A Relaxed-Projection Splitting Algorithm for Variational Inequalities in Hilbert Spaces

J.Y. Bello Cruz*   R. Díaz Millán†

December 13, 2013

Abstract

We introduce a relaxed-projection splitting algorithm for solving variational inequalities in Hilbert spaces for the sum of nonsmooth maximal monotone operators, where the feasible set is defined by a nonlinear and nonsmooth continuous convex function inequality. In our scheme, the orthogonal projections onto the feasible set are replaced by projections onto separating hyperplanes. Furthermore, each iteration of the proposed method consists of simple subgradient-like steps, which does not demand the solution of a nontrivial subproblem, using only individual operators, which explores the structure of the problem. Assuming monotonicity of the individual operators and the existence of solutions, we prove that the generated sequence converges weakly to a solution.

Keywords: Point-to-set operator, Projection method, Relaxed method, Splitting methods, Variational inequality problem, Weak convergence.

Mathematical Subject Classification (2008): 90C47, 49J35.

1 Introduction

Let $\mathcal{H}$ be a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. For C, a nonempty, convex and closed subset of $\mathcal{H}$, we define the orthogonal projection of x onto C, $P_C(x)$, as the unique point in C such that $\|P_C(x) - x\| \leq \|y - x\|$ for all $y \in C$. Recall that an operator $T : \mathcal{H} \rightrightarrows \mathcal{H}$ is monotone if, for all $(x, u), (y, v) \in \operatorname{Gr}(T)$, we have $\langle x - y, u - v \rangle \geq 0$, and it is maximal if T has no proper monotone extension in the graph inclusion sense.

In this paper, we present a relaxed-projection splitting algorithm for solving the variational inequality problem for T and C, with T as a sum of m nonsmooth maximal monotone operators, i.e, $T = T_1 + T_2 + \cdots + T_m$ where $T_i : \mathcal{H} \rightrightarrows \mathcal{H}$, $(i = 1, 2, \ldots, m)$ and C is given of the following

*IME - Federal University of Goiás, Goiânia, GO, Brazil. E-mail: yunier.bello@gmail.com
†Federal Institute of Education, Science and Technology, Goiânia, GO, Brazil. E-mail: rdiazmillan@gmail.com

form: $C := \{x \in \mathcal{H} : c(x) \leq 0\}$ where $c : \mathcal{H} \to \mathbb{R}$ is a continuous and convex function, possibly nondifferentiable.

It is clear that, if T_i, $(i = 1, \dots, m)$ are monotone, then $T = T_1 + T_2 + \cdots + T_m$ is also monotone. But if T_i, $(i = 1, 2, \dots, m)$ are maximal, it does not necessarily imply, that T is maximal even when $\mathrm{dom}(T)$ is nonempty. Some additional condition is needed, since for example the graph of T can be even empty (as happens when $\mathrm{dom}(T) = \mathrm{dom}(T_1) \cap \mathrm{dom}(T_2) \cap \cdots \cap \mathrm{dom}(T_m) = \varnothing$).

The problem of determining conditions under which the sum is maximal, turns out to be of fundamental importance in the theory of monotone operators. Results in this directions were proved in [38]. It is clear that in our case ($\mathrm{dom}(T_i) = \mathcal{H}$, $i = 1, 2, \dots, m$) all these sufficient conditions for establishing the maximality of T, are satisfied.

Now, we recall the formulation of the variational inequality problem for T and C, namely:

$$\text{Find } \; x_* \in C \; \text{ such that } \; \exists\, u_* \in T(x_*), \text{ with } \; \langle u_*, x - x_* \rangle \geq 0 \qquad \forall\, x \in C. \tag{1}$$

The solution set of problem (1) is denoted by S_*. This problem has been a classical subject in economics, operations research and mathematical physics, particularly in the calculus of variations associated with the minimization of infinite-dimensional functionals; see, for instance, [29] and the references therein. It is closely related to many problems of nonlinear analysis, such as optimization, complementarity and equilibrium problems and finding fixed points; see, for instance, [22, 29, 44]. Many methods have been proposed to solve problem (1), with T point-to-point; see [27, 31, 42, 43], and for T point-to-set; see [3, 25, 26, 30]. An excellent survey of methods for variational inequality problems can be found in [21].

Variational inequality problems are related to inclusion problems. In fact, when the feasible set is the whole space, the variational inequality problem may be formulated as an inclusion problem. In this work, we are interested in methods that explore the structure of T. This kind of methods are called splitting, since each iteration involves only the individual operators, but not the sum. Many splitting algorithms have been proposed in order to solve this type of inclusion problem; see [2, 18, 19, 33, 34, 36, 45, 47] and the references therein. However, in all of them, the resolvent operator of any individual operator, must be evaluated in each iteration. It is important to mention, that this proximal-like iteration is a nontrivial problem, which demands hard work from computational point of view. Our algorithm avoids this difficulty, replacing proximal-like iterations by subgradient-like projection steps. This represents a significant advantage in both implementational and theoretical senses.

Another weakness which appears in almost all methods solving problem (1), is the necessity to compute the exact projection onto C. This limits the applicability of the methods, especially when such a projection is computationally difficult to evaluate. It is well known that only in a few specific instances the projection onto a convex set has an explicit formula. When the feasible set of problem (1) is a general closed convex set, C, we must solve a nontrivial quadratical problem, in order to compute the projection onto C. This difficulty also appears when the feasible set of problem (1) is expressed as the solution set of another problem, as in this paper. In this kind of problems, it is very hard to find the projection onto the feasible set or even find a feasible point. One option for avoiding this difficulty, consists in replacing at each iteration, the projection onto C, by the projection onto halfspaces containing the given set C and not the current point.

For variational inequality problems, the above approach was introduced in [23], for point-to-point and strongly monotone operators. Other schemes have been proposed, in order to improve the convergence results, without doing exact projections onto C: in [7, 10] for point-to-set and paramonotone operators; in [9] for point-to-point and monotone operators. Algorithms using similar ideas may be found in [15, 16].

In this work, we propose a new splitting scheme for solving problem (1), in which the orthogonal projections onto the feasible set, are replaced by projections onto separating hyperplanes and only simple subgradient-like steps are performed. Assuming maximal monotonicity of the individual operators and existence of solutions, we establish the weak convergence to a solution, of the whole generated sequence.

Our method was inspired by the incremental subgradient method for nondifferentiable optimization, proposed in [35], and it uses an idea similar to the ideas exposed in [5, 7, 10]. In the case of only one operator, it is known that a natural extension of the subgradient iteration (one step), the convergence fails in general for monotone operators; see [9, 10]. However, as will be shown, we introduce an extra step in order to prove the weak convergence of the sequence generated by our algorithm.

The paper is organized as follows. The next section provides some notation and preliminary results that will be used in the remainder of this paper. The relaxed splitting method is presented in Section 3 and Subsection 3.1 contains the convergence analysis of the algorithm. Section 4 contains some discussion on the assumptions with examples showing the effectiveness of our scheme.

2 Preliminary Result

In this section, we present some definitions and results that are needed for the convergence analysis of the proposed method. First, we state two well known facts on orthogonal projections.

Lemma 2.1. *Let S be any nonempty closed and convex set in $\mathcal{H}$, and P_S the orthogonal projection onto S. For all $x, y \in \mathcal{H}$ and $z \in S$, the following properties hold:*

(a) $\|P_S(x) - P_S(y)\| \leq \|x - y\|$.

(b) $\langle x - P_S(x), z - P_S(x)\rangle \leq 0$.

Proof. See Lemmas 1.1 and 1.2 in [46]. □

We next deal with the so called quasi-Fejér convergence and its properties.

Definition 1. *Let S be a nonempty subset of $\mathcal{H}$. A sequence $(x^k)_{k\in\mathbb{N}}$ in $\mathcal{H}$ is said to be quasi-Fejér convergent to S if and only if for all $x \in S$ there exist $k_0 \geq 0$ and a sequence $(\delta_k)_{k\in\mathbb{N}}$ in $\mathbb{R}_+$ such that $\sum_{k=0}^{\infty} \delta_k < \infty$ and*

$$\|x^{k+1} - x\|^2 \leq \|x^k - x\|^2 + \delta_k,$$

for all $k \geq k_0$.

This definition originates in [20] and has been elaborated further in [17, 28].

Proposition 2.1. *If $(x^k)_{k\in\mathbb{N}}$ is quasi-Fejér convergent to S, then:*

(a) *$(x^k)_{k\in\mathbb{N}}$ is bounded.*

(b) *If all weak cluster point of $(x^k)_{k\in\mathbb{N}}$ belong to S, then the sequence $(x^k)_{k\in\mathbb{N}}$ is weakly convergent.*

Proof. See Theorem 4.1 in [28]. □

The next lemma will be useful for proving that the sequence generated by our algorithm, converges weakly to some point belonging to the solution set of problem (1).

Lemma 2.2. *If a sequence $(x^k)_{k\in\mathbb{N}}$ is quasi-Fejér convergent to a closed and convex set S, then the sequence $(P_S(x^k))_{k\in\mathbb{N}}$ is strongly convergent.*

Proof. See Lemma 2 in [7]. □

We also need the following results on maximal monotone operators and monotone variational inequalities.

Lemma 2.3. *Let $T : \mathcal{H} \rightrightarrows \mathcal{H}$ be a maximal monotone operator and C a closed and convex set. Then, the solution set of problem (1), S_*, is closed and convex.*

Proof. See Lemma 2.4(ii) in [11]. □

The next lemma will be useful for proving that all weak cluster points of the sequence generated by our algorithm belong to the solution set of problem (1).

Lemma 2.4. *Consider the variational inequality problem for T and C. If $T : \mathcal{H} \rightrightarrows \mathcal{H}$ is maximal monotone, then*

$$S_* = \{x \in C \,:\, \langle v, y - x\rangle \geq 0\,,\, \forall\, y \in C,\;\; \forall\, v \in T(y)\}.$$

Proof. See Lemma 3 in [41]. □

The next lemma provides a computable upper bound for the distance from a point to the feasible set C.

Lemma 2.5. *Let $c : \mathcal{H} \to \mathbb{R}$ be a convex function and $C := \{x \in \mathcal{H} : c(x) \leq 0\}$. Assume that there exists $w \in C$ such that $c(w) < 0$. Then, for all $y \in \mathcal{H}$ such that $c(y) > 0$, we have*

$$\operatorname{dist}(y, C) \leq \frac{\|y - w\|}{c(y) - c(w)} \, c(y) \, .$$

Proof. See Lemma 4 in [7]. □

Now, the next proposition will be useful for calculating the projections onto the halfspaces that will appear in our algorithm.

Proposition 2.2. *Let $f : \mathcal{H} \to \mathcal{H}$ a convex function and given $y, z, w \in \mathcal{H}$ and $v \in \partial f(z)$. Define $W_{z,w} := \{x \in \mathcal{H} : \langle x - z, w - z \rangle \leq 0\}$ and $C_z := \{x \in \mathcal{H} : f(z) + \langle v, x - z \rangle \leq 0\}$. Then,*

$$P_{C_z \cap W_{z,w}}(w) = w + \max\{0, \lambda_1\} v + \lambda_2 (w - z),$$

where λ_1, λ_2 are solution of the linear system:

$$\begin{aligned} \lambda_1 \|v\|^2 + \lambda_2 \langle v, w - z \rangle &= -\langle v, w - z \rangle - f(z) \\ \lambda_1 \langle v, w - z \rangle + \lambda_2 \|w - z\|^2 &= -\|w - z\|^2 \end{aligned}$$

and

$$P_{C_z}(y) = y - \max\left\{0, \frac{f(z) + \langle v, y - z \rangle}{\|v\|^2}\right\} v.$$

Proof. See Proposition 28.19 in [4]. □

Finally, we will need the following elementary result on sequence averages.

Proposition 2.3. *Let $(p^k)_{k \in \mathbb{N}} \subset \mathcal{H}$ be a sequence strongly convergent to $\tilde{p}$. Take nonnegative real numbers $(\zeta_{k,j})_{k \in \mathbb{N}, 0 \leq j \leq k}$ such that $\lim_{k \to \infty} \zeta_{k,j} = 0$ for all $0 \leq j \leq k$ and $\sum_{j=0}^{k} \zeta_{k,j} = 1$ for all $k \in \mathbb{N}$. Define*

$$w^k := \sum_{j=0}^{k} \zeta_{k,j} p^j .$$

Then $(w^k)_{k \in \mathbb{N}}$ also converges strongly to $\tilde{p}$.

Proof. See Proposition 3 in [7].

□

3 A Relaxed-Projection Splitting Method

In this section, we introduce an algorithm for solving the variational inequality problem when $T = T_1 + T_2 + \ldots + T_m$ and C is of the form

$$C = \{x \in \mathcal{H} \,:\, c(x) \le 0\}, \tag{2}$$

where $c : \mathcal{H} \to \mathbb{R}$ is a continuous convex function. Differentiability of c is not assumed and therefore the representation (2) is rather general, since any system of inequalities $c_j(x) \le 0$ with $j \in J$, where each c_j is convex, may be represented as in (2) with $c(x) = \sup\{c_j(x) \,:\, j \in J\}$.
Consider an exogenous sequence $(\alpha_k)_{k\in\mathbb{N}}$ in $\mathbb{R}_{++}$. Then the algorithm is defined as follows.

Algorithm A

Initialization step: Take $x^0 \in \mathcal{H}$ and $\theta > 0$. Define $z^0 := x^0$ and $\sigma_0 := \alpha_0$.

Iterative step: Given z^k. If $c(z^k) \le 0$, then take $z_0^k := z^k$ and $C_k = \{x \in \mathcal{H} : \langle g^k, x - z_0^k\rangle \le 0\}$ where $g^k \in \partial c^+(z_0^k)$ with $c^+(x) = \max\{0, c(x)\}$. Else, perform the following inner loop:

Inner Loop: Take $y^{k,0} = z^k$. Compute the points $y^{k,1}$, $y^{k,2}$, ... Defining

$$y^{k,j+1} := P_{C_{k,j}\cap W_{k,j}}(y^{k,0}), \tag{3}$$

where

$$C_{k,j} := \{x \in \mathcal{H} \,:\, c(y^{k,j}) + \langle g^{k,j}, x - y^{k,j}\rangle \le 0\}, \tag{4}$$

and

$$W_{k,j} := \{x \in \mathcal{H} \,:\, \langle x - y^{k,j}, y^{k,0} - y^{k,j}\rangle \le 0\}, \tag{5}$$

with $g^{k,j} \in \partial c(y^{k,j})$. Stop the inner loop when $\operatorname{dist}(y^{k,j+1}, C) \le \theta\alpha_k$, put $j = j(k)$.

Set

$$C_k := C_{k,j(k)}, \tag{6}$$

and

$$z_0^k := y^{k,j(k)+1}. \tag{7}$$

Compute the cycle, from $i = 1, 2, \ldots, m$, as follows

$$z_i^k = P_{C_k}\left(z_{i-1}^k - \alpha_k u_i^k\right), \tag{8}$$

where $u_i^k \in T_i(z_{i-1}^k)$. The vector z^{k+1} is obtained doing $z^{k+1} = z_m^k$. Define

$$\sigma_k := \sigma_{k-1} + \alpha_k, \tag{9}$$

and

$$x^{k+1} := \left(1 - \frac{\alpha_k}{\sigma_k}\right) x^k + \frac{\alpha_k}{\sigma_k} z^{k+1}. \tag{10}$$

Before the formal analysis of the convergence properties of Algorithm A, we make some comments and discuss about the assumptions. First, unlike other projection methods, Algorithm A generates a sequence $(x^k)_{k\in\mathbb{N}}$ which is not necessarily contained in the set C. As will be shown in the next subsection, the generated sequence is asymptotically feasible and, in fact, converges to some point in the solution set. We observe that the inner loop in (3)-(7), starts with the point z^k and ends with a point z_0^k close to C, in fact $\operatorname{dist}(z_0^k, C) \leq \theta\alpha_k$, this is possible since the inner loop, in the step k, is a direct application of Algorithm A in [8], with $C = \mathcal{H}$, $x^0 = y^{k,0}$, $f(x) = c^+(x) := \max\{0, c(x)\}$ and $f_* := \inf_{x\in\mathcal{H}} c^+(x) = 0$. Recently has been proposed in [6] a restricted memory level bundle method improving the convergence result of [8], which it to be used into the inner loop accelerating its convergence.

It might seem that this inner loop can be replaced by any finite procedure leading to an approximation of $P_C(z^k)$, say a point z_0^k such that $\|z_0^k - P_C(z^k)\|$ is sufficiently small. This is not the case: In the first place, depending on the location of the intermediate hyperplanes $C_{k,j}$ and $W_{k,j}$, the sequence $(y^{k,j})_{j\in\mathbb{N}}$ may approach points in C far from $P_C(z^k)$; in fact the computational cost of our inner loop is lower than the computation of an inexact orthogonal projection of z^k onto C. On the other hand, it is not the case that any point z close enough to $P_C(z^k)$ will do the job. The crucial relation for convergence of our method is $\|z_0^k - x\| \leq \|z^k - x\|$ for all $x \in C$, which may fail if we replace z_0^k by points z arbitrarily closed to $P_C(z^k)$.

Algorithm A is easily implementer, since $P_{W_{k,j}\cap C_{k,j}}$ and P_{C_k}, given in (3) and (7) respectively, have easy formulae by proposition 2.2.
Hence, by Proposition 2.2 the projections onto $C_{k,j} \cap W_{k,j}$ in the inner loop (3)-(7) and C_k in (8), can be calculated explicitly. Therefore, Algorithm A may be considered as an explicit method, since it do not solve a nontrivial subproblem.
We need the following boundedness assumptions on ∂c.

(H1) ∂c is bounded on bounded sets.

In finite-dimensional spaces, this assumption is always satisfied in view of Theorem 4.6.1(ii) in [14], due to the maximality of ∂c. The maximality has been proved in [39]. For some equivalences with condition (H1), see for instance Proposition 16.17 in [4]. Moreover in the literature, (H1) has been considered as the convergence analysis of various methods solving optimization problems in infinite-dimensional spaces; see, for instance, [1,8,37]. We only use this assumption for establishing the well definition of the inner loop.

(H2) Define

$$\eta_k := \max_{1\leq i\leq m} \{1, \|u_i^k\|\}, \tag{11}$$

with $u_i^k \in T_i(z_{i-1}^k)$. Then, assuming that the stepsize sequence, $(\alpha_k)_{k\in\mathbb{N}}$, satisfies:

$$\sum_{k=0}^{\infty} \alpha_k = \infty, \tag{12}$$

$$\sum_{k=0}^{\infty} (\eta_k \alpha_k)^2 < \infty. \tag{13}$$

We mention that in the analysis of [35], a stronger condition than (H2) is required for proving convergence of the incremental subgradient method. Recently, similar assumptions have been used in the convergence analysis in [2, 12].

The condition (12) (divergent-series) on the stepsizes has been used widely for the convergence of classical projected subgradient methods; see [1, 37].

The condition (13) is used for establishing Proposition 3.3, which implies boundedness of the sequence $(z^k)_{k\in\mathbb{N}}$. When $(\alpha_k)_{k\in\mathbb{N}}$ is in $\ell_2(\mathbb{N})$, the condition (13) holds, assuming that the image of T_i, $(i = 1, 2, \dots, m)$ are bounded. Furthermore, it is possible to assume a weaker sufficient condition for (13) as for example: if $(\alpha_k)_{k\in\mathbb{N}} = (1/k)_{k\in\mathbb{N}}$, then the sequence $(\eta_k)_{k\in\mathbb{N}}$, defined in (11), may be unlimited like the sequence $(k^s)_{k\in\mathbb{N}}$ for any $s \in (0, 1/2)$.

3.1 Convergence Analysis

Before establishing convergence of Algorithm A, we need to ascertain the validity of the stopping criterion as well as the fact that Algorithm A is well defined.

Proposition 3.1. *Take C, $C_{k,j}$, $W_{k,j}$, W_k and C_k defined by Algorithm A. Then,*

(a) $C \subseteq C_{k,j} \cap W_{k,j}$, $C \subseteq C_k$ and $z_i^k \in C_k$ for all k, j and $i = 0, 1, \dots, m$.

(b) $j(k)$ is well defined.

Proof. (a): It follows from (4) and the definition of the subdifferential that $C \subseteq C_{k,j}$ for all k and j. Note that for all $y^{k,j} \notin C$, we have $\partial f(y^{k,j}) = \partial c(y^{k,j})$, thus $C \subseteq C_k$ by (6). Using Proposition 4 and Corollary 1 of [8], with $C = \mathcal{H}$, $f(x) = c^+(x) := \max\{0, c(x)\}$ implying that $f_* = 0$, since our $C \neq \varnothing$, we get $C \subseteq W_{k,j}$ for all k, j. By (3) and (7), we have $z_0^k \in C_k$ and by (8), $z_i^k \in C_k$ for all k, $(i = 1, \dots, m)$.

(b): Regarding the projection step in (3) and (8), item (a) shows that the projections onto $C_{k,j} \cap W_{k,j}$ and C_k are well defined. Using Theorem 2 of [8], with $C = \mathcal{H}$, $x^0 = y^{k,0}$, $f(x) = c^+(x) := \max\{0, c(x)\}$, we have that $(y^{k,j})_{j\in\mathbb{N}}$ converges strongly to $P_C(y^{k,0}) \in C$. Thus, $j(k)$ is well defined. □

A useful proposition for the convergence of algorithm is:

Proposition 3.2. *Let $(z^k)_{k\in\mathbb{N}}$ and $(z_i^k)_{k\in\mathbb{N}}$, with $i = 0, 1, \cdots, m$ be sequences generated by Algorithm A. Then*

(a) $\|z_j^k - z_i^k\| \le (j-i)\eta_k\alpha_k$, *for all $k \in \mathbb{N}$ and $0 \le i \le j \le m$.*

(b) *For any $x \in C$ and $u \in T(x)$ such that $u = \sum_{i=1}^m u_i$ with $u_i \in T_i(x)$, $(i = 1, 2, \ldots, m)$. Then,*

$$\|z^{k+1} - x\|^2 \le \|z^k - x\|^2 + m\left[(\eta_k\alpha_k)^2 + (m-1)\eta\eta_k\alpha_k^2\right] - 2\alpha_k\langle u, z_0^k - x\rangle,$$

where $\eta := \max_{1\le i\le m}\|u_i\|$.

Proof. (a): Since $z_i^k \in C_k$ for all $0 \le i \le m$ and all k, taking any $u_j^k \in T_j(z_{j-1}^k)$ and using the Cauchy-Schwarz inequality, we have

$$\begin{aligned} \|z_j^k - z_i^k\| &= \left\|P_{C_k}\left(z_{j-1}^k - \alpha_k u_j^k\right) - P_{C_k}\left(z_i^k\right)\right\| \\ &\le \|z_{j-1}^k - z_i^k - \alpha_k u_{j,k}\| \le \|z_{j-1}^k - z_i^k\| + \|u_j^k\|\alpha_k \le \cdots \le (j-i)\,\eta_k\alpha_k. \end{aligned}$$

(b): Take $(x, u) \in \mathrm{Gr}(T)$ with $u = \sum_{i=1}^m u_i$ and $u_i \in T_i(x)$, $(i = 1, 2, \ldots, m)$. Using Lemma 2.1(a) in the first inequality, and the monotonicity of each component operator T_i in the latter, we obtain,

$$\begin{aligned} \|z_i^k - x\|^2 &= \left\|P_{C_k}\left(z_{i-1}^k - \alpha_k u_i^k\right) - P_{C_k}(x)\right\|^2 \le \left\|\left(z_{i-1}^k - \alpha_k u_i^k\right) - x\right\|^2 \\ &= \|z_{i-1}^k - x\|^2 + (\|u_i^k\|\alpha_k)^2 - 2\alpha_k\,\langle u_i^k, z_{i-1}^k - x\rangle \\ &\le \|z_{i-1}^k - x\|^2 + (\|u_i^k\|\alpha_k)^2 - 2\alpha_k\,\langle u_i, z_{i-1}^k - x\rangle, \end{aligned}$$

for all k and $i = 1, 2, \ldots, m$.

By summing the above inequalities over $i = 1, 2, \ldots, m$ and using (11), we get

$$\begin{aligned} \|z^{k+1} - x\|^2 &\le \|z_0^k - x\|^2 + m(\eta_k\alpha_k)^2 - 2\,\alpha_k\sum_{i=1}^m\langle u_i, z_{i-1}^k - x\rangle \\ &= \|z_0^k - x\|^2 + m(\eta_k\alpha_k)^2 - 2\alpha_k\sum_{i=1}^m\Big(\langle u_i, z_0^k - x\rangle + \langle u_i, z_{i-1}^k - z_0^k\rangle\Big) \\ &= \|z_0^k - x\|^2 + m(\eta_k\alpha_k)^2 - 2\alpha_k\langle u, z_0^k - x\rangle - 2\alpha_k\sum_{i=1}^m\langle u_i, z_{i-1}^k - z_0^k\rangle. \end{aligned}$$

Using the Cauchy-Schwarz inequality and item (a) for $j = m$ and $i = 0$, we have

$$\begin{aligned}\|z^{k+1} - x\|^2 &\le \|z_0^k - x\|^2 + m(\eta_k\alpha_k)^2 - 2\alpha_k\langle u, z_0^k - x\rangle + 2\alpha_k \sum_{i=1}^m \|u_i\|\|z_{i-1}^k - z_0^k\| \\ &\le \|z_0^k - x\|^2 + m(\eta_k\alpha_k)^2 + 2\alpha_k^2 \sum_{i=1}^m (i-1)\|u_i\|\eta_k - 2\alpha_k\langle u, z^k - x\rangle \\ &\le \|z^k - x\|^2 + m(\eta_k\alpha_k)^2 + m(m-1)\eta\eta_k\alpha_k^2 - 2\alpha_k\langle u, z_0^k - x\rangle,\end{aligned}$$

where the last inequality is a direct consequence of the fact that z_0^k is obtained by the inner loop (3)-(7) and defining $\eta = \max_{1\le i\le m}\|u_i\|$ thus proving the proposition. □

We continue by proving the quasi-Fejér properties of the sequences $(z^k)_{k\in\mathbb{N}}$ generated by Algorithm A. From now on, we assume that the solution set, S_*, of problem (1) is nonempty.

Proposition 3.3. *The sequences $(z^k)_{k\in\mathbb{N}}$ are quasi-Fejér convergent to S_*.*

Proof. Take $\bar{x} \in S_*$. Then, there exists $\bar{u} \in T(\bar{x})$ such that

$$\langle \bar{u}, x - \bar{x}\rangle \ge 0 \quad \forall\, x \in C, \tag{14}$$

where $\bar{u} = \sum_{i=1}^m \bar{u}_i$, with $\bar{u}_i \in T_i(\bar{x})$, $(i = 1, 2, \ldots, m)$. Using now Proposition 3.2(b), taking $\bar{\eta} = \max_{1\le i\le m}\|\bar{u}_i\|$, we get

$$\begin{aligned}\|z^{k+1} - \bar{x}\|^2 &\le \|z^k - \bar{x}\|^2 + m\left[(\eta_k\alpha_k)^2 + (m-1)\bar{\eta}\eta_k\alpha_k^2\right] - 2\alpha_k\langle \bar{u}, z_0^k - x\rangle \\ &= \|z^k - \bar{x}\|^2 + m\left[(\eta_k\alpha_k)^2 + (m-1)\bar{\eta}\eta_k\alpha_k^2\right] - 2\alpha_k(\langle \bar{u}, z_0^k - P_C(z_0^k)\rangle + \langle \bar{u}, P_C(z_0^k) - \bar{x}\rangle) \\ &\le \|z^k - \bar{x}\|^2 + m\left[(\eta_k\alpha_k)^2 + (m-1)\bar{\eta}\eta_k\alpha_k^2\right] + 2\alpha_k\|\bar{u}\|\operatorname{dist}(z_0^k, C) \\ &\le \|z^k - \bar{x}\|^2 + m\left[(\eta_k\alpha_k)^2 + (m-1)\bar{\eta}\eta_k\alpha_k^2\right] + 2\theta\|\bar{u}\|\alpha_k^2,\end{aligned} \tag{15}$$

where we used (14) and the Cauchy-Schwarz inequality in the second inequality and the last inequality is a consequence of the fact that z_0^k is obtained by the inner loop (3)-(7). It follows from (15), (11) and (H2) that $(z^k)_{k\in\mathbb{N}}$ is quasi-Fejér convergent to S_*.

□

Next we establish some convergence properties of Algorithm A.

Proposition 3.4. *Let $(z^k)_{k\in\mathbb{N}}$ and $(x^k)_{k\in\mathbb{N}}$ be the sequences generated by Algorithm A. Then,*

(a) $x^{k+1} = \dfrac{1}{\sigma_k}\sum_{j=0}^k \alpha_j z^{j+1}$, *for all k;*

(b) $(x^k)_{k\in\mathbb{N}}$ *are bounded;*

(c) $\lim_{k\to\infty} \operatorname{dist}(x^k, C) = 0$;

(d) *all weak cluster points of* $(x^k)_{k\in\mathbb{N}}$ *belong to* C.

Proof. (a): We proceed by induction on k. For $k = 0$, using (10) and that $\sigma_0 = \alpha_0$, we have that $x^1 = z^1$. By hypothesis of induction, assume that

$$x^k = \frac{1}{\sigma_{k-1}} \sum_{j=0}^{k-1} \alpha_j z^{j+1}. \tag{16}$$

Using (9) and (10), we obtain

$$x^{k+1} = \frac{\sigma_{k-1}}{\sigma_k} x^k + \frac{\alpha_k}{\sigma_k} z^{k+1}.$$

By (16) and the above equation, we get

$$x^{k+1} = \frac{1}{\sigma_k} \sum_{j=0}^{k-1} \alpha_j z^{j+1} + \frac{\alpha_k}{\sigma_k} z^{k+1} = \frac{1}{\sigma_k} \sum_{j=0}^{k} \alpha_j z^{j+1},$$

proving the assertion.

(b): Using Proposition 3.3 and Proposition 2.1(a), we have the boundedness of $(z^k)_{k\in\mathbb{N}}$. We assume that there exists $R > 0$ such that $\|z^k\| \le R$, for all k. By the previous item,

$$\|x^k\| \le \frac{1}{\sigma_{k-1}} \sum_{j=0}^{k-1} \alpha_j \|z^{j+1}\| \le R,$$

for all k.

(c): It follows from definition of z_0^k that

$$\operatorname{dist}(z_0^k, C) \le \theta \alpha_k. \tag{17}$$

Define

$$\tilde{x}^{k+1} := \frac{1}{\sigma_k} \sum_{j=0}^{k} \alpha_j P_C(z_0^j). \tag{18}$$

Since $\frac{1}{\sigma_k}\sum_{j=0}^{k}\alpha_j = 1$ by (9), we get from the convexity of C, that $\tilde{x}^{k+1} \in C$. Thus,

$$\begin{aligned}
\operatorname{dist}(x^{k+1}, C) &\leq \|x^{k+1} - \tilde{x}^{k+1}\| = \left\| \frac{1}{\sigma_k}\sum_{j=0}^{k}\alpha_j \left(z^{j+1} - P_C(z_0^j)\right)\right\| \leq \frac{1}{\sigma_k}\sum_{j=0}^{k}\alpha_j \left\|z^{j+1} - P_C(z_0^j)\right\| \\
&\leq \frac{1}{\sigma_k}\sum_{j=0}^{k}\alpha_j \left(\|z^{j+1} - z_0^j\| + \|z_0^j - P_C(z_0^j)\|\right) \leq \frac{1}{\sigma_k}\sum_{j=0}^{k}\alpha_j (m\,\eta_j\alpha_j + \operatorname{dist}(z_0^j, C)) \\
&\leq \frac{1}{\sigma_k}\sum_{j=0}^{k}(m\,\eta_j\alpha_j^2 + \theta\alpha_j^2), \qquad (19)
\end{aligned}$$

using the fact that $\tilde{x}^{k+1}$ belongs to C in the first inequality, (b) and (18) in the equality, convexity of $\|\cdot\|$ in the second inequality, Proposition 3.2(a), with $j = m$ and $i = 0$, in the fourth inequality and (17) in the last one. Taking limits in (19) and using (9) and (12), we get $\lim_{k\to\infty} \operatorname{dist}(x^{k+1}, C) = 0$, establishing (c).

(d): Follows directly from (c). □

Next we prove optimality of the cluster points of $(x^k)_{k\in\mathbb{N}}$.

Theorem 3.1. *All weak cluster points of the sequence $(x^k)_{k\in\mathbb{N}}$ generated by Algorithm A solve problem* (1).

Proof. Using Proposition 3.2(b), we get, for any $x \in C$, $u \in T(x)$ and for all $j > 0$,

$$\begin{aligned}
\|z^{j+1} - x\|^2 &\leq \|z^j - x\|^2 + m\left[(\eta_j\alpha_j)^2 + (m-1)\eta\eta_j\alpha_j^2\right] - 2\alpha_j\langle u, z_0^j - x\rangle, \\
&= \|z^j - x\|^2 + m\left[(\eta_j\alpha_j)^2 + (m-1)\eta\eta_j\alpha_j^2\right] - 2\alpha_j\langle u, z_0^j - z^{j+1}\rangle - 2\alpha_j\langle u, z^{j+1} - x\rangle \\
&\leq \|z^j - x\|^2 + m\left[(\eta_j\alpha_j)^2 + (m-1)\eta\eta_j\alpha_j^2\right] + 2\alpha_j\|u\|\|z_0^j - z^{j+1}\| - 2\alpha_j\langle u, z^{j+1} - x\rangle \\
&\leq \|z^j - x\|^2 + m\left[(\eta_j\alpha_j)^2 + (m-1)\eta\eta_j\alpha_j^2\right] + 2m\|u\|\eta_j\alpha_j^2 - 2\alpha_j\langle u, z^{j+1} - x\rangle, \qquad (20)
\end{aligned}$$

using the Cauchy-Schwarz inequality in the second inequality and Proposition 3.2(a), with $j = m$ and $i = 0$, in the last one. Rewriting and summing (20) from $j = 0$ to $j = k$ and dividing by σ_k, we obtain from Proposition 3.4(a) that

$$\frac{1}{\sigma_k}\sum_{j=0}^{k}\left(\|z^{j+1} - x\|^2 - \|z^j - x\|^2 - m\left[(\eta_j\alpha_j)^2 - (m-1)\eta\eta_j\alpha_j^2 - 2\|u\|\eta_j\alpha_j^2\right]\right) \leq 2\langle u, x - x^{k+1}\rangle.$$

After rearrangements, we have

$$\frac{1}{\sigma_k}(\|z^{k+1}-x\|^2-\|z^0-x\|^2-m\sum_{j=0}^{\infty}\left[(\eta_j\alpha_j)^2+(m-1)\eta\eta_j\alpha_j^2+2\|u\|\eta_j\alpha_j^2\right])\leq 2\langle u,x-x^{k+1}\rangle. \quad (21)$$

Let $\hat{x}$ be a weak cluster point of $(x^k)_{k\in\mathbb{N}}$. Existence of $\hat{x}$ is guaranteed by Proposition 3.4(b). Note that $\hat{x}\in C$ by Proposition 3.4(d). Taking limits in (21), using (13), boundedness of $(z^k)_{k\in\mathbb{N}}$ by Proposition 3.3 and (12), we obtain that $\langle u, x-\hat{x}\rangle \geq 0$ for all $x\in C$ and $u\in T(x)$. Using Lemma 2.4, we get that $\hat{x}\in S_*$. Hence, all weak cluster points of $(x^k)_{k\in\mathbb{N}}$ solve problem (1). □

Finally, we state and prove the weak convergence of the main sequence generated by Algorithm A.

Theorem 3.2. *Define $x_* = \lim_{k\to\infty} P_{S_*}(z^k)$. Then $(x^k)_{k\in\mathbb{N}}$ converges weakly to x_*.*

Proof. Define $p^k := P_{S_*}(z^k)$ the orthogonal projection of z^k onto S_*. Note that p^k exists, since the solution set S_* is nonempty, closed and convex by Lemma 2.3. By Proposition 2.1, $(z^k)_{k\in\mathbb{N}}$ is quasi-Fejér convergent to S_*. Therefore, it follows from Lemma 2.2 that $(P_{S_*}(z^k))_{k\in\mathbb{N}}$ is strongly convergent. Set

$$x_* := \lim_{k\to\infty} P_{S_*}(z^k) = \lim_{k\to\infty} p^k. \quad (22)$$

By Proposition 3.4(b), $(x^k)_{k\in\mathbb{N}}$ is bounded and by Theorem 3.1 each of its weak cluster points belong to S_*. Let $(x^{i_k})_{k\in\mathbb{N}}$ be any weakly convergent subsequence of $(x^k)_{k\in\mathbb{N}}$, and let $\bar{x}\in S_*$ be its weak limit. In order to establish the weak convergence of $(x^k)_{k\in\mathbb{N}}$, it suffices to show that $\bar{x}=x_*$.

By Lemma 2.1(ii) we have that $\langle \bar{x}-p^j, z^j-p^j\rangle \leq 0$ for all j. Let $\xi = \sup_{0\leq j\leq\infty}\|z^j-p^j\|$. Since $(z^k)_{k\in\mathbb{N}}$ is bounded by Proposition 2.1(a), we get that $\xi<\infty$. Using the Cauchy-Schwarz inequality,

$$\langle \bar{x}-x_*, z^j-p^j\rangle \leq \langle p^j-x_*, z^j-p^j\rangle \leq \xi\,\|p^j-x_*\|, \quad (23)$$

for all j. Multiplying (23) by $\dfrac{\alpha_{j-1}}{\sigma_{k-1}}$ and summing from $j=1$ to $k-1$, we get from Proposition 3.4(a),

$$\left\langle \bar{x}-x_*, x^k-\frac{1}{\sigma_{k-1}}\sum_{j=1}^{k-1}\alpha_{j-1}p^j\right\rangle \leq \frac{\xi}{\sigma_{k-1}}\sum_{j=1}^{k-1}\alpha_{j-1}\|p^j-x_*\|. \quad (24)$$

Define

$$\zeta_{k,j} := \frac{\alpha_j}{\sigma_k} \qquad (k\geq 0, \quad 0\leq j\leq k).$$

It follows from the definition of σ_k, that $\lim_{k\to\infty}\zeta_{k,j}=0$ for all j and $\sum_{j=0}^{k}\zeta_{k,j}=1$ for all k. Using (22) and Proposition 2.3 with $w^k=\displaystyle\sum_{j=1}^{k}\zeta_{k-1,j-1}p^j=\frac{1}{\sigma_{k-1}}\sum_{j=0}^{k-1}\alpha_j p^{j+1}$, we have

$$x_* = \lim_{k\to\infty} p^k = \lim_{k\to\infty} \frac{1}{\sigma_k} \sum_{j=0}^{k} \alpha_j p^{j+1}, \tag{25}$$

and

$$\lim_{k\to\infty} \frac{1}{\sigma_k} \sum_{j=0}^{k} \alpha_j \|p^{j+1} - x_*\| = 0. \tag{26}$$

Taking limits in (24) over the subsequence $(i_k)_{k\in\mathbb{N}}$, and using (25) and (26), we get

$$\langle \bar{x} - x_*, \bar{x} - x_* \rangle \le 0,$$

implying that $\bar{x} = x_*$. □

4 Final Remarks

In this section we discuss the assumptions of our scheme, showing examples as well as some alternatives for changing these assumptions.

One problem in establishing the well definition of the sequence generated by Algorithm A may be the difficulty of choosing stepsizes satisfying Assumption (H2); see (12)-(13). This is clear in the important case, where the operators have bounded range, the sequence $(\eta_k)_{k\in\mathbb{N}}$, defined in (11), is bounded. Hence, any sequence $(\alpha_k)_{k\in\mathbb{N}}$ in $\ell_2(\mathbb{N}) \setminus \ell_1(\mathbb{N})$ may be used satisfying (H2). Now we present some examples showing that (H2) is verified for some different instances.

Example 1. *Consider the variational inequality problem in a Hilbert space $\mathcal{H}$, for T and $S = \operatorname{argmin}_{x\in\mathcal{H}} f(x)$, where $T : \mathcal{H} \rightrightarrows \mathcal{H}$ is a maximal monotone operator and $f : \mathcal{H} \to \mathbb{R}$ is a continuous convex function which satisfies (H1). This problem is equivalent to problem (1), with $m = 1$, $T_1 = T$, $c = f - f_*$, where $f_* = \min_{x\in\mathcal{H}} f(x)$ and Algorithm A may be rewritten as follows:*

Algorithm A1

Initialization step: *Take $x^0 \in \mathcal{H}$ and $\theta > 0$. Define $z^0 := x^0$ and $\sigma_0 := \beta_0$.*

Iterative step: *Given z^k. If $f(z^k) \le f_*$, then take $z_0^k := z^k$ and $S_k = \{x \in \mathcal{H} : \langle g^k, x - z_0^k \rangle \le 0\}$ where $g^k \in \partial f(z_0^k)$. Else, perform the inner loop (3)-(7) of Algorithm A beginning with z^k and finishing with z_0^k and S_k. Compute*

$$z^{k+1} = P_{S_k}\left(z_0^k - \frac{\beta_k}{\eta_k} u^k\right),$$

where $u^k \in T(z_0^k)$ and $\eta_k = \max\{1, \|u^k\|\}$. Define

$$\sigma_k := \sigma_{k-1} + \frac{\beta_k}{\eta_k},$$

$$x^{k+1} := \left(1 - \frac{\beta_k}{\sigma_k}\right) x^k + \frac{\beta_k}{\sigma_k} z^{k+1}.$$

Algorithm A1 is the point-to-set version of the algorithm proposed in [7]. Assuming that the problem has solutions and that $(\beta_k)_{k\in\mathbb{N}}$ in $\ell_2(\mathbb{N}) \setminus \ell_1(\mathbb{N})$, the analysis of the convergence follows directly from the analysis in [7]. (The inner loop is slightly different however the convergence proof remains essentially unchanged.)

Example 2. *We consider the optimization problem of the form*

$$\min_{x\in X} \quad \phi_1(L(x)) + \phi_2(x), \tag{27}$$

where $L : X \to Y$ is a continuous linear operators, with closed range, X and Y are two Hilbert spaces and $\phi_1 : Y \to \mathbb{R}$, $\phi_2 : X \to \mathbb{R}$ are convex lower semicontinuous functions.

This is a classical problem which appears in many applications in mechanics and economics; see, for instance, [24]. Denote $K = \{(x, y) \in X \times Y \,:\, L(x) - y = 0\}$, and

$$A = \begin{pmatrix} \partial\phi_1 & 0 \\ 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 0 \\ 0 & \partial\phi_2 \end{pmatrix}.$$

A and B are maximal monotone and (27) is equivalent to problem (1), with $m = 2$, $T_1 = A$, $T_2 = B$ and $C = K$ in $\mathcal{H} = X \times Y$.

In this case our algorithm does not require the inner loop, since the feasible set K is a linear and closed subspace. Thus, the projection onto K is easy to compute; in effect, the set K can be rewritten as

$$K = \{(x, y) \in X \times Y \,:\, c(x, y) = \frac{1}{2}\|L(x) - y\|^2 \leq 0\},$$

and $\nabla c(x, y) = \begin{pmatrix} L^*(L(x) - y) \\ y - L(x) \end{pmatrix}$ and hence,

$$P_K(x, y) = (x, y) - \frac{1}{2}\|L(x) - y\|^2 \frac{\nabla c(x, y)}{\|\nabla c(x, y)\|^2}.$$

Algorithm A may be rewritten as follows:

Algorithm A2

Initialization step: *Take $x^0 := (x_1^0, x_2^0) \in K$ and $\theta > 0$. Define $z^0 := x^0$ and $\sigma_0 := \beta_0$.*

Iterative step: *Given $z^k = (z_1^k, z_2^k)$. Compute*

$$(z_{1,1}^{k+1}\,,\, z_{1,2}^{k+1}) = P_K\left(z_1^k - \frac{\beta_k}{\eta_k} u_1^k \,,\, z_2^k\right),$$

and

$$(z_{2,1}^{k+1}\,,\,z_{2,2}^{k+1}) = P_K\left(z_{1,1}^{k+1}\,,\,z_{1,2}^{k+1} - \frac{\beta_k}{\eta_k}u_2^k\right),$$

where $u_1^k \in \partial\phi_1(z_1^k)$*,* $u_2^k \in \partial\phi_2(z_{1,2}^{k+1})$ *and* $\eta_k = \max\{1, \|u_1^k\|\,\|u_2^k\|\}$*. Set* $z^{k+1} = (z_{2,1}^{k+1}, z_{2,2}^{k+1})$*, and define*

$$\sigma_k := \sigma_{k-1} + \frac{\beta_k}{\eta_k},$$

$$x_1^{k+1} := \left(1 - \frac{\beta_k}{\sigma_k}\right)x_1^k + \frac{\beta_k}{\sigma_k}z_{2,1}^{k+1}.$$

and

$$x_2^{k+1} := \left(1 - \frac{\alpha_k}{\sigma_k}\right)x_2^k + \frac{\beta_k}{\sigma_k}z_{2,2}^{k+1}.$$

Set

$$x^{k+1} = \left(x_1^{k+1}, x_2^{k+1}\right).$$

Example 3. *Consider the minimax problem:*

$$\min_{x_1\in X}\max_{x_2\in X}\{\phi_1(x_1) - \phi_2(x_2) + \langle x_2, L(x_1)\rangle\}, \tag{28}$$

where $L : X \to X$ *is a self adjoint and continuous linear operator,* X *is a Hilbert space,* $\phi_1 : X \to \mathbb{R}$*,* $\phi_2 : X \to \mathbb{R}$ *are convex lower semicontinuous functions and* ϕ_2 *is Gâteaux differentiable. This problem was presented in [38] and under a suitable constraint qualification, this problem is equivalent to problem* (1)*, with* $m = 2$*,* $\mathcal{H} = C = X \times X$*, and*

$$T_1(x_1, x_2) = A(x_1, x_2) = (\partial\phi_1(x_1), 0)$$

and

$$T_2(x_1, x_2) = B(x_1, x_2) = (L(x_2), \nabla\phi_2(x_2) - L(x_1)),$$

which are maximal monotone operators. Algorithm A can be rewritten as follows:

Algorithm A3

Initialization step: *Take* $x^0 := (x_1^0, x_2^0) \in X \times X$ *and* $\theta > 0$*. Define* $z^0 := x^0$ *and* $\sigma_0 := \beta_0$*.*

Iterative step: *Given* $z^k = (z_1^k, z_2^k)$*. Compute*

$$z_{1,1}^{k+1} = z_1^k - \frac{\beta_k}{\eta_k}u_1^k \qquad z_{2,1}^{k+1} = z_2^k,$$

$$z_{1,2}^{k+1} = z_{1,1}^{k+1} - \frac{\beta_k}{\eta_k}L(z_{2,1}^{k+1}) \qquad z_{2,2}^{k+1} = z_{2,1}^k - \frac{\beta_k}{\eta_k}\left(\nabla\phi_2(z_{2,1}^{k+1}) - L(z_{1,1}^{k+1})\right),$$

where $u_1^k \in \partial\phi_1(z_1^k)$ *and* $\eta_k = \max\left\{1, \|u_1^k\|, \|L(z_{2,1}^{k+1})\|, \|\nabla\phi_2(z_{2,1}^{k+1}) - L(z_{1,1}^{k+1})\|\right\}$*. Set* $z^{k+1} := (z_1^{k+1}, z_2^{k+1}) = (z_{1,2}^{k+1}, z_{2,2}^{k+1})$*, and define*

$$\sigma_k := \sigma_{k-1} + \frac{\beta_k}{\eta_k},$$

$$x_1^{k+1} := \left(1 - \frac{\beta_k}{\sigma_k}\right) x_1^k + \frac{\beta_k}{\sigma_k} z_1^{k+1}.$$

$$x_2^{k+1} := \left(1 - \frac{\beta_k}{\sigma_k}\right) x_2^k + \frac{\beta_k}{\sigma_k} z_2^{k+1}.$$

Set

$$x^{k+1} = (x_1^{k+1}, x_2^{k+1}).$$

In Algorithms A2 and A3, presented in the above examples, the stepsize is $\alpha_k = \dfrac{\beta_k}{\eta_k}$, for all k and hence $\eta_k\alpha_k = \beta_k$ and (13) is equivalent to $(\beta_k)_{k\in\mathbb{N}}$ in $\ell_2(\mathbb{N})$, which by Proposition 3.3 implies, the boundedness of the sequence $(z^k)_{k\in\mathbb{N}}$. Moreover as a consequence of Proposition 3.2(a) the sequences $(z_{i-1}^k)_{k\in\mathbb{N}}$, $(i = 1, 2, \ldots, m)$ are bounded. Now, condition (12) becomes in

$$\sum_{k=0}^{\infty} \frac{\beta_k}{\eta_k} = \infty, \tag{29}$$

where $\eta_k = \max_{1\le i\le m}\{1, \|u_i^k\|\}$, with $u_i^k \in T_i(z_{i-1}^k)$. Thus, a sufficient condition for (29) is that the image of T_i, $(i = 1, 2, \ldots, m)$ is bounded on bounded sets (since the sequences $(z_{i-1}^k)_{k\in\mathbb{N}}$, $(i = 1, 2, \ldots, m)$ are bounded). Moreover, $(\eta_k)_{k\in\mathbb{N}}$ may be an unlimited sequence as for example $(k^s)_{k\in\mathbb{N}}$ for any $s \in (0, 1/2)$.

Therefore Assumption (H2) turns into "$(\beta_k)_{k\in\mathbb{N}}$ lies in $\ell_2(\mathbb{N}) \setminus \ell_1(\mathbb{N})$", which is a requirement widely used in the literature. The convergence analysis of Algorithms A2 and A3 follows from the convergence analysis of Algorithm A.

Another important point on Algorithm A, is that the inner loop (3)-(7) uses the distance function. It is clear that this is weakest that compute the exact projection for almost all instances. Furthermore inside of the inner loop, we may only check the condition related with the distance on selected index. In connection we may include the following assumption:

(H3) Assume that a Slater point is available, i.e. there exists a point $w \in \mathcal{H}$ such that $c(w) < 0$.

If Assumption (H3) holds, by Lemma 2.5 we can obtain an explicit algorithm for a quite general convex set C, replacing the inequality $\operatorname{dist}(y^{k,j+1}, C) \le \theta\alpha_k$ in the Inner Loop of Algorithm A by

$\tilde{c}(y^{k,j+1}) \leq \theta\alpha_k$, where

$$\tilde{c}(x) = \begin{cases} \dfrac{\|x-w\|c(x)}{c(x)-c(w)} & \text{if } x \notin C \\ \\ 0 & \text{if } x \in C. \end{cases}$$

All our convergence results are preserved. (H3) is a hard assumption in Hilbert spaces and the point w is almost always unavailable. Hence, such assumptions can be replaced by a rather weaker one, namely:

(H3*) There exists an easily computable and continuous $\tilde{c} : \mathcal{H} \to \mathbb{R}$, such that $\text{dist}(x, C) \leq \tilde{c}(x)$ for all $x \in \mathcal{H}$, and $\tilde{c}(x) = 0$ if and only if $c(x) = 0$.

There are examples of sets C for which no Slater point is available, while (H3*) holds, including instances in which C has an empty interior. An exhaustive discussion about weak constraint qualifications for getting error-bound can be found in [32, 40].

ACKNOWLEDGMENT

JYBC and RDM were partially supported by CNPq, by project PROCAD-nf - UFG/UnB/IMPA and by project CAPES-MES-CUBA 226/2012 "Modelos de Otimização e Aplicações". RDM was supported by his scholarship for his doctoral studies, granted by CAPES.

The authors would like to thanks to Professor Dr. Heinz H. Bauschke and to Professor Dr. Ole Peter Smith whose suggestions helped us to improve the presentation of this paper.

References

[1] Alber, Ya.I., Iusem, A.N., Solodov, M.V. On the projected subgradient method for nonsmooth convex optimization in a Hilbert space. *Mathematical Programming* **81**, 23–37 (1998).

[2] Attouch, H., Czarnecki, M.-O., Peypouquet, J. Coupling forward-backward with penalty schemes and parallel splitting for constrained variational inequalities. *SIAM Journal on Optimization* **21**, 1251–1274 (2011).

[3] Bao, T.Q., Khanh, P.Q. A projection-type algorithm for pseudomonotone nonlipschitzian multivalued variational inequalities. *Nonconvex Optimization and its Applications* **77**, 113–129 (2005).

[4] Bauschke, H.H., Combettes, P.L. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York (2011).

[5] Bello Cruz, J.Y., Díaz Millán, R. A direct splitting method for nonsmooth variational inequalities. *Journal of Optimization Theory and Applications*, doi: 10.1007/s10957-013-0478-2 (2013).

[6] Bello Cruz, J.Y., de Oliveira. W. Level bundle-like algorithms for convex optimization. *Journal of Global Optimization*, doi: 10.1007/s10898-013-0096-4 (2013).

[7] Bello Cruz, J.Y., Iusem, A.N. An explicit algorithm for monotone variational inequalities. *Optimization* **61**, 855–871 (2012).

[8] Bello Cruz, J.Y., Iusem, A.N. A strongly convergent method for nonsmooth convex minimization in Hilbert spaces. *Numerical Functional Analysis and Optimization* **32**, 1009–1018 (2011).

[9] Bello Cruz, J.Y., Iusem, A.N. Full convergence of an approximate projection method for nonsmooth variational inequalities. *Mathematics and Computers in Simulation*, doi: 10.1016/j.matcom.2010.05.026 (2010).

[10] Bello Cruz, J.Y., Iusem, A.N. Convergence of direct methods for paramonotone variational inequalities. *Computational Optimization and Application* **46**, 247–263 (2010).

[11] Bello Cruz, J.Y., Iusem, A.N. A strongly convergent direct method for monotone variational inequalities in Hilbert spaces. *Numerical Functional Analysis and Optimization* **30**, 23–36 (2009).

[12] Boţ, R.I., Hendrich, C. A Douglas-Rachford type primal-dual method for solving inclusions with mixtures of composite and parallel-sum type monotone operators. http://arxiv.org/abs/1212.0326 (2012).

[13] Brezis, H. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, New York (2010).

[14] Burachik, R.S., Iusem, A.N. *Set-Valued Mappings and Enlargements of Monotone Operators*. Springer, Berlin (2008).

[15] Burachik, R.S., Lopes, J.O., Svaiter, B.F. An outer approximation method for the variational inequality problem. *SIAM Journal on Control and Optimization* **43**, 2071–2088 (2005).

[16] Censor, Y., Gibali, A., Reich, S. The Subgradient Extragradient Method for Solving Variational Inequalities in Hilbert Space. *Journal of Optimization Theory and Applications* **148**, 318–335 (2011).

[17] Combettes, P.L. Quasi-Fejérian analysis of some optimization algorithms: in Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications, (D. Butnariu, Y. Censor, and S. Reich, Eds.) 115–152 Elsevier, New York (2001).

[18] Combettes, P.L., Pesquet, J.-C. Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators. *Set-Valued and Variational Analysis* **20**, 307–330 (2012).

[19] Eckstein, J., Svaiter, B.F. General projective splitting methods for sums of maximal monotone operators. *SIAM Journal on Control and Optimization* **48**, 787–811 (2009).

[20] Ermoliev, Yu.M. On the method of generalized stochastic gradients and quasi-Fejér sequences. *Cybernetics* **5**, 208–220 (1969).

[21] Facchinei, F., Pang, J.S. *Finite-dimensional Variational Inequalities and Complementarity Problems*. Springer, Berlin (2003).

[22] Fang, S.C., Petersen, E.L. Generalized variational inequalities. *Journal of Optimization Theory and Applications* **38**, 363–383 (1982).

[23] Fukushima, M. A Relaxed projection for variational inequalities. *Mathematical Programming* **35**, 58–70 (1986).

[24] Gabay, D., Mercier, B. A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Computers & Mathematics with Applications* **2**, 17–40 (1976).

[25] He, B.S. A new method for a class of variational inequalities. *Mathematical Programming* **66**, 137–144 (1994).

[26] Iusem, A.N., Lucambio Pérez, L.R. An extragradient-type method for non-smooth variational inequalities. *Optimization* **48**, 309–332 (2000).

[27] Iusem, A.N., Svaiter, B.F. A variant of Korpelevich's method for variational inequalities with a new search strategy. *Optimization* **42**, 309–321 (1997); *Addendum Optimization* **43**, 85 (1998).

[28] Iusem, A.N., Svaiter, B.F., Teboulle, M.: Entropy-like proximal methods in convex programming. *Mathematics of Operations Research* **19**, 790–814 (1994).

[29] Kinderlehrer, D., Stampacchia, G. *An Introduction to Variational Inequalities and Their Applications*. Academic Press, New York (1980).

[30] Konnov, I.V. *Combined Relaxation Methods for Variational Inequalities*. Springer-Verlag, Berlin (2001).

[31] Korpelevich, G.M. The extragradient method for finding saddle points and other problems. *Ekonomika i Matematcheskie Metody* **12**, 747–756 (1976).

[32] Lewis, A. S., Pang, J.-S. Error bounds for convex inequality systems. Generalized convexity, generalized monotonicity: recent results: Nonconvex Optimization and Its Applications **27**, 75–110 Kluwer Acad. Publ., Dordrecht (1998).

[33] Lions, P.L., Mercier, B. Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal of Numererical Analysis* **16**, 964–979 (1979).

[34] Moudafi, A. On the convergence of splitting proximal methods for equilibrium problems in Hilbert spaces. *Journal of Mathematical Analysis and Applications* **359**, 508–513 (2009).

[35] Nedic, A., Bertsekas, D. Incremental subgradient methods for nondifferentiable optimization. *SIAM Journal of Optimization* **12**, 109–138 (2001).

[36] Passty, G.B. Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. *Journal of Mathematical Analysis and Applications* **72**, 383–390 (1979).

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

[38] Rockafellar, R.T. On the maximality of sums of nonlinear monotone operators. *Transactions of the American Mathematical Society* **149**, 75–88 (1970).

[39] Rockafellar, R.T. On the maximal monotonicity of subdifferential mapping. *Pacific Journal of Mathematics* **33**, 209–216 (1970).

[40] Sien, D. Computable error bounds for convex inequality systems in reflexive Banach spaces. *SIAM Journal on Optimization* **1**, 274–279 (1997).

[41] Shih, M.H., Tan, K.K. Browder-Hartmann-Stampacchia variational inequalities for multi-valued monotone operators. *Journal of Mathematical Analysis and Applications* **134**, 431–440 (1988).

[42] Solodov, M.V., Svaiter, B.F. A new projection method for monotone variational inequality problems. *SIAM Journal on Control and Optimization* **37**, 765–776 (1999).

[43] Solodov, M.V., Tseng, P. Modified projection-type methods for monotone variational inequalities. *SIAM Journal on Control and Optimization* **34**, 1814–1830 (1996).

[44] Todd, M.J. *The Computations of Fixed Points and Applications.* Springer, Berlin (1976)

[45] Tseng, P. A modified forward-backward splitting method for maximal monotone mappings. *SIAM Journal on Control and Optimization* **38**, 431–446 (2000).

[46] Zarantonello, E.H. *Projections on convex sets in Hilbert space and spectral theory.* in Contributions to Nonlinear Functional Analysis, E. Zarantonello, ed., Academic Press, New York 237-424 (1971).

[47] Zhang, H., Cheng, L. Projective splitting methods for sums of maximal monotone operators with applications. *Journal of Mathematical Analysis and Applications* **406**, 323–334 (2013).