^{k+1}$ we denote, respectively, Problem (5) with $\delta_k = 0$ and its solution.

The criterion for the approximate calculation of $\bar{x}^{k+1}$ inserted in $(P^k_{\delta_k})$ is not suitable for a straightforward use, but it permits one to extend the convergence results, obtained in this paper, to related algorithms with more reasonable criteria. So, in [12], Eckstein has analyzed different accuracy conditions on the iterates of Bregman–function-based methods for the inclusion

$$\text{find } z \in \mathbb{R}^n : 0 \in \mathcal{T}z, \tag{6}$$

with $\mathcal{T} : \mathbb{R}^n \to 2^{\mathbb{R}^n}$ a maximal monotone operator.

It is well-known that this problem with $\mathcal{T} : X \to 2^{X'}$ is equivalent to the Problem (P) if $\mathcal{T} \equiv \mathcal{Q} + \mathcal{N}_K$ and the operator $\mathcal{Q} + \mathcal{N}_K$ is maximal monotone. The method studied in [12] has the form

$$0 \in \chi_k^{-1}\mathcal{T}(x^{k+1}) + \nabla g(x^{k+1}) - \nabla g(x^k) + e^{k+1}, \tag{7}$$

under rather standard assumptions on a Bregman function g (here, e^{k+1} is an error vector). Such a relaxation of the exact inclusion ((7) given with $e^{k+1} = \mathbf{0}$) is considered in [12] as to be preferable for numerical implementations. Convergence of the iterates x^k generated in (7) to a solution of (6) is established obeying the following conditions on a sequence of errors $\{e^k\}$:

$$\sum_{k=1}^{\infty} \|e^k\| < \infty \tag{8}$$

and

$$\sum_{k=1}^{\infty} \langle e^k, x^k\rangle < \infty. \tag{9}$$

Considering in the sequel method (7) with $\mathcal{T} = \mathcal{Q} + \mathcal{N}_K$ on a pair (X, X'), we have to take $\|\cdot\|_{X'}$ in (8).

- If $K \cap \bar{S}$ is a bounded set*, then (9) follows immediately from (8). In this case, one can easily see that (7) implies the validity of (5) given with an appropriate α in (2) and $\delta_k = \alpha^{-1/2}\chi_k\|e^{k+1}\|_{X'}$, such that (8) provides the condition

* This can be supposed for $K \cap D(\mathcal{Q}) \cap \bar{S}$ instead of $K \cap \bar{S}$.

$\sum_{k=1}^{\infty} \chi_k^{-1}\delta_k < \infty$ in A5 (naturally, for this comparison, we assume that $\mathcal{B} = \mathbf{0}$, $K^k \equiv K$ and $g = h$; but the same arguments suit if $g = h$ in (7) satisfies A3, B1–B3 with $\mathcal{B} \neq \mathbf{0}$).

Indeed, taking $\alpha < (2\,\mathrm{diam}(K \cap \bar{S}))^{-2}$, the inequality

$$\Gamma_1(x, y) = \alpha\|x - y\|^2 \quad \forall x \in K \cap \bar{S},\ y \in K \cap S$$

follows from the definition of Γ_1. Due to the definition of $\mathcal{N}_K$, one can rewrite (7) as

$$\langle \chi_k e^{k+1} + q(x^{k+1}) + \chi_k(\nabla h(x^{k+1}) - \nabla h(x^k)), x - x^{k+1}\rangle \geqslant 0 \quad \forall x \in K,$$

with $q(x^{k+1}) \in \mathcal{Q}(x^{k+1})$, hence

$$\begin{aligned}&\langle q(x^{k+1}) + \chi_k(\nabla h(x^{k+1}) - \nabla h(x^k)), x - x^{k+1}\rangle \\ &\quad \geqslant -\chi_k\|e^{k+1}\|_{X'}\|x - x^{k+1}\| \quad \forall x \in K.\end{aligned}$$

In view of the assumption $\{x^k\} \subset S$ made in [12] (see also Section 3 below), $\Gamma_1(x, x^{k+1}) = \alpha\|x - x^{k+1}\|^2$ holds for $x \in K \cap \bar{S}$, and together with the last inequality this yields

$$\begin{aligned}&\langle q(x^{k+1}) + \chi_k(\nabla h(x^{k+1}) - \nabla h(x^k)), x - x^{k+1}\rangle \\ &\quad \geqslant -\alpha^{-1/2}\chi_k\|e^{k+1}\|_{X'}\sqrt{\Gamma_1(x, x^{k+1})} \quad \forall x \in K \cap \bar{S}.\end{aligned} \tag{10}$$

Therefore, the claim above follows immediately.

– Now, let us trace the situation when K is not bounded, nevertheless condition (8) ensures (5) with $\sum_{k=1}^{\infty} \chi_k^{-1}\delta_k < \infty$ (hence, condition (9) is superfluous). In this case, however, the use of a suitable operator $\mathcal{B} \neq \mathbf{0}$ is in essence.

We suppose that $X = H^1(\Omega)$ (where Ω is an open domain in $\mathbb{R}^n$), that $K \subset \{x \in X : \|x\|_{L_2(\Omega)} \leqslant 1\}$ is an unbounded set in X and $\mathcal{B} : X \to X'$ is given by $\langle \mathcal{B}x, x\rangle = \|\nabla x\|^2_{L_2(\Omega)}$. A similar choice of $\mathcal{B}$ is quite realistic for elliptic problems. In this situation, for any functional h satisfying B1–B3,

$$\begin{aligned}\Gamma(x, y) + 1 &\geqslant \frac{1}{4}\|x - y\|^2_{L_2(\Omega)} + \langle \mathcal{B}(x - y), x - y\rangle \\ &\geqslant \frac{1}{4}\|x - y\|^2_{H^1(\Omega)} \quad \forall x \in K \cap \bar{S},\ y \in K \cap S,\end{aligned}$$

is valid, hence setting $\alpha \leqslant 1/4$ in (2), one gets

$$\Gamma_1(x, y) = \alpha\|x - y\|^2_{H^1(\Omega)} \quad \forall x \in K \cap \bar{S},\ y \in K \cap S.$$

Now, the same arguments as in the case of a bounded set K enable us to conclude that (7) implies (5) with $\delta_k = \alpha^{-1/2}\chi_k\|e^{k+1}\|_{X'}$.

– Return to the general process, assuming that $g = h$ and

$$\Gamma(x, y) \geqslant m\|x - y\|^2 \quad \forall x \in \bar{S},\ y \in S \tag{11}$$

is valid with $m > 0$. Here, boundedness of K is not supposed. However, $\alpha \leqslant m$ ensures

$$\Gamma_1(x, y) = \alpha\|x - y\|^2, \tag{12}$$

and relation (10) follows as above. Hence, condition (8) suffices to conclude that (5) is valid with $\sum_{k=1}^{\infty} \chi_k^{-1}\delta_k < \infty$.

Note that in the particular case $h(x) = \frac{1}{2}\|x\|^2$ we deal with the classical proximal point method, and (8) is nothing else as the known criterion (A') in [27]. Relation (12) holds also true if the functional h is chosen as in methods with weak regularization or regularization on a subspace (see [18, 19]). Thus, the convergence results established below can be applied to these methods in the form (7), (8).

REMARK 2. Eckstein connects the discrepancy in the convergence conditions (8) (for the classical method) and (8),(9) (for non-quadratic proximal methods) with the fact that 'no triangle inequality applies' to Bregman distance $D(x, y) = h(x) - h(y) - \langle \nabla h(y), x - y\rangle$ with non-quadratic h. However, the analysis above shows that probably the 'sufficiency' of criterion (8) depends more on the fulfillment of relation (12) for some $\alpha > 0$.

3. Solvability of Problem $(P^k_{\delta_k})$

In this section we show existence and uniqueness of a solution of Problem (P^k_0), and the inclusion $x^{k+1} \in S$ for a solution of $(P^k_{\delta_k})$.

According to B1, the subdifferential operator $\partial\eta$ is maximal monotone. The conditions B1–B3 and B9 provide that $D(\partial\eta) = S$. Indeed, the inclusion $D(\partial\eta) \supset S$ follows from B2, and assuming that $\partial\eta(x) \neq \emptyset$ holds for some $x \in \bar{S}\backslash S$, in view of B3 we obtain

$$\langle \nabla\eta(y) - \xi(x), y - x\rangle > 0 \quad \forall y \in S,\ \xi(x) \in \partial\eta(x).$$

But, for a fixed $\xi(x) \in \partial\eta(x)$, due to B9, there exists $y \in S$ such that $\nabla\eta(y) = \xi(x)$, in contradiction with the last inequality.

In turns, the conclusion $D(\partial\eta) = S$ means that $D(\nabla h) = S$, and the both operators $\nabla\eta$ and ∇h are maximal monotone.

Thus, if Problem (P^k_0) is solvable, then it has a unique solution, here denoted by $\bar{x}^{k+1}$ (the strict monotonicity of $\mathcal{Q} + \chi_k\nabla h$ on $S \cap K^k \cap D(\mathcal{Q})$ follows immediately from A3 and B3), and $\bar{x}^{k+1} \in S$. Then, of course, the solution x^{k+1} of Problem $(P^k_{\delta_k})$ exists, and $D(\nabla h) = S$ provides $x^{k+1} \in S$.

Because the operator ∇h is maximal monotone and S is an open set, the maximal monotonicity of the operators $\mathcal{Q} + \chi_k\nabla h + \mathcal{N}_{K^k}$ and $x \to \mathcal{Q}(x) + \chi_k\nabla h(x) + \mathcal{N}_{K^k}(x) - \chi_k\nabla h(x^k)$ follows from A1, A2 and A5 according to Theorem 1 in [26].

Since $K^k \cap S \neq \emptyset$, the Moreau-Rockafellar theorem yields

$$\mathcal{N}_{K^k \cap \bar{S}} = \mathcal{N}_{K^k} + \mathcal{N}_{\bar{S}}.$$

Taking into account that $\bar{x}^{k+1}$ (if it exists) belongs to S, this permits to transform Problem (P_0^k) into the inclusion

$$\begin{aligned} 0 &\in \mathcal{Q}(x) + \chi_k \nabla h(x) + \mathcal{N}_{K^k}(x) - \chi_k \nabla h(x^k) \\ &= \mathcal{Q}(x) - \chi_k \mathcal{B}x + \mathcal{N}_{K^k}(x) + \chi_k \mathcal{B}x^k + \chi_k \left(\nabla \eta(x) - \nabla \eta(x^k)\right), \end{aligned}$$

and with regard to A1, A3 and $0 < \chi_k \leqslant 1$, the operator

$$x \to \mathcal{Q}(x) - \chi_k \mathcal{B}x + \mathcal{N}_{K^k}(x) + \chi_k \mathcal{B}x^k$$

is maximal monotone (see Proposition 2.6 in [25]). Now, applying Lemma 5 in [6], one can conclude the solvability of Problem (P_0^k). So, the following statement is proved.

THEOREM 1. *Let the conditions A1–A3, A5 and B1–B3, B9 be valid. Then Problem (P_0^k) is uniquely solvable (for each k), the sequence $\{x^k\}$ is well defined and contained in S.*

REMARK 3. Using instead of B9 the condition (see [14])

$$\{v^k\} \subset S,\ v^k \rightharpoonup v \in \bar{S} \backslash S \implies \lim_{k \to \infty} \langle \nabla h(v^k), y - v^k \rangle = -\infty \quad \forall y \in S,$$

the conclusion $D(\partial \eta) = S$ can be obtained from Lemma 1 in [6], and a result on solvability like Theorem 2 in [6] holds also true.

4. Convergence analysis

In the sequel, we need the following assertion proved in [16].

LEMMA 1. *Let $C \subset X$ be a convex closed set, the operators $\mathcal{A} : X \to 2^{X'}$, $\mathcal{A} + \mathcal{N}_C$ be maximal monotone and $D(\mathcal{A}) \cap C$ be a convex set. Moreover, assume that the operator*

$$\mathcal{A}_C : v \to \begin{cases} \mathcal{A}(v) & if\ v \in C \\ \emptyset & otherwise \end{cases}$$

is locally hemi-bounded at each point $v \in D(\mathcal{A}) \cap C$ and that, for some $u \in D(\mathcal{A}) \cap C$ and each $v \in D(\mathcal{A}) \cap C$, there exists $\zeta(v) \in \mathcal{A}(v)$ satisfying

$$\langle \zeta(v), v - u \rangle \geqslant 0.$$

Then, with some $\zeta \in \mathcal{A}(u)$, the inequality

$$\langle \zeta, v - u \rangle \geqslant 0$$

holds for all $v \in C$.

REMARK 4. Here, a weakened notion of the local hemi-boundedness is supposed. We call an operator $\mathcal{M} : X \to 2^{X'}$ locally hemi-bounded at a point v^0, if for each $v,\ v \neq v^0$, there exists a number $t_0(v^0, v) > 0$ such that the set

$$\bigcup_{0<t\leqslant t_0(v^0,v)} \mathcal{M}(v^0 + t(v - v^0)) \text{ is bounded in } X'.$$

The standard notion supposes the boundedness of

$$\bigcup_{0\leqslant t\leqslant t_0(v^0,v)} \mathcal{M}(v^0 + t(v - v^0)).$$

This relaxation may be significant, see - for instance - the following example: $\mathcal{M} = \mathcal{N}_C$, with $C = \{v \in \mathbb{R}^n : \sum_{i=1}^n v_i^2 \leqslant 1\}$, $n > 1$.

LEMMA 2. *Let the sequence $\{x^k\}$, generated by the GPPM, belong to S and assume that, for some $x^* \in X^* \cap \bar{S}$, condition A6 is valid. Moreover, let the conditions A3, A5 and B1, B2, B5 be fulfilled. Then the sequence $\{\Gamma(x^*, x^k)\}$ is convergent, $\{x^k\}$ is bounded, and* $\lim_{k\to\infty} \Gamma(x^{k+1}, x^k) = 0$.

Proof. We rewrite

$$\Gamma(x^*, x^{k+1}) - \Gamma(x^*, x^k) = s_1 + \chi_k^{-1} s_2 + s_3,$$

with

$$\begin{aligned}
s_1 &= h(x^k) - h(x^{k+1}) + \langle \nabla h(x^k), x^{k+1} - x^k \rangle,\\
s_2 &= \chi_k \langle \nabla h(x^k) - \nabla h(x^{k+1}), x^* - x^{k+1} \rangle,\\
s_3 &= \frac{1}{2}\langle \mathcal{B}(x^{k+1} - x^*), x^{k+1} - x^* \rangle - \frac{1}{2}\langle \mathcal{B}(x^k - x^*), x^k - x^* \rangle.
\end{aligned}$$

In view of $\Gamma_1(x, y) \leqslant \Gamma(x, y) + 1$, A5 and $K^k \supset K$, the inequality

$$\langle q(x^{k+1}) + \chi_k(\nabla h(x^{k+1}) - \nabla h(x^k)), x^* - x^{k+1} \rangle \geqslant -\delta_k \sqrt{\Gamma(x^*, x^{k+1}) + 1}$$

follows immediately from (5). Together with A3, this yields

$$\begin{aligned}
s_2 &\leqslant \langle q(x^{k+1}), x^* - x^{k+1} \rangle + \delta_k \sqrt{\Gamma(x^*, x^{k+1}) + 1}\\
&\leqslant \langle q^*(x^*), x^* - x^{k+1} \rangle - \langle \mathcal{B}(x^* - x^{k+1}), x^* - x^{k+1} \rangle + \delta_k \sqrt{\Gamma(x^*, x^{k+1}) + 1}\\
&\leqslant \langle q^*(x^*), w^{k+1} - x^{k+1} \rangle + \langle q^*(x^*), x^* - w^{k+1} \rangle\\
&\quad - \langle \mathcal{B}(x^* - x^{k+1}), x^* - x^{k+1} \rangle + \delta_k \sqrt{\Gamma(x^*, x^{k+1}) + 1}, \qquad (13)
\end{aligned}$$

where we take $q^*(x^*)$ and $w^{k+1} = w^{k+1}(x^{k+1})$ according to A6. From the definition of x^* and $q^*(x^*)$, one gets

$$\langle q^*(x^*), x^* - w^{k+1} \rangle \leqslant 0, \qquad (14)$$

and (13), (14) and A6 permit us to conclude

$$\begin{aligned} s_2 \leqslant & c(\Gamma(x^*, x^{k+1}) + 1)\varphi_k - \langle \mathcal{B}(x^* - x^{k+1}), x^* - x^{k+1} \rangle \\ & + (1 + \Gamma(x^*, x^{k+1}))\delta_k. \end{aligned}$$

With regard to $0 < \chi_k \leqslant 1$, this leads to

$$\begin{aligned} \chi_k^{-1} s_2 + s_3 \leqslant {} & \Gamma(x^*, x^{k+1}) \left(\frac{c\varphi_k}{\chi_k} + \frac{\delta_k}{\chi_k} \right) + \frac{c\varphi_k}{\chi_k} + \frac{\delta_k}{\chi_k} \\ & - \frac{1}{2} \langle \mathcal{B}(x^* - x^{k+1}), x^* - x^{k+1} \rangle - \frac{1}{2} \langle \mathcal{B}(x^* - x^k), x^* - x^k \rangle. \end{aligned}$$

But

$$\begin{aligned} & \langle \mathcal{B}(x^* - x^{k+1}), x^* - x^{k+1} \rangle + \langle \mathcal{B}(x^* - x^k), x^* - x^k \rangle \\ & \times \frac{1}{2} \langle \mathcal{B}(x^{k+1} - x^k), x^{k+1} - x^k \rangle \end{aligned}$$

is obvious, and taking into account the definition of Γ and the convexity of h,

$$\begin{aligned} s_1 + \chi_k^{-1} s_2 + s_3 \leqslant {} & \Gamma(x^*, x^{k+1}) \left(\frac{c\varphi_k}{\chi_k} + \frac{\delta_k}{\chi_k} \right) + \frac{c\varphi_k}{\chi_k} + \frac{\delta_k}{\chi_k} \\ & - \frac{1}{4} \langle \mathcal{B}(x^{k+1} - x^k), x^{k+1} - x^k \rangle \\ & - \left[h(x^{k+1}) - h(x^k) - \langle \nabla h(x^k), x^{k+1} - x^k \rangle \right] \\ \leqslant {} & \Gamma(x^*, x^{k+1}) \left(\frac{c\varphi_k}{\chi_k} + \frac{\delta_k}{\chi_k} \right) + \frac{c\varphi_k}{\chi_k} + \frac{\delta_k}{\chi_k} - \frac{1}{2} \Gamma(x^{k+1}, x^k). \end{aligned}$$

Hence,

$$\Gamma(x^*, x^{k+1}) \left(1 - \frac{c\varphi_k}{\chi_k} - \frac{\delta_k}{\chi_k} \right) \leqslant \Gamma(x^*, x^k) + \frac{c\varphi_k}{\chi_k} + \frac{\delta_k}{\chi_k} - \frac{1}{2} \Gamma(x^{k+1}, x^k) \tag{15}$$

is valid. But, condition A5 implies the existence of an index k_0 such that

$$\frac{c\varphi_k}{\chi_k} + \frac{\delta_k}{\chi_k} \leqslant \frac{1}{2} \quad \text{for } k \geqslant k_0,$$

i.e.,

$$1 \leqslant \left(1 - \frac{c\varphi_k}{\chi_k} - \frac{\delta_k}{\chi_k} \right)^{-1} \leqslant 1 + 2 \left(\frac{c\varphi_k}{\chi_k} + \frac{\delta_k}{\chi_k} \right) \leqslant 2.$$

Thus, for $k \geqslant k_0$, we obtain from (15) and $\Gamma(x^{k+1}, x^k) \geqslant 0$ that

$$\begin{aligned} \Gamma(x^*, x^{k+1}) \leqslant {} & \left[1 + 2 \left(\frac{c\varphi_k}{\chi_k} + \frac{\delta_k}{\chi_k} \right) \right] \Gamma(x^*, x^k) \\ & + 2 \left(\frac{c\varphi_k}{\chi_k} + \frac{\delta_k}{\chi_k} \right) - \frac{1}{2} \Gamma(x^{k+1}, x^k), \end{aligned} \tag{16}$$

and, in view of condition A5, Lemma 2.2.2 in [23] provides that the sequence $\{\Gamma(x^*, x^k)\}$ is convergent. Now, boundedness of $\{x^k\}$ follows from B5, and using again (16) one gets $\lim_{k\to\infty} \Gamma(x^{k+1}, x^k) = 0$. □

In the sequel, we deal in particular with the case that, besides the usual property of maximal monotonicity, the operator $\mathcal{Q}$ is pseudomonotone and paramonotone. Just this class of operators is considered in [6].

Let us recall the corresponding notions.

DEFINITION 1. The operator $\mathcal{A} : X \to 2^{X'}$ is called *pseudomonotone* if it satisfies the following condition: if $\{v^k\} \subset D(\mathcal{A})$ converges weakly to $v \in D(\mathcal{A})$ and

$$\overline{\lim}_{k\to\infty}\langle w^k, v^k - v\rangle \leqslant 0$$

holds with $w^k \in \mathcal{A}(v^k)$, then for each $y \in D(\mathcal{A})$ there exists $w \in \mathcal{A}(v)$ such that

$$\langle w, v - y\rangle \leqslant \underline{\lim}_{k\to\infty}\langle w^k, v^k - y\rangle.$$

Note that this notion of the pseudomonotonicity (see [21], Sect. 2.2.4 for single-valued operators, where boundedness of $\mathcal{A}$ is also supposed) should not be mixed up with those used in a couple of recent papers on variational inequalities (see, for instance [10]).

DEFINITION 2. (see [9]) The operator $\mathcal{A} : X \to 2^{X'}$ is called *paramonotone*⋆ in a set $C \subset X$ if it is monotone and

$$\langle z - z', v - v'\rangle = 0 \text{ with } v, v' \in C,\ z \in \mathcal{A}(v),\ z' \in \mathcal{A}(v')$$

implies $z \in \mathcal{A}(v'),\ z' \in \mathcal{A}(v)$.

We will use the following property of a paramonotone operator $\mathcal{A}$ in C (cf. [6]): if x^* solves the variational inequality

$$\langle \mathcal{A}(x), y - x\rangle \geqslant 0 \quad \forall y \in C \tag{17}$$

and for $\bar{x} \in C$ there exist $\bar{z} \in \mathcal{A}(\bar{x})$ with $\langle \bar{z}, x^* - \bar{x}\rangle \geqslant 0$, then $\bar{x}$ is also a solution of (17).

LEMMA 3. *Let the assumptions of Lemma 2 as well as the conditions A1, A4 and B6, B7 be valid. Moreover, suppose that one of the following assumptions*⋆⋆ *is fulfilled:*

⋆ Operators with this property have been considered earlier by Bruck [5].

⋆⋆ For a motivation of the inclusion $S \supset \overline{D(\mathcal{Q})}$, which excludes the usual choice of a function h leading to interior point methods, see [11]. In the case D2, condition A4 can be weakened assuming that each limit point of $\{v^k\}$ belongs to K (in place of $K \cap D(\mathcal{Q})$).

(D1) $S \supset \overline{D(\mathcal{Q})}$, *$\nabla h$ is Lipschitz continuous on closed and bounded subsets of S and the conditions of Lemma 1 hold with $\mathcal{A} := \mathcal{Q},\ C := K \cap \bar{S}$;*

(D2) $\mathcal{Q}$ is the subdifferential of a proper convex lower semicontinuous functional f, and f is continuous at some $x \in K$;

(D3) $\mathcal{Q} : D(\mathcal{Q}) \to 2^{X'}$ is pseudomonotone operator and $\mathcal{Q}$ is a paramonotone in $\bar{S}$.

Then the sequence $\{x^k\}$, generated by the GPPM, is bounded and each weak limit point is a solution of Problem (P).

Proof. According to Lemma 2, the sequence $\{x^k\}$ is bounded, hence, there exists a weakly convergent subsequence $\{x^{j_k}\}$, $x^{j_k} \rightharpoonup_{k\to\infty} \bar{x}$. In view of $\{x^k\} \subset S$, A4 and the convexity of S, the inclusion $\bar{x} \in \bar{S} \cap K \cap D(\mathcal{Q})$ ($\bar{x} \in \bar{S} \cap K$ in the case D2) is valid. Due to $\lim_{k\to\infty} \Gamma(x^{k+1}, x^k) = 0$, one can use condition B7 with $v^k := x^{j_k+1},\ y^k := x^{j_k}$. This leads to

$$\lim_{k\to\infty} \|x^{j_k+1} - x^{j_k}\| = 0. \tag{18}$$

If D1 holds, then with regard to the boundedness of $\{x^k\}$, $\{x^k\} \subset D(\mathcal{Q})$, A5 and (18), the relation

$$\lim_{k\to\infty} \chi_{j_k} \langle \nabla h(x^{j_k+1}) - \nabla h(x^{j_k}), x - x^{j_k+1} \rangle = 0 \quad \forall x \in X \tag{19}$$

follows immediately.

Now, take (5) with an arbitrary $x \in K \cap \bar{S}$ and replace $q(x^{k+1})$ by $q(x) \in \mathcal{Q}(x)$ (this is possible in view of the monotonicity of $\mathcal{Q}$). Then, passing to the limit in the obtained inequality with $k := j_k,\ k \to \infty$, due to the boundedness of $\{x^k\}$, the definition of Γ_1, A5 and (19), we obtain

$$\langle q(x), x - \bar{x} \rangle \geqslant 0 \quad \forall x \in K \cap \bar{S}.$$

The conditions A1 and $S \supset \overline{D(\mathcal{Q})}$ guarantee the maximal monotonicity of the operator $\mathcal{Q} + \mathcal{N}_{K\cap\bar{S}}$ (in fact, $\mathcal{Q} + \mathcal{N}_{K\cap\bar{S}}$ coincides with $\mathcal{Q} + \mathcal{N}_K$). Thus, we are able to apply Lemma 1 with $C := K \cap \bar{S},\ \mathcal{A} := \mathcal{Q}$, which ensures that

$$\exists\, q(\bar{x}) \in \mathcal{Q}(\bar{x}) : \langle q(\bar{x}), y - \bar{x} \rangle \geqslant 0 \quad \forall y \in K \cap \bar{S}.$$

This yields

$$0 \in q(\bar{x}) + \mathcal{N}_{K\cap\bar{S}}(\bar{x}) \subset \mathcal{Q}(\bar{x}) + \mathcal{N}_{K\cap\bar{S}}(\bar{x}),$$

hence $0 \in \mathcal{Q}(\bar{x}) + \mathcal{N}_K(\bar{x})$ holds, proving $\bar{x} \in X^*$.

Suppose now that D2 is valid and take x^* as in A6. With regard to the symmetry of $\mathcal{B}$ a straightforward calculation gives

$$\begin{aligned} &- \langle \nabla h(x^k) - \nabla h(x^{k+1}), x^* - x^{k+1} \rangle = \Gamma(x^*, x^k) - \Gamma(x^*, x^{k+1}) \\ &- \Gamma(x^{k+1}, x^k) - \langle \mathcal{B}(x^{k+1} - x^k), x^* - x^{k+1} \rangle. \end{aligned} \tag{20}$$

Using Lemma 2, (18) and $0 < \chi_k \leqslant 1$, we infer from (20) that

$$\lim_{k\to\infty} \chi_{j_k} \langle \nabla h(x^{j_k+1}) - \nabla h(x^{j_k}), x^* - x^{j_k+1} \rangle = 0. \tag{21}$$

But, relation (5) given with $x = x^*$ and $k := j_k$ implies, due to the convexity of f, that

$$\begin{aligned} &\delta_{j_k} \sqrt{\Gamma_1(x^*, x^{j_k+1})} + \chi_{j_k} \langle \nabla h(x^{j_k+1}) - \nabla h(x^{j_k}), x^* - x^{j_k+1} \rangle \\ &\quad \geqslant f(x^{j_k+1}) - f(x^*). \end{aligned} \tag{22}$$

Taking the limit in (22) as $k \to \infty$, due to Lemma 2, (21), A5 and the lower semicontinuity of f, one gets $f(\bar{x}) \leqslant f(x^*)$. Thus, $0 \in \partial\, (f(\bar{x}) + \delta(\bar{x}|K))$, $\delta(\cdot|K)$ the indicator functional of K, and the Moreau–Rockafellar theorem provides $\bar{x} \in X^*$.

Finally, let us consider the case D3. Using equality (20) with $k := j_k$ and $\bar{x}$ in place of x^*, from (18), Lemma 2 and condition B6 for $v^k := x^{j_k+1}$, $y^k := x^{j_k}$, we conclude that

$$\lim_{k\to\infty} \langle \nabla h(x^{j_k+1}) - \nabla h(x^{j_k}), \bar{x} - x^{j_k+1} \rangle = 0.$$

Thus, (5) taken with $x = \bar{x}$ implies

$$\overline{\lim}_{k\to\infty} \langle q(x^{j_k+1}), x^{j_k+1} - \bar{x} \rangle \leqslant 0,$$

and the pseudomonotonicity of $\mathcal{Q}$ provides the existence of $q(\bar{x}) \in \mathcal{Q}(\bar{x})$ such that

$$\langle q(\bar{x}), \bar{x} - x^* \rangle \leqslant \underline{\lim}_{k\to\infty} \langle q(x^{j_k+1}), x^{j_k+1} - x^* \rangle.$$

Now, from (5) and relation (21), which is true also in this case, one gets $\langle q(\bar{x}), \bar{x} - x^* \rangle \leqslant 0$. Therefore, the above mentioned property of paramonotonicity permits to conclude that $\bar{x} \in X^*$. □

THEOREM 2. *Let the conditions A1–A5 and B1–B9 be valid, and condition A6 hold for each $x \in X^* \cap \bar{S}$ (constant c in A6 may depend on x). Moreover, let the operator $\mathcal{Q}$ possess one of the properties D1, D2 or D3 in Lemma 3. Then the sequence $\{x^k\}$, generated by the GPPM, converges weakly to a solution of Problem (P).*

Proof. The existence of the sequence $\{x^k\}$ and the inclusion $\{x^k\} \subset S$ are guaranteed by Theorem 1. Denote

$$d_k(x) = \Gamma(x, x^k) - \frac{1}{2} \langle \mathcal{B}x, x \rangle - h(x).$$

According to Lemma 2, the sequence $\{\Gamma(x, x^k)\}$ converges for each $x \in X^* \cap \bar{S}$, hence, the sequence $\{d_k(x)\}$ possesses the same property.

Boundedness of $\{x^k\}$ was proved in Lemma 2, and Lemma 3 yields that each weak limit point of $\{x^k\}$ belongs to $X^* \cap \bar{S}$.

Assume that $\{x^{j_k}\}$ and $\{x^{i_k}\}$ converge weakly to $\bar{x}, \ \tilde{x}$, correspondingly. Then it holds $\bar{x}, \tilde{x} \in X^* \cap \bar{S}$. Let

$$l_1 = \lim_{k\to\infty} d_k(\bar{x}), \quad l_2 = \lim_{k\to\infty} d_k(\tilde{x}).$$

Obviously,

$$l_1 - l_2 = \lim_{k\to\infty} (d_k(\bar{x}) - d_k(\tilde{x})) = \lim_{k\to\infty} \langle \nabla h(x^k) + \mathcal{B}x^k, \tilde{x} - \bar{x}\rangle.$$

Considering the latter equality now for the subsequences $\{x^{j_k}\}$ and $\{x^{i_k}\}$, one can conclude that

$$\lim_{k\to\infty} \langle \nabla h(x^{j_k}) + \mathcal{B}x^{j_k} - \nabla h(x^{i_k}) - \mathcal{B}x^{i_k}, \tilde{x} - \bar{x}\rangle = 0. \tag{23}$$

A comparison of (23) and B8 (given with $v^k := x^{j_k}, \ y^k := x^{i_k}$) indicates $\tilde{x} = \bar{x}$, proving uniqueness of the weak limit point of $\{x^k\}$. □

REMARK 5. Theorem 1 remains true if condition B9 is replaced by any other condition guaranteeing that $\{x^k\}$ is well defined and $\{x^k\} \subset S$ (see, in particular, Remark 3). If we replace $S \supset \overline{D(\mathcal{Q})}$ by the weaker requirement that

$$S \supset \overline{D(\mathcal{Q}) \cap (\cup_{k\geqslant k_0} K^k)}$$

is valid for an arbitrary large k_0, then — with evident technical alterations — the proofs of Lemma 3 and Theorem 2 hold true.

Acknowledgments

The authors are grateful to an anonymous referee for his valuable and helpful comments and suggestions.

References

1. Alber, Y., R. Burachik, and A. Iusem (1997), A proximal method for non-smooth convex optimization problems in Banach spaces. *Abstr. Appl. Anal.* 2 (1-2), 97–120.
2. Auslender, A. and M. Haddou (1995), An interior-proximal method for convex linearly constrained problems and its extension to variational inequalities. *Math. Programming* 71, 77–100.
3. Auslender, A., M. Teboulle, and S. Ben-Tiba (1999), Interior proximal and multiplier methods based on second order homogeneous kernels. *Mathematics of Oper. Res.* 24, 645–668.
4. Ben-Tal, A. and M. Zibulevski (1997), Penalty-barrier methods for convex programming problems. *SIAM J. Optim.* 7, 347–366.

5. Bruck, R. E. (1975), An iterative solution of a variational inequality for certain monotone operators in Hilbert space. *Bulletin of the American Mathem. Society* 81, 890–892.
6. Burachik, R. and A. Iusem (1998), A generalized proximal point algorithm for the variational inequality problem in a Hilbert space. *SIAM J. Optim.* 8, 197–216.
7. Butnariu, D. and A. Iusem (1997), On a proximal point method for convex optimization in Banach spaces. *Numer. Funct. Anal. and Optimiz.* 18(7-8), 723–744.
8. Butnariu, D. and A. Iusem (2000), *Totally Convex Functions for Fixed Points Computation and Infinite Dimensional Optimization*. Kluwer, Dordrecht-Boston-London.
9. Censor, Y., A. Iusem, and S. A. Zenios (1998), An interior point method with Bregman functions for the variational inequality problem with paramonotone operators. *Math. Programming* 81, 373–400.
10. Crouzeix, J.-P. (1997), Pseudo-monotone variational inequality problems: Existence of solutions. *Math. Programming* 78, 305–314.
11. Eckstein, J. (1993), Nonlinear proximal point algorithms using Bregman functions, with application to convex programming. *Math. Oper. Res.* 18, 202–226.
12. Eckstein, J. (1998), Approximate iterations in Bregman-function-based proximal algorithms. *Math. Programming* 83, 113–123.
13. Glowinski, R., J. L. Lions, and R. Trémolières (1981), *Numerical Analysis of Variational Inequalities*. North-Holland, Amsterdam.
14. Iusem, A. (1995), On some properties of generalized proximal point methods for quadratic and linear programming. *JOTA* 85, 593–612.
15. Kaplan, A. and R. Tichatschke, A general view on proximal point methods to variational inequalities in Hilbert spaces - iterative regularization and approximation. *J. of Nonlinear and Convex Analysis* (to appear).
16. Kaplan, A. and R. Tichatschke (2000), Auxiliary problem principle and the approximation of variational inequalities with non-symmetric multi-valued operators. *CMS Conference Proc.* 27, 185-209.
17. Kaplan, A. and R. Tichatschke (1994), *Stable Methods for Ill-Posed Variational Problems - Prox-Regularization of Elliptic Variational Inequalities and Semi-Infinite Optimization Problems*. Akademie Verlag Berlin.
18. Kaplan, A. and R. Tichatschke (1997), Prox-regularization and solution of ill-posed elliptic variational inequalities. *Applications of Mathematics* 42, 111–145.
19. Kaplan, A. and R. Tichatschke (2000), Proximal point approach and approximation of variational inequalities. *SIAM J. Control Optim.* 39, 1136–1159.
20. Kiwiel, K. (1999), A projection-proximal bundle method for convex nondifferentiable minimization. In: M. Thera and R. Tichatschke (eds.): *Ill-posed Variational Problems and Regularization Techniques*. pp. 137–150.
21. Lions, J. L. (1969), *Quelques Méthodes de Résolution des Problèmes aux Limites non Linéaires*. Dunod, Gauthier-Villars, Paris.
22. Martinet, B. (1970), Régularisation d'inéquations variationelles par approximations successives. *RIRO* 4, 154–159.
23. Polyak, B. T. (1987), *Introduction to Optimization*. Optimization Software, Inc. Publ. Division, New York.
24. Polyak, R. and M. Teboulle (1997), Nonlinear rescaling and proximal-like methods in convex optimization. *Math. Programming* 76, 265–284.
25. Renaud, A. and G. Cohen (1997), An extension of the auxiliary problem principle to nonsymmetric auxiliary operators. *ESAIM: Control, Optimization and Calculus of Variations* 2, 281–306.
26. Rockafellar, R. T. (1970), On the maximality of sums of nonlinear monotone operators. *Trans. Amer. Math. Soc.* 149, 75–88.

27. Rockafellar, R. T. (1976), Monotone operators and the proximal point algorithm. *SIAM J. Control Optim.* 14, 877–898.
28. Teboule, M. (1992), Entropic proximal mappings with application to nonlinear programming. *Mathematics of Oper. Res.* 17, 670–690.
29. Tseng, P. and D. Bertsekas (1993), On the convergence of the exponential multiplier method for convex programming. *Math. Programming* 60, 1–19.