A Reflected Forward-Backward Splitting Method for Monotone Inclusions Involving Lipschitzian Operators

Volkan Cevher1 · Bàng Công Vũ1

Received: 23 August 2019 / Accepted: 19 March 2020 / Published online: 30 March 2020 © Springer Nature B.V. 2020

Abstract
In this paper, we propose a novel splitting method for finding a zero point of the sum of two monotone operators where one of them is Lipschitzian. The weak convergence the method is proved in real Hilbert spaces. Applying the proposed method to composite monotone inclusions involving parallel sums yields a new primal-dual splitting which is different from the existing methods. Connections to existing works are clearly stated. We also provide an application of the proposed method to the image denoising by the total variation.

Keywords Monotone inclusion · Monotone operator · Operator splitting · Cocoercive · Forward-backward-forward method · Forward-backward algorithm · Composite operator · Duality · Primal-dual algorithm

Mathematics Subject Classification (2010) 47H05 · 49M29 · 49M27 · 90C25

1 Introduction

The forward-backward-forward splitting method (FBFS) or Tseng’s splitting method first appeared in [17]. This method was proposed to find a zero point of the sum of two monotone operators acting on a real Hilbert space \( (\mathcal{H}, \langle \cdot, \cdot \rangle) \), namely,

\[
\text{find } \bar{x} \in \mathcal{H} \text{ such that } 0 \in A\bar{x} + B\bar{x}.
\]

(1.1)

under the assumption that \( A : \mathcal{H} \to 2^{\mathcal{H}} \) is a maximally monotone, \( B : \mathcal{H} \to \mathcal{H} \) be a monotone and \( \mu \)-Lipschitzian, i.e.,

\[
(\forall x \in \mathcal{H})(\forall y \in \mathcal{H}) \| Bx - By \| \leq \mu \| x - y \|,
\]

(1.2)
and that such a solution exists. The FBFS method operates according to the routine

\[
\begin{align*}
y_n &= x_n - \gamma Bx_n \\
z_n &= (\text{Id} + \gamma A)^{-1} y_n \\
r_n &= z_n - \gamma Bz_n \\
x_{n+1} &= x_n + r_n - y_n.
\end{align*}
\]  

(1.3)

The weak convergence of \((x_n)_{n \in \mathbb{N}}\) to a solution of Eq. 1.1 was proved under the condition \(0 < \gamma < 1/\mu\). Inexact version of the FBF method was investigated in [4]. Then, variable metric version and the stochastic version the FBF method are in [20] and [18], respectively. One of the most important example of \(B\) is the case when \(B\) is a linear skew operator [4] where monotone plus skew model plays a central role in solving primal-dual monotone inclusions and primal-dual convex optimization problems. The main idea of [4] was then developed in [6]. Several developments and extensions of [6] are in [3, 5, 8, 20].

The advantage of this framework is its generality and the main disadvantage of Eq. 1.3 is that it requires two calls of \(B\) per one iteration. This issue was recently resolved in [11]. Specifically, they propose a forward-reflected-backward splitting method (FRBS) for solving (1.1), namely,

\[
\gamma \in [0, +\infty[, \quad x_{n+1} = (\text{Id} + \gamma A)^{-1} (x_n - 2\gamma Bx_n + \gamma Bx_{n-1}).
\]

(1.4)

The weak convergence of the iterates generated by Eq. 1.4 is proved under the condition \(\gamma \in [0, 1/(2\mu)]\). If \(B\) is linear and \(A\) is the normal cone of some non-empty closed convex set \(K\), the FRBS method admits the same structure as the reflected projected gradient methods for variational inequalities [9], namely,

\[
\gamma \in [0, +\infty[, \quad x_{n+1} = (\text{Id} + \gamma N_K)^{-1} (x_n - \gamma B(2x_n - x_{n-1})).
\]

(1.5)

For any \(\mu\)-Lipschitzian monotone operator \(B\), the weak convergence of the iterates generated by Eq. 1.5 is proved under the condition \(\gamma \in \left[0, (\sqrt{2} - 1)/\mu\right]\). When \(N_K\) is replaced by a subdifferential of some proper lower semicontinuous convex function \(f\), line-search versions (1.5) are proposed in [10].

The objective of this paper is to investigate the convergence of Eq. 1.5 for the problem (1.1) for any maximally monotone operator \(A\), i.e., we propose to investigate the convergence of the following reflected forward-backward splitting method (RFBS) for Eq. 1.1:

\[
\begin{align*}
y_n &= 2x_n - x_{n-1} \\
x_{n+1} &= (\text{Id} + \gamma A)^{-1} (x_n - \gamma By_n),
\end{align*}
\]

(1.6)

where \(\gamma > 0\).

In Section 2, we prove the weak convergence of Eq. 1.6 and provide an application to composite monotone inclusions involving the parallel sums and Lipschitzian monotone operators. We compare, in Section 3, the proposed method to several existing methods, for the image denoising by the total variation.

**Notations** (See [1]) The scalar products and the associated norms of all Hilbert spaces used in this paper are denoted respectively by \(\langle \cdot, \cdot \rangle\) and \(\| \cdot \|\). We denote by \(\mathcal{B} (\mathcal{H}, \mathcal{G})\) the space of all bounded linear operators from \(\mathcal{H}\) to \(\mathcal{G}\). The symbols \(\rightharpoonup\) and \(\to\) denote respectively weak and strong convergence. Let \(A : \mathcal{H} \to 2^{\mathcal{H}}\) be a set-valued operator. The domain of \(A\) is denoted by \(\text{dom}(A)\) that is a set of all \(x \in \mathcal{H}\) such that \(Ax \neq \emptyset\). The range of \(A\) is \(\text{ran}(A) = \{u \in \mathcal{H} \mid (\exists x \in \mathcal{H}) u \in Ax\}\). The graph of \(A\) is \(\text{gra}(A) = \{(x, u) \in \mathcal{H} \times \mathcal{H} \mid u \in Ax\}\).
The inverse of $A$ is $A^{-1}: u \mapsto \{ x \mid u \in Ax \}$. The zero set of $A$ is $\text{zer}(A) = A^{-1}0$. We say that $A$ is monotone if

$$\langle \forall u \in Ax \rangle (\forall (y, v) \in \text{gra} \ A) \quad \langle x - y \mid u - v \rangle \geq 0, \quad (1.7)$$

and it is maximally monotone if there exists no monotone operator $B$ such that $\text{gra}(B)$ properly contains $\text{gra}(A)$. The resolvent of $A$ is

$$J_A = (\text{Id} + A)^{-1}, \quad (1.8)$$

where $\text{Id}$ denotes the identity operator on $\mathcal{H}$. A single-valued operator $B: \mathcal{H} \to \mathcal{H}$ is $\beta$-cocoercive, for some $\beta \in ]0, +\infty[$, if

$$\langle \forall (x, y) \in \mathcal{H}^2 \rangle \langle x - y \mid Bx - By \rangle \geq \beta \|Bx - By\|^2. \quad (1.9)$$

The parallel sum of two operator $A: \mathcal{H} \to 2^{\mathcal{H}}$ and $B: \mathcal{H} \to 2^{\mathcal{H}}$ is $A \square B = (A^{-1} + B^{-1})^{-1}$. The class of all lower semicontinuous convex functions $f: \mathcal{H} \to ]-\infty, +\infty]$ such that $\text{dom} \ f = \{ x \in \mathcal{H} \mid f(x) < +\infty \} \neq \emptyset$ is denoted by $\Gamma(\mathcal{H})$. Now, let $f \in \Gamma(\mathcal{H})$. The subdifferential of $f \in \Gamma(\mathcal{H})$ is the maximally monotone operator

$$\partial f: \mathcal{H} \to 2^{\mathcal{H}}: x \mapsto \{ u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \quad \langle y - x \mid u \rangle + f(x) \leq f(y) \} \quad (1.10)$$

Moreover, the proximity operator of $f$ is

$$\text{prox}_f = J_{\partial f}: \mathcal{H} \to \mathcal{H}: x \mapsto \arg\min_{y \in \mathcal{H}} f(y) + \frac{1}{2} \|x - y\|^2. \quad (1.11)$$

Various closed-form expressions of the proximity operators are in [1, Chapter 24].

## 2 Weak Convergence

We first prove an auxiliary result which will be used to prove the weak convergence of the sequence generated by the reflected forward-backward splitting.

**Lemma 2.1** Let $(x_n)_{n \in \mathbb{N}}$ and $(y_n)_{n \in \mathbb{N}}$ be generated by Eq. 1.6. Set

$$\langle \forall n \in \mathbb{N} \rangle \quad p_{n+1} = x_n - \gamma By_n - x_{n+1}. \quad (2.1)$$

Suppose that $B$ is $\beta$-cocoercive. Then, for every $x \in \text{zer}(A + B)$,

$$\|x_{n+1} - x\|^2 + \|x_{n+1} - y_n\|^2 + \|x_{n+1} - x_n\|^2 + \|p_{n+1} + \gamma Bx\|^2 + \gamma(2\beta - \gamma)\|By_n - Bx\|^2$$

$$\leq \|x_n - x\|^2 + \|p_n + \gamma Bx\|^2 - \gamma^2 \|By_{n-1} - Bx\|^2 + 2\gamma \langle By_{n-1} - Bx, x_{n+1} - y_n \rangle. \quad (2.2)$$

If $B$ is monotone, we remain have Eq. 2.2 with $\beta = 0$.

**Proof** Let $x \in \text{zer}(A + B)$ and $n \in \mathbb{N}$. By the definition of the resolvent, $p_{n+1} \in \gamma A_{x_{n+1}}$. We have

$$\langle x_n + \gamma By_{n-1} - x_{n-1} \mid x_{n+1} - y_n \rangle = - \langle p_n \mid x_{n+1} - y_n \rangle. \quad (2.3)$$

Let us recall that

$$y_n = 2x_n - x_{n-1}$$

which is equivalent to $x_n - x_{n-1} = y_n - x_n$. (2.4)

Then it follows that

$$\langle y_n - x_n \mid x_{n+1} - y_n \rangle = - \langle p_n + \gamma By_{n-1} \mid x_{n+1} - y_n \rangle, \quad (2.5)$$

and

$$\langle x_{n+1} - x_n \mid x - x_{n+1} \rangle = - \langle p_{n+1} + \gamma By_n \mid x - x_{n+1} \rangle. \quad (2.6)$$
We have
\[
\begin{align*}
2 \langle y_n - x_n, x_{n+1} - y_{n+1} \rangle &= \|x_n - x_{n+1}\|^2 - \|x_n - y_n\|^2 - \|x_{n+1} - y_n\|^2 \\
2 \langle x_{n+1} - x_n, x - x_{n+1} \rangle &= \|x_n - x\|^2 - \|x_n - x_{n+1}\|^2 - \|x_{n+1} - x\|^2.
\end{align*}
\] (2.7)

In turn,
\[
\|x_{n+1} - x\|^2 + \|x_n - y_n\|^2 + \|x_{n+1} - y_n\|^2 = \|x_n - x\|^2 + 2\Gamma_n + 2\gamma \langle By_{n+1} - By_n, x_{n+1} - y_n \rangle, \tag{2.8}
\]
where
\[
\Gamma_n = \langle p_n + \gamma By_n, x_{n+1} - y_n \rangle + \langle p_{n+1} + \gamma By_n, x - x_{n+1} \rangle = \langle p_n, x_{n+1} - y_n \rangle + \langle p_{n+1}, x - x_{n+1} \rangle + \gamma \langle By_n, x - y_n \rangle. \tag{2.9}
\]

Since \(\gamma A\) is monotone and \(-\gamma Bx \in \gamma Ax\), we obtain
\[
\langle p_{n+1}, x - x_{n+1} \rangle \leq \langle p_{n+1}, x - x_{n+1} \rangle + \langle -\gamma Bx - p_{n+1}, x - x_{n+1} \rangle = \gamma \langle Bx, x_{n+1} - x \rangle. \tag{2.10}
\]

Since \(B\) is \(\beta\)-cocoercive, we also have
\[
\gamma \langle By_n, x - y_n \rangle \leq \gamma \langle Bx, x - y_n \rangle - \gamma \beta \|By_n - Bx\|^2. \tag{2.11}
\]

Adding (2.10) and (2.11), and using Eq. 2.4, the monotonicity of \(\gamma A\), we get
\[
\Gamma_n \leq \langle p_n, x_{n+1} - y_n \rangle + \gamma \langle Bx, x_{n+1} - y_n \rangle - \gamma \beta \|By_n - Bx\|^2
= \langle p_n + \gamma Bx, x_{n+1} - x_n \rangle - \langle p_n + \gamma Bx, x_n - x_{n-1} \rangle - \gamma \beta \|By_n - Bx\|^2
\leq \langle p_{n+1} + \gamma Bx, x_{n+1} - x_n \rangle - \langle p_{n+1} + \gamma Bx, x_n - x_{n-1} \rangle - \gamma \beta \|By_n - Bx\|^2. \tag{2.12}
\]

Let us set
\[
T_n = \|x_n - x\|^2 - 2 \langle p_n + \gamma Bx, x_n - x_{n-1} \rangle. \tag{2.13}
\]

It follows from Eqs. 2.12 and 2.8 that
\[
T_{n+1} + \|x_n - y_n\|^2 + \|x_{n+1} - y_n\|^2 + 2\gamma \beta \|By_n - Bx\|^2 \leq T_n + 2\gamma \langle By_{n+1} - By_n, x_{n+1} - y_n \rangle. \tag{2.14}
\]

By Eq. 2.1, we have \(p_{n+1} + x_{n+1} - x_n = -\gamma By_n\), and hence,
\[
-2 \langle p_{n+1} + \gamma Bx, x_{n+1} - x_n \rangle = \|x_{n+1} - x_n\|^2 + \|p_{n+1} + \gamma Bx\|^2
- \|p_{n+1} + \gamma Bx, x_{n+1} - x_n\|^2
= \|x_{n+1} - x_n\|^2 + \|p_{n+1} + \gamma Bx\|^2 - \gamma^2 \|By_n - Bx\|^2, \tag{2.15}
\]
which implies that
\[
T_{n+1} = \|x_{n+1} - x\|^2 + \|x_{n+1} - x_n\|^2 + \|p_{n+1} + \gamma Bx\|^2 - \gamma^2 \|By_n - Bx\|^2. \tag{2.16}
\]

Therefore, using Eq. 2.4 again, Eq. 2.14 becomes
\[
\|x_{n+1} - x\|^2 + \|x_{n+1} - y_n\|^2 + \|x_{n+1} - x_n\|^2 + \|p_{n+1} + \gamma Bx\|^2 + (2\beta - \gamma) \|By_n - Bx\|^2
\leq \|x_n - x\|^2 + \|p_n + \gamma Bx\|^2 - \gamma^2 \|By_n - Bx\|^2 + 2\gamma \langle By_{n+1} - By_n, x_{n+1} - y_n \rangle, \tag{2.17}
\]
which proves the desired result. In the case when \(B\) is monotone, setting \(\beta = 0\), we obtain the second conclusion. \(\square\)

The main result of this paper is stated in the following theorem where we prove the weak convergence of the sequence generated by the reflected forward-backward splitting to a zero point of \(A + B\).
**Theorem 2.2** Let \((x_n)_{n \in \mathbb{N}}\) and \((y_n)_{n \in \mathbb{N}}\) be generated by Eq. 1.6. Suppose that \(\gamma \in \left[0, (\sqrt{2} - 1)/\mu\right]\), then \(x_n \to \bar{x} \in \text{zer}(A + B)\) and \(y_n \to \bar{x} \in \text{zer}(A + B)\).

**Proof** Since \(B\) is monotone, by Lemma 2.1, we have
\[
\|x_{n+1} - x\|^2 + \|p_{n+1} + \gamma Bx\|^2 - \gamma^2 \|B y_n - Bx\|^2 \\
\leq \|x_n - x\|^2 + \|p_n + \gamma Bx\|^2 - \gamma^2 \|B y_{n-1} - Bx\|^2 + 2\gamma \langle B y_{n-1} - B y_n, x_{n+1} - y_n \rangle. \tag{2.18}
\]
Since \(B\) is \(\mu\)-Lipschitz continuous, for any \(\tau_1 > 0\) and \(\tau_2 > 0\), we obtain
\[
2 \langle B y_n - B y_{n-1}, y_n - x_{n+1} \rangle \leq 2\mu \|y_n - y_{n-1}\| \|y_n - x_{n+1}\| \\
\leq \frac{\mu}{\tau_1} \|y_n - y_{n-1}\|^2 + \mu \tau_1 \|y_n - x_{n+1}\|^2 \\
\leq \frac{\mu}{\tau_1} \left( 1 + \frac{1}{\tau_2} \right) \|y_n - x_n\|^2 + \left( 1 + \tau_2 \right) \|x_n - y_{n-1}\|^2 + \mu \tau_1 \|y_n - x_{n+1}\|^2 \\
= \alpha_1 \|y_n - x_n\|^2 + \alpha_2 \|x_n - y_{n-1}\|^2 + \mu \tau_1 \|y_n - x_{n+1}\|^2, \tag{2.19}
\]
where we set
\[\alpha_1 = \frac{\mu}{\tau_1} \left( 1 + \frac{1}{\tau_2} \right) \text{ and } \alpha_2 = \frac{\mu}{\tau_1} \left( 1 + \tau_2 \right). \tag{2.20}\]
Therefore, we derive from Eq. 2.18 that
\[
\|x_{n+1} - x\|^2 + \gamma \alpha_2 \|x_{n+1} - y_n\|^2 + \|x_{n+1} - x_n\|^2 + \|p_{n+1} + \gamma Bx\|^2 - \gamma^2 \|B y_n - Bx\|^2 \\
\leq \|x_n - x\|^2 + \|x_{n-1} - x_n\|^2 + \|p_n + \gamma Bx\|^2 - \gamma^2 \|B y_{n-1} - Bx\|^2 \\
- \left( 1 - \gamma \alpha_1 \right) \|x_n - y_n\|^2 + \gamma \alpha_2 \|x_n - y_{n-1}\|^2 - \left( 1 - \gamma \left( \alpha_2 + \mu \tau_1 \right) \right) \|x_{n+1} - y_n\|^2. \tag{2.21}
\]
Therefore, we need the following condition on \(\gamma\),
\[
0 < \gamma < \frac{1}{\alpha_1} = \frac{\tau_1 \tau_2}{\mu(1 + \tau_2)} \text{ and } 0 < \gamma < \frac{1}{\alpha_2 + \mu \tau_1} = \frac{\tau_1}{\mu(1 + \tau_2 + \tau_1^2)}. \tag{2.22}
\]
The optimal bound of \(\gamma\) happens when
\[
\frac{\tau_1 \tau_2}{\mu(1 + \tau_2)} = \frac{\tau_1}{\mu(1 + \tau_2 + \tau_1^2)}, \tag{2.23}
\]
which implies that
\[
\tau_2 \tau_1^2 = 1 - \tau_2^2. \tag{2.24}
\]
Hence,
\[
0 < \gamma^2 < \frac{\tau_2 \tau_1^2}{\mu^2(1 + \tau_2)^2} = \frac{\tau_2(1 - \tau_2)}{\mu^2(1 + \tau_2)} = \frac{(1 + \tau_2 - 1)(2 - (1 + \tau_2))}{\mu^2(1 + \tau_2)}. \tag{2.25}
\]
We express the last term as
\[
\frac{(1 + \tau_2 - 1)(2 - (1 + \tau_2))}{\mu^2(1 + \tau_2)} = \frac{1}{\mu^2} \left( 3 - 1 - \tau_2 - \frac{2}{1 + \tau_2} \right). \tag{2.26}
\]
which attains the maximum when \(1 + \tau_2 = \sqrt{2}\) and hence \(\tau_1 = \sqrt{2}\). It follows that \(\gamma^2 < \mu^{-2}(3 - 2\sqrt{2})\). Thus
\[
0 < \gamma < \frac{\sqrt{2} - 1}{\mu}. \tag{2.27}
\]
We next set
\[
E_n = \|x_n - x\|^2 + \|x_{n-1} - x_n\|^2 + \|p_n + \gamma Bx\|^2 + \gamma \alpha_2 \|x_n - y_{n-1}\|^2 - \gamma^2 \|B y_{n-1} - Bx\|^2. \tag{2.28}
\]
Then, we can rewrite (2.21) as
\[ E_{n+1} \leq E_n - (1 - \gamma \mu(1 + \sqrt{2}))\|x_n - y_n\|^2 - (1 - \gamma \mu(1 + \sqrt{2}))\|x_{n+1} - y_n\|^2. \] (2.29)
We have
\[
\gamma^2 \|By_{n-1} - Bx\|^2 \leq 2\gamma^2 \|By_{n-1} - Bx_n\|^2 + 2\gamma^2 \|Bx_n - Bx\|^2 \\
\leq 2\gamma^2 \mu^2 \|y_{n-1} - x_n\|^2 + 2\gamma^2 \mu^2 \|x_n - x\|^2.
\] (2.30)
Thus, since \( \gamma < (\sqrt{2} - 1)/\mu \), we have \( \gamma < 2/\mu \) and hence there exists \( \epsilon > 0 \) such that
\[
E_n \geq \gamma \mu(1 - 2\gamma \mu)\|x_n - y_n - 1\|^2 + (1 - 2\gamma \mu^2)\|x_n - x\|^2 \\
\geq \epsilon (\|y_{n-1} - x_n\|^2 + \|x_n - x\|^2) \\
\geq 0.
\] (2.31)
In turn, we derive from Eq. 2.29 that
\[
\begin{cases}
E_n \to \bar{\zeta} \in \mathbb{R} \\
\sum_{n \in \mathbb{N}} \|x_n - y_n\|^2 < +\infty \\
\sum_{n \in \mathbb{N}} \|x_{n+1} - y_n\|^2 < +\infty
\end{cases}
\] (2.32)
Since \((E_n)_{n \in \mathbb{N}}\) converges, it is bounded and therefore, it follows from Eq. 2.31 that \((\|x_n - x\|)_{n \in \mathbb{N}}\) and \((x_n)_{n \in \mathbb{N}}\) are bounded. Hence, \((y_n)_{n \in \mathbb{N}}\) and \((Bx_{n-1} - Bx)_{n \in \mathbb{N}}\) are also bounded. Since \(x_n - y_n \to 0\), we obtain \((y_n - x_n \mid Bx_{n-1} - Bx) \to 0\). We have
\[
\|p_n + \gamma Bx\|^2 - \gamma^2 \|By_{n-1} - Bx\|^2 = \|x_{n-1} - x_n\|^2 - 2\gamma \langle x_{n-1} - x_n \mid By_{n-1} - Bx \rangle \\
= \|y_n - x_n\|^2 + 2\gamma \langle y_n - x_n \mid By_{n-1} - Bx \rangle \\
\to 0.
\] (2.33)
Therefore, \(\|x_n - x\| \to \bar{\zeta}\). Let \(\bar{x}\) be a weak cluster point of \((x_n)_{n \in \mathbb{N}}\), then there exists \(x_{k_n} \to \bar{x}\). Note that \(x_{k_n} - x_{k_n+1} - \gamma By_{k_n} + \gamma Bx_{k_n+1} \to 0\) and
\[
x_{k_n} - x_{k_n+1} - \gamma By_{k_n} + \gamma Bx_{k_n+1} = p_{k_n+1} + \gamma Bx_{k_n+1} \in \gamma(A + B)x_{k_n+1}.
\] (2.34)
Since \(A + B\) is maximally monotone, its graph is closed in \(\mathcal{H}^{strong} \times \mathcal{H}^{weak}\). Therefore, it follows from Eq. 2.34 that \(\bar{x} \in \text{zer}(A + B)\). Using Opial’s result [12, 13], we obtain \(x_n \to \bar{x} \in \text{zer}(A + B)\).

**Proposition 2.3** If \(B\) is \(\beta\)-cocoercive for some \(\beta \in ]0, \infty[\), then the conclusion of Theorem 2.2 remains valid for \(\gamma < \beta/2\).

**Proof** By Lemma 2.1, we get
\[
\|x_{n+1} - x\|^2 + \|x_{n+1} - y_n\|^2 + \|x_{n+1} - x_n\|^2 + \|p_{n+1} + \gamma Bx\|^2 + \gamma(2\beta - \gamma)\|By_n - Bx\|^2 \\
\leq \|x_n - x\|^2 + \|p_n + \gamma Bx\|^2 - \gamma^2 \|By_{n-1} - Bx\|^2 + 2\gamma \langle By_{n-1} - By\mid x_{n+1} - y_n \rangle.
\] (2.35)
Let us estimate the term \(q_n = 2\gamma \langle By_{n-1} - By\mid x_{n+1} - y_n \rangle\). Let \(\epsilon \in ]0, \infty[\) be such that \(0 < \gamma < (1 - \epsilon)\beta/2\). We have
\[
q_n = 2\gamma \langle By_{n-1} - Bx\mid x_{n+1} - y_n \rangle + 2\gamma \langle Bx - By\mid x_{n+1} - y_n \rangle \\
\leq \frac{2\gamma^2}{1 - \epsilon} \|By_{n-1} - Bx\|^2 + \frac{2\gamma^2}{1 - \epsilon} \|By_n - Bx\|^2 + (1 - \epsilon)\|x_{n+1} - y_n\|^2,
\] (2.36)
and thus, we derive from Eq. 2.35 that
\[
\|x_{n+1} - x\|^2 + \|x_{n+1} - x_n\|^2 + \|p_{n+1} + \gamma Bx\|^2 + \gamma^2 \frac{1 + \epsilon}{1 - \epsilon} \|By_n - Bx\|^2 + \epsilon \|x_{n+1} - y_n\|^2 \\
\leq \|x_n - x\|^2 + \|p_n + \gamma Bx\|^2 + \gamma^2 \frac{1 + \epsilon}{1 - \epsilon} \|By_{n-1} - Bx\|^2 - \gamma(2\beta - \frac{2\gamma^2}{1 - \epsilon})\|By_n - Bx\|^2.
\] (2.37)
Since \( \gamma < (1 - \epsilon)\beta / 2 \), we obtain

\[
\begin{align*}
\|x_n - x\|^2 + \|p_n + \gamma Bx\|^2 + \gamma^2 \frac{1 + \epsilon}{1 - \epsilon}\|By_{n-1} - Bx\|^2 &\to \overline{\xi} \in \mathbb{R}, \\
\sum_{n \in \mathbb{N}} \|By_n - Bx\|^2 &< \infty, \\
\sum_{n \in \mathbb{N}} \|x_n - x_{n+1}\|^2 &< +\infty, \\
\sum_{n \in \mathbb{N}} \|y_n - x_{n+1}\|^2 &< +\infty.
\end{align*}
\] (2.38)

Since \( p_{n+1} + \gamma Bx = \gamma (Bx - By_n) + x_n - x_{n+1} \to 0 \), it follows that \( \|x_n - x\|^2 \to \overline{\xi} \in \mathbb{R} \)

and hence \((x_n)_{n \in \mathbb{N}}\) is bounded. Let \( \overline{x} \) be a weak cluster point of \((x_n)_{n \in \mathbb{N}}\), then there exists a subsequence \((x_{k_n})_{n \in \mathbb{N}}\) of \((x_n)_{n \in \mathbb{N}}\) such that \( x_{k_n} \to \overline{x} \). Note that \( By_{k_n} \to Bx \) and \( x_n - y_n = x_{n-1} - x_n \to 0 \). Since \( B \) is maximally monotone, its graph is closed in \( \mathcal{H}^{\text{strong}} \times \mathcal{H}^{\text{weak}} \), we obtain \( Bx = B\overline{x} \) and thus \( By_{k_n} \to B\overline{x} \). Since \( A \) is maximally monotone, its graph is closed in \( \mathcal{H}^{\text{strong}} \times \mathcal{H}^{\text{weak}} \), passing limit from

\[
\frac{x_{k_n} - x_{k_n+1}}{\gamma} - By_{k_n} = p_{k_n+1}/\gamma \in Ax_{k_n+1},
\] (2.39)

we obtain \( \overline{x} \in \text{zer}(A + B) \). Therefore, using Opial’s result [12], we obtain \( x_n \to \overline{x} \in \text{zer}(A + B) \).

**Remark 2.4** Here are some remarks.

(i) A special case of Theorem 2.2 is in [9] when \( A = N_S \), the normal cone operator to a closed convex set \( S \subset \mathcal{H} \). The line-search versions for the case where \( A = \partial f \) for some \( f \in \Gamma_0(\mathbb{R}^d) \) are in [10]. The connection to the existing work concerning with solving variational inequalities can be found in [9, 10]. For the compactness, we do not cite all of them here.

(ii) In the case, \( B \) is \( \beta \)-cocoercive, \( \mu = 1/\beta \) and the range of the step size is relaxed from \( [0, (\sqrt{2} - 1)/\mu] \) to \( [0, 0.5\beta] \) which is relatively small in comparison to the standard forward-backward splitting [1].

(iii) In the case, \( B \) is linear, Eq. 1.6 is exactly the same as the one in [11] where the convergence is proved under the condition \( \gamma \in [0, 0.5/\mu] \). The computational cost of Eq. 1.6 and [11] are much cheaper than that of FBFS in [17].

**Corollary 2.5** Let \( B : \mathcal{H} \to \mathcal{H} \) be a monotone and \( \mu_0 \)-Lipschitzian, \( \mu_0 \in ]-\infty, +\infty[\), and \( A : \mathcal{H} \to 2^{\mathcal{H}} \) be maximally monotone. Let \( m \) be a strictly positive integer and let \((G_i)_{1 \leq i \leq m}\) be real Hilbert spaces. For every \( i \in \{1, \ldots, m\} \), let \( A_i : G_i \to 2^{G_i} \) be a maximally monotone, and let \( B_i : G_i \to 2^{G_i} \) be a maximally monotone such that \( B_i^{-1} \) is \( \mu_i \)-Lipschitzian operator for some \( \mu_i \in ]-\infty, +\infty[\). let \( L_i : \mathcal{H} \to G_i \) be a bounded linear operator such that \( 0 \neq \sum_{i=1}^{m} \|L_i\|^2 \). Suppose that

\[
0 \in \text{ran } (A + \sum_{i=1}^{m} L_i^* (A_i \Box B_i) L_i + B).
\] (2.40)

The primal inclusion is to find \( \overline{x} \) such that

\[
0 \in A\overline{x} + \sum_{i=1}^{m} L_i^* (A_i \Box B_i) L_i \overline{x} + B \overline{x},
\] (2.41)
and the dual inclusion is to find \((\bar{v}_i)_{1 \leq i \leq m} \in (\mathcal{G}_i)_{1 \leq i \leq m}\) such that

\[
(\forall i \in \{1, \ldots, m\}) \quad 0 \in L_i(A + B)^{-1}(- \sum_{i=1}^{m} L_i^* \bar{v}_i) + A_i^{-1} \bar{v}_i + B_i^{-1} \bar{v}_i.
\] (2.42)

Set

\[
\mu = \max\{\mu_0, \ldots, \mu_m\} + \sqrt{\sum_{i=1}^{m} \|L_i\|^2}.
\] (2.43)

Let \(\gamma \in \left[0, \frac{\sqrt{2} - 1}{\mu}\right]\). (2.44)

Then \(x_n \rightharpoonup \bar{x}\) solves (2.41) and \((v_{1,n}, \ldots, v_{m,n}) \rightharpoonup (\bar{v}_1, \ldots, \bar{v}_m)\) solves (2.42).

**Proof** We use the technique in [6]. Let \(\mathcal{K} = \mathcal{H} \oplus \mathcal{G}_1 \oplus \ldots \oplus \mathcal{G}_m\) be the Hilbert direct sum of the Hilbert spaces \(\mathcal{H}\) and \((\mathcal{G}_i)_{1 \leq i \leq m}\), where the scalar product and the associated norm of \(\mathcal{K}\) are respectively defined as, for any \((x, v) = (x, v_1, \ldots, v_m) \in \mathcal{K}\) and \((y, w) = (x, w_1, \ldots, w_m) \in \mathcal{K}\),

\[
\langle x, y \rangle + \sum_{i=1}^{m} \langle v_i, w_i \rangle,
\] (2.45)

and

\[
\|x\| = \left(\|x\|^2 + \sum_{i=1}^{m} \|v_i\|^2\right)^{1/2}.
\] (2.46)

Let us define

\[
\begin{align*}
B: \mathcal{K} &\to \mathcal{K}: (x, v_1, \ldots, v_m) \mapsto (Bx + \sum_{i=1}^{m} L_i^* v_i, -L_1 x + B_1^{-1} v_1, \ldots, -L_m x + B_m^{-1} v_m), \\
A: \mathcal{K} &\to 2^{\mathcal{K}}: (x, v_1, \ldots, v_m) \mapsto A x \times A_1^{-1} v_1 \times \ldots \times A_m^{-1} v_m.
\end{align*}
\] (2.47)

It is shown in [6, Eq. (3.12)] and [6, Eq. (3.13)] that under the condition (2.40), \(\text{zer}(A + B) \neq \emptyset\). Furthermore, [6, Eq. (3.21)] and [6, Eq. (3.22)] yield

\[
(\bar{x}, \bar{v}_1, \ldots, \bar{v}_m) \in \text{zer}(A + B) \Rightarrow \bar{x} \text{ solves (2.41) and } (\bar{v}_1, \ldots, \bar{v}_m) \text{ solves (2.42)}.
\] (2.48)

It is show in [6] that \(\mathcal{B}\) is monotone and \(\mu\)-Lipschitzian and using [1, Proposition 20.23], \(A\) is also a maximally monotone operator. Furthermore, it follows from [1, Proposition 23.16] that

\[
(\forall x = (x, v_1, \ldots, v_m) \in \mathcal{K})(\forall \gamma \in [0, +\infty]) \quad J_{y,A} x = \left(J_{y,A} x, J_{y,A_1^{-1} v_1} \ldots, J_{y,A_m^{-1} v_m}\right).
\] (2.49)

For every \(n \in \mathbb{N}\), set \(x_n = (x_n, v_{1,n}, \ldots, v_{m,n})\). Then the proposed algorithm can be rewritten in the space \(\mathcal{K}\) as follows

\[
x_{n+1} = J_{y,A}(x_n - \gamma B(2x_n - x_{n-1})).
\] (2.50)

In view of Theorem 2.2(ii), \((x_n)_{n \in \mathbb{N}}\) converges weakly to \(\bar{x} = (\bar{x}, \bar{v}_1, \ldots, \bar{v}_m)\) in \(\text{zer}(A + B)\). By Eq. 2.48, it follows that \(x_n \rightharpoonup \bar{x}\) solves Eq. 2.41 and \((v_{1,n}, \ldots, v_{m,n}) \rightharpoonup (\bar{v}_1, \ldots, \bar{v}_m)\) solves (2.42).
Remark 2.6 Here are some remarks:

(i) The structured primal-dual monotone inclusions (2.41)–(2.42) is firstly introduced in [6] and then [2, 8, 16, 18, 20]. Various special can be found in the literature [6, 14, 15].

(ii) The iteration Eq. 2.44 is different from the one in [6] and Eq. 2.44 requires only one call of $B_i, (B_i)_{1 \leq i \leq m}, (L_i)_{1 \leq i \leq m}$ per iteration.

(iii) Even when $B_i, (B_i)_{1 \leq i \leq m}$ are restricted to be cocoercive, Eq. 2.44 is different from the one in [19].

(iv) Using the same idea as in [6], concretes applications to minimization problem involving the parallel sums are straightforward and we omit them here.

3 Numerical Experiment

We provide an application of the proposed method to the image denoising by the total variation which was investigated in [7, Problem 4.16]. Let $z_0 \in \mathbb{R}^{N \times N}$ be the ideal image and $z$ be its corrupted observation of the form

$$z = z_0 + w,$$  \hspace{1cm} (3.1)

where $w$ is the realization of a noise process. The goal is to recover $z_0$ from $z$. Let us recall some notations in [7, Section 4] and the references therein. The discrete gradient operator is defined by

$$\nabla : \mathbb{R}^{N \times N} \to \mathbb{R}^{N \times N} \oplus \mathbb{R}^{N \times N} : (x_{k,l})_{1 \leq k,l \leq N} \mapsto (\eta^{(1)}_{k,l}, \eta^{(2)}_{k,l})_{1 \leq k,l \leq N},$$  \hspace{1cm} (3.2)

Fig. 1 Convergence results for the parrotgray image.
where
\[
(\forall (k, l) \in \{1, \ldots, N\}^2) \quad \begin{cases} 
\eta^{(1)}_{k,l} = x_{k+1,l} - x_{k,l}, & \text{if } k < N; \\
\eta^{(1)}_{N,l} = 0; \\
\eta^{(2)}_{k,l} = x_{k,l+1} - x_{k,l}, & \text{if } l < N; \\
\eta^{(2)}_{k,N} = 0.
\end{cases}
\] (3.3)

Then the discrete total variation function is defined by
\[
tv: \mathbb{R}^{N \times N} \rightarrow \mathbb{R}: x \mapsto \|\nabla x\|_{2,1} = \sum_{1 \leq k,l \leq N} \sqrt{|\eta^{(1)}_{k,l}|^2 + |\eta^{(2)}_{k,l}|^2}.
\] (3.4)

Now let \( X \) be a nonempty closed convex set that model prior knowledge on the ideal image \( z_0 \). To recover \( z_0 \) from \( z \), we focus on solving the following problem
\[
\text{minimize } \xi \ tv(x) + \frac{1}{2} \|x - z\|^2;
\] (3.5)
where \( \xi \) is a strictly positive number, together with its dual problem
\[
\text{minimize } \tilde{\sigma}_X(z + \xi \ \text{div } \nu),
\] (3.6)
where \( \text{div} = -\nabla^* \) [7, Eq. (4.45)], \( \tilde{\sigma}_X \) is the Moreau envelope of the support function \( \sigma_X \) [7, Eq. (2.13)], and \( Y \) is given by
\[
Y = \left\{ (\eta^{(1)}_{k,l}, \eta^{(2)}_{k,l})_{1 \leq k,l \leq N} \in \mathbb{R}^{N \times N} \oplus \mathbb{R}^{N \times N} \mid \max_{1 \leq k,l \leq N} \sqrt{|\eta^{(1)}_{k,l}|^2 + |\eta^{(2)}_{k,l}|^2} \leq 1 \right\}. \] (3.7)

It is shown in [7] that the set of solutions to Eq. 3.6 is nonempty. Moreover, if \( \tilde{\nu} \) is a solution to Eq. 3.6, then \( x^* = P_X(z + \xi \ \text{div } \tilde{\nu}) \) is the solution to Eq. 3.5. Furthermore, the function \( \nu \mapsto \tilde{\sigma}_X(z + \xi \ \text{div } \nu) \) is the differentiable convex function with \( 8\xi \)-Lipschitz continuous gradient. Therefore, the dual problem can be solved by various methods in the literature. In this section, we will compare the following methods with \( \xi = 0.1 \) and \( X = \{ x \in \mathbb{R}^{N \times N} \mid (\forall (k, l) \in \{1, \ldots, N\}^2) \ 0 \leq x_{k,l} \leq 1 \} \).

(i) Tseng’s splitting method (FBFS) with \( \gamma = 0.99/(8\xi) \).
(ii) Forward-reflected-backward splitting method (FRBS) with \( \gamma = 0.49/(8\xi) \).
(iii) Reflected forward-backward splitting method (RFBS) with \( \gamma = 0.49/(8\xi) \).
(iv) Dual forward-backward splitting (DFBS) with \( \gamma = 1.99/(8\xi) \) [7].

We use the parrotgray image with \( N = 256 \) and \( w = \mathcal{N}(0, 0.1) \). The optimal solution \( x^* \) is found by the primal-dual method in [19]. The results, which are implemented by Octave 4.1, are presented in Figs. 1, 2 and 3.

Fig. 2 Original parrotgray image (left) and noisy image (right) with \( w = \mathcal{N}(0, 0.1) \)
Fig. 3 Images denoising: RFBS (top left), FRBS (top right), FBFS (bottom left), DFBS (bottom right)

Acknowledgments We thank the referees for their suggestions and correction which helped to improve the first version of the manuscript. The work of B. Cong Vu and Volkan Cevher were supported by European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement no 725594 - time-data).

References

1. Bauschke, H.H., Combettes, P.L., Analysis, Convex. Convex Analysis and Monotone Operator Theory in Hilbert Spaces, 2nd edn. Springer, New York (2017)
2. 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. SIAM J. Optim. 23, 2541–2565 (2013)
3. Boţ, R.I., Hendrich, C.: Convergence analysis for a Primal-Dual monotone + skew splitting algorithm with applications to total variation minimization. J. Math. Imaging Vis. 49, 551–568 (2014)
4. Briceño-Arias, L.M., Combettes, P.L.: A monotone+skew splitting model for composite monotone inclusions in duality. SIAM J. Optim. 21, 1230–1250 (2011)
5. Combettes, P.L.: Systems of structured monotone inclusions: duality, algorithms, and applications. SIAM J. Optim. 23, 2420–2447 (2013)
6. 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 Var. Anal. 20, 307–330 (2012)
7. Combettes, P.L., Dün, D., Vü, B.C.: Dualization of Signal Recovery Problems. Set-Valued Var. Anal., (18) 373–404 (2010)
8. Dün, D., Vü, B.C.: A splitting algorithm for system of composite monotone inclusions. Vietnam J. Maths. 43, 323–341 (2015)
9. Malitsky, Y.: Projected reflected gradient methods for monotone variational inequalities. SIAM J. Control Optim. 25, 502–520 (2015)
10. Malitsky, Y.: Proximal extrapolated gradient methods for variational inequalities. Optim. Methods Softw. 33, 140–164 (2018)
11. Malitsky, Y., Tam, M.K.: A Forward-Backward Splitting Method for Monotone Inclusions Without Cocoercivity, arXiv preprint (2018)
12. Opial, Z.: Weak convergence of the sequence of successive approximations for nonexpansive mappings. Bull. Amer. Math. Soc. 73, 591–597 (1967)
13. Passty, G.B.: Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. J. Math. Anal. Appl. 72, 383–390 (1979)
14. Pham, M.Q., Duval, L., Chaux, C., Pesquet, J.-C.: A primal-dual proximal algorithm for sparse template-based adaptive filtering: Application to seismic multiple removal. IEEE Trans. Signal Process 62, 4256–4269 (2014)
15. Repetti, A., Chouzenoux, E., Pesquet, J.-C.: A penalized weighted least squares approach for restoring data corrupted with signal-dependent noise. In: Proceedings of the 20th European Signal Processing (SIPCO 2012), 1553-1557, Bucharest, Romania (2012)
16. Ryu, E.K., Vù, B.C.: Finding the Forward-Douglas-Rachford-Forward Method. J. Optim. Theory Appl. 184, 858–876 (2020)
17. Tseng, P.: A modified forward-backward splitting method for maximal monotone mappings. SIAM J. Control Optim. 38, 431–446 (2000)
18. Vù, B.C.: Almost sure convergence of the forward-backward-forward splitting algorithm. Optim. Lett. 10, 781–803 (2016)
19. Vù, B.C.: A splitting algorithm for dual monotone inclusions: involving cocoercive operators. Adv. Comput. Math. 38, 667–681 (2013)
20. Vù, B.C.: A variable metric extension of the forward–backward–forward algorithm for monotone operators. Numer. Funct. Anal. Optim. 34, 1050–1065 (2013)

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