Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH

Open Access

The forward–backward splitting methods for variational inequalities and minimization problems in Banach spaces

Wei-Bo Guan[1*] and Wen Song[1]

*Correspondence:
guanweibo0451@163.com
[1] School of Mathematical Sciences,
Harbin Normal University, Harbin,
P.R. China

Abstract

This paper is concerned with the study of a class of forward–backward splitting methods based on Lyapunov distance for variational inequalities and convex minimization problem in Banach spaces.

MSC: 49M30; 46N10; 90C25; 90C30

Keywords: Banach space; Forward–backward splitting methods; Variational inequalities; Maximal monotone operators; Duality mapping

1 Introduction

Let X be a reflexive, strictly convex and smooth Banach space with the dual space X^*, $A : X \rightrightarrows X^*$ a general maximal monotone operator, and C a closed convex set in X. We denote by N_C the normal cone to C. In this work, we study the following variational inequality problem: find x in a Banach space X such that

$$0 \in A(x) + N_C(x). \tag{1}$$

Problem (1) is a very important format for certain concrete problems in machine learning, linear inverse and many nonlinear problems such as convex programming, split feasibility problem, see [3, 4, 15, 17]. The set of solutions of (1) is supposed to be nonempty and is denoted by $\mathcal{S}$. In [3], the authors provided a generic framework, the so called backward–backward splitting method, for solving (1) in a Hilbert space:

$$x_{n+1} = (I + \lambda_n \beta_n \partial \Psi)^{-1} (I + \lambda_n A)^{-1}(x_n), \tag{2}$$

where Ψ is a penalization function and λ_n, β_n are sequences of positive parameters. In [3], convergence results have been obtained for the backward–backward splitting method (2) under the key Fenchel conjugate assumption that

$$\sum_{n=1}^{+\infty} \lambda_n \beta_n \left[\Psi^* \left(\frac{p^*}{\beta_n} \right) - \sigma_C \left(\frac{p^*}{\beta_n} \right) \right] < +\infty, \quad \forall p^* \in R(N_C),$$

where Ψ^* is the Fenchel conjugate of Ψ, σ_C is the support function of C and $R(N_C)$ denotes the range of N_C. This condition somehow relates the growth of the sequence $\{\beta_n\}$ to the shape of Ψ around its minimizing set C. The reader is referred to [14] for further discussion.

When the penalization function Ψ is Gâteaux differentiable, it is rather natural to consider the following forward–backward splitting method (FBS):

$$x_{n+1} = (I + \lambda_n A)^{-1}(x_n - \lambda_n \beta_n \nabla\Psi(x_n)). \tag{3}$$

The forward–backward method has the advantage of being easier to compute than the backward–backward method, which ensures enhanced applicability to real-life problems. Iterations have lower computational cost and can be computed exactly. A special case of this method is the projected subgradient algorithm aimed at solving constrained minimization problems. There have been many works concerning the problem of finding zero points of the sum of two monotone operators. For further information on forward–backward splitting methods and their applications, see [4, 6, 9, 11, 12].

Let $X = H$ be a Hilbert space. If $A = \partial\Phi$ is the subdifferential of a proper, lower-semicontinuous and convex function $\Phi : H \to (-\infty, +\infty]$, the variational inequality problem (1) becomes the following minimization problem:

$$x \in \text{Argmin}\{\Phi(z) : z \in \text{Argmin}\,\Psi\}.$$

It is convenient to reformulate method (3) as

$$\frac{x_n - x_{n+1}}{\lambda_n} - \beta_n \nabla\Psi(x_n) \in \partial\Phi(x_{n+1}).$$

This is equivalent to

$$x_{n+1} = \underset{y\in X}{\text{argmin}}\left\{\frac{1}{2}\|x_n - y\|^2 + \lambda_n \beta_n \langle\nabla\Psi(x_n), y\rangle + \lambda_n \Phi(y)\right\}.$$

In [4], the authors prove that every sequence generated by the forward–backward splitting method converges weakly to a solution of the minimization problem if either the penalization function or the objective function is inf-compact. However, this inf-compactness assumption is not necessary. In [13], the authors prove that every sequence generated by the forward–backward splitting method converges weakly to a solution without the inf-compactness assumption.

A generalization of this method from Hilbert to Banach space is not immediate. The main difficulties are due to the fact that the inner product structure of a Hilbert space is missing in a Banach space. In [18], the authors prove that every sequence generated by a projection iterative method converges strongly to a common minimum norm solution of a variational inequality problem for an inverse strongly monotone mapping in Banach spaces.

In this paper, we extend the forward–backward splitting method (3) to a Banach space, that is,

$$x_{n+1} = (cJ + \lambda_n A)^{-1}(cJ(x_n) - \lambda_n \beta_n \nabla\Psi(x_n)), \tag{4}$$

where J is the duality mapping and c is constant. If $A = \partial\Phi$ is the subdifferential of a proper, lower-semicontinuous and convex function $\Phi : X \to (-\infty, +\infty]$, the forward–backward splitting method (4) becomes

$$\frac{cJ(x_n) - cJ(x_{n+1})}{\lambda_n} - \beta_n \nabla\Psi(x_n) \in \partial\Phi(x_{n+1}). \tag{5}$$

Iterative formula (5) can be rewritten as

$$x_{n+1} = \operatorname*{argmin}_{y \in X} \left\{ \frac{c}{2} W(x_n, y) + \lambda_n\beta_n \langle\nabla\Psi(x_n), y\rangle + \lambda_n\Phi(y) \right\}.$$

Throughout this paper, let $A : X \rightrightarrows X^*$ be a maximal monotone operator and let the monotone operator $T_{A,C} = A + N_C$ be also maximal monotone. Let $\Psi : X \to (-\infty, +\infty]$ be a proper, lower-semicontinuous and convex function with $C = \operatorname{argmin}(\Psi) \neq \emptyset$ and $\min(\Psi) = 0$. We assume that Ψ is Gâteaux differentiable and $\nabla\Psi$ is L-Lipschitz continuous on the domain of Ψ. We also assume that there exists $c > 0$ such that

$$cW(x, y) \geq \|x - y\|^2, \quad \forall x, y \in X. \tag{6}$$

The objective of the present paper is to propose a forward–backward splitting method to solve problem (1), which is so far limited to Hilbert spaces, in the general framework of Banach spaces. The paper is organized as follows. In Sect. 2, we provide some preliminary results. We present the forward–backward splitting method and prove its convergence in Sect. 3. Section 4 is devoted to an application of our result to convex minimization problem. Finally, in Sect. 5, we also prove a convergence result without Fenchel conjugate assumption.

2 Preliminaries

Let f be a proper, lower-semicontinuous and convex function on a Banach space X. The subdifferential of f at $x \in X$ is the convex set

$$\partial f(x) = \left\{ x^* \in X^* : \langle x^*, y - x\rangle \leq f(y) - f(x), \forall y \in X \right\}.$$

It is easy to verify that $0 \in \partial f(\hat{x})$ if and only if $f(\hat{x}) = \min_{x \in X} f(x)$. We denote by f^* the Fenchel conjugate of f:

$$f^*\left(x^*\right) = \sup_{x \in X} \left\{ \langle x^*, x\rangle - f(x) \right\}.$$

Given a nonempty closed convex set $C \subset X$, its indicator function is defined as $\delta_C(x) = 0$ if $x \in C$, and $+\infty$ otherwise. The support function of C at a point x^* is $\sigma_C(x^*) = \sup_{y \in C} \langle x^*, y\rangle$. Then the normal cone of C at $x \in X$ is $N_C(x) = \partial\delta_C(x) = \{x^* \in X^* | \langle x^*, y - x\rangle \leq 0, \forall y \in C\}$.

The duality mapping $J : X \rightrightarrows X^*$ is defined by

$$J(x) = \left\{ x^* \in X^* | \langle x^*, x\rangle = \|x^*\|\|x\|, \|x^*\| = \|x\| \right\}, \quad \forall x \in X.$$

The Hahn–Banach theorem guarantees that $J(x) \neq \emptyset$ for every $x \in X$. It is clear that $J(x) = \partial(\frac{1}{2}\|\cdot\|^2)(x)$ for all $x \in X$. It is well known that if X is smooth, then J is single-valued and is

norm-to-weak star continuous. It is also well known that if a Banach space X is reflexive, strictly convex and smooth, then the duality mapping J^* from X^* into X is the inverse of J, that is, $J^{-1} = J^*$. Properties of the duality mapping have been given in [1, 2, 8, 17].

Let X be a smooth Banach space. Alber [1, 2] considered the following Lyapunov distance function:

$$W(x,y) = \|x\|^2 - 2\langle J(x), y\rangle + \|y\|^2, \quad \forall x, y \in X.$$

It is obvious from the definition of W that

$$\left(\|x\| - \|y\|\right)^2 \leq W(x,y) \leq \left(\|x\| + \|y\|\right)^2, \quad \forall x, y \in X.$$

We also know that

$$W(x,y) = W(x,z) + W(z,y) + 2\langle J(x) - J(z), z - y\rangle, \quad \forall x, y, z \in X. \tag{7}$$

A set-valued mapping $A : X \rightrightarrows X^*$ is said to be a monotone operator if $\langle x^* - y^*, x - y\rangle \geq 0$, for all $x^* \in A(x)$ and for all $y^* \in A(y)$. It is maximal monotone if its graph is not properly contained in the graph of any other monotone operator. The subdifferential of a proper, lower-semicontinuous and convex function is maximal monotone. The inverse $A^{-1} : X^* \rightrightarrows X$ of A is defined by $x \in A^{-1}(x^*) \Leftrightarrow x^* \in A(x)$. The operator $J + \lambda A$ is surjective for any maximal monotone operator $A : X \rightrightarrows X^*$ and for any $\lambda > 0$ by Minty's Theorem. The operator $(J + \lambda A)^{-1}$ is nonexpansive and everywhere defined. If X is a reflexive, strictly convex and smooth Banach space and A is a maximal monotone operator, then for each $\lambda > 0$ and $x \in X$, there is a unique element $\bar{x}$ satisfying $J(x) \in J(\bar{x}) + \lambda A(\bar{x})$ (see [16]). An operator $A : X \rightrightarrows X^*$ is strongly monotone with parameter $\alpha > 0$ if

$$\langle x^* - y^*, x - y\rangle \geq \alpha \|x - y\|^2, \quad \forall x^* \in A(x), y^* \in A(y).$$

Observe that the set of zeroes of a maximal monotone operator which is strongly monotone must contain exactly one element.

Let $A : X \rightrightarrows X^*$ be a maximal monotone operator. Suppose the operator $T_{A,C} = A + N_C$ is maximal monotone and $\mathcal{S} = (T_{A,C})^{-1}(0) \neq \emptyset$. By maximal monotonicity of $T_{A,C}$, we know that

$$z \in \mathcal{S} \iff \langle 0 - \omega^*, z - u\rangle \geq 0, \quad \forall (u, \omega^*) \in T_{A,C},$$

that is,

$$z \in \mathcal{S} \iff \langle 0 - \omega^*, z - u\rangle \geq 0, \quad \forall u \in \operatorname{dom}(T_{A,C}) = C \cap \operatorname{dom} A, \forall \omega^* \in T_{A,C}(u).$$

If $A = \partial\Phi$ is the subdifferential of a proper, lower-semicontinuous and convex function $\Phi : X \to (-\infty, +\infty]$ and if $u \in \mathcal{S}$, then there exists $u^* \in N_C(u)$ such that $-u^* \in \partial\Phi(u)$. Hence, by $u^* \in N_C(u) \Rightarrow \sigma_C(u^*) = \langle u^*, u\rangle$, one has

$$\Phi(x) \geq \Phi(u) + \langle -u^*, x - u\rangle = \Phi(u) + \sigma_C\left(u^*\right) - \langle u^*, x\rangle \geq \Phi(u), \quad \forall x \in C.$$

Thus, when $A = \partial \Phi$, the maximal monotonicity of $T_{A,C}$ implies

$$\mathcal{S} = \text{Argmin}\{\Phi(x) : x \in C\}.$$

The following result will be used subsequently.

Lemma 2.1 ([4]) *Let $\{a_n\}, \{b_n\}$ and $\{\epsilon_n\}$ be real sequences. Assume that $\{a_n\}$ is bounded from below, $\{b_n\}$ is nonnegative, $\sum_{n=1}^{\infty} |\epsilon_n| < +\infty$ and $a_{n+1} - a_n + b_n \leq \epsilon_n$. Then $\{a_n\}$ converges and $\sum_{n=1}^{\infty} b_n < +\infty$.*

3 The FBS method for variational inequalities

In this section, we firstly extend Baillon–Haddad theorem (see [5]) to Banach space.

Lemma 3.1 *Let $\Psi : X \to (-\infty, +\infty]$ be a proper, lower-semicontinuous and convex function and let Ψ be Gâteaux differentiable on the domain of Ψ. The following are equivalent:*

   (i) $\nabla \Psi$ *is Lipschitz continuous with constant L.*
   (ii) $\Psi(y) - \Psi(x) - \langle \nabla \Psi(x), y - x \rangle \leq \frac{L}{2} \|y - x\|^2, \forall x, y \in \text{dom}\,\Psi$.
   (iii) $\Psi(x) + \langle \nabla \Psi(x), y - x \rangle + \frac{L}{2} \|\nabla \Psi(x) - \nabla \Psi(y)\|^2 \leq \Psi(y), \forall x, y \in \text{dom}\,\Psi$.
   (iv) $\nabla \Psi$ *is $\frac{1}{L}$ cocoercive, that is,*

$$\langle \nabla \Psi(x) - \nabla \Psi(y), x - y \rangle \geq \frac{1}{L} \|\nabla \Psi(x) - \nabla \Psi(y)\|^2, \quad \forall x, y \in \text{dom}\,\Psi.$$

Proof (i) $\Rightarrow$ (ii). By the mean value theorem, we have

$$
\begin{aligned}
\Psi(y) - \Psi(x) - \langle \nabla \Psi(x), y - x \rangle &= \int_0^1 \langle \nabla \Psi(x) - \nabla \Psi(x + t(y - x)), x - y \rangle \, dt \\
&\leq \int_0^1 \|\nabla \Psi(x + t(y - x)) - \nabla \Psi(x)\| \|x - y\| \, dt \\
&\leq \int_0^1 L \|x - y\|^2 t \, dt \\
&\leq \frac{L}{2} \|x - y\|^2.
\end{aligned}
$$

(ii) $\Rightarrow$ (iii). Let us fix $x_0 \in \text{dom}\,\Psi$. Consider the function

$$F(y) = \Psi(y) - \langle \nabla \Psi(x_0), y \rangle.$$

Note that F is a proper, lower-semicontinuous, convex and Gâteaux differentiable function and ∇F is Lipschitz continuous on the $\text{dom}\,F = \text{dom}\,\Psi$ with constant L. Therefore, by (i) $\Rightarrow$ (ii), we have

$$F(u) - F(v) - \langle \nabla F(v), u - v \rangle \leq \frac{L}{2} \|u - v\|^2, \quad \forall u, v \in \text{dom}\,F. \tag{8}$$

By the definition of F, we have $x_0 \in \text{Argmin}_{x \in X} F(x)$. Then, by (8), we have

$$F(x_0) \leq F\left(y - \frac{1}{L} J^{-1} \nabla F(y)\right) \leq F(y) - \frac{L}{2} \left\|\left(y - \frac{1}{L} J^{-1} \nabla F(y)\right) - y\right\|^2 = F(y) - \frac{1}{2L} \|\nabla F(y)\|^2.$$

Hence, by $\nabla F(y) = \nabla \Psi(y) - \nabla \Psi(x_0)$, we get (iii).

(iii) $\Rightarrow$ (iv). For any $x, y \in \text{dom}\,\Psi$, by (iii), we have

$$\Psi(x) + \langle \nabla\Psi(x), y - x \rangle + \frac{L}{2}\|\nabla\Psi(x) - \nabla\Psi(y)\|^2 \le \Psi(y)$$

and

$$\Psi(y) + \langle \nabla\Psi(y), x - y \rangle + \frac{L}{2}\|\nabla\Psi(x) - \nabla\Psi(y)\|^2 \le \Psi(x).$$

Adding the two inequalities, we get

$$\langle \nabla\Psi(x) - \nabla\Psi(y), x - y \rangle \ge \frac{1}{L}\|\nabla\Psi(x) - \nabla\Psi(y)\|^2.$$

(iv) $\Rightarrow$ (i). By Cauchy–Schwartz inequality, we get $\|\nabla\Psi(x) - \nabla\Psi(y)\| \le L\|x - y\|$. $\square$

Iterative Method 3.1 *Given $x_0 \in X$, set*

$$x_{n+1} = (cJ + \lambda_n A)^{-1}\big(cJ(x_n) - \lambda_n \beta_n \nabla\Psi(x_n)\big), \tag{9}$$

where $\{\lambda_n\}, \{\beta_n\}$ are two sequences of positive real numbers with $\sum_{n=1}^{\infty} \lambda_n = +\infty$, $\sum_{n=1}^{\infty} \lambda_n^2 < +\infty$ and $\lambda_n \beta_n < \frac{2c}{L}$.

For our results, the following Fenchel conjugate assumption will be used subsequently:

$$\sum_{n=1}^{+\infty} \lambda_n \beta_n \left[\Psi^*\left(\frac{p^*}{\beta_n}\right) - \sigma_C\left(\frac{p^*}{\beta_n}\right) \right] < +\infty, \quad \forall p^* \in R(N_C). \tag{10}$$

Remark 3.1 Since $\Psi(x) \le \delta_C(x)$ for all $x \in X$, we obtain, $\Psi^*(x^*) - \sigma_C(x^*) \ge 0$ for all $x^* \in X^*$. Hence, the terms in the sum are nonnegative.

Remark 3.2 If $\Psi(x) = \frac{1}{2}\text{dist}(x, C)^2$, then we have $\Psi^*(x^*) - \sigma_C(x^*) = \frac{1}{2}\|x^*\|^2$ for all $x^* \in X^*$ and so

$$\sum_{n=1}^{+\infty} \lambda_n \beta_n \left[\Psi^*\left(\frac{p^*}{\beta_n}\right) - \sigma_C\left(\frac{p^*}{\beta_n}\right) \right] < +\infty, \quad \forall p^* \in R(N_C) \quad \Longleftrightarrow \quad \sum_{n=1}^{+\infty} \frac{\lambda_n}{\beta_n} < +\infty.$$

It is easy to see that, if the sequence $\{\beta_n\}$ is chosen so that $\limsup_{n \to +\infty} \lambda_n \beta_n < +\infty$ and $\liminf_{n \to +\infty} \lambda_n \beta_n > 0$, then

$$\sum_{n=1}^{+\infty} \frac{\lambda_n}{\beta_n} < +\infty \quad \Longleftrightarrow \quad \sum_{n=1}^{+\infty} \lambda_n^2 < +\infty.$$

Proposition 3.1 *Let $\{x_n\}$ be a sequence generated by iterative formula* (9). *Take $u \in C \cap \text{dom}\,A$ and $v^* \in A(u)$. Then for all $t \ge 0$, we have*

$$cW(x_{n+1}, u) - cW(x_n, u) + cW(x_n, x_{n+1}) - \frac{1}{1+t}\|x_n - x_{n+1}\|^2 + \frac{2t}{1+t}\lambda_n \beta_n \Psi(x_n)$$

$$\le \lambda_n \beta_n \left((1+t)\lambda_n \beta_n - \frac{2}{L(1+t)} \right)\|\nabla\Psi(x_n)\|^2 + 2\lambda_n \langle v^*, u - x_{n+1} \rangle. \tag{11}$$

Proof Since $v^* \in A(u)$ and $cJ(x_n) - cJ(x_{n+1}) - \lambda_n \beta_n \nabla\Psi(x_n) \in \lambda_n A(x_{n+1})$, the monotonicity of A implies

$$\langle cJ(x_n) - cJ(x_{n+1}) - \lambda_n \beta_n \nabla\Psi(x_n) - \lambda_n v^*, x_{n+1} - u \rangle \geq 0, \tag{12}$$

and so

$$\langle cJ(x_n) - cJ(x_{n+1}), u - x_{n+1} \rangle \leq \langle \lambda_n \beta_n \nabla\Psi(x_n) + \lambda_n v^*, u - x_{n+1} \rangle,$$

which in turn gives

$$cW(x_{n+1}, u) - cW(x_n, u) + cW(x_n, x_{n+1}) \leq 2\lambda_n \langle \beta_n \nabla\Psi(x_n) + v^*, u - x_{n+1} \rangle.$$

Hence, we have that

$$cW(x_{n+1}, u) - cW(x_n, u) + cW(x_n, x_{n+1})$$
$$\leq 2\lambda_n \langle \beta_n \nabla\Psi(x_n), u - x_n \rangle + 2\lambda_n \langle \beta_n \nabla\Psi(x_n), x_n - x_{n+1} \rangle + 2\lambda_n \langle v^*, u - x_{n+1} \rangle. \tag{13}$$

On the one hand, since $\nabla\Psi$ is Lipschitz continuous, by Lemma 3.1, we have

$$\langle \nabla\Psi(x_n) - \nabla\Psi(u), x_n - u \rangle \geq \frac{1}{L} \|\nabla\Psi(x_n) - \nabla\Psi(u)\|^2.$$

Hence, by $\nabla\Psi(u) = 0$, we obtain

$$\langle \nabla\Psi(x_n), u - x_n \rangle \leq -\frac{1}{L} \|\nabla\Psi(x_n)\|^2. \tag{14}$$

Since $\Psi(u) = 0$, by $\Psi(u) \geq \Psi(x_n) + \langle \nabla\Psi(x_n), u - x_n \rangle$, we have

$$\langle \nabla\Psi(x_n), u - x_n \rangle \leq -\Psi(x_n). \tag{15}$$

For any $t \geq 0$, by taking a convex combination of inequalities (14) and (15), we have

$$2\lambda_n \beta_n \langle \nabla\Psi(x_n), u - x_n \rangle \leq -\frac{2}{L(1+t)} \lambda_n \beta_n \|\nabla\Psi(x_n)\|^2 - \frac{2t}{1+t} \lambda_n \beta_n \Psi(x_n). \tag{16}$$

On the other hand, for the remaining term $2\lambda_n \beta_n \langle \nabla\Psi(x_n), x_n - x_{n+1} \rangle$, we have

$$2\lambda_n \beta_n \langle \nabla\Psi(x_n), x_n - x_{n+1} \rangle$$
$$= 2\left\langle \lambda_n \beta_n \sqrt{1+t} \nabla\Psi(x_n), \frac{1}{\sqrt{1+t}}(x_n - x_{n+1}) \right\rangle$$
$$\leq 2\left\| \lambda_n \beta_n \sqrt{1+t} \nabla\Psi(x_n) \right\| \left\| \frac{1}{\sqrt{1+t}}(x_n - x_{n+1}) \right\|$$
$$\leq (1+t)\lambda_n^2 \beta_n^2 \|\nabla\Psi(x_n)\|^2 + \frac{1}{(1+t)} \|x_n - x_{n+1}\|^2. \tag{17}$$

Inequalities (13), (16) and (17) together give

$$
cW(x_{n+1}, u) - cW(x_n, u) + cW(x_n, x_{n+1}) - \frac{1}{1+t}\|x_n - x_{n+1}\|^2 + \frac{2t}{1+t}\lambda_n\beta_n\Psi(x_n)
$$

$$
\leq \lambda_n\beta_n\left((1+t)\lambda_n\beta_n - \frac{2}{L(1+t)}\right)\|\nabla\Psi(x_n)\|^2 + 2\lambda_n\langle v^*, u - x_{n+1}\rangle. \qquad \square
$$

Proposition 3.2 *Let $\{x_n\}$ be a sequence generated by iterative formula* (9). *Then, there exist $a > 0$ and $b > 0$, such that for any $u \in C \cap \mathrm{dom}\,A$ and any $v^* \in A(u)$, we have*

$$
cW(x_{n+1}, u) - cW(x_n, u) + a\left(\|x_n - x_{n+1}\|^2 + \lambda_n\beta_n\Psi(x_n) + \lambda_n\beta_n\|\nabla\Psi(x_n)\|^2\right)
$$

$$
\leq b\lambda_n^2\|v^*\|^2 + 2\lambda_n\langle v^*, u - x_n\rangle. \tag{18}
$$

Proof Using the fact that $\langle p^*, p\rangle \leq \frac{s}{2}\|p^*\|^2 + \frac{1}{2s}\|p\|^2$ for any $p^* \in X^*$, $p \in X$ and any $s > 0$ yields

$$
2\lambda_n\langle v^*, u - x_{n+1}\rangle
$$

$$
= 2\lambda_n\langle v^*, x_n - x_{n+1}\rangle + 2\lambda_n\langle v^*, u - x_n\rangle
$$

$$
\leq \frac{t}{2(1+t)}\|x_n - x_{n+1}\|^2 + \frac{2(1+t)}{t}\lambda_n^2\|v^*\|^2 + 2\lambda_n\langle v^*, u - x_n\rangle.
$$

Then, we get from (11) that

$$
cW(x_{n+1}, u) - cW(x_n, u) + cW(x_n, x_{n+1})
$$

$$
- \frac{1}{1+t}\|x_n - x_{n+1}\|^2 + \frac{2t}{1+t}\lambda_n\beta_n\Psi(x_n) + \frac{t}{1+t}\lambda_n\beta_n\|\nabla\Psi(x_n)\|^2
$$

$$
\leq \lambda_n\beta_n\left((1+t)\lambda_n\beta_n - \frac{2}{L(1+t)} + \frac{t}{1+t}\right)\|\nabla\Psi(x_n)\|^2
$$

$$
+ \frac{t}{2(1+t)}\|x_n - x_{n+1}\|^2 + \frac{2(1+t)}{t}\lambda_n^2\|v^*\|^2 + 2\lambda_n\langle v^*, u - x_n\rangle.
$$

Hence, by $cW(x_n, x_{n+1}) \geq \|x_n - x_{n+1}\|^2$, we have

$$
cW(x_{n+1}, u) - cW(x_n, u) + \frac{t}{2+2t}\|x_n - x_{n+1}\|^2 + \frac{2t}{1+t}\lambda_n\beta_n\Psi(x_n)
$$

$$
+ \frac{t}{1+t}\lambda_n\beta_n\|\nabla\Psi(x_n)\|^2
$$

$$
\leq \lambda_n\beta_n\left((1+t)\lambda_n\beta_n - \frac{2}{L(1+t)} + \frac{t}{1+t}\right)\|\nabla\Psi(x_n)\|^2 + \frac{2(1+t)}{t}\lambda_n^2\|v^*\|^2
$$

$$
+ 2\lambda_n\langle v^*, u - x_n\rangle.
$$

Since $\lambda_n\beta_n < \frac{2c}{L}$, we have

$$
\lim_{t\to 0}\lambda_n\beta_n\left((1+t)\lambda_n\beta_n - \frac{2}{L(1+t)} + \frac{t}{1+t}\right) = \lambda_n\beta_n\left(\lambda_n\beta_n - \frac{2}{L}\right) < 0.
$$

Therefore, it suffices to take $t_0 > 0$ small enough, then set

$$
a = \frac{t_0}{2(1 + t_0)}, \qquad b = \frac{2(1 + t_0)}{t_0}
$$

to obtain (18). □

Proposition 3.3 *Let $\{x_n\}$ be a sequence generated by iterative formula* (9) *and let $u \in C \cap \operatorname{dom} A$. Take $\omega^* \in T_{A,C}(u), v^* \in A(u)$ and $p^* \in N_C(u)$, such that $v^* = \omega^* - p^*$. The following inequality holds*:

$$
cW(x_{n+1}, u) - cW(x_n, u) + a\left(\|x_n - x_{n+1}\|^2 + \frac{\lambda_n \beta_n}{2} \Psi(x_n) + \lambda_n \beta_n \|\nabla \Psi(x_n)\|^2 \right)
$$
$$
\leq b\lambda_n^2 \|v^*\|^2 + 2\lambda_n \langle \omega^*, u - x_n \rangle + \frac{a\lambda_n \beta_n}{2}\left(\Psi^*\left(\frac{4p^*}{a\beta_n} \right) - \sigma_C\left(\frac{4p^*}{a\beta_n} \right) \right). \tag{19}
$$

Proof First observe that

$$
2\lambda_n \langle v^*, u - x_n \rangle - \frac{a\lambda_n \beta_n}{2} \Psi(x_n)
$$
$$
= 2\lambda_n \langle \omega^*, u - x_n \rangle + 2\lambda_n \langle p^*, x_n \rangle - \frac{a\lambda_n \beta_n}{2} \Psi(x_n) - 2\lambda_n \langle p^*, u \rangle
$$
$$
= 2\lambda_n \langle \omega^*, u - x_n \rangle + \frac{a\lambda_n \beta_n}{2}\left(\left\langle \frac{4}{a\beta_n} p^*, x_n \right\rangle - \Psi(x_n) - \left\langle \frac{4}{a\beta_n} p^*, u \right\rangle \right)
$$
$$
\leq 2\lambda_n \langle \omega^*, u - x_n \rangle + \frac{a\lambda_n \beta_n}{2}\left(\Psi^*\left(\frac{4p^*}{a\beta_n} \right) - \left\langle \frac{4p^*}{a\beta_n}, u \right\rangle \right).
$$

Since $\frac{4p^*}{a\beta_n} \in N_C(u)$, the support function satisfies

$$
\sigma_C\left(\frac{4p^*}{a\beta_n} \right) = \left\langle \frac{4p^*}{a\beta_n}, u \right\rangle,
$$

whence

$$
2\lambda_n \langle v^*, u - x_n \rangle
$$
$$
\leq \frac{a\lambda_n \beta_n}{2} \Psi(x_n) + 2\lambda_n \langle \omega^*, u - x_n \rangle + \frac{a\lambda_n \beta_n}{2}\left(\Psi^*\left(\frac{4p^*}{a\beta_n} \right) - \sigma_C\left(\frac{4p^*}{a\beta_n} \right) \right). \tag{20}
$$

Hence by (18) and (20), we obtain (19). □

Theorem 3.1 *Let $\{x_n\}$ be a sequence generated by iterative formula* (9). *Then, we have the following*:

　(i) *For each $u \in \mathcal{S}, \lim_{n \to +\infty} W(x_n, u)$ exists.*

　(ii) *The series $\sum_{n=1}^{+\infty} \|x_n - x_{n+1}\|^2, \sum_{n=1}^{+\infty} \lambda_n \beta_n \Psi(x_n)$ and $\sum_{n=1}^{+\infty} \lambda_n \beta_n \|\nabla \Psi(x_n)\|^2$ are convergent.*

In particular, $\lim_{n \to +\infty} \|x_n - x_{n+1}\| = 0$. If, moreover, $\liminf_{n \to +\infty} \lambda_n \beta_n > 0$, then $\lim_{n \to +\infty} \Psi(x_n) = \lim_{n \to +\infty} \|\nabla \Psi(x_n)\| = 0$ and every weak cluster point of $\{x_n\}$ lies in C.

Proof Since $u \in \mathcal{S}$ one can take $\omega^* = 0$ in (19). By hypothesis the right-hand side is summable, and all the conclusions follow using Lemma 2.1. □

Theorem 3.2 *Let* $\{x_n\}$ *be a sequence generated by iterative formula* (9)*, and let* $\{z_k\}$ *be the sequence of weighted averages*

$$z_k = \frac{1}{\gamma_k} \sum_{n=1}^{k} \lambda_n x_n, \quad \text{where } \gamma_k = \sum_{n=1}^{k} \lambda_n.$$

Then every weak cluster of $\{z_k\}$ *lies in* $\mathcal{S}$.

Proof Let $u \in C \cap \operatorname{dom} A$. Take $\omega^* \in T_{A,C}(u)$, $v^* \in A(u)$ and $p^* \in N_C(u)$, so that $v^* = \omega^* - p^*$. By Proposition 3.3, we have

$$cW(x_{n+1}, u) - cW(x_n, u)$$

$$\leq b\lambda_n^2 \|v^*\|^2 + 2\lambda_n \langle \omega^*, u - x_n \rangle + \frac{a\lambda_n \beta_n}{2} \left(\Psi^* \left(\frac{4p^*}{a\beta_n} \right) - \sigma_C \left(\frac{4p^*}{a\beta_n} \right) \right).$$

Hence, we obtain

$$-c\frac{W(x_1, u)}{2\gamma_k}$$

$$\leq \frac{\sum_{n=1}^{k} b\lambda_n^2 \|v^*\|^2 + \sum_{n=1}^{k} \frac{a\lambda_n \beta_n}{2} (\Psi^*(\frac{4p^*}{a\beta_n}) - \sigma_C(\frac{4p^*}{a\beta_n}))}{2\gamma_k} + \frac{2\sum_{n=1}^{k} \langle \omega^*, \lambda_n u - \lambda_n x_n \rangle}{2\gamma_k}$$

$$= \frac{\sum_{n=1}^{k} b\lambda_n^2 \|v^*\|^2 + \sum_{n=1}^{k} \frac{a\lambda_n \beta_n}{2} (\Psi^*(\frac{4p^*}{a\beta_n}) - \sigma_C(\frac{4p^*}{a\beta_n}))}{2\gamma_k} + \left\langle \omega^*, u - \frac{\sum_{n=1}^{k} \lambda_n x_n}{\gamma_k} \right\rangle. \quad (21)$$

Then by (10), (21) and using that $\gamma_k \to +\infty$ as $k \to +\infty$, we obtain

$$\liminf_{k \to +\infty} \langle \omega^*, u - z_k \rangle \geq 0.$$

Finally, if z is any weak sequential cluster point of the sequence $\{z_k\}$, then $\langle \omega^*, u - z \rangle \geq 0$. Since $\omega^* \in T_{A,C}(u)$ and $T_{A,C}$ is maximal monotone, we obtain that $z \in \mathcal{S}$. □

Theorem 3.3 *Let* $\{x_n\}$ *be a sequence generated by iterative formula* (9) *and let A be a maximal monotone and strongly monotone operator. Then the sequence* $\{x_n\}$ *converges strongly as* $n \to +\infty$ *to a point in* $\mathcal{S}$.

Proof Take $u \in \mathcal{S} \subset C \cap \operatorname{dom} A$, $v^* \in A(u)$, $\omega^* \in T_{A,C}(u)$ and $p^* \in N_C(u)$, so that $v^* = \omega^* - p^*$. Since $v^* \in A(u)$ and $cJ(x_n) - cJ(x_{n+1}) - \lambda_n \beta_n \nabla \Psi(x_n) \in \lambda_n A(x_{n+1})$, the strong monotonicity of A implies

$$\langle cJ(x_n) - cJ(x_{n+1}) - \lambda_n \beta_n \nabla \Psi(x_n) - \lambda_n v^*, x_{n+1} - u \rangle \geq \lambda_n \alpha \|x_{n+1} - u\|^2.$$

We follow the arguments in the proof of Proposition 3.3 to obtain successively

$$2\lambda_n \alpha \|x_{n+1} - u\|^2 + cW(x_{n+1}, u) - cW(x_n, u)$$

$$+ a\left(\|x_n - x_{n+1}\|^2 + \frac{\lambda_n \beta_n}{2} \Psi(x_n) + \lambda_n \beta_n \|\nabla \Psi(x_n)\|^2 \right)$$

$$\leq b\lambda_n^2 \|v^*\|^2 + 2\lambda_n \langle \omega^*, u - x_n \rangle + \frac{a\lambda_n \beta_n}{2}\left(\Psi^*\left(\frac{4p^*}{a\beta_n}\right) - \sigma_C\left(\frac{4p^*}{a\beta_n}\right) \right). \tag{22}$$

Since $u \in S$, one can take $\omega^* = 0$ in (22). By $a(\|x_n - x_{n+1}\|^2 + \frac{\lambda_n \beta_n}{2}\Psi(x_n) + \lambda_n \beta_n \|\nabla \Psi(x_n)\|^2) \geq 0$, we have

$$2\lambda_n \alpha \|x_{n+1} - u\|^2 + cW(x_{n+1}, u) - cW(x_n, u)$$

$$\leq b\lambda_n^2 \|v^*\|^2 + \frac{a\lambda_n \beta_n}{2}\left(\Psi^*\left(\frac{4p^*}{a\beta_n}\right) - \sigma_C\left(\frac{4p^*}{a\beta_n}\right) \right).$$

Summation gives

$$2\alpha \sum_{n=1}^{+\infty} \lambda_n \|x_{n+1} - u\|^2$$

$$\leq cW(x_1, u) + b\sum_{n=1}^{+\infty} \lambda_n^2 \|v^*\|^2 + \sum_{n=1}^{+\infty} \frac{a\lambda_n \beta_n}{2}\left(\Psi^*\left(\frac{4p^*}{a\beta_n}\right) - \sigma_C\left(\frac{4p^*}{a\beta_n}\right) \right).$$

Since $\sum_{n=1}^{+\infty} \lambda_n = +\infty$, there exists subsequence $\{x_{n,k}\} \subset \{x_n\}$, such that $\lim_{k\to+\infty} \|x_{n,k} - u\| = 0$. Then, $\lim_{k\to+\infty} W(x_{n,k}, u) = 0$. Since $\lim_{n\to+\infty} W(x_n, u)$ exists by Theorem 3.1(i), we must have $\lim_{n\to+\infty} W(x_n, u) = 0$. Hence, by $cW(x_n, u) \geq \|x_n - u\|^2$, we have $\lim_{n\to+\infty} \|x_n - u\| = 0$. $\square$

4 The FBS method for the minimization

In this section, we consider the forward–backward splitting method in the special case where $A = \partial\Phi$ is the subdifferential of a proper, lower-semicontinuous and convex function $\Phi : X \to (-\infty, +\infty]$. The solution set $\mathcal{S}$ is equal to

$$(\partial\Phi + N_C)^{-1}(0) = \underset{C}{\text{Argmin}}\,\Phi.$$

Iterative Method 4.1 *Given* $x_0 \in X$, *set*

$$x_{n+1} = (cJ + \lambda_n \partial\Phi)^{-1}\left(cJx_n - \lambda_n \beta_n \nabla\Psi(x_n) \right), \tag{23}$$

where $\{\lambda_n\}, \{\beta_n\}$ *are two sequences of positive real numbers with* $\sum_{n=1}^{\infty} \lambda_n = +\infty$, $\sum_{n=1}^{\infty} \lambda_n^2 < +\infty$, $\beta_{n+1} - \beta_n \leq K$, $K > 0$, $0 < \bar{c} \leq \lambda_n \beta_n < \frac{2c}{L}$.

We also shall make the following Fenchel conjugate assumption:

$$\sum_{n=1}^{+\infty} \lambda_n \beta_n \left[\Psi^*\left(\frac{p^*}{\beta_n}\right) - \sigma_C\left(\frac{p^*}{\beta_n}\right) \right] < +\infty, \quad \forall p^* \in R(N_C).$$

The analysis relies on the study of the sequence $\{H_n(x_n)\}$, where H_n is the penalized function given by $H_n = \Phi + \beta_n \Psi$ for $n \geq 1$.

Proposition 4.1 *Let $\{x_n\}$ be a sequence generated by iterative formula* (23). *Then the sequence $\{H_n(x_n)\}$ converges as $n \to +\infty$.*

Proof Recall that $\frac{cJ(x_n) - cJ(x_{n+1})}{\lambda_n} - \beta_n \nabla \Psi(x_n) \in \partial \Phi(x_{n+1})$. The subdifferential inequality for Φ gives

$$\Phi(x_n) \geq \Phi(x_{n+1}) + \left\langle \frac{cJ(x_n) - cJ(x_{n+1})}{\lambda_n} - \beta_n \nabla \Psi(x_n), x_n - x_{n+1} \right\rangle,$$

and so

$$\Phi(x_{n+1}) - \Phi(x_n) + \frac{1}{2\lambda_n} \big(cW(x_{n+1}, x_n) + cW(x_n, x_{n+1}) \big) \leq \beta_n \langle \nabla \Psi(x_n), x_n - x_{n+1} \rangle. \qquad (24)$$

Then, by Lemma 3.1(ii), we have

$$\Psi(x_{n+1}) \leq \Psi(x_n) + \langle \nabla \Psi(x_n), x_{n+1} - x_n \rangle + \frac{L}{2} \|x_{n+1} - x_n\|^2, \qquad (25)$$

whence

$$\beta_{n+1} \Psi(x_{n+1}) - \beta_n \Psi(x_n)$$
$$\leq \beta_n \langle \nabla \Psi(x_n), x_{n+1} - x_n \rangle + \frac{L\beta_n}{2} \|x_{n+1} - x_n\|^2 + (\beta_{n+1} - \beta_n) \Psi(x_{n+1}). \qquad (26)$$

Adding (24) and (26), we obtain

$$H_{n+1}(x_{n+1}) - H_n(x_n) + \frac{1}{2\lambda_n} \big(cW(x_{n+1}, x_n) + cW(x_n, x_{n+1}) \big) - \frac{L\beta_n}{2} \|x_{n+1} - x_n\|^2$$
$$\leq (\beta_{n+1} - \beta_n) \Psi(x_{n+1}). \qquad (27)$$

Since $\lambda_n \beta_n < \frac{2c}{L}$ and $cW(x, y) \geq \|x - y\|^2$, $\forall x, y \in X$, we have

$$\frac{1}{2\lambda_n} \big(cW(x_{n+1}, x_n) + cW(x_n, x_{n+1}) \big) - \frac{L\beta_n}{2} \|x_{n+1} - x_n\|^2 \geq 0. \qquad (28)$$

Since $\beta_{n+1} - \beta_n \leq K$, by (27) and (28),

$$H_{n+1}(x_{n+1}) - H_n(x_n) \leq K\Psi(x_{n+1}).$$

By Theorem 3.1(i), we deduce that $\{x_n\}$ is bounded and $\{\Phi(x_n)\}$ is therefore bounded from below. Hence, the sequence $\{H_n(x_n)\}$ is also bounded from below. The right-hand side is summable by Theorem 3.1(ii), whence Lemma 2.1 implies that $\lim_{n \to +\infty} H_n(x_n)$ exists. $\square$

Proposition 4.2 *Let $\{x_n\}$ be a sequence generated by iterative formula* (23). *For each $u \in C$, we have $\sum_{n=1}^{+\infty} \lambda_n(H_{n+1}(x_{n+1}) - \Phi(u)) < +\infty$.*

Proof First observe that

$$H_{n+1}(x_{n+1}) - \Phi(u)$$

$$= \Phi(x_{n+1}) + \beta_n \Psi(x_n) - \Phi(u) + (\beta_{n+1} - \beta_n)\Psi(x_{n+1}) + \beta_n\big(\Psi(x_{n+1}) - \Psi(x_n)\big)$$

$$\leq \Phi(x_{n+1}) + \beta_n \Psi(x_n) - \Phi(u) + K\Psi(x_{n+1}) + \beta_n\big(\Psi(x_{n+1}) - \Psi(x_n)\big). \tag{29}$$

Using (25), we obtain

$$\beta_n\big(\Psi(x_{n+1}) - \Psi(x_n)\big) \leq \beta_n\big\langle \nabla\Psi(x_n), x_{n+1} - x_n \big\rangle + \frac{L\beta_n}{2}\|x_{n+1} - x_n\|^2$$

$$\leq \beta_n\|\nabla\Psi(x_n)\|\|x_{n+1} - x_n\| + \frac{L\beta_n}{2}\|x_{n+1} - x_n\|^2$$

$$\leq \frac{\beta_n}{2}\|\nabla\Psi(x_n)\|^2 + \frac{(L+1)\beta_n}{2}\|x_{n+1} - x_n\|^2. \tag{30}$$

Inequalities (29) and (30) give

$$\lambda_n\big(H_{n+1}(x_{n+1}) - \Phi(u)\big)$$

$$\leq \lambda_n\big(\Phi(x_{n+1}) + \beta_n\Psi(x_n) - \Phi(u)\big)$$

$$+ \lambda_n K\Psi(x_{n+1}) + \frac{\lambda_n\beta_n}{2}\|\nabla\Psi(x_n)\|^2 + \frac{(L+1)\lambda_n\beta_n}{2}\|x_{n+1} - x_n\|^2.$$

Since the sequence $\{\lambda_n\}$ is bounded and $0 < \bar{c} \leq \lambda_n\beta_n < \frac{2}{L}$, Theorem 3.1 implies

$$\sum_n^{+\infty}\left(\lambda_n K\Psi(x_{n+1}) + \frac{\lambda_n\beta_n}{2}\|\nabla\Psi(x_n)\|^2 + \frac{(L+1)\lambda_n\beta_n}{2}\|x_{n+1} - x_n\|^2\right) < +\infty.$$

On the other hand, the subdifferential inequality for Φ at points u and x_{n+1} gives

$$\Phi(u) \geq \Phi(x_{n+1}) + \left\langle \frac{cJx_n - cJx_{n+1}}{\lambda_n} - \beta_n\nabla\Psi(x_n), u - x_{n+1} \right\rangle. \tag{31}$$

Since $\Psi(u) = 0$, the subdifferential inequality for Ψ at points u and x_n gives

$$0 \geq \Psi(x_n) + \big\langle \nabla\Psi(x_n), u - x_n \big\rangle = \Psi(x_n) + \big\langle \nabla\Psi(x_n), u - x_{n+1} \big\rangle + \big\langle \nabla\Psi(x_n), x_{n+1} - x_n \big\rangle. \tag{32}$$

Combining (31) and (32), we obtain

$$2\lambda_n\big(\Phi(x_{n+1}) + \beta_n\Psi(x_n) - \Phi(u)\big) \leq 2\langle cJx_n - cJx_{n+1}, x_{n+1} - u \rangle + 2\lambda_n\beta_n\big\langle \nabla\Psi(x_n), x_n - x_{n+1} \big\rangle.$$

However,

$$2\langle cJx_n - cJx_{n+1}, x_{n+1} - u \rangle = cW(x_n, u) - cW(x_{n+1}, u) - cW(x_n, x_{n+1})$$

and

$$2\lambda_n\beta_n\big\langle \nabla\Psi(x_n), x_n - x_{n+1} \big\rangle \leq \frac{4}{L^2}\|\nabla\Psi(x_n)\|^2 + \|x_n - x_{n+1}\|^2.$$

Hence,

$$2\lambda_n\big(\Phi(x_{n+1}) + \beta_n\Psi(x_n) - \Phi(u)\big)$$

$$\leq cW(x_n, u) - cW(x_{n+1}, u) - cW(x_n, x_{n+1}) + \frac{4}{L^2}\left\|\nabla\Psi(x_n)\right\|^2 + \|x_n - x_{n+1}\|^2$$

$$\leq cW(x_n, u) - cW(x_{n+1}, u) + \frac{4}{L^2}\left\|\nabla\Psi(x_n)\right\|^2.$$

We conclude that

$$\sum_{n=1}^{m} 2\lambda_n\left(\Phi(x_{n+1}) + \beta_n\Psi(x_n) - \Phi(u)\right) \leq cW(x_1, u) - cW(x_{m+1}, u) + \frac{4}{L^2}\sum_{n=1}^{m}\left\|\nabla\Psi(x_n)\right\|^2$$

for $m \geq 1$. In view of Theorem 3.1, this show

$$\sum_{n=1}^{+\infty} \lambda_n\left(\Phi(x_{n+1}) + \beta_n\Psi(x_n) - \Phi(u)\right) < +\infty,$$

and completes the proof. □

The duality mapping J is said to be weakly continuous on a smooth Banach space if $x_n \rightharpoonup x$ implies $J(x_n) \rightharpoonup J(x)$. This happens, for example, if X is a Hilbert space, or finite-dimensional and smooth, or $l^p, 1 < p < +\infty$. This property of Banach spaces was introduced by Browder [7]. More information can be found in [10].

Theorem 4.1 *Let $\{x_n\}$ be a sequence generated by iterative formula* (23). *Then every weak cluster point of $\{x_n\}$ lies in $\mathcal{S}$. If the duality mapping J is weakly continuous, then $\{x_n\}$ convergence weakly as $n \to +\infty$ to a point in $\mathcal{S}$.*

Proof Since $\sum_{n=1}^{+\infty} \lambda_n = +\infty$, Propositions 4.1 and 4.2 imply $\lim_{n\to+\infty} H_n(x_n) \leq \Phi(u)$ whenever $u \in C$. Suppose that a subsequence $\{x_{n,k}\}$ of $\{x_n\}$ converges weakly to some $\hat{x}$ as $k \to +\infty$. Then $\hat{x} \in C$ by Theorem 3.1. The weak lower-semicontinuity of Φ and $\Phi = H_n - \beta_n\Psi \leq H_n$ then gives

$$\Phi(\hat{x}) \leq \liminf_{k\to+\infty} \Phi(x_{n,k}) \leq \liminf_{k\to+\infty} H_{n,k}(x_{n,k}) = \lim_{n\to+\infty} H_n(x_n) \leq \Phi(u).$$

Therefore, $\hat{x}$ minimizes Φ on C, and so $\hat{x} \in \mathcal{S}$.

Clearly, the sequence $\{x_n\}$ is bounded (see Theorem 3.1(i)). The space being reflexive, it suffices to prove that $\{x_n\}$ has only one weak cluster point as $n \to +\infty$. Suppose otherwise that $x_{n,l} \rightharpoonup \bar{x}$ and $x_{n,k} \rightharpoonup \hat{x}$. Since

$$2\langle J(x_n), \bar{x} - \hat{x}\rangle = W(x_n, \hat{x}) - W(x_n, \bar{x}) - \|\hat{x}\|^2 + \|\bar{x}\|^2,$$

we deduce the existence of $\lim_{n\to+\infty} 2\langle J(x_n), \bar{x} - \hat{x}\rangle$. Hence,

$$\lim_{l\to+\infty}\langle J(x_{n,l}), \bar{x} - \hat{x}\rangle - \lim_{k\to+\infty}\langle J(x_{n,k}), \bar{x} - \hat{x}\rangle = 0.$$

Since the duality mapping J is weakly continuous, we have

$$\langle J(\bar{x}) - J(\hat{x}), \bar{x} - \hat{x}\rangle = 0.$$

Since X is strictly convex, we have that $\bar{x} = \hat{x}$. □

If $\Phi : X \to (-\infty, +\infty]$ is also a strongly convex function, that is, there exists $\lambda > 0$, for any $0 < t < 1$, any $x, y \in \mathrm{dom}\,\Phi$ such that $t\Phi(x) + (1-t)\Phi(y) \geq \Phi(tx + (1-t)y) + \lambda t(1-t)\|x - y\|^2$, then $\partial\Psi$ is strong monotone. Hence, the following theorem follows immediately from Theorem 3.3.

Theorem 4.2 *Let $\{x_n\}$ be a sequence generated by iterative formula* (23) *and let Φ be a proper, lower semicontinuous and strongly convex function. Then the sequence $\{x_n\}$ converges strongly as $n \to +\infty$ to a point in $\mathcal{S}$.*

5 Additional result

The purpose of this section is to prove a convergence result without Fenchel conjugate assumption.

Iterative Method 5.1 *Given $x_0 \in X$, set*

$$x_{n+1} = (cJ + \lambda_n A)^{-1}(cJx_n - \lambda_n \nabla\Psi(x_n)), \tag{33}$$

where $\{\lambda_n\}$ is a sequence of positive real numbers with $\sum_{n=1}^{\infty} \lambda_n = +\infty$, $\sum_{n=1}^{\infty} \lambda_n^{\frac{4}{3}} < +\infty$.

Keeping the notations of the preceding section, set $z_k = \frac{1}{\gamma_k} \sum_{n=1}^{k} \lambda_n x_n$, where $\gamma_k = \sum_{n=1}^{k} \lambda_n$. The following gives the weak ergodic convergence of the sequence $\{x_n\}$ given by (33).

Proposition 5.1 *Let $\{x_n\}$ be a sequence generated by iterative formula* (33). *Assume that the sequence $\{\lambda_n^{\frac{1}{3}} \nabla\Psi(x_n)\}$ is bounded. Then every weak cluster of $\{z_k\}$ lies in $\mathcal{S}$.*

Proof Take $u \in C \cap \mathrm{dom}\,A$, $v^* \in A(u)$ such that $v^* = v^* + 0 \in A(u) + N_C(u) = T_{A,C}(u)$. Since $v^* \in A(u)$ and $cJ(x_n) - cJ(x_{n+1}) - \lambda_n \nabla\Psi(x_n) \in \lambda_n A(x_{n+1})$, the monotonicity of A implies

$$\langle cJ(x_n) - cJ(x_{n+1}) - \lambda_n \nabla\Psi(x_n) - \lambda_n v^*, x_{n+1} - u \rangle \geq 0,$$

and so

$$\langle cJ(x_n) - cJ(x_{n+1}), u - x_{n+1} \rangle \leq \langle \lambda_n \nabla\Psi(x_n) + \lambda_n v^*, u - x_{n+1} \rangle.$$

Then, we get from (7) that

$$cW(x_{n+1}, u) - cW(x_n, u) + cW(x_n, x_{n+1}) \leq 2\lambda_n \langle \nabla\Psi(x_n) + v^*, u - x_{n+1} \rangle.$$

By developing the right-hand side, we deduce the following inequality:

$$2\lambda_n \langle \nabla\Psi(x_n) + v^*, x_n - u \rangle + 2\lambda_n \langle \nabla\Psi(x_n) + v^*, x_{n+1} - x_n \rangle$$
$$\leq cW(x_n, u) - cW(x_{n+1}, u) - cW(x_n, x_{n+1}). \tag{34}$$

Now, combing the facts that

$$2\lambda_n \langle \nabla\Psi(x_n) + v^*, x_{n+1} - x_n \rangle \geq -\|x_{n+1} - x_n\|^2 - \lambda_n^2 \|\nabla\Psi(x_n) + v^*\|^2$$

and

$$\langle \nabla\Psi(x_n) + v^*, x_n - u\rangle = \langle \nabla\Psi(x_n), x_n - u\rangle + \langle v^*, x_n - u\rangle,$$

we derive from (34) that

$$2\lambda_n\langle \nabla\Psi(x_n), x_n - u\rangle + 2\lambda_n\langle v^*, x_n - u\rangle - \lambda_n^2\|\nabla\Psi(x_n) + v^*\|^2$$
$$\leq cW(x_n, u) - cW(x_{n+1}, u) - cW(x_n, x_{n+1}) + \|x_{n+1} - x_n\|^2. \tag{35}$$

Since $u \in C \cap \mathrm{dom}\,A$ and $C = \mathrm{argmin}(\Psi)$, we have $\nabla\Psi(u) = 0$. Hence,

$$\langle \nabla\Psi(x_n), x_n - u\rangle = \langle \nabla\Psi(x_n) - \nabla\Psi(u), x_n - u\rangle \geq 0. \tag{36}$$

Since $cW(x_n, x_{n+1}) \geq \|x_{n+1} - x_n\|^2$, then by (35) and (36) we have

$$2\lambda_n\langle v^*, x_n - u\rangle - \lambda_n^2\|\nabla\Psi(x_n) + v^*\|^2 \leq cW(x_n, u) - cW(x_{n+1}, u).$$

Summing up these inequalities over n from 1 to k, and dividing by γ_k gives

$$2\langle v^*, z_k - u\rangle \leq \frac{cW(x_1, u)}{\gamma_k} + \frac{\sum_{n=1}^{k}\lambda_n^2\|\nabla\Psi(x_n) + v^*\|^2}{\gamma_k}. \tag{37}$$

Since $\{\lambda_n^{\frac{1}{3}}\nabla\Psi(x_n)\}$ is bounded, due to $\sum_{n=1}^{\infty}\lambda_n^{\frac{4}{3}} < +\infty$ and $\{\lambda_n^2\|\nabla\Psi(x_n)\|^2\} = \{\lambda_n^{\frac{4}{3}}\lambda_n^{\frac{2}{3}}\|\nabla\Psi(x_n)\|^2\}$, we have

$$\sum_{n=1}^{+\infty}\lambda_n^2\|\nabla\Psi(x_n) + v^*\|^2 \leq 2\left(\sum_{n=1}^{+\infty}\lambda_n^2\|\nabla\Psi(x_n)\|^2 + \sum_{n=1}^{+\infty}\lambda_n^2\|v^*\|^2\right) < +\infty.$$

Finally, since $\gamma_k \to +\infty$ as $k \to +\infty$, we conclude that

$$\lim_{k\to+\infty}\frac{\sum_{n=1}^{k}\lambda_n^2\|\nabla\Psi(x_n) + v^*\|^2}{\gamma_k} = 0.$$

Consequently, if z is any weak sequential cluster point of the sequence $\{z_k\}$, letting $k \to +\infty$ on both side of (37) yields

$$\langle v^*, z - u\rangle \leq 0.$$

Then by maximal monotonicity of $A + N_C$, we conclude that $0 \in (A + N_C)(z)$, that is, $z \in \mathcal{S}$.   □

6 Concluding remarks

In this paper, we considered a class of forward–backward splitting methods based on Lyapunov distance for variational inequalities and convex minimization problem in a reflexive, strictly convex and smooth Banach space. Weak and strong convergence results have been obtained for the forward–backward splitting method under the key Fenchel conjugate assumption. Finally, we have also obtained a weak convergence result without Fenchel conjugate assumption.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Albert, Y.: Iterative regularization in Banach spaces. Sov. Math. **30**, 1–8 (1986)
2. Albert, Y.: Decomposition theorem in Banach spaces. Fields Inst. Commun. **25**, 77–93 (2000)
3. Attouch, H., Czarnecki, M.O., Peypouquet, J.: Prox-penalization and splitting methods for constrained variational problems. SIAM J. Optim. **21**(1), 149–173 (2011)
4. Attouch, H., Czarnecki, M.O., Peypouquet, J.: Coupling forward–backward with penalty schemes and parallel splitting for constrained variational inequalities. SIAM J. Optim. **21**(4), 1251–1274 (2011)
5. Baillon, J.B., Haddad, G.: Quelques propriétés des opérateurs angle-bornés et n-cycliquement monotones. Israel J. Math. **26**, 137–150 (1977)
6. Bredies, K.: A forward–backward splitting algorithm for the minimization of nonsmooth convex functionals in Banach space. Inverse Probl. **25**, 015005 (2008)
7. Browder, F.E.: Fixed point therems for nonlinear semicontractive mappings in Banach spaces. Arch. Ration. Mech. Anal. **21**, 259–269 (1965)
8. Cioranescu, I.: Geometry of Banach Spaces, Duality Mappings and Nonlinear Problems. Kluwer Academic Publishers, Dordrecht (1990)
9. Combettes, P.L., Wajs, V.R.: Signal recovery by proximal forward–backward splitting. Multiscale Model. Simul. **4**, 1168–1200 (2005)
10. Gossez, J.P., Dozo, E.L.: Some geometric properties related to the fixed point theory for nonexpansive mappings. Pacific J. Math. **40**, 565–573 (1972)
11. Guan, W.B., Song, W.: The generalized forward–backward splitting method for the minimization of the sum of two functions in Banach spaces. Numer. Funct. Anal. Optim. **36**, 867–886 (2015)
12. Moudafi, A.: On the convergence of the forward–backward algorithm for null-point problems. J. Nonlinear Var. Anal. **2**, 263–268 (2018)
13. Noun, N., Peypouquet, J.: Forward–backward penalty scheme for constrained convex minimization without inf-compactness. J. Optim. Theory Appl. **158**, 787–795 (2013)
14. Peypouquet, J.: Coupling the gradient method with a general exterior penalization scheme for convex minimization. J. Optim. Theory Appl. **153**(1), 123–138 (2012)
15. Yao, Y., Postolache, M., Yao, J.C.: An iterative algorithm for solving the generalized variational inequalities and fixed points problems. Mathematics **7**, Article ID 61 (2019)
16. Yao, Y., Shahzad, N.: Strong convergence of a proximal point algorithm with general errors. Optim. Lett. **6**, 621–628 (2012)
17. Yuan, H.: A splitting algorithm in a uniformly convex and 2-uniformly smooth Banach space. J. Nonlinear Funct. Anal. **2018**, Article ID 26 (2018)
18. Zegeye, H., Shahzad, N., Yao, Y.H.: Minimum-norm solution of variational inequality and fixed point problem in Banach spaces. Optimization **64**, 453–471 (2015)