



The forward–backward splitting method for non-Lipschitz continuous minimization problems in Banach spaces

Wei-Bo Guan¹ · Wen Song¹

Received: 2 March 2021 / Accepted: 7 December 2021 / Published online: 8 March 2022

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

In this paper we propose the forward–backward splitting methods with linesearches for solving nonsmooth optimization problems without the standard assumption of the Lipschitz continuity of the gradient in Banach spaces. We prove the weak convergence of the iterative sequence generated by these methods, and further prove convergence with asymptotic rate $\frac{1}{n}$ to the optimal value under the assumption of the boundedness of the iterative sequence.

Keywords Banach space · Forward–backward splitting method · Minimization problems · Error bound · Linesearch

1 Introduction

Let X be a reflexive, strictly convex and smooth Banach space with dual space X^* . Consider the optimization problem:

$$\min_{x \in X} \Phi(x) := f(x) + g(x), \quad (1)$$

where $f : X \rightarrow (-\infty, +\infty]$ is proper, lower semicontinuous and convex, $g : X \rightarrow (-\infty, +\infty)$ is convex Gâteaux differentiable. We denote by Φ^* the infimum value of the problem (1). The set of solutions to problem (1) is denoted by \mathcal{S} . We shall assume in what follows that the set \mathcal{S} is nonempty. Since Φ is proper, lower semicontinuous and convex, \mathcal{S} is a nonempty closed and convex set. Despite its simple form, problem

✉ Wen Song
wsong@hrbnu.edu.cn

Wei-Bo Guan
guanweibo0451@163.com

¹ School of Mathematical Sciences, Harbin Normal University, Harbin 150025, People's Republic of China

(1) has been shown to cover a wide range of apparently unrelated signal recovery formulations (see [12–14, 18]).

The forward–backward splitting method is an effective method to solve (1), which allows to decouple the contributions of the functions f and g in a gradient descent step determined by f and in a backward implicit step induced by g . Forward–backward methods belong to the class of proximal splitting methods. These methods require the computation of the proximity operator and the approximation of proximal points (see [14, 20]).

The forward–backward splitting iteration procedure proposed in [14] is given in Hilbert space H and is governed by the updating rule:

$$x_{n+1} \in \operatorname{argmin}_{y \in H} \left\{ \frac{1}{2} \|y - x_n\|^2 + t_n (\langle \nabla g(x_n), y \rangle + f(y)) \right\}. \quad (2)$$

Generalization of this method from Hilbert spaces to Banach spaces 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. The proposed forward–backward splitting method for solving problem (1) in Banach space suggested in this work aims at establishing a bridge between the well known forward–backward splitting in Hilbert spaces [14, 20] and that in Banach spaces. The use of a general regularizing function $\|\cdot\|^p$ instead of the square of the norm is rather natural in a Banach space, where the square of the norm loses the privileged role it enjoys in Hilbert spaces [4]. For instance, it is not hard to verify that in the spaces $X = l_p$ or $X = L_p$ ($1 < p < +\infty$), $x_{n+1} \in \operatorname{argmin}_{y \in X} \left\{ \frac{1}{p} \|y - x_n\|^p + t_n (\langle \nabla g(x_n), y \rangle + f(y)) \right\}$ becomes simpler than $x_{n+1} \in \operatorname{argmin}_{y \in X} \left\{ \frac{1}{2} \|y - x_n\|^2 + t_n (\langle \nabla g(x_n), y \rangle + f(y)) \right\}$. In [10], the following generalization of the forward–backward iteration procedure was proposed in reflexive Banach spaces X :

$$x_0 \in X, \quad x_{n+1} \in \operatorname{argmin}_{y \in X} \left\{ \frac{1}{p} \|y - x_n\|^p + t_n (\langle \nabla g(x_n), y \rangle + f(y)) \right\}, \quad (3)$$

where, the gradient operator ∇g is $(p-1)$ Hölder-continuous on X , i.e., there exists a constant L such that

$$\|\nabla g(x) - \nabla g(y)\| \leq L \|x - y\|^{p-1}, \quad \forall x, y \in X.$$

In [16], Guan and Song proposed another type generalization of the forward–backward method in reflexive Banach spaces

$$x_{n+1} = \operatorname{argmin}_{y \in X} \left\{ \frac{1}{p} \|y\|^p - \langle J_p(x_n), y \rangle + t_n (\langle \nabla g(x_n) + J_p(z_n), y \rangle + f(y)) \right\}, \quad (4)$$

where $J_p : X \rightarrow X^*$ is the p -duality mapping and $\{z_n\}$ is absolutely summable sequence.

In particular, if X is a Hilbert space and $p = 2$, then formula (3) and (4) reduces to formula (2). In [16], it is shown that the sequence of functional values converges with the convergence rate n^{1-p} to the optimal value of Problem (1) under appropriate assumptions. In [17], Guan and Song further extended forward–backward splitting method (3) to more general case, i.e., by taking a convex combination of the current step and the previous step:

$$\begin{cases} y_n = \operatorname{argmin}_{y \in X} \left\{ \frac{1}{p} \|x_n - y\|^p + t_n (\langle \nabla g(x_n) + J_p(z_n), y \rangle + f(y)) \right\}, \\ x_{n+1} = (1 - \lambda_n)x_n + \lambda_n y_n, \end{cases} \quad (5)$$

and proved that the sequence of the functional values converges with an asymptotic rate n^{1-p} to the optimal value. They also proved that the sequence of the functional values converges with an asymptotic rate under an error bound assumption.

The convergence of the forward–backward splitting method to an optimal solution of (1) is usually established under the assumption that the gradient of g is Lipschitz continuous and the stepsize t_n is taken less than some constant related with Lipschitz modulus. When ∇g is Lipschitz continuous but somehow the Lipschitz constant is not known or ∇g is not Lipschitz continuous, finding the stepsize t_n that guarantees the convergence of (2) would be a challenge.

Beck and Teboulle [6] proposed a proximal gradient method with backtracking step-size rules overcome this inconvenience. As far as we observe, the theory of convergence and complexity for the proximal forward–backward is almost complete under such a Lipschitz assumption. However, the Lipschitz condition fails in many natural circumstance; see, e.g., [15]. Bello Cruz and Nghia [7] introduced two new linesearches into the frame of the forward–backward splitting method and prove the convergence analysis and complexity results of cost values. These two linesearch rules were also recently studied in [19,21,22] in conjunction with the forward–backward splitting algorithm for convex minimization problems without the assumption of the Lipschitz continuous gradient.

Bauschke, Bolte and Teboulle [5] introduced Bregman-based proximal gradient methods which share most of the convergence properties and complexity of the classical proximal-gradient, instead of the restrictive condition of Lipschitz continuity of the gradient of the differentiable they assuming a more general and flexible convexity condition. [8] further extended the above approach to nonconvex composite model with smooth adaptable functions and proved the global convergence to a critical point under natural assumptions on the problems data.

In this paper, following the lines of [7,21], we propose the forward–backward splitting method with linesearches in the framework of Banach spaces. The main advantage of our algorithms is that the Lipschitz constants of the gradient of functions do not require in computation. The paper is organized as follows. The next section presents some preliminary results that will be used throughout the paper. Section 4 is devoted to the study of the forward–backward splitting method (3) with **Linesearch 1**. We prove that the sequence of the functional values converges with an asymptotic rate $\frac{1}{n}$ to the optimal value of the minimization Problem (1). In Sect. 5, we further study forward–backward splitting method (5) with **Linesearch 2**. We prove that the

sequence of the functional values converges with an asymptotic rate $\frac{1}{n}$ to the optimal value. We also prove that the sequence of the functional values Q -linearly converges under an error bound assumption if $t_n \geq \bar{t} > 0$.

2 Preliminaries

In this section we present some definitions and results needed for our paper. Let f be a lower semi-continuous proper convex function from X to $(-\infty, +\infty]$. We denote the domain of f by $\text{dom } f := \{x \in X \mid f(x) < +\infty\}$. The subdifferential of f at $x \in X$ is the convex set

$$\partial f(x) = \{x^* \in X^* : \langle x^*, y - x \rangle \leq f(y) - f(x), \forall y \in X\}. \quad (6)$$

It follows from definition (6) that a point \hat{x} is a minimizer of f if and only if $0 \in \partial f(\hat{x})$. The subdifferential mapping $x \rightarrow \partial f(x)$ has following property of monotonicity, i.e.,

$$\langle x_1^* - x_2^*, x_1 - x_2 \rangle \geq 0, \quad \forall x_1, x_2 \in X, \forall x_1^* \in \partial f(x_1), \forall x_2^* \in \partial f(x_2).$$

The subdifferential operator ∂f is maximal monotone [2]. Moreover, the graph of ∂f is demiclosed [2], i.e., if $\{(x_n, x_n^*)\} \subset \text{Gph}(\partial f)$ satisfies that $\{x_n\}$ converges weakly to x and $\{x_n^*\}$ converges strongly to x^* , then $(x, x^*) \in \text{Gph}(\partial f)$. If f and g are two lower semicontinuous proper convex functions and if the regularity condition $0 \in \text{int}(\text{dom } f - \text{dom } g)$ holds, then for any $\bar{x} \in \text{dom } f \cap \text{dom } g$, we have (see [9]) $\partial(f + g)(\bar{x}) = \partial f(\bar{x}) + \partial g(\bar{x})$.

The p -duality mapping $J_p : X \rightarrow X^*$ is defined by

$$J_p(x) = \{x^* \in X^* \mid \langle x^*, x \rangle = \|x^*\| \|x\|, \|x^*\| = \|x\|^{p-1}\}, \quad \forall x \in X.$$

The Hahn-Banach theorem guarantees that $J_p(x) \neq \emptyset$ for every $x \in X$. It is clear that $J_p(x) = \partial(\frac{1}{p}\|\cdot\|^p)(x)$ for all $x \in X$. It is well known that if X is smooth, then J_p is single valued and is norm-to-weak star continuous. Properties of the duality mapping have been given in [1, 11, 23].

Let $\{x_n\}$ be a sequence in X that converges to \bar{x} . We say that the convergence is Q -linear if there exists a constant $r \in (0, 1)$ such that

$$\|x_{n+1} - \bar{x}\| \leq r \|x_n - \bar{x}\|,$$

for all n sufficiently large. We say that the convergence is R -linear if there exists a sequence of nonnegative scalars $\{\alpha_n\}$ such that

$$\|x_n - \bar{x}\| \leq \alpha_n, \quad \forall n \geq 1,$$

and $\{\alpha_n\}$ converges Q -linearly to zero.

Definition 1 [2] A functional f is called lower semicontinuous at the point $x_0 \in \text{dom } f$ if for any sequence $x_n \in \text{dom } f$ such that $x_n \rightarrow x_0$ there holds the inequality

$$f(x_0) \leq \liminf_{n \rightarrow \infty} f(x_n). \tag{7}$$

If the inequality (7) occurs with the condition that the convergence of x_n to x_0 is weak, then the functional f is called weakly lower semicontinuous at x_0 .

Lemma 1 [2] *Let f be a convex and lower semicontinuous functional. Then it is weakly lower semicontinuous.*

Lemma 2 [3] *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 $\sum_{n=1}^{\infty} b_n < +\infty$.*

Definition 2 [10] Let $f : X \rightarrow (-\infty, +\infty]$ be proper, convex and lower semicontinuous. f is called totally convex at $\bar{x} \in X$, if for each $x^* \in \partial f(\bar{x})$ and each sequence $\{x_n\}$, the following implication holds

$$f(x_n) - f(\bar{x}) - \langle x^*, x_n - \bar{x} \rangle \rightarrow 0 \implies \|x_n - \bar{x}\| \rightarrow 0.$$

The following standing assumptions on the data of problem (1) will be used throughout the paper:

Assumption 1 The gradient ∇g is uniformly continuous on any bounded subset of X and maps any bounded subset of X to a bounded set in X^* .

3 The FBS Method with Linesearch 1

In this section, we shall consider the convergence and convergence rate of the following forward–backward splitting method:

Linesearch 1(x, α, θ, β).
 Given $x \in \text{dom } f$, $p > 1$, $\alpha > 0$, $\theta \in (0, 1)$ and $\beta \in (\frac{1}{2}, 1)$.
Input.
 Set $t = \alpha$ and $\hat{x}_t = \text{argmin}_{y \in X} \{ \frac{1}{p} \|x - y\|^p + t(\langle \nabla g(x), y \rangle + f(y)) \}$.
While $t \|\nabla g(\hat{x}_t) - \nabla g(x)\| > \beta \|\hat{x}_t - x\|^{p-1}$
 do $t = \theta t$.
End While
Output. t .

Iterative Method 3.1. Given $x_0 \in X$, for every $n \in \mathbb{N}$, set

$$x_{n+1} = \text{argmin}_{y \in X} \{ \frac{1}{p} \|x_n - y\|^p + t_n(\langle \nabla g(x_n), y \rangle + f(y)) \},$$

where $t_n = \text{Linesearch 1}(x_n, \alpha, \theta, \beta)$.

Lemma 3 [17] For any $x \in X$ and $t > 0$, let

$$\hat{x}_t = \operatorname{argmin}_{y \in X} \left\{ \frac{1}{p} \|y - x\|^p + t(\langle \nabla g(x), y \rangle + f(y)) \right\}.$$

Then, for any $0 < t_1 \leq t_2$, we have

$$\|x - \hat{x}_{t_1}\| \leq \|x - \hat{x}_{t_2}\| \leq \left(\frac{t_2}{t_1}\right)^{\frac{1}{p-1}} \|x - \hat{x}_{t_1}\|. \tag{8}$$

Lemma 4 If $x \in \operatorname{dom} f$, then **Linesearch 1**(x, α, θ, β) stops after finitely many steps.

Proof If $x \in \mathcal{S}$, then $x = \hat{x}_\alpha$. Thus the linesearch stops at zero step and gives us the output α . If $x \notin \mathcal{S}$, by contradiction suppose that for all $t_k = \alpha\theta^k, k \in \mathbf{N}$

$$t_k \|\nabla g(\hat{x}_{t_k}) - \nabla g(x)\| > \beta \|\hat{x}_{t_k} - x\|^{p-1}. \tag{9}$$

It follows from Lemma 3 that $\|x - \hat{x}_{t_k}\| \leq \|x - \hat{x}_\alpha\|$, that is, $\{\hat{x}_{t_k}\}$ is bounded. Thus we get from (9) that $\|\hat{x}_{t_k} - x\| \rightarrow 0$ as $k \rightarrow +\infty$ thanks to Assumption 1. The latter implies $\|\nabla g(\hat{x}_{t_k}) - \nabla g(x)\| \rightarrow 0$ as $k \rightarrow +\infty$ by Assumption 1 again. From (9) we also obtain

$$\lim_{k \rightarrow +\infty} \frac{\|\hat{x}_{t_k} - x\|^{p-1}}{t_k} = 0. \tag{10}$$

The optimality of \hat{x}_{t_k} implies

$$0 \in \partial \left(\frac{1}{p} \|x - \cdot\|^p + t_k(\langle \nabla g(x), \cdot \rangle + f(\cdot)) \right) (\hat{x}_{t_k}).$$

Then, we have

$$\frac{J_p(x - \hat{x}_{t_k})}{t_k} - \nabla g(x) \in \partial f(\hat{x}_{t_k}).$$

By letting $k \rightarrow +\infty$ in the above inclusion and using (10), we get from the demiclosedness of $\operatorname{Gph}(\partial f)$ that $0 \in \partial f(x) + \nabla g(x)$. This contradicts the assumption that x is not an optimal solution to problem (1) and completes the proof of the lemma. \square

Remark 1 We observe from Lemma 4 that for finding the stepsize t_n in the above scheme is finite. Hence the choice of sequence $\{x_n\}$ in Iterative Method 3.1 is well defined. Another important feature from the definition of **Linesearch 1** useful for our analysis is the following inequality

$$t_n \|\nabla g(x_{n+1}) - \nabla g(x_n)\| \leq \beta \|x_{n+1} - x_n\|^{p-1}. \tag{11}$$

Proposition 1 *Let $\{x_n\}$ be a sequence generated by Iterative Method 3.1 and define*

$$h(x_n) := f(x_n) - f(x_{n+1}) + \langle \nabla g(x_n), x_n - x_{n+1} \rangle.$$

Then, we have

- (i) $\|x_n - x_{n+1}\|^p \leq t_n h(x_n)$.
- (ii) $\Phi(x_{n+1}) \leq \Phi(x_n) - (1 - \beta)h(x_n)$.
- (iii) $\Phi(x_n)$ converges and $\sum_{n=1}^\infty h(x_n) < +\infty$.

Proof (i) The optimality of x_{n+1} implies

$$0 \in \partial \left(\frac{1}{p} \|x_n - \cdot\|^p + t_n (\langle \nabla g(x_n), \cdot \rangle + f(\cdot)) \right) (x_{n+1}).$$

Then, we have

$$\frac{J_p(x_n - x_{n+1})}{t_n} - \nabla g(x_n) \in \partial f(x_{n+1}). \tag{12}$$

Hence,

$$\left\langle \frac{J_p(x_n - x_{n+1})}{t_n} - \nabla g(x_n), x_n - x_{n+1} \right\rangle \leq f(x_n) - f(x_{n+1}),$$

and so that

$$\|x_n - x_{n+1}\|^p \leq t_n f(x_n) - t_n f(x_{n+1}) + t_n \langle \nabla g(x_n), x_n - x_{n+1} \rangle = t_n h(x_n).$$

(ii) Using the definition of $h(x_n)$ and (11) we have

$$\begin{aligned} \Phi(x_n) - \Phi(x_{n+1}) &= h(x_n) + g(x_n) - g(x_{n+1}) - \langle \nabla g(x_n), x_n - x_{n+1} \rangle \\ &\geq h(x_n) + \langle \nabla g(x_{n+1}), x_n - x_{n+1} \rangle - \langle \nabla g(x_n), x_n - x_{n+1} \rangle \\ &\geq h(x_n) - \langle \nabla g(x_n) - \nabla g(x_{n+1}), x_n - x_{n+1} \rangle \\ &\geq h(x_n) - \|\nabla g(x_n) - \nabla g(x_{n+1})\| \|x_n - x_{n+1}\| \\ &\geq h(x_n) - \frac{\beta}{t_n} \|x_n - x_{n+1}\|^p \\ &\geq h(x_n) - \beta h(x_n) \\ &= (1 - \beta)h(x_n). \end{aligned}$$

(iii) The conclusion follows using Lemma 2 and Proposition 1 (ii). □

Proposition 2 *Let $\{x_n\}$ be a sequence generated by Iterative Method 3.1. Assume the sequence $\{x_n\}$ is bounded. Then,*

- (i) $\lim_{n \rightarrow \infty} \Phi(x_n) = \Phi^*$.
- (ii) all weak accumulation points of $\{x_n\}$ belong to \mathcal{S} .

Proof (i) Let $\hat{x} \in \mathcal{S}$. Since $g(x_n) - g(\hat{x}) \leq \langle \nabla g(x_n), x_n - \hat{x} \rangle$ and $\Phi(x_n) - \Phi^* = f(x_n) - f(\hat{x}) + g(x_n) - g(\hat{x})$, we have that

$$\begin{aligned} \Phi(x_n) - \Phi^* &\leq f(x_n) - f(\hat{x}) + \langle \nabla g(x_n), x_n - \hat{x} \rangle \\ &= h(x_n) + f(x_{n+1}) - f(\hat{x}) + \langle \nabla g(x_n), x_{n+1} - \hat{x} \rangle. \end{aligned} \tag{13}$$

Then, by (12) and (13), we have

$$\begin{aligned} \Phi(x_n) - \Phi^* &\leq h(x_n) - \left\langle \frac{J_p(x_n - x_{n+1})}{t_n}, \hat{x} - x_{n+1} \right\rangle \\ &\leq h(x_n) + \frac{1}{t_n} \|x_n - x_{n+1}\|^{p-1} \|\hat{x} - x_{n+1}\|. \end{aligned} \tag{14}$$

Now let us split our further analysis into two distinct cases.

Case 1 Suppose that there exists $\bar{t} > 0$ such that $t_n \geq \bar{t}$ for all $n \in \mathbf{N}$. Then, by Proposition 1 and (14), we have

$$\begin{aligned} \Phi(x_n) - \Phi^* &\leq \frac{\Phi(x_n) - \Phi(x_{n+1})}{1 - \beta} + \frac{1}{t_n} \|x_n - x_{n+1}\|^{p-1} \|\hat{x} - x_{n+1}\| \\ &\leq \frac{\Phi(x_n) - \Phi(x_{n+1})}{1 - \beta} + \left(\frac{\Phi(x_n) - \Phi(x_{n+1})}{1 - \beta} \right)^{\frac{p-1}{p}} \frac{1}{t_n^{\frac{1}{p}}} \|\hat{x} - x_{n+1}\|. \end{aligned}$$

Since $\{x_n\}$ is bounded and $\alpha \geq t_n \geq \bar{t} > 0$, there exists $c_1 \geq 0$ such that $t_n^{-\frac{1}{p}} \|\hat{x} - x_{n+1}\| \leq c_1$. Hence,

$$\Phi(x_n) - \Phi^* \leq \frac{\Phi(x_n) - \Phi(x_{n+1})}{1 - \beta} + \left(\frac{\Phi(x_n) - \Phi(x_{n+1})}{1 - \beta} \right)^{\frac{p-1}{p}} c_1. \tag{15}$$

Since $\{\Phi(x_n) - \Phi^*\}$ is bounded by Proposition 1(iii), there exist $c_2 \geq 0$ such that

$$(\Phi(x_n) - \Phi(x_{n+1}))^{\frac{1}{p}} + (1 - \beta)^{\frac{1}{p}} c_1 \leq c_2. \tag{16}$$

Then, by (15) and (16), we have

$$\begin{aligned} &(1 - \beta)(\Phi(x_n) - \Phi^*) \\ &\leq (\Phi(x_n) - \Phi(x_{n+1})) + (\Phi(x_n) - \Phi(x_{n+1}))^{\frac{p-1}{p}} (1 - \beta)^{\frac{1}{p}} c_1 \\ &= (\Phi(x_n) - \Phi(x_{n+1}))^{\frac{p-1}{p}} \left((\Phi(x_n) - \Phi(x_{n+1}))^{\frac{1}{p}} + (1 - \beta)^{\frac{1}{p}} c_1 \right) \\ &\leq (\Phi(x_n) - \Phi(x_{n+1}))^{\frac{p-1}{p}} c_2. \end{aligned} \tag{17}$$

Using (17), we get that

$$\frac{1 - \beta}{c_2} (\Phi(x_n) - \Phi^*) \leq (\Phi(x_n) - \Phi(x_{n+1}))^{\frac{p-1}{p}},$$

and so that,

$$(\Phi(x_{n+1}) - \Phi^*) \leq (\Phi(x_n) - \Phi^*) - \left(\frac{1 - \beta}{c_2}\right)^{\frac{p}{p-1}} (\Phi(x_n) - \Phi^*)^{\frac{p}{p-1}}. \tag{18}$$

Hence, by (18) and Lemma 2, we easily obtain, $\lim_{n \rightarrow \infty} \Phi(x_n) = \Phi^*$.

Case 2 Suppose now that $\{t_k\} \subset \{t_n\}$ and $\lim_{k \rightarrow +\infty} t_k = 0$. Since $\{x_n\}$ is bounded, the set of its weak accumulation points is nonempty. Without loss of generality, we assume that $\{x_k\}$ weakly converging to \bar{x} . Define $\hat{t}_k = \frac{t_k}{\theta} > t_k > 0$ and

$$\hat{x}_{\hat{t}_k} = \operatorname{argmin}_{y \in X} \left\{ \frac{1}{p} \|x_k - y\|^p + \hat{t}_k (\langle \nabla g(x_k), y \rangle + f(y)) \right\}.$$

Due to Lemma 3 we have

$$\|x_k - \hat{x}_{\hat{t}_k}\| \leq \left(\frac{\hat{t}_k}{t_k}\right)^{\frac{1}{p-1}} \|x_k - x_{k+1}\| = \left(\frac{1}{\theta}\right)^{\frac{1}{p-1}} \|x_k - x_{k+1}\|,$$

which combines with the boundedness of $\{x_k\}$ to show that the sequence $\{\hat{x}_{\hat{t}_k}\}$ is also bounded. It follows from the definition of **Linesearch 1** that

$$\hat{t}_k \|\nabla g(x_k) - \nabla g(\hat{x}_{\hat{t}_k})\| > \beta \|x_k - \hat{x}_{\hat{t}_k}\|^{p-1}. \tag{19}$$

Since $\lim_{k \rightarrow +\infty} \hat{t}_k = 0$ and both $\{x_k\}$ and $\{\hat{x}_{\hat{t}_k}\}$ are bounded, (19) together with Assumption 1 tell us that $\lim_{k \rightarrow +\infty} \|x_k - \hat{x}_{\hat{t}_k}\| = 0$ and thus $\{\hat{x}_{\hat{t}_k}\}$ also weakly converges to \bar{x} . Thanks to Assumption 1 again, we have

$$\lim_{k \rightarrow +\infty} \|\nabla g(x_k) - \nabla g(\hat{x}_{\hat{t}_k})\| = 0. \tag{20}$$

This and (19) imply that

$$\lim_{k \rightarrow +\infty} \frac{1}{\hat{t}_k} \|x_k - \hat{x}_{\hat{t}_k}\|^{p-1} = 0. \tag{21}$$

Since

$$\frac{J_p(x_k - \hat{x}_{\hat{t}_k}) - \hat{t}_k \nabla g(x_k)}{\hat{t}_k} + \nabla g(\hat{x}_{\hat{t}_k}) \in \partial f(\hat{x}_{\hat{t}_k}) + \nabla g(\hat{x}_{\hat{t}_k}) = \partial(f + g)(\hat{x}_{\hat{t}_k}).$$

By letting $k \rightarrow +\infty$, we get from (20) and (21) that $0 \in \partial(f + g)(\bar{x})$, which means $\bar{x} \in \mathcal{S}$. It remains to verify (i) in this case. Indeed, we get from Lemma 3 that

$$\|x_k - \hat{x}_{t_k}\| \geq \|x_k - x_{k+1}\|.$$

This together with (21) yields

$$\lim_{k \rightarrow +\infty} \frac{1}{t_k} \|x_k - x_{k+1}\|^{p-1} = 0.$$

Hence, by Proposition 1(iii) and (14), we have $\Phi(x_n) \rightarrow \Phi^*$.

(ii) Since $\lim_{n \rightarrow \infty} \Phi(x_n) = \Phi^*$, the sequence $\{x_n\}$ is minimizing sequence, thus, due to the weak lower semicontinuity of Φ (see lemma 1), all weak accumulation points of $\{x_n\}$ belong to \mathcal{S} . □

Proposition 3 *Let $\{x_n\}$ be a sequence generated by Iterative Method 3.1. Suppose that the sequence $\{x_n\}$ is bounded and there exists $\bar{t} > 0$ such that $t_n \geq \bar{t}$. Then, $\Phi(x_n) - \Phi^* \leq \lambda n^{1-p}$ for some $\lambda > 0$.*

Proof By the mean value theorem, we have

$$\begin{aligned} & \frac{1}{(\Phi(x_{n+1}) - \Phi^*)^{q-1}} - \frac{1}{(\Phi(x_n) - \Phi^*)^{q-1}} \\ &= \frac{(\Phi(x_n) - \Phi^*)^{q-1} - (\Phi(x_{n+1}) - \Phi^*)^{q-1}}{(\Phi(x_{n+1}) - \Phi^*)^{q-1}(\Phi(x_n) - \Phi^*)^{q-1}} \\ &= \frac{(q-1)\xi^{q-2}[(\Phi(x_n) - \Phi^*) - (\Phi(x_{n+1}) - \Phi^*)]}{(\Phi(x_{n+1}) - \Phi^*)^{q-1}(\Phi(x_n) - \Phi^*)^{q-1}}, \end{aligned}$$

where, q is the dual exponent, i.e., $\frac{1}{p} + \frac{1}{q} = 1$ and $\Phi(x_{n+1}) - \Phi^* \leq \xi \leq \Phi(x_n) - \Phi^*$. Thus,

$$\xi^{q-2} \geq (\Phi(x_{n+1}) - \Phi^*)^{q-1}(\Phi(x_n) - \Phi^*)^{-1}$$

and, by (18),

$$\begin{aligned} & \frac{1}{(\Phi(x_{n+1}) - \Phi^*)^{q-1}} - \frac{1}{(\Phi(x_n) - \Phi^*)^{q-1}} \\ & \geq \frac{(q-1)\bar{c}[(\Phi(x_{n+1}) - \Phi^*)^{q-1}(\Phi(x_n) - \Phi^*)^{q-1}]}{(\Phi(x_{n+1}) - \Phi^*)^{q-1}(\Phi(x_n) - \Phi^*)^{q-1}} \\ & = (q-1)\bar{c}, \end{aligned}$$

where, $\bar{c} = \left(\frac{1-\beta}{c_2}\right)^{\frac{p}{p-1}}$. Summing up then yields

$$\begin{aligned} & \frac{1}{(\Phi(x_n) - \Phi^*)^{q-1}} - \frac{1}{(\Phi(x_0) - \Phi^*)^{q-1}} \\ &= \sum_{i=0}^{n-1} \frac{1}{(\Phi(x_{i+1}) - \Phi^*)^{q-1}} - \frac{1}{(\Phi(x_i) - \Phi^*)^{q-1}} \\ &\geq n(q - 1)\bar{c}. \end{aligned}$$

and consequently,

$$(\Phi(x_n) - \Phi^*)^{q-1} \leq ((\Phi(x_0) - \Phi^*)^{1-q} + n(q - 1)\bar{c})^{-1}.$$

Hence, there exists $\lambda > 0$, such that $\Phi(x_n) - \Phi^* \leq \lambda n^{1-p}$. □

4 The FBS Method with Linesearch 2

Linesearch 2(x, θ). Given $x \in \text{dom } f$, $\theta \in (0, 1)$ and $p > 1$.
Input. Set $t = 1$ and $\hat{x} = \text{argmin}_{y \in X} \{\frac{1}{p}\|x - y\|^p + \langle \nabla g(x), y \rangle + f(y)\}$.
While
 $(f + g)(x - t(x - \hat{x}))$
 $> (f + g)(x) - t(f(x) - f(\hat{x})) - t\langle \nabla g(x), x - \hat{x} \rangle + \frac{t}{p}\|\hat{x} - x\|^p$
do $t = \theta t$.
End While
Output. t .

Lemma 5 *If $x \in \text{dom } f$, then **Linesearch 2**(x, θ) stops after finitely many steps.*

Proof If $x \in \mathcal{S}$, then $x = \hat{x}$. Thus the linesearch immediately give us the output 1 without proceeding any step. If $x \notin \mathcal{S}$, by contradiction let us assume that **Linesearch 2** does not stop after finitely many steps. Thus for all $t \in \{1, \theta, \theta^2, \dots\}$, we obtain

$$\begin{aligned} (f + g)(x - t(x - \hat{x})) &> (f + g)(x) - t(f(x) - f(\hat{x})) - t\langle \nabla g(x), x - \hat{x} \rangle \\ &+ \frac{t}{p}\|\hat{x} - x\|^p. \end{aligned}$$

Since $f(x - t(x - \hat{x})) \leq (1 - t)f(x) + tf(\hat{x})$, we have

$$\frac{g(x - t(x - \hat{x})) - g(x)}{t} + \langle \nabla g(x), x - \hat{x} \rangle > \frac{1}{p}\|\hat{x} - x\|^p.$$

Then we get that

$$0 = \lim_{t \rightarrow 0} \frac{g(x - t(x - \hat{x})) - g(x) + t\langle \nabla g(x), x - \hat{x} \rangle}{t} \geq \frac{1}{p}\|\hat{x} - x\|^p.$$

Hence we have $\hat{x} = x$, which readily implies that $0 \in \partial f(x) + \nabla g(x)$. This contradicts the assumption that x is not an optimal solution to problem and completes the proof of the lemma. \square

In this section, we shall consider the convergence and convergence rate of the following forward–backward splitting method:

Iterative Method 4.1.

Given $x_0 \in X$, for every $n \in \mathbb{N}$, set

$$\begin{cases} y_n = \operatorname{argmin}_{y \in X} \{ \frac{1}{p} \|x_n - y\|^p + \langle \nabla g(x_n), y \rangle + f(y) \}, \\ x_{n+1} = (1 - t_n)x_n + t_n y_n, \end{cases}$$

where $t_n = \mathbf{Linesearch\ 2}(x_n, \theta)$.

Remark 2 Let $\{x_n\}$ and $\{y_n\}$ be sequences generated by Iterative Method 4.1. It follows from **Linesearch 2**(x_n, θ) that

$$\begin{aligned} (f + g)(x_{n+1}) &\leq (f + g)(x_n) - t_n(f(x_n) - f(y_n)) \\ &\quad - t_n \langle \nabla g(x_n), x_n - y_n \rangle + \frac{t_n}{p} \|x_n - y_n\|^p. \end{aligned}$$

If we remove the f term from **Linesearch 2**(x_n, θ), we can still get this result from the convexity of f . However, if f is a strictly convex or strongly convex function, the number of iteration step for **Linesearch 2** may be reduced.

Next we obtain some similar results for Iterative Method 4.1 to the ones in Sect. 4 for Iterative Method 3.1.

Proposition 4 Let $\{x_n\}$ be a sequence generated by Iterative Method 4.1 and define

$$\rho(x_n) := f(x_n) - f(y_n) + \langle \nabla g(x_n), x_n - y_n \rangle.$$

Then, we have

- (i) $\frac{1}{t_n^p} \|x_n - x_{n+1}\|^p \leq \rho(x_n)$.
- (ii) $\Phi(x_{n+1}) \leq \Phi(x_n) - t_n(1 - \frac{1}{p})\rho(x_n)$.
- (iii) $\Phi(x_n)$ converges and $\sum_{n=1}^\infty t_n \rho(x_n) < +\infty$.

Proof (i) The optimality of y_n implies

$$0 \in \partial \left(\frac{1}{p} \|x_n - \cdot\|^p + \langle \nabla g(x_n), \cdot \rangle + f(\cdot) \right) (y_n).$$

Then, we have

$$J_p(x_n - y_n) - \nabla g(x_n) \in \partial f(y_n). \tag{22}$$

Hence, by $y_n = \frac{1}{t_n}x_{n+1} + (1 - \frac{1}{t_n})x_n$, we have

$$\left\langle J_p\left(\frac{x_n - x_{n+1}}{t_n}\right) - \nabla g(x_n), x_n - \left(\frac{1}{t_n}x_{n+1} + \left(1 - \frac{1}{t_n}\right)x_n\right) \right\rangle \leq f(x_n) - f(y_n),$$

and so that

$$\left\| \frac{x_n - x_{n+1}}{t_n} \right\|^p \leq f(x_n) - f(y_n) + \langle \nabla g(x_n), x_n - y_n \rangle = \rho(x_n). \tag{23}$$

(ii) Using the definition of $\rho(x_n)$ and Remark 2, we get that

$$\begin{aligned} \Phi(x_n) - \Phi(x_{n+1}) &\geq t_n(f(x_n) - f(y_n)) + t_n \langle \nabla g(x_n), x_n - y_n \rangle - \frac{t_n}{p} \|x_n - y_n\|^p \\ &= t_n \rho(x_n) - \frac{1}{p t_n^{p-1}} \|x_n - x_{n+1}\|^p. \end{aligned} \tag{24}$$

It follows from (23) and (24) that

$$\Phi(x_{n+1}) \leq \Phi(x_n) - t_n \left(1 - \frac{1}{p}\right) \rho(x_n).$$

(iii) The conclusion follow using Lemma 2 and Proposition 4 (ii). □

Proposition 5 *Let $\{x_n\}$ and $\{y_n\}$ be sequences generated by Iterative Method 4.1. Assume the sequence $\{x_n\}$ is bounded. Then,*

- (i) *all weak accumulation points of $\{x_n\}$ belong to \mathcal{S} .*
- (ii) *if there exists $\{t_k\} \subset \{t_n\}$ such that $t_k \rightarrow \bar{t} > 0$, then $\lim_{n \rightarrow \infty} \Phi(x_n) = \Phi^*$.*
- (iii) *if $t_n \rightarrow 0$ and f is uniformly continuous on any bounded subset of $\text{dom } f$, then $\lim_{n \rightarrow \infty} \Phi(x_n) = \Phi^*$.*

Proof (i) Let $\hat{x} \in \mathcal{S}$ and $\{x_k\} \subset \{x_n\}$ weakly converging to \bar{x} . Since

$$g(x_n) - g(\hat{x}) \leq \langle \nabla g(x_n), x_n - \hat{x} \rangle$$

and

$$\Phi(x_n) - \Phi^* = f(x_n) - f(\hat{x}) + g(x_n) - g(\hat{x}),$$

we have that

$$\begin{aligned} \Phi(x_n) - \Phi^* &\leq f(x_n) - f(\hat{x}) + \langle \nabla g(x_n), x_n - \hat{x} \rangle \\ &= \rho(x_n) + f(y_n) - f(\hat{x}) + \langle \nabla g(x_n), y_n - \hat{x} \rangle. \end{aligned} \tag{25}$$

Then, by (22) and (25), we have

$$\begin{aligned}
 & \Phi(x_n) - \Phi^* \\
 & \leq \rho(x_n) - \langle J_p(\frac{x_n - x_{n+1}}{t_n}) - \nabla g(x_n), \hat{x} - y_n \rangle + \langle \nabla g(x_n), y_n - \hat{x} \rangle \\
 & = \rho(x_n) - \langle J_p(\frac{x_n - x_{n+1}}{t_n}), \hat{x} - y_n \rangle \\
 & \leq \rho(x_n) + (\frac{1}{t_n})^{p-1} \|x_n - x_{n+1}\|^{p-1} \|\hat{x} - (\frac{1}{t_n}x_{n+1} + (1 - \frac{1}{t_n})x_n)\|. \tag{26}
 \end{aligned}$$

Now let us split our further analysis into two distinct cases.

Case 1 Without loss of generality, we assume that $t_k \rightarrow \bar{t} > 0$. Then, by Proposition 4, we have $\lim_{n \rightarrow \infty} \rho(x_n) = 0$ and $\lim_{n \rightarrow \infty} \|x_n - x_{n+1}\| = 0$. Then by (26), we have $\lim_{n \rightarrow \infty} \Phi(x_n) = \Phi^*$. Hence, the sequence $\{x_k\}$ is minimizing sequence, thus, due to the weak lower semicontinuity of Φ , $\bar{x} \in \mathcal{S}$.

Case 2 Suppose now that $\lim_{k \rightarrow +\infty} t_k = 0$. Define $\hat{t}_k = \frac{t_k}{\theta} > t_k > 0$ and

$$\begin{cases} y_k = \operatorname{argmin}_{y \in X} \{ \frac{1}{p} \|x_k - y\|^p + \langle \nabla g(x_k), y \rangle + f(y) \}, \\ \hat{x}_{k+1} = (1 - \hat{t}_k)x_k + \hat{t}_k y_k, \end{cases} \tag{27}$$

It follows from the definition of **Linesearch 2**(x_k, θ) that

$$\Phi(x_k) - \Phi(\hat{x}_{k+1}) \leq \hat{t}_k(f(x_k) - f(y_k)) + \hat{t}_k \langle \nabla g(x_k), x_k - y_k \rangle - \frac{\hat{t}_k}{p} \|x_k - y_k\|^p.$$

This together with (6) and (27) gives us that

$$\begin{aligned}
 0 & \geq \Phi(x_k) - \Phi(\hat{x}_{k+1}) - \hat{t}_k(f(x_k) - f(y_k)) - \hat{t}_k \langle \nabla g(x_k), x_k - y_k \rangle \\
 & \quad + \frac{\hat{t}_k}{p} \|x_k - y_k\|^p \\
 & = f(x_k) - f(\hat{x}_{k+1}) + g(x_k) - g(\hat{x}_{k+1}) \\
 & \quad - \hat{t}_k(f(x_k) - f(y_k)) - \hat{t}_k \langle \nabla g(x_k), x_k - y_k \rangle + \frac{\hat{t}_k}{p} \|x_k - y_k\|^p \\
 & \geq -\hat{t}_k \langle \nabla g(x_k), x_k - y_k \rangle + \langle \nabla g(\hat{x}_{k+1}), x_k - \hat{x}_{k+1} \rangle + \frac{\hat{t}_k}{p} \|x_k - y_k\|^p \\
 & \quad + f(x_k) - \hat{t}_k f(y_k) - (1 - \hat{t}_k)f(x_k) - \hat{t}_k(f(x_k) - f(y_k)) \\
 & = \hat{t}_k \langle \nabla g(\hat{x}_{k+1}) - \nabla g(x_k), x_k - y_k \rangle + \frac{\hat{t}_k}{p} \|x_k - y_k\|^p.
 \end{aligned}$$

We obtain from the latter that

$$\frac{\hat{t}_k}{p} \|x_k - y_k\|^p \leq \hat{t}_k \|\nabla g(\hat{x}_{k+1}) - \nabla g(x_k)\| \|x_k - y_k\|,$$

which yields

$$\frac{1}{p} \|x_k - y_k\|^{p-1} \leq \|\nabla g(\hat{x}_{k+1}) - \nabla g(x_k)\|. \tag{28}$$

On the other hand,

$$\langle J_p(x_k - y_k) - \nabla g(x_k), y_0 - y_k \rangle \leq f(y_0) - f(y_k)$$

and

$$\langle J_p(x_0 - y_0) - \nabla g(x_0), y_k - y_0 \rangle \leq f(y_k) - f(y_0).$$

Adding the two inequalities, we get

$$\langle J_p(x_k - y_k) - J_p(x_0 - y_0), y_0 - y_k \rangle \leq \langle \nabla g(x_0) - \nabla g(x_k), y_k - y_0 \rangle.$$

Then we have

$$\begin{aligned} \|x_k - y_k\|^p &\leq \langle J_p(x_k - y_k), x_k - y_0 \rangle + \langle J_p(x_0 - y_0), y_0 - y_k \rangle \\ &\quad + \langle \nabla g(x_0) - \nabla g(x_k), y_k - y_0 \rangle. \end{aligned}$$

Hence

$$\begin{aligned} \|x_k - y_k\|^p &\leq \|x_k - y_k\|^{p-1} \|y_0 - x_k\| + \|x_0 - y_0\|^{p-1} \|y_0 - y_k\| \\ &\quad + \|\nabla g(x_0) - \nabla g(x_k)\| \|y_k - y_0\|. \end{aligned}$$

Due to Assumption 1, $p > 1$ and the boundedness of $\{x_n\}$, the latter tells us that $\{y_k\}$ is also bounded. This together with (27) and the fact $\lim_{k \rightarrow +\infty} t_k = 0$ implies $\lim_{k \rightarrow +\infty} \|\hat{x}_{k+1} - x_k\| = 0$. Since ∇g is uniformly continuous on bounded sets, we get $\lim_{k \rightarrow +\infty} \|\nabla g(\hat{x}_{k+1}) - \nabla g(x_k)\| = 0$ and derive from (28) that

$$\lim_{k \rightarrow +\infty} \|x_k - y_k\| = 0. \tag{29}$$

Since ∇g is uniformly continuous on bounded sets, (29) implies

$$\lim_{k \rightarrow +\infty} \|\nabla g(x_k) - \nabla g(y_k)\| = 0. \tag{30}$$

Since $J_p(x_k - y_k) - \nabla g(x_k) + \nabla g(y_k) \in \partial f(y_k) + \nabla g(y_k) = \partial(f + g)(y_k)$. By passing to the limit over the subsequence $\{x_k\}$ in the above inclusion, we get from (29) and (30) that $0 \in \partial(f + g)(\bar{x})$, which means $\bar{x} \in \mathcal{S}$.

(ii) Suppose now that $\{t_k\} \subset \{t_n\}$ and $t_k \rightarrow \bar{t} > 0$. Hence, by Proposition 4 and (26), we have $\lim_{k \rightarrow \infty} \Phi(x_k) = \Phi^*$. Furthermore, we have $\lim_{n \rightarrow \infty} \Phi(x_n) = \Phi^*$.

(iii) By (29), we have

$$\lim_{n \rightarrow +\infty} \|x_n - y_n\| = 0. \tag{31}$$

Since f is uniformly continuous on any bounded set and ∇g is uniformly continuous on bounded sets, (31) implies

$$\lim_{n \rightarrow +\infty} \rho(x_n) = \lim_{n \rightarrow +\infty} f(x_n) - f(y_n) + \langle \nabla g(x_n), x_n - y_n \rangle = 0.$$

Hence, by Proposition 4(i) and (26), we have $\lim_{n \rightarrow \infty} \Phi(x_n) = \Phi^*$. □

Proposition 6 *Let $\{x_n\}$ and $\{y_n\}$ be sequences generated by Iterative Method 4.1. Suppose that the sequence $\{x_n\}$ is bounded and there exists $\bar{t} > 0$ such that $t_n \geq \bar{t}$ for all $n \in \mathbb{N}$. Then, $\Phi(x_n) - \Phi^* \leq \lambda n^{1-p}$ for some $\lambda > 0$.*

Proof By (26) and Proposition 4, we have

$$\begin{aligned} & \Phi(x_n) - \Phi^* \\ & \leq \rho(x_n) + \left(\frac{1}{t_n}\right)^{p-1} \|x_n - x_{n+1}\|^{p-1} \|\hat{x} - \left(\frac{1}{t_n}x_{n+1} + \left(1 - \frac{1}{t_n}\right)x_n\right)\| \\ & \leq \frac{1}{t_n} \frac{\Phi(x_n) - \Phi(x_{n+1})}{1 - \frac{1}{p}} + \left\| \frac{x_n - x_{n+1}}{t_n} \right\|^{p-1} \|\hat{x} - \left(\frac{1}{t_n}x_{n+1} + \left(1 - \frac{1}{t_n}\right)x_n\right)\| \\ & \leq \frac{1}{t_n} \frac{\Phi(x_n) - \Phi(x_{n+1})}{1 - \frac{1}{p}} \\ & \quad + \left(\frac{\Phi(x_n) - \Phi(x_{n+1})}{1 - \frac{1}{p}} \right)^{\frac{p-1}{p}} t_n^{-\frac{p-1}{p}} \|\hat{x} - \left(\frac{1}{t_n}x_{n+1} + \left(1 - \frac{1}{t_n}\right)x_n\right)\|. \end{aligned}$$

Since $\{x_n\}$ is bounded, there exist $c_1 \geq 0$ such that

$$t_n^{-\frac{p-1}{p}} \|\hat{x} - \left(\frac{1}{t_n}x_{n+1} + \left(1 - \frac{1}{t_n}\right)x_n\right)\| \leq c_1.$$

Hence,

$$\Phi(x_n) - \Phi^* \leq \frac{1}{t_n} \frac{\Phi(x_n) - \Phi(x_{n+1})}{1 - \frac{1}{p}} + \left(\frac{\Phi(x_n) - \Phi(x_{n+1})}{1 - \frac{1}{p}} \right)^{\frac{p-1}{p}} c_1. \tag{32}$$

Since $\{\Phi(x_n) - \Phi^*\}$ is bounded by Proposition 4(iii), there exist $c_2 \geq 0$ such that

$$\frac{1}{t_n} (\Phi(x_n) - \Phi(x_{n+1}))^{\frac{1}{p}} + \left(1 - \frac{1}{p}\right)^{\frac{1}{p}} c_1 \leq c_2. \tag{33}$$

Then, by (32) and (33), we have

$$\begin{aligned}
 & \left(1 - \frac{1}{p}\right)(\Phi(x_n) - \Phi^*) \\
 & \leq \frac{1}{t_n}(\Phi(x_n) - \Phi(x_{n+1})) + (\Phi(x_n) - \Phi(x_{n+1}))^{\frac{p-1}{p}} \left(1 - \frac{1}{p}\right)^{\frac{1}{p}} c_1 \\
 & = (\Phi(x_n) - \Phi(x_{n+1}))^{\frac{p-1}{p}} \left(\frac{1}{t_n}(\Phi(x_n) - \Phi(x_{n+1}))^{\frac{1}{p}} + \left(1 - \frac{1}{p}\right)^{\frac{1}{p}} c_1 \right) \\
 & \leq c_2(\Phi(x_n) - \Phi(x_{n+1}))^{\frac{p-1}{p}}. \tag{34}
 \end{aligned}$$

Using (34), we get that

$$\frac{p-1}{pc_2}(\Phi(x_n) - \Phi^*) \leq (\Phi(x_n) - \Phi(x_{n+1}))^{\frac{p-1}{p}}$$

and so that,

$$(\Phi(x_{n+1}) - \Phi^*) \leq (\Phi(x_n) - \Phi^*) - \left(\frac{p-1}{pc_2}\right)^{\frac{p}{p-1}} (\Phi(x_n) - \Phi^*)^{\frac{p}{p-1}}. \tag{35}$$

By the mean value theorem, we have

$$\begin{aligned}
 & \frac{1}{(\Phi(x_{n+1}) - \Phi^*)^{q-1}} - \frac{1}{(\Phi(x_n) - \Phi^*)^{q-1}} \\
 & = \frac{(\Phi(x_n) - \Phi^*)^{q-1} - (\Phi(x_{n+1}) - \Phi^*)^{q-1}}{(\Phi(x_{n+1}) - \Phi^*)^{q-1}(\Phi(x_n) - \Phi^*)^{q-1}} \\
 & = \frac{(q-1)\xi^{q-2}[(\Phi(x_n) - \Phi^*) - (\Phi(x_{n+1}) - \Phi^*)]}{(\Phi(x_{n+1}) - \Phi^*)^{q-1}(\Phi(x_n) - \Phi^*)^{q-1}}
 \end{aligned}$$

with $\Phi(x_{n+1}) - \Phi^* \leq \xi \leq \Phi(x_n) - \Phi^*$ and $\frac{1}{q} + \frac{1}{p} = 1$. Thus,

$$\xi^{q-2} \geq (\Phi(x_{n+1}) - \Phi^*)^{q-1}(\Phi(x_n) - \Phi^*)^{-1}$$

and, by (35),

$$\begin{aligned}
 & \frac{1}{(\Phi(x_{n+1}) - \Phi^*)^{q-1}} - \frac{1}{(\Phi(x_n) - \Phi^*)^{q-1}} \\
 & \geq \frac{(q-1)\bar{c}[(\Phi(x_{n+1}) - \Phi^*)^{q-1}(\Phi(x_n) - \Phi^*)^{q-1}]}{(\Phi(x_{n+1}) - \Phi^*)^{q-1}(\Phi(x_n) - \Phi^*)^{q-1}} \\
 & = (q-1)\bar{c},
 \end{aligned}$$

where, $\bar{c} = \left(\frac{p-1}{pc_2}\right)^{\frac{p}{p-1}}$. Summing up then yields

$$\begin{aligned} & \frac{1}{(\Phi(x_n) - \Phi^*)^{q-1}} - \frac{1}{(\Phi(x_0) - \Phi^*)^{q-1}} \\ &= \sum_{i=0}^{n-1} \frac{1}{(\Phi(x_{i+1}) - \Phi^*)^{q-1}} - \frac{1}{(\Phi(x_i) - \Phi^*)^{q-1}} \\ &\geq n(q-1)\bar{c}. \end{aligned}$$

and consequently,

$$(\Phi(x_n) - \Phi^*)^{q-1} \leq ((\Phi(x_0) - \Phi^*)^{1-q} + n(q-1)\bar{c})^{-1}.$$

Hence, there exists $\lambda > 0$, such that $\Phi(x_n) - \Phi^* \leq \lambda n^{1-p}$. □

Definition 3 [17] We say that the sequence of iterates generated by Iterative Method 4.1 satisfies the error bound property, if there exists $\kappa > 0$ such that

$$\min_{y \in \mathcal{S}} \|x_n - y\| \leq \kappa \|x_n - y_n\|. \tag{36}$$

Proposition 7 Let $\{x_n\}$ and $\{y_n\}$ be sequences generated by Iterative Method 4.1. Assume that the error bound property (36) holds for $\{x_n\}$ and the sequence $\{x_n\}$ is bounded. If there exists $\bar{t} > 0$ such that $t_n \geq \bar{t}$ for all $n \in \mathbf{N}$, then the sequence $\{\Phi(x_n)\}$ converges Q -linearly, that is, there exists $0 < \epsilon < 1$ such that

$$\Phi(x_{n+1}) - \Phi^* \leq \epsilon(\Phi(x_n) - \Phi^*).$$

Proof Since \mathcal{S} is nonempty, closed and convex and X is reflexive, there exists $\hat{x}_n \in \mathcal{S}$, such that $\|x_n - \hat{x}_n\| = \min_{y \in \mathcal{S}} \|x_n - y\|$ and $\Phi(\hat{x}_n) = \Phi^*$. Since $\{x_n\}$ is bounded, $\{\hat{x}_n\}$ is also bounded. Since

$$g(x_n) - g(\hat{x}_n) \leq \langle \nabla g(x_n), x_n - \hat{x}_n \rangle$$

and

$$\Phi(x_n) - \Phi^* = f(x_n) - f(\hat{x}_n) + g(x_n) - g(\hat{x}_n),$$

we have that

$$\begin{aligned} \Phi(x_n) - \Phi^* &\leq f(x_n) - f(\hat{x}_n) + \langle \nabla g(x_n), x_n - \hat{x}_n \rangle \\ &= \rho(x_n) + f(y_n) - f(\hat{x}_n) + \langle \nabla g(x_n), y_n - \hat{x}_n \rangle. \end{aligned} \tag{37}$$

Then, by (22) and (37), we have

$$\begin{aligned}
 \Phi(x_n) - \Phi^* &\leq \rho(x_n) - \langle J_p(\frac{x_n - x_{n+1}}{t_n}) - \nabla g(x_n), \hat{x}_n - y_n \rangle + \langle \nabla g(x_n), y_n - \hat{x}_n \rangle \\
 &= \rho(x_n) - \langle J_p(\frac{x_n - x_{n+1}}{t_n}), \hat{x}_n - y_n \rangle \\
 &\leq \rho(x_n) + (\frac{1}{t_n})^{p-1} \|x_n - x_{n+1}\|^{p-1} \|\hat{x}_n - (\frac{1}{t_n}x_{n+1} + (1 - \frac{1}{t_n})x_n)\|. \tag{38}
 \end{aligned}$$

By (38), we have

$$\begin{aligned}
 \Phi(x_n) - \Phi^* &\leq \rho(x_n) + (\frac{1}{t_n})^{p-1} \|x_n - x_{n+1}\|^{p-1} \|\hat{x}_n - x_n - \frac{1}{t_n}(x_{n+1} - x_n)\| \\
 &\leq \rho(x_n) + (\frac{1}{t_n})^{p-1} \|x_n - x_{n+1}\|^{p-1} (\|\hat{x}_n - x_n\| + \frac{1}{t_n} \|x_{n+1} - x_n\|).
 \end{aligned}$$

Now we use the error bound condition $\min_{y \in \mathcal{S}} \|x_n - y\| \leq \kappa \|x_n - y_n\|$ together with Proposition 4 to obtain

$$\begin{aligned}
 \Phi(x_n) - \Phi^* &\leq \frac{p}{t_n(p-1)} \times (\Phi(x_n) - \Phi(x_{n+1})) \\
 &\quad + (\frac{1}{t_n})^{p-1} \|x_n - x_{n+1}\|^{p-1} (\kappa \|x_n - y_n\| + \frac{1}{t_n} \|x_{n+1} - x_n\|) \\
 &= \frac{p}{t_n(p-1)} \times (\Phi(x_n) - \Phi(x_{n+1})) + \frac{1 + \kappa}{t_n^p} \|x_n - x_{n+1}\|^p \\
 &\leq \frac{(2 + \kappa)p}{t_n(p-1)} \times (\Phi(x_n) - \Phi(x_{n+1})) \\
 &\leq \frac{(2 + \kappa)p}{\bar{t}(p-1)} \times (\Phi(x_n) - \Phi(x_{n+1})). \tag{39}
 \end{aligned}$$

It follows from (39) that

$$\begin{aligned}
 \Phi(x_{n+1}) - \Phi^* &\leq \left(\frac{(2 + \kappa)p}{\bar{t}(p-1)} - 1\right) \left(\frac{(2 + \kappa)p}{\bar{t}(p-1)}\right)^{-1} (\Phi(x_n) - \Phi^*) \\
 &= \left(1 - \frac{\bar{t}(p-1)}{(2 + \kappa)p}\right) (\Phi(x_n) - \Phi^*) \\
 &= \epsilon (\Phi(x_n) - \Phi^*),
 \end{aligned}$$

where, $0 < \epsilon = \left(1 - \frac{\bar{t}(p-1)}{(2 + \kappa)p}\right) < 1$. Hence, the sequence $\{\Phi(x_n)\}$ converges Q -linearly. □

Corollary 1 *Let $\{x_n\}$ and $\{y_n\}$ be sequences generated by Iterative Method 4.1. Assume that the error bound property (36) holds for $\{x_n\}$ and the sequence $\{x_n\}$ is bounded.*

If f is totally convex and there exists $\bar{t} > 0$ such that $t_n \geq \bar{t}$ for all $n \in \mathbf{N}$, then $\{x_n\}$ converges R -linearly to the unique minimizer \bar{x} .

Proof Consider the following Bregman-like distance to the solution \bar{x} of the problem (1):

$$R(x) := f(x) - f(\bar{x}) + \langle \nabla g(\bar{x}), x - \bar{x} \rangle$$

which is non-negative since $-\nabla g(\bar{x}) \in \partial f(\bar{x})$. Note that if $\partial f(\bar{x})$ consists of one point, R is indeed the Bregman distance. By optimality of \bar{x} and by Proposition 7, the iterates x_n satisfy

$$0 \leq R(x_n) = f(x_n) - f(\bar{x}) + \langle \nabla g(\bar{x}), x_n - \bar{x} \rangle \leq \Phi(x_n) - \Phi^* \rightarrow 0.$$

Hence, by Definition 2, $x_n \rightarrow \bar{x}$.

On the other hand, If f is totally convex, then $\mathcal{S} = \{\bar{x}\}$ and

$$\min_{y \in \mathcal{S}} \|x_n - y\| = \|x_n - \bar{x}\|.$$

Using the error bound condition (36), we have

$$\|x_n - \bar{x}\| \leq \kappa \|x_n - y_n\| = \frac{\kappa}{t_n} \|x_n - x_{n+1}\|.$$

Then, by Proposition 4 and $\Phi(x_n) - \Phi(\bar{x}) \geq \Phi(x_n) - \Phi(x_{n+1})$, we have $\{x_n\}$ is bounded and

$$\Phi(x_n) - \Phi(\bar{x}) \geq \frac{p-1}{p t_n^{p-1}} \|x_n - x_{n+1}\|^p \geq \frac{(p-1)t_n}{p \kappa^p} \|x_n - \bar{x}\|^p \geq \frac{(p-1)\bar{t}}{p \kappa^p} \|x_n - \bar{x}\|^p$$

and consequently

$$\left(\frac{\Phi(x_n) - \Phi(\bar{x})}{\mu} \right)^{\frac{1}{p}} \geq \|x_n - \bar{x}\|, \tag{40}$$

where, $\mu = \frac{(p-1)\bar{t}}{p \kappa^p}$. Since the sequence $\{\Phi(x_n)\}$ is converges Q -linearly by Proposition 7, that is, $\Phi(x_{n+1}) - \Phi^* \leq \epsilon (\Phi(x_n) - \Phi^*)$, then we have

$$\left(\frac{\Phi(x_{n+1}) - \Phi(\bar{x})}{\mu} \right)^{\frac{1}{p}} \leq \epsilon^{\frac{1}{p}} \left(\frac{\Phi(x_n) - \Phi(\bar{x})}{\mu} \right)^{\frac{1}{p}}.$$

Hence, the sequence $\left\{ \left(\frac{\Phi(x_n) - \Phi(\bar{x})}{\mu} \right)^{\frac{1}{p}} \right\}$ is also converges Q -linearly. Thus, by (40), $\{x_n\}$ converges R -linearly to an optimal solution in \mathcal{S} . □

5 Conclusion

In this work, we discuss the modified forward–backward splitting method involving new linesearches for solving minimization problems of two convex functions in Banach space. Our algorithms do not require the Lipschitz constant of the gradient of functions. We proved the weak convergence of the iterative sequence generated by these methods, and further proved convergence with asymptotic rate $\frac{1}{n}$ to the optimal value under the assumption of the boundedness of the iterative sequence.

Acknowledgements The work was supported by PHD research startup foundation of Harbin Normal University(No. XKB201804) and the National Natural Sciences Grant(No. 11871182)

References

1. Albert, Y.: Decomposition theorem in Banach spaces. *Field Inst. Commun.* **25**, 77–93 (2000)
2. Albert, Y., Ryazantseva, I.: *Nonlinear Ill-Posed Problems of Monotone Type*. Springer, New York (2006)
3. 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)
4. Bauschke, H.H., Combettes, P.L.: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York (2011)
5. Bauschke, H.H., Bolte, J., Teboulle, M.: A descent lemma beyond Lipschitz gradient continuity: first order methods revisited and applications. *Math. Oper. Res.* **42**, 330–348 (2016)
6. Beck, A., Teboulle, M.: Gradient-based algorithms with applications to signal recovery problems. In: Palomar, D., Eldar, Y. (eds.) *Convex Optimization in Signal Processing and Communications*, pp. 42–88. University Press, Cambridge (2010)
7. Bello Cruz, J.Y., Nghia, T.T.: On the convergence of the forward–backward splitting method with linesearches. *Optim. Method Softw.* **7**, 1–30 (2016)
8. Bolte, J., Sabach, S., Teboulle, M., Vaisbourd, Y.: First order methods beyond convexity and Lipschitz gradient continuity with applications to quadratic inverse problems. *SIAM J. Optim.* **28**, 2131–2151 (2018)
9. Bonnans, J.F., Shapiro, A.: *Perturbation Analysis of Optimization Problems*. Springer, New York (2000)
10. Bredies, K.: A forward–backward splitting algorithm for the minimization of nonsmooth convex functionals in Banach space. *Inverse Probl.* **25**, 015005 (2009)
11. Cioranescu, I.: *Geometry of Banach Spaces, Duality Mappings and Nonlinear Problems*. Mathematics and Its Applications, vol. 62. Kluwer Academic Publishers, Dordrecht, The Netherlands (1990)
12. Combettes, P.L.: Inconsistent signal feasibility problems: least-squares solutions in a product space. *IEEE Trans. Signal Process.* **42**, 2955–2966 (1994)
13. Combettes, P.L., Dũng, D., Vũ, B.C.: Dualization of signal recovery problems. *Set-Valued Anal.* **18**, 373–404 (2010)
14. Combettes, P.L., Wajs, V.R.: Signal recovery by proximal forward–backward splitting. *Multiscale Model. Simul.* **4**, 1168–1200 (2005)
15. Combettes, P.L., Pesquet, J.C.: A Douglas–Rachford splitting approach to nonsmooth convex variational signal recovery. *IEEE J. Sel. Top. Signal Process.* **1**, 564–574 (2007)
16. 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.* **7**, 867–886 (2015)
17. Guan, W.B., Song, W.: The forward–backward splitting method and its convergence rate for the minimization of the sum of two functions in Banach space. *Optim. Lett.* **2**, 1–24 (2020)
18. Kankam, K., Pholasa, N., Cholamjiak, P.: Hybrid forward–backward algorithms using linesearch rule for minimization problem. *Thai J. Math.* **17**(3), 607–625 (2019)

19. Kankam, K., Pholasa, N., Chalamjiak, P.: On convergence and complexity of the modified forward–backward method involving new linesearches for convex minimization. *Math. Meth. Appl. Sci.* **42**(5), 1352–1362 (2019)
20. Lions, P.L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.* **16**, 964–979 (1979)
21. Salzo, S.: The variable metric forward–backward splitting algorithm under mild differentiability assumptions. *SIAM J. Optim.* **4**, 2153–2181 (2017)
22. Suantai, S., Noor, M.A., Kankam, K., Chalamjiak, P.: Novel forward–backward algorithms for optimization and applications to compressive sensing and image inpainting. *Adv. Differ. Equ.-ny.* **2021**, 265 (2021)
23. Takahashi, W.: *Convex Analysis and Approximation of Fixed Points*. Mathematical Analysis Series, vol. 2. Yokohama Publishers, Yokohama, Japan (2000)

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