Four-Operator Splitting via a Forward–Backward–Half-Forward Algorithm with Line Search

Luis M. Briceño-Arias · Fernando Roldán

Received: 13 May 2022 / Accepted: 8 July 2022 / Published online: 9 September 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
In this article, we provide a splitting method for solving monotone inclusions in a real Hilbert space involving four operators: a maximally monotone, a monotone-Lipschitzian, a cocoercive, and a monotone-continuous operator. The proposed method takes advantage of the intrinsic properties of each operator, generalizing the forward–backward–half-forward splitting and the Tseng’s algorithm with line search. At each iteration, our algorithm defines the step size by using a line search in which the monotone-Lipschitzian and the cocoercive operators need only one activation. We also derive a method for solving nonlinearly constrained composite convex optimization problems in real Hilbert spaces. Finally, we implement our algorithm in a nonlinearly constrained least-square problem and we compare its performance with available methods in the literature.

Keywords Convex optimization · Cocoercive operator · Lipschitzian operator · Monotone operator theory · Splitting algorithms

Mathematics Subject Classification 47H05 · 47H10 · 65K05 · 65K15 · 90C25 · 49M29
1 Introduction

In this paper, we aim at finding a zero of the sum of a maximally monotone, a cocoercive, a monotone-Lipschitzian, and a monotone-continuous operators, which is also in a convex closed subset $X$ of a Hilbert space. This inclusion encompasses several problems in partial differential equations coming from mechanical models [24–26], differential inclusions [1, 36], game theory [11], among other disciplines. When $X$ is the whole Hilbert space, the algorithms proposed in [3–7, 10, 13, 14, 16–23, 27, 28, 30–32, 34, 39] can solve this kind of problems under additional assumptions or without exploiting the intrinsic properties of the involved operators. Indeed, the algorithms in [5–7, 10, 16, 23] need to compute the resolvents of all the monotone operators, which are not explicit in general, or they can be numerically expensive. The schemes proposed in [3, 17, 22, 27] take advantage of the properties of the monotone-Lipschitzian operator, but the cocoercivity and the continuity of the others are not leveraged. In fact, the algorithms in [3, 17, 22, 27] need to activate the continuous operator via its resolvent and to explicitly implement the cocoercive and the monotone-Lipschitzian operators twice by iteration. In contrast, the algorithms in [14, 19, 30, 32, 34] activate the cocoercive and the monotone-Lipschitzian operators only once by iteration, but they need to store in the memory the two past iterations and the step size is reduced significantly. Furthermore, the methods in [14, 19, 30] consider only one maximally monotone operator and, hence, it needs to compute the resolvent of the sum of the maximally monotone and the monotone-continuous operator. On the other hand, methods in [32, 34] need to calculate the resolvent of the monotone-continuous operator, which is not simple in general. In addition, methods proposed in [4, 18, 20, 21, 28, 31, 39] take advantage of the cocoercive operator by activating it once by iteration, but they do not exploit continuity nor the monotone-Lipschitzian property of the operators and they need to compute their resolvents. The method in [13] exploits the properties of the cocoercive and the monotone-Lipschitzian operators, but it does not take into advantage the monotonicity and continuity of one of the operators and need to compute its resolvent. Other methods solving this kind of problems including a normal cone to a closed vector subspace, when the continuous operator is zero and either the cocoercive or the monotone-Lipschitzian operator is zero, are discussed in [8, 9, 37].

In the case when $X$ is a subset of the domain of the maximally monotone operator, methods exploiting the monotonicity and continuity property of one of the operators involved in our problem are proposed in [12, 38]. In particular, the algorithm in [38] solves our problem by using a line search procedure involving the sum of the cocoercive, the monotone-Lipschitzian, and the monotone-continuous operators. On the other hand, the forward–backward–half-forward splitting (FBHF) method, proposed in [12], solves our problem via a line search which activates the sum of the monotone-Lipschitzian and the monotone-continuous operators. As perceived in [12], to reduce the activation of monotone operators in the line search procedure can reduce the number of iterations significantly (see, e.g., [12, Table 3] in which this reduction is around a 20%). Moreover, when the explicit implementation of these operators is expensive (e.g., high-dimensional problems), the computational time can be much larger because line search procedures need to activate those operators several times by iteration.
In this paper, we propose a fully split method for solving our monotone inclusion, which takes advantage of the intrinsic properties of each operator involved in the inclusion. More precisely, the proposed method activates the cocoercive operator once, the monotone-Lipschitzian operator twice by iteration, and uses a line search only implementing the monotone-continuous operator. We also ensure the weak convergence of our algorithm under hypotheses on the set \( X \) which are independent of the domain of the maximally monotone operator, generalizing some results in [12, 38]. We explore an interesting example in optimization, involving nonlinear inequality constraints governed by convex Gâteaux differentiable functions. We provide conditions on this function in order to guarantee that the saddle operator obtained from the Lagrangian of the problem is monotone and continuous and satisfies the hypotheses of our main convergence theorem. Finally, we provide a numerical experiment which illustrates the efficiency of our algorithm.

The paper is organized as follows. In Sect. 2, we set our notation. In Sect. 3, we provide some technical lemmas, our splitting method for solving Problem 3.1, and our convergence result. In Sect. 4, we derive an algorithm for solving a nonlinearly constrained composite convex optimization problem. Finally, in Sect. 5 we provide a numerical experiment illustrating the efficiency of the method proposed in Sect. 4.

2 Preliminaries

Throughout this paper, \( \mathcal{H} \) and \( \mathcal{G} \) are real Hilbert spaces. We denote the scalar product by \( \langle \cdot \mid \cdot \rangle \) and the associated norm by \( \| \cdot \| \). The symbols \( \rightharpoonup \) and \( \rightarrow \) denote the weak and strong convergence, respectively. Given a linear bounded operator \( M : \mathcal{H} \rightarrow \mathcal{G} \), we denote its adjoint by \( M^* : \mathcal{G} \rightarrow \mathcal{H} \). \( \text{Id} \) denotes the identity operator on \( \mathcal{H} \). Let \( D \subset \mathcal{H} \) be nonempty, let \( T : D \rightarrow \mathcal{H} \), and let \( \beta \in ]0, +\infty[ \). The operator \( T \) is \( \beta \)-cocoercive if

\[
(\forall x \in D)(\forall y \in D) \quad \langle x - y \mid Tx - Ty \rangle \geq \beta \|Tx - Ty\|^2, \tag{2.1}
\]

and it is \( \beta \)-Lipschitzian if

\[
(\forall x \in D)(\forall y \in D) \quad \|Tx - Ty\| \leq \beta \|x - y\|. \tag{2.2}
\]

Let \( A : \mathcal{H} \rightarrow 2^\mathcal{H} \) be a set-valued operator. The domain, range, graph, and the zeros of \( A \) are, respectively, \( \text{dom} \ A = \{x \in \mathcal{H} \mid Ax \neq \emptyset\} \), \( \text{ran} \ A = \{u \in \mathcal{H} \mid (\exists x \in \mathcal{H}) u \in Ax\} \), \( \text{gra} \ A = \{(x, u) \in \mathcal{H} \times \mathcal{H} \mid u \in Ax\} \), and \( \text{zer} \ A = \{x \in \mathcal{H} \mid 0 \in Ax\} \). The inverse of \( A \) is \( A^{-1} : \mathcal{H} \rightarrow 2^\mathcal{H} : u \mapsto \{x \in \mathcal{H} \mid u \in Ax\} \), and the resolvent of \( A \) is \( J_A = (\text{Id} + A)^{-1} \). The operator \( A \) is monotone if

\[
(\forall (x, u) \in \text{gra} \ A)(\forall (y, v) \in \text{gra} \ A) \quad \langle x - y \mid u - v \rangle \geq 0. \tag{2.3}
\]

Moreover, \( A \) is maximally monotone if it is monotone and there exists no monotone operator \( B : \mathcal{H} \rightarrow 2^\mathcal{H} \) such that \( \text{gra} \ B \) properly contains \( \text{gra} \ A \), i.e., for every \( (x, u) \in \mathcal{H} \times \mathcal{H} \),

\[ \square \]
\((x, u) \in \text{gra } A \iff (\forall (y, v) \in \text{gra } A) \langle x - y | u - v \rangle \geq 0. \) \tag{2.4}

A is locally bounded at \(x \in \mathcal{H}\), if there exists \(\delta \in ]0, +\infty[\) such that \(A(B(x; \delta))\) is bounded, where \(B(x; \delta)\) stands for the ball centered at \(x\) with radius \(\delta\). Moreover, \(A\) is locally bounded in \(\emptyset \neq D \subset \mathcal{H}\) if, for every \(x \in D, A\) is locally bounded at \(x\).

We denote by \(\Gamma_0(\mathcal{H})\) the class of proper lower-semicontinuous convex functions \(f : \mathcal{H} \rightarrow ]-\infty, +\infty[\). Let \(f \in \Gamma_0(\mathcal{H})\). The Fenchel conjugate of \(f\) is defined by \(f^* : u \mapsto \sup_{x \in \mathcal{H}} (\langle x | u \rangle - f(x))\), which is a function in \(\Gamma_0(\mathcal{H})\). The subdifferential of \(f\) is the maximally monotone operator \(\partial f : \mathcal{H} \rightarrow 2^{\mathcal{H}} : x \mapsto \{ u \in \mathcal{H} \, | \, (\forall y \in \mathcal{H}) \, f(x) + \langle y - x | u \rangle \leq f(y) \}\).

It turns out that \((\partial f)^{-1} = \partial f^*\) and that \(\text{zer } \partial f\) is the set of minimizers of \(f\), which is denoted by \(\arg \min_{x \in \mathcal{H}} f(x)\). We denote the proximity operator of \(f\) by

\[
\text{prox}_f : x \mapsto \arg \min_{y \in \mathcal{H}} \left( f(y) + \frac{1}{2} \| x - y \|^2 \right). \tag{2.5}
\]

We have \(\text{prox}_f = J_{\partial f}\). Moreover, it follows from [2, Theorem 14.3] that

\[
(\forall \gamma > 0) \quad \text{prox}_{\gamma f} + \gamma \text{prox}_{f^* / \gamma} \circ (\text{Id} / \gamma) = \text{Id}. \tag{2.6}
\]

We denote by \(\text{lev}_{< 0} f = \{ x \in \mathcal{H} | f(x) < 0 \}\) and by \(\text{lev}_{\leq 0} f = \{ x \in \mathcal{H} | f(x) \leq 0 \}\). Given a nonempty closed convex set \(C \subset \mathcal{H}\), we denote by \(P_C\) the projection onto \(C\) and by \(\iota_C \in \Gamma_0(\mathcal{H})\) the indicator function of \(C\), which takes the value 0 in \(C\) and \(+\infty\) otherwise. For further properties of monotone operators, nonexpansive mappings, and convex analysis, the reader is referred to [2].

### 3 Main Results

We aim at solving the following monotone inclusion problem.

**Problem 3.1** Let \(X\) be a nonempty closed convex subset of a real Hilbert space \(\mathcal{H}\), let \(A : \mathcal{H} \rightarrow 2^{\mathcal{H}}\) be a maximally monotone operator, let \(B_1 : \mathcal{H} \rightarrow \mathcal{H}\) be a \(\beta\)-cocoercive operator for some \(\beta > 0\), let \(B_2 : \mathcal{H} \rightarrow \mathcal{H}\) be a monotone and \(L\)-Lipschitzian operator for some \(L > 0\), and let \(B_3 : \mathcal{H} \rightarrow 2^{\mathcal{H}}\) be a maximally monotone operator such that \(B_3\) is single-valued and continuous in \(\text{dom } A \cup X \subset \text{dom } B_3\). Moreover, assume that \(A + B_3\) is maximally monotone. The problem is to

\[
\text{find } x \in X \text{ such that } 0 \in Ax + B_1 x + B_2 x + B_3 x, \tag{3.1}
\]

under the assumption that the set of solutions to (3.1) is nonempty.

We first study some properties of the monotone operators involved in Problem 3.1, which ensure the finite termination of the backtracking procedure in our method.
Lemma 3.2  In the context of Problem 3.1, let $z$ and $y$ in $\mathcal{H}$, and define

\[
(\forall \gamma > 0) \quad x_{z,y}(\gamma) = J_A(z - \gamma y) \quad \text{and} \quad \varphi_{z,y}(\gamma) = \frac{\|z - x_{z,y}(\gamma)\|}{\gamma}.
\]  

(3.2)

Then, the following statements hold:

1. $\varphi_{z,y}$ is nonincreasing.
2. $(\forall z \in \text{dom } A) \quad \lim_{\gamma \downarrow 0} \varphi_{z,y}(\gamma) = \|(A + y)^0 z\| = \min_{w \in Az + y} \|w\|.$
3. Set
   
   \[
   C = \{z \in \mathcal{H} \mid \lim_{\gamma \downarrow 0} \varphi_{z,y}(\gamma) < +\infty\}.
   \]  

(3.3)

Then, $\text{dom } A \subset C \subset \overline{\text{dom } A}$.
4. Suppose that one of the following holds:
   a. $z \in C$.
   b. $z \in \text{dom } B_3 \setminus C$, $y = (B_1 + B_2 + B_3)z$, and $B_3$ is locally bounded at $P_{\text{dom } A} z$.
   c. $z \in \text{dom } B_3 \setminus C$, $y = (B_1 + B_2 + B_3)z$, and $\text{dom } A \subset \text{int } \text{dom } B_3$.

Then, for every $\theta \in ]0, 1[$, there exists $\gamma(z) > 0$ such that, for every $\gamma \in ]0, \gamma(z)]$,

\[
\gamma\|B_3z - B_3x_{z,y}(\gamma)\| \leq \theta\|z - x_{z,y}(\gamma)\|.
\]  

(3.4)

Proof  Let $z \in \mathcal{H}$. Note that if $z \in \text{zer } (A + y)$,

\[
(\forall \gamma > 0) \quad 0 \in Az + y \iff z - \gamma y \in \gamma Az + z \\
\iff z = x_{z,y}(\gamma) \\
\iff \varphi_{z,y}(\gamma) = 0.
\]  

(3.5)

In this case, 1, 2, and 4 are clear. Henceforth, assume $z \in \mathcal{H} \setminus \text{zer } (A + y)$. It follows from (3.2) that

\[
\frac{z - x_{z,y}(\gamma)}{\gamma} - y \in Ax_{z,y}(\gamma).
\]  

(3.6)

1: For every $\gamma_1$ and $\gamma_2$ in $]0, +\infty[$, (3.6) and the monotonicity of $A$ yield

\[
0 \leq \left( \frac{z - x_{z,y}(\gamma_1)}{\gamma_1} - \frac{z - x_{z,y}(\gamma_2)}{\gamma_2} \right) \left( x_{z,y}(\gamma_1) - x_{z,y}(\gamma_2) \right)
\]

\[
= \left( \frac{z - x_{z,y}(\gamma_1)}{\gamma_1} - \frac{z - x_{z,y}(\gamma_2)}{\gamma_2} \right) \left( x_{z,y}(\gamma_1) - z - (x_{z,y}(\gamma_2) - z) \right)
\]

\[
= - \frac{1}{\gamma_1} \|z - x_{z,y}(\gamma_1)\|^2 + \left( \frac{1}{\gamma_1} + \frac{1}{\gamma_2} \right) (z - x_{z,y}(\gamma_1) \mid z - x_{z,y}(\gamma_2))
\]

\[
- \frac{1}{\gamma_2} \|z - x_{z,y}(\gamma_2)\|^2.
\]

Hence, we obtain
\[\gamma_1 \varphi_{z,y}(\gamma_1)^2 + \gamma_2 \varphi_{z,y}(\gamma_2)^2 \leq (\gamma_1 + \gamma_2) \left( \frac{z - x_{z,y}(\gamma_1)}{\gamma_1} \left| \frac{z - x_{z,y}(\gamma_2)}{\gamma_2} \right| \right) \]
\[\leq \frac{(\gamma_1 + \gamma_2)}{2} (\varphi_{z,y}(\gamma_1)^2 + \varphi_{z,y}(\gamma_2)^2),\]

which yields \((\gamma_1 - \gamma_2)(\varphi_{z,y}(\gamma_1)^2 - \varphi_{z,y}(\gamma_2)^2) \leq 0\) and 1 follows.

2: It follows from the monotonicity of \(A\) and (3.6) that, for every \(w \in Az + y\) and \(\gamma \in [0, +\infty[\),
\[0 \leq \left( \frac{z - x_{z,y}(\gamma)}{\gamma} \right) - w \left| x_{z,y}(\gamma) - z \right|,
\]
which yields
\[\frac{1}{\gamma} \|z - x_{z,y}(\gamma)\|^2 \leq (w \mid z - x_{z,y}(\gamma)) \]
\[\leq \|w\| \|z - x_{z,y}(\gamma)\|. \tag{3.7}\]

Thus, \(\varphi_{z,y}(\gamma) \leq \|w\|\). Therefore, since [2, Proposition 20.36] implies that, for every \(z \in \text{dom } A\), \(Az + y\) is nonempty, closed, and convex, [2, Theorem 11.10] yields
\[(\forall \gamma \in [0, +\infty[\) \(\varphi_{z,y}(\gamma) \leq \min_{w \in Az + y} \|w\|. \tag{3.8}\]

Hence, since \(\varphi_{z,y} \geq 0, 1\) implies that \(\lim_{\gamma \downarrow 0} \varphi_{z,y}(\gamma)\) exists. In turn, since \(z \in \mathcal{H} \setminus \text{zer } (A + y)\), it follows from (3.8) and (3.5) that
\[0 < \varphi_{z,y}(1) \leq \lim_{\gamma \downarrow 0} \varphi_{z,y}(\gamma) \leq \min_{w \in Az + y} \|w\|, \tag{3.9}\]

which, in view of (3.2), implies
\[\lim_{\gamma \downarrow 0} x_{z,y}(\gamma) = z. \tag{3.10}\]

Since \((\varphi_{z,y}(\gamma))(\gamma > 0)\) is bounded, by [2, Lemma 2.45], there exists a sequence \((\gamma_k)_{k \in \mathbb{N}} \subseteq ]0, +\infty[\) and \(\overline{w} \in \mathcal{H}\) such that \(\gamma_k \downarrow 0\) and \(\frac{z - x_{z,y}(\gamma_k)}{\gamma_k} \rightarrow \overline{w}\) as \(k \rightarrow +\infty\). Therefore, (3.6), (3.10), and [2, Proposition 20.38(i)] imply \(\overline{w} \in Az + y\). Hence, noting that
\[\frac{(\varphi_{z,y}(\gamma_k))^2}{\gamma_k^2} = \left( \frac{z - x_{z,y}(\gamma_k)}{\gamma_k} - \frac{w}{\overline{w}} \right)^2 + \|w\|^2 + 2 \left( \frac{z - x_{z,y}(\gamma_k)}{\gamma_k} - \frac{w}{\overline{w}} \right) \frac{\left| z - x_{z,y}(\gamma_k) \right|}{\gamma_k} - \frac{w}{\overline{w}} \right), \tag{3.11}\]

we deduce
\[
\lim_{\gamma \downarrow 0} (\varphi_{z,y}(\gamma))^2 = \lim_{k \to +\infty} (\varphi_{z,y}(\gamma_k))^2 \geq \|w\|^2 \geq \min_{w \in Az+y} \|w\|^2.
\]

Therefore, we obtain from (3.9) that
\[
\lim_{\gamma \downarrow 0} \varphi_{z,y}(\gamma) = \min_{w \in Az+y} \|w\|,
\]
and 2 follows.

3: It follows from (3.3) and 2 that \( \text{dom } A \subseteq C \). Let \( z \in \mathcal{H} \setminus \text{dom } A \) and let \( \gamma > 0 \). The firm nonexpansiveness of \( J_{\gamma}A \) [2, Proposition 23.8(ii)] implies
\[
\|x_{z,y}(\gamma) - P_{\text{dom } A}z\| = \|J_{\gamma}A(z - \gamma y) - J_{\gamma}A\gamma z - P_{\text{dom } A}z\| \\
\leq \|J_{\gamma}A(z - \gamma y) - J_{\gamma}A\gamma z\| + \|J_{\gamma}A\gamma z - P_{\text{dom } A}z\| \\
\leq \gamma \|y\| + \|J_{\gamma}A\gamma z - P_{\text{dom } A}z\|.
\]

Hence, by taking \( \gamma \downarrow 0 \) in (3.13) we conclude from [2, Theorem 23.48] that \( x_{z,y}(\gamma) \to P_{\text{dom } A}z \) as \( \gamma \downarrow 0 \). Then, by the continuity of the norm and \( z \notin \text{dom } A \), we deduce
\[
\lim_{\gamma \downarrow 0} \|z - x_{z,y}(\gamma)\| = \|z - P_{\text{dom } A}z\| > 0.
\]

Therefore, \( \varphi_{z,y}(\gamma) = \|z - x_{z,y}(\gamma)\|/\gamma \to +\infty \) as \( \gamma \downarrow 0 \) and, hence, it follows from (3.3) that \( z \in \mathcal{H} \setminus C \).

4a: If \( z \in C \), it follows from 1 that
\[
0 < \varphi_{z,y}(1) \leq \lim_{\gamma \downarrow 0} \varphi_{z,y}(\gamma) < +\infty.
\]

Therefore, \( \lim_{\gamma \downarrow 0} x_{z,y} = z \) and the continuity of \( B_3 \) implies
\[
\lim_{\gamma \downarrow 0} B_3x_{z,y} = B_3z.
\]

The result follows from (3.14) and (3.15).

4b: Set \( B = B_1 + B_2 + B_3 \) and let \( p = P_{\text{dom } A}z \). Since \( B_3 \) is locally bounded at \( p \), there exists \( \delta_p \in ]0, +\infty[ \) such that \( B_3(B(p; \delta_p)) \) is bounded. Now, since \( y = Bz \) and
\[
\frac{z - J_{\gamma}A\gamma z}{\gamma} \in AJ_{\gamma}A\gamma z,
\]

it follows from (3.6) and the monotonicity of \( A \) that
\[
0 \leq \left( \frac{z - x_{z,y}(\gamma)}{\gamma} - Bz - \frac{z - J_{\gamma}A\gamma z}{\gamma} \right) x_{z,y}(\gamma) - J_{\gamma}A\gamma z \\
= -\frac{1}{\gamma} \|x_{z,y}(\gamma) - J_{\gamma}A\gamma z\|^2 + \langle Bz \mid J_{\gamma}A\gamma z - x_{z,y}(\gamma) \rangle
\]
\[
\leq - \frac{1}{\gamma} \|x_{z,y}(\gamma) - J_{Y_A}z\|^2 + \|Bz\| \|J_{Y_A}z - x_{z,y}(\gamma)\|.
\]

Hence, we obtain
\[
\|x_{z,y}(\gamma) - J_{Y_A}z\| \leq \gamma \|Bz\|. \tag{3.17}
\]

Additionally, by [2, Theorem 23.48], there exists \(\gamma_1\) such that, for every \(\gamma < \gamma_1\),
\[
\|J_{Y_A}z - p\| \leq \delta_p/2.
\]

By defining
\[
\gamma := \begin{cases} 
\gamma_1, & \text{if } Bz = 0; \\
\min\{\delta_p/(2\|Bz\|), \gamma_1\}, & \text{if } Bz \neq 0,
\end{cases} \tag{3.18}
\]

it follows from (3.17) that, for every \(\gamma < \gamma\),
\[
\|x_{z,y}(\gamma) - p\| \leq \|x_{z,y}(\gamma) - J_{Y_A}z\| + \|J_{Y_A}z - p\| \\
\leq \gamma \|Bz\| + \frac{\delta_p}{2} < \delta_p,
\]

which yields \((x_{z,y}(\gamma))_{0<\gamma<\gamma} \subset B(p, \delta_p)\) and, thus, \((\|Bz - B_3x_{z,y}(\gamma)\|)_{0<\gamma<\gamma}\) is bounded. Therefore, since \(z \in H \setminus C\) implies \(\varphi_{z,y}(\gamma) \to +\infty\) as \(\gamma \downarrow 0\), the result follows.

4c: Since \(p = P_{\text{dom} A}z \notin \text{bdry dom } B_3\), \(B_3\) is locally bounded at \(p\) [2, Theorem 21.18] and the result follows from 4b. \(\square\)

**Remark 3.1**

1. In the case \(B_2 = 0\), by setting \(y = (B_1 + B_3)z\) in Lemma 3.2(1&2), we recover [12, Lemma 2.2(1)].

2. Realizing that [12, Lemma 2.2(2)] is valid for every \(z \in \text{dom } A\), it is a particular case of Lemma 3.2(3&4a).

Now, we state our main result.

**Theorem 3.3** *In the context of Problem 3.1, suppose that one of the following holds:*

1. \(X \subset \text{dom } A\).
2. \(\overline{\text{dom } A} \subset \text{dom } B_3\) and \(B_3\) is locally bounded in \(\text{dom } B_3\).
3. \(\overline{\text{dom } A} \subset \text{int } \text{dom } B_3\).

*Let \(\varepsilon \in ]0, 1[, \text{ set } \rho = \min\{2\beta \varepsilon, \sqrt{1 - \varepsilon}/L\}, \text{ let } \sigma \in ]0, 1[, \text{ let } \theta \in ]0, \sqrt{1 - \varepsilon} - L\rho\sigma[, \text{ let } z_0 \in \text{dom } B_3, \text{ and consider the sequence } (z_n)_{n \in \mathbb{N}} \text{ defined by the recurrence}

\[
\begin{array}{l}
(\forall n \in \mathbb{N}) \quad x_n = J_{Y_A}(z_n - \gamma_n(B_1 + B_2 + B_3)z_n) \\
z_{n+1} = P_X(x_n + \gamma_n(B_2 + B_3)z_n - \gamma_n(B_2 + B_3)x_n),
\end{array} \tag{3.19}
\]

where, for every \(n \in \mathbb{N}\), \(\gamma_n\) is the largest \(\gamma \in \{\rho \sigma, \rho \sigma^2, \rho \sigma^3, \ldots\}\) satisfying

\[
\gamma \|B_3z_n - B_3J_{Y_A}(z_n - \gamma(B_1 + B_2 + B_3)z_n)\| \\
\leq \theta \|z_n - J_{Y_A}(z_n - \gamma(B_1 + B_2 + B_3)z_n)\|. \tag{3.20}
\]

Moreover, assume that at least one of the following additional statements holds:

\[\square\] Springer
(i) \( \liminf_{n \to \infty} \gamma_n = \delta > 0 \).

(ii) \( B_3 \) is uniformly continuous in any weakly compact subset of \( \overline{\text{conv}}(\text{dom } A \cup X) \).

Then, \( (z_n)_{n \in \mathbb{N}} \) converges weakly to a solution to Problem 3.1.

**Proof** Set \( B = B_1 + B_2 + B_3 \) and fix \( n \in \mathbb{N} \). If \( z_n \in C \), where \( C \) is defined in (3.3), then \( \gamma_n \) is well defined in view of Lemma 3.2(4a). In particular, if 1 holds, \( \gamma_n \) is well defined in view of Lemma 3.2(3). Now, suppose that \( z_n \in \mathcal{H} \setminus C \). If \( n = 0 \), it is clear that \( z_0 \in \text{dom } B_3 \setminus C \). Otherwise, since \( X \subset \text{dom } B_3 \), we have \( z_n \in \text{dom } B_3 \setminus C \).

Now, if we assume 2, then \( B_3 \) is locally bounded in \( P_{\text{dom } A} z_n \in \text{dom } A \) and \( \gamma_n \) is well defined from Lemma 3.2(4b). Similarly, if we assume 3, \( \gamma_n \) is well defined from Lemma 3.2(4c).

Now, let \( z^* \in \text{zer}(A + B) \cap X \). Note that the maximal monotonicity of \( A + B_3 \), the full domain of \( B_1 \) and \( B_2 \), and [2, Corollary 25.5(i)] imply that \( A + B_2 + B_3 \) and \( A + B \) are maximally monotone. Then, since \( B_2 + B_3 \) is continuous and single-valued in \( \text{dom } (B_2 + B_3) = \text{dom } B_3 \supset \text{dom } A \cup X \) and \( B_1 \) is \( \beta \)-cocoercive, it follows from [12, Proposition 2.1(1)&(2)] that, for every \( n \in \mathbb{N} \), we have

\[
\begin{align*}
\|z_{n+1} - z^*\|^2 &\leq \|z_n - z^*\|^2 - (1 - \epsilon)\|z_n - x_n\|^2 + \gamma_n^2 \|(B_2 + B_3)z_n - (B_2 + B_3)x_n\|^2 \\
&\quad - \frac{\gamma_n}{\epsilon} (2\beta \epsilon - \gamma_n)\|B_1 z_n - B_1 z^*\|^2.
\end{align*}
\] (3.21)

Note that the Lipschitz property of \( B_2 \), (3.20) and (2.30) yield

\[
\begin{align*}
\gamma_n^2 \|(B_2 + B_3)z_n - (B_2 + B_3)x_n\|^2 &\leq (L \gamma_n \|z_n - x_n\| + \gamma_n \|B_3z_n - B_3x_n\|)^2 \\
&\leq (L \gamma_n + \theta)^2 \|z_n - x_n\|^2 \\
&\leq (L \rho \sigma + \theta)^2 \|z_n - x_n\|^2.
\end{align*}
\] (3.22)

Hence, it follows from (3.21) and (3.22) that

\[
(\forall n \in \mathbb{N}) \quad \|z_{n+1} - z^*\|^2 \leq \|z_n - z^*\|^2 - ((1 - \epsilon) - (L \rho \sigma + \theta)^2) \|z_n - x_n\|^2.
\]

Therefore, since \((1 - \epsilon - (L \rho \sigma + \theta)^2) > 0, [15, Lemma 3.1(i)]\) implies that \( (\|z_n - z^*\|)_{n \in \mathbb{N}} \) is a convergent sequence and

\[
z_n - x_n \to 0.
\] (3.23)

Let \( z \in \mathcal{H} \) be a weak limit point of the subsequence \( (z_n)_{n \in K} \) for some \( K \subset \mathbb{N} \). Then, \( z \) is also a weak limit point of \( (x_n)_{n \in K} \) in view of (3.23). Since \( X \) is closed and convex, and \( (z_n)_{n \in K} \) is a sequence in \( X \), we conclude that \( z \in X \). Let us prove that \( z \in \text{zer}(A + B) \).

(i) Assume that \( \liminf_{n \to +\infty} \gamma_n = \delta > 0 \). Then, there exists \( n_0 \in \mathbb{N} \) such that \( \inf_{n \geq n_0} \gamma_n \geq \delta \). Hence, (3.20), (3.19), the Lipschitz continuity of \( B_2 \), and the cocoercivity of \( B_1 \) yield

\[
(\forall n \geq n_0) \quad \|Bz_n - Bx_n\| \leq \|B_1 z_n - B_1 x_n\| + \|B_2 z_n - B_2 x_n\| + \|B_3 z_n - B_3 x_n\|.
\]
\[ \leq \left( \frac{1}{\beta} + L + \frac{\theta}{\delta} \right) \| z_n - x_n \|, \quad (3.24) \]

which implies \( Bz_n - Bx_n \to 0 \) in view of (3.23). Hence, it follows from (3.19) that

\[ (\forall n \in \mathbb{N}) \quad u_n := \frac{z_n - x_n}{\gamma_n} - Bz_n + Bx_n \in (A + B)x_n, \quad (3.25) \]

and (3.23) and \( \lim \inf_{n \to +\infty} \gamma_n = \delta > 0 \) imply that \( u_n \to 0 \). Therefore, since \( x_n \to z, \ n \in K \), the weak–strong closure of the graph of the maximally monotone operator \( A + B \) and (3.25) yield \( z \in \text{zer} (A + B) \). The convergence follows from [2, Lemma 2.47].

(ii) Without loss of generality, suppose that \( \lim_{n \to \infty, n \in K} \gamma_n = 0 \). Our choice of \( \gamma_n \) guarantee that, for every \( n \in K \), we have

\[ \tilde{\gamma}_n \| B_3 z_n - B_3 J_{\tilde{\gamma}_n} A (z_n - \tilde{\gamma}_n Bz_n) \| > \theta \| z_n - J_{\tilde{\gamma}_n} A (z_n - \tilde{\gamma}_n Bz_n) \|, \quad (3.26) \]

where \( \tilde{\gamma}_n := \frac{\gamma_n}{\alpha} > \gamma_n \). Now, by the nonincreasing property of \( \gamma \mapsto \frac{1}{\gamma} \| z - J_{\gamma} A (z - \gamma Bz) \| \) provided by Lemma 3.2(1) with \( y = Bz \), we have

\[ \frac{1}{\tilde{\gamma}_n} \| z_n - J_{\tilde{\gamma}_n} A (z_n - \tilde{\gamma}_n Bz_n) \| \leq \frac{1}{\gamma_n} \| z_n - J_{\gamma_n} A (z_n - \gamma_n Bz_n) \|, \quad (3.27) \]

which is equivalent to

\[ \| z_n - J_{\tilde{\gamma}_n} A (z_n - \tilde{\gamma}_n Bz_n) \| \leq \frac{1}{\sigma} \| z_n - x_n \|. \quad (3.28) \]

Thus, by defining

\[ (\forall n \in K) \quad \tilde{x}_n = J_{\tilde{\gamma}_n} A (z_n - \tilde{\gamma}_n Bz_n), \quad (3.29) \]

(3.23) implies that

\[ z_n - \tilde{x}_n \to 0 \quad \text{as} \quad n \to \infty, \ n \in K. \quad (3.30) \]

Therefore, since \( z_n \to z, \ n \in K \), we have \( \tilde{x}_n \to z \) as \( n \to \infty, \ n \in K \). Furthermore, (3.29) yields

\[ \frac{z_n - \tilde{x}_n}{\tilde{\gamma}_n} + B\tilde{x}_n - Bz_n \in (A + B)\tilde{x}_n. \quad (3.31) \]

Since \( \{z\} \cup \bigcup_{n \in K} [\tilde{x}_n, z_n] \) is a weakly compact subset of \( \overline{\text{conv}} (\text{dom} A \cup X) \) [35, Lemma 3.2], it follows from the uniform continuity of \( B_3 \) and (3.30) that

\[ B_3 \tilde{x}_n - B_3 z_n \to 0 \quad \text{as} \quad n \to \infty, \ n \in K, \quad (3.32) \]

which, combined with (3.26), yields \( (z_n - \tilde{x}_n)/\tilde{\gamma}_n \to 0 \) as \( n \to \infty, \ n \in K \). Moreover, the Lipschitz continuity of \( B_1 + B_2 \), (3.30), and (3.32) imply

\[ B\tilde{x}_n - Bz_n \to 0 \quad \text{as} \quad n \to \infty, \ n \in K. \quad (3.33) \]
Altogether, the convergence follows, as in the case (i), from (3.31), the weak–strong closedness of the graph of the maximally monotone operator $A + B$, and [2, Lemma 2.47].

**Remark 3.2**

1. In Theorem 3.3, if $B_3 = 0$, we have $\text{dom } B_3 = \mathcal{H}$ and, for all $n \in \mathbb{N}$, $\gamma_n = \sigma \rho = \sigma \min\{2\beta \varepsilon, \sqrt{1 - \varepsilon/L}\}$. Since in this case $(\gamma_n)_{n \in \mathbb{N}}$ is constant, the largest step size is obtained by taking $\varepsilon = \varepsilon(L, \beta) = 2/(1 + \sqrt{1 + 16\beta^2L^2})$, which satisfies $2\beta \varepsilon = \sqrt{1 - \varepsilon/L} = \chi(L, \beta)$, where

$$\chi(L, \beta) = \frac{4\beta}{1 + \sqrt{1 + 16\beta^2L^2}}, \quad (3.34)$$

and $\gamma_n \equiv \gamma = \sigma \chi(L, \beta) \in [0, \chi(L, \beta)]$. Hence, we recover the result in [12, Theorem 2.3(1)] for constant step sizes. Additionally, if $B_2 = 0$ and $X = \mathcal{H}$, we have $\varepsilon(L, \beta) \to 1$ and $\chi(L, \beta) \to 2\beta$ as $L \to 0$ and $\gamma_n \equiv 2\beta \sigma \in [0, 2\beta]$, recovering the forward–backward algorithm [29]. On the other hand, if $B_1 = 0$, we have $\chi(L, \beta) \to 1/L$ as $\beta \to \infty$ and $\gamma_n \equiv \sigma/L \in [0, 1/L]$, recovering the result in [38] for constant step sizes.

2. Suppose that $B_2 = 0$ and $X \subset \text{dom } A$. Then by taking $L \to 0$, we have $\rho \to 2\beta \varepsilon$ and $\theta \in [0, \sqrt{1 - \varepsilon}]$. Hence, Theorem 3.3 recovers [12, Theorem 2.3(2)] noting that the uniform continuity in weakly compact subsets of $\text{conv}(\text{dom } A \cup X) = \text{dom } A$ is needed. We hence generalize [12, Theorem 2.3(1)&(2)] to the case when $X \not\subset \text{dom } A$.

### 4 Application to Convex Optimization with Nonlinear Constraints

In this section, we consider the following optimization problem.

**Problem 4.1** Let $f \in \Gamma_0(\mathcal{H})$, let $g \in \Gamma_0(\mathcal{G})$, let $h : \mathcal{H} \to \mathbb{R}$ be a convex Gâteaux differentiable function such that $\nabla h$ is $\beta^{-1}$-Lipschitzian for some $\beta \in [0, +\infty[$, let $M : \mathcal{H} \to \mathcal{G}$ be a bounded linear operator, and let $e : \mathcal{H} \to ]-\infty, +\infty[^p : x \mapsto (e_i(x))_{1 \leq i \leq p}$ be such that, for every $i \in \{1, \ldots, p\}$, $e_i$ is convex and Gâteaux differentiable in $\text{int } \text{dom } e_i$, $\text{dom } e_i$ is closed, $\cap_{i=1}^p \text{int } \text{dom } e_i \neq \emptyset$, and $\text{dom } f \subset \cap_{i=1}^p \text{int } \text{dom } e_i$. Assume that $0 \in \text{sri } (\text{dom } g - M(\text{dom } f))$ and that

$$\begin{align*}
(\forall i \in \{1, \ldots, p\}) & \quad \text{lev}_{\leq 0} e_i \subset \text{int } \text{dom } e_i; \\
\text{dom } (f + g \circ M) \cap \cap_{i=1}^p \text{lev}_{\leq 0} e_i & \neq \emptyset.
\end{align*} \quad (4.1)
$$

The problem is to

$$\min_{e(x) \in [-\infty, 0]^p} f(x) + g(Mx) + h(x), \quad (4.2)$$

and we assume that solutions exist.
It follows from (4.1) and [2, Proposition 27.21] that \( \hat{x} \in \mathcal{H} \) is a solution to Problem 4.1 if and only if there exists \( \hat{v} \in [0, +\infty]^p \) such that

\[
- \sum_{i=1}^p \hat{v}_i \nabla e_i(\hat{x}) \in \partial (f \circ g \circ M + h) \quad \text{and} \quad (\forall i \in \{1, \ldots, p\}) \begin{cases} e_i(\hat{x}) \leq 0, \\
\hat{v}_i e_i(\hat{x}) = 0. \end{cases}
\]

(4.3)

Hence, by [2, Example 16.13 & Example 6.42(i)], we deduce that

\[
\begin{cases}
0 \in \partial (f \circ g \circ M + h)(\hat{x}) + \sum_{i=1}^p \hat{v}_i \nabla e_i(\hat{x}) \\
0 \in N_{[0, +\infty]^p}(\hat{v}) - e(\hat{x}).
\end{cases}
\]

(4.4)

Then, we deduce from [2, Theorem 16.47] that there exists \( \hat{u} \in \mathcal{G} \) such that (4.4) reduces to

\[
\begin{cases}
0 \in \partial f(\hat{x}) + M^* \hat{u} + \nabla h(\hat{x}) + \sum_{i=1}^p \hat{v}_i \nabla e_i(\hat{x}) \\
0 \in \partial g^*(\hat{u}) - M \hat{x} \\
0 \in N_{[0, +\infty]^p}(\hat{v}) - e(\hat{x}),
\end{cases}
\]

which is equivalent to

\[
\begin{pmatrix}
0 \\
0 \\
0
\end{pmatrix} \in \begin{pmatrix}
\partial f(\hat{x}) & \nabla h(\hat{x}) & M^* \hat{u} \\
\partial g^*(\hat{u}) & 0 & -M \hat{x} \\
N_{[0, +\infty]^p}(\hat{v}) & 0 & -e(\hat{x})
\end{pmatrix}.
\]

(4.5)

**Proposition 4.2** *In the context of Problem 4.1, let \( X = X_1 \times X_2 \times X_3 \subset dom \partial f \times dom \partial g^* \times [0, +\infty]^p \) be nonempty, closed, and convex, and define the operator

\[
B_3 : \mathcal{H} \times \mathcal{G} \times \mathbb{R} \to 2^{\mathcal{H} \times \mathcal{G} \times \mathbb{R}} \]

\[
(x, u, v) \mapsto \begin{cases}
\left\{ \left( \sum_{i=1}^p v_i \nabla e_i(x), 0, -e(x) \right) \right\}, & \text{if } v \in [0, +\infty]^p \text{ and } x \in \bigcap_{i=1}^p \text{int} dom e_i; \\
\emptyset, & \text{otherwise.}
\end{cases}
\]

(4.6)

Then, the following hold:

1. \( B_3 \) is maximally monotone.
2. Suppose that one of the following holds:
   
   (a) \( (\nabla e_i)_{1 \leq i \leq p} \) are bounded and uniformly continuous in every weakly compact subset of \( dom \partial f \).
   
   (b) \( \mathcal{H} \) is finite-dimensional, and \( (\nabla e_i)_{1 \leq i \leq p} \) are continuous in every compact subset of \( dom \partial f \).

Then, \( B_3 \) is uniformly continuous in every compact subset of \( \overline{dom \partial f} \times \overline{dom \partial g^*} \times [0, +\infty]^p \).

\[ \square \] Springer
Proof 1: Consider the saddle function

\[ \ell: \mathcal{H} \times \mathcal{G} \times \mathbb{R}^p \to [-\infty, +\infty] \]

\[ (x, u, v) \mapsto \begin{cases} 
  e(x) \cdot v, & \text{if } v \in [0, +\infty[^p \text{ and } x \in \bigcap_{i=1}^p \text{dom } e_i; \\
  +\infty, & \text{if } v \in [0, +\infty[^p \text{ and } x \not\in \bigcap_{i=1}^p \text{dom } e_i; \\
  -\infty, & \text{if } v \not\in [0, +\infty[^p.
\end{cases} \tag{4.7} \]

Note that if \( v \in [0, +\infty[^p, \)

\[ \ell:(x, u, v) \mapsto \begin{cases} 
  e(x) \cdot v, & \text{if } x \in \bigcap_{i=1}^p \text{dom } e_i; \\
  +\infty, & \text{otherwise}
\end{cases} \]

and if \( v \not\in [0, +\infty[^p, \ell(\cdot, \cdot, v) \equiv -\infty. \) Hence, for every \( v \in \mathbb{R}^p, \ell(\cdot, \cdot, v) \) is lower-semicontinuous. Additionally, if \( (x, u) \in \bigcap_{i=1}^p \text{dom } e_i \times \mathcal{G}, \) we have

\[ -\ell:(x, u, v) \mapsto \begin{cases} 
  -e(x) \cdot v, & \text{if } v \in [0, +\infty[^p; \\
  +\infty, & \text{otherwise}
\end{cases} \tag{4.8} \]

and if \( x \not\in \bigcap_{i=1}^p \text{dom } e_i \) and \( u \in \mathcal{G}, \)

\[ -\ell:(x, u, v) \mapsto \begin{cases} 
  -\infty, & \text{if } v \in [0, +\infty[^p; \\
  +\infty, & \text{otherwise}. \tag{4.9} \end{cases} \]

Therefore, for every \( (x, u) \in \mathcal{H} \times \mathcal{G}, \) \(-\ell(x, u, \cdot)\) is lower-semicontinuous. Furthermore,

\[ (\forall (x, u) \in \mathcal{H} \times \mathcal{G}) (\forall v \in \mathbb{R}^p) \quad B_3(x, u, v) = \partial \ell(\cdot, \cdot, v)(x, u) \times \partial (-\ell(x, u, \cdot))(v). \tag{4.10} \]

The result follows from [33, Corollary 1].

2: First, assume 2a. Let \( Y = Y_1 \times Y_2 \times Y_3 \subset \text{dom } \partial f \times \text{dom } \partial g^* \times [0, +\infty[^p \) be a weakly compact set. Let \( x = (x_1, u_1, v_1) \) and \( y = (x_2, u_2, v_2) \) in \( Y, \) fix \( i \in \{1, \ldots, p\}, \) define \( \rho_i: [0, 1] \to \mathbb{R} : t \mapsto e_i(x_1 + t(x_2 - x_1)), \) which is differentiable in \( ]0, 1[. \) Since \( Y_1 \) is weakly compact, by [2, Theorem 3.37], \( \text{conv } Y_1 \) is also weakly compact. Moreover, we deduce from the boundedness of \( \nabla e_i \) in \( \text{conv } Y_1 \subset \text{dom } \partial f \) that there exists \( K_i > 0 \) such that \( \sup_{x \in \text{conv } Y_1} ||\nabla e_i(x)|| \leq K_i. \) Therefore, since \( \rho_i': t \mapsto (\nabla e_i(x_1 + t(x_2 - x_1)) \mid x_2 - x_1), \) we obtain

\[ |e_i(x_2) - e_i(x_1)| = |\rho_i(1) - \rho_i(0)| \]

\[ = \left| \int_0^1 \rho_i'(t) dt \right| \]

\[ = \left| \int_0^1 \langle \nabla e_i(x_1 + t(x_2 - x_1)) \mid x_2 - x_1 \rangle dt \right| \]
\[ \leq \int_{0}^{1} \|\nabla e_i(x_1 + t(x_2 - x_1))\| x_2 - x_1 \| dt \]
\[ \leq K_i \| x_2 - x_1 \|. \]

Thus, we conclude \( |e_i(x_1) - e_i(x_2)| \leq K_i \| x_2 - x_1 \| \) and therefore
\[ \| e(x_2) - e(x_1)\| \leq K \| x_2 - x_1 \|, \] (4.11)

where \( K = \max_{i \in \{1, \ldots, p\}} K_i \). Since \( Y_3 \) is weakly compact, there exists \( V > 0 \) such that \( \sup_{v \in Y_3} \| v \| \leq V \) [2, Lemma 2.36]. Let \( \varepsilon > 0 \). The uniform continuity of \( (\nabla e_i)_{1 \leq i \leq p} \) implies the existence of \( \delta > 0 \) such that
\[ (\forall i \in \{1, \ldots, p\})(\forall (z_1, z_2) \in Y_1^2) \| z_1 - z_2 \| < \delta \Rightarrow \| \nabla e_i(z_1) - \nabla e_i(z_2) \|^2 \leq \frac{\varepsilon^2}{4pV^2}. \] (4.12)

Now, suppose that
\[ \| x - y \|^2 \leq \min \left\{ \frac{\varepsilon^2}{4pK^2}, \delta^2 \right\}. \]

Then, (4.11), the convexity of \( \| \cdot \|^2 \), and (4.12) imply
\[ \| B_3 x - B_3 y \|^2 = \left| \sum_{i=1}^{p} v_{1,i} \nabla e_i(x_1) - v_{2,i} \nabla e_i(x_2) \right|^2 + \| e(x_1) - e(x_2) \|^2 \]
\[ \leq 2 \left| \sum_{i=1}^{p} v_{1,i} \nabla e_i(x_1) \right|^2 + 2 \left| \sum_{i=1}^{p} v_{2,i} \nabla e_i(x_1) - \nabla e_i(x_2) \right|^2 + K^2 \| x_1 - x_2 \|^2 \]
\[ \leq 2p \sum_{i=1}^{p} |v_{1,i} - v_{2,i}| \| \nabla e_i(x_1) \|^2 + 2p \sum_{i=1}^{p} |v_{2,i}| \| \nabla e_i(x_1) - \nabla e_i(x_2) \|^2 + K^2 \| x_1 - x_2 \|^2 \]
\[ \leq 2pK^2 \| v_1 - v_2 \|^2 + 2pV^2 \sum_{i=1}^{p} \| \nabla e_i(x_1) - \nabla e_i(x_2) \|^2 + K^2 \| x_1 - x_2 \|^2 \]
\[ \leq 2pK^2 \| x - y \|^2 + \frac{\varepsilon^2}{2} \]
\[ \leq \varepsilon^2. \]

Therefore, \( B_3 \) is uniformly continuous in \( Y \).

Now, assume 2b. Since \( \mathcal{H} \) is finite-dimensional, the weak and strong topologies coincide [2, Fact 2.33]. Hence, since \( \nabla e_i \) is continuous, it is bounded and uniformly continuous in every compact subset of \( X \). The result follows from 2a. \( \Box \)

**Remark 4.1** Note that if, for every \( i \in \{1, \ldots, p\} \), \( \nabla e_i \) is bounded and uniformly continuous in every weakly compact subset of \( \text{dom } f \), since \( \overline{\text{dom } \partial f} \subset \text{dom } f \), \( B_3 \) is uniformly continuous in every compact subset of \( \overline{\text{dom } \partial f \times \text{dom } \partial g^* \times [0, +\infty[p} \) in view of Proposition 4.2.

**Proposition 4.3** In Problem 4.1, let \( X = X_1 \times X_2 \times X_3 \subset \text{dom } \partial f \times \text{dom } \partial g^* \times [0, +\infty[p \) be nonempty, closed, and convex, let \( \varepsilon \in ]0, 1[ \), set \( \rho = \min \{2\beta \varepsilon, \sqrt{1 - \varepsilon} / \sqrt{\varepsilon} \}. \)
\[ \|M\|, \text{let } \sigma \in [0, 1], \text{and let } \theta \in ]0, \sqrt{1-\epsilon} - \|M\|\rho\sigma[. \text{ For every } z = (z^1, z^2, z^3) \in H \times G \times \mathbb{R}^p, \text{ define } \Phi_z: y \mapsto (\Phi^1_z(y), \Phi^2_z(y), \Phi^3_z(y)), \text{ where}
\]

\[
\phi^1_z: y \mapsto \text{prox}_{\gamma f} \left( z^1 - \gamma \left( \nabla h(z^1) + M^*z^3 + \sum_{i=1}^p z^3_i \nabla e_i(z^1) \right) \right)
\]

\[
\phi^2_z: y \mapsto \text{prox}_{\gamma g^*}(z^2 + \gamma Mz^1)
\]

\[
\phi^3_z: y \mapsto P_{[0, +\infty]^p}(z^3 + \gamma e(z^1)).
\]

(4.13)

Let \( z_0 = (z^1_0, z^2_0, z^3_0) \in H \times G \times \mathbb{R}^p \) and consider the recurrence, for every \( n \in \mathbb{N} \),

\[
\begin{align*}
  x^n_1 &= \phi^1_{z^n}(y^n) \\
  x^n_2 &= \phi^2_{z^n}(y^n) \\
  x^n_3 &= \phi^3_{z^n}(y^n) \\
  z^n_{1+1} &= P_{X_1}(x^n_1 + \gamma_n (M^*z^n_2 + \sum_{i=1}^p z^n_3 \nabla e_i(x^n_1) - \gamma_n (M^*x^n_2 + \sum_{i=1}^p x^n_3 \nabla e_i(x^n_1))) \\
  z^n_{2+1} &= P_{X_2}(x^n_2 - \gamma_n Mz^n_1 + \gamma_n Mx^n_1) \\
  z^n_{3+1} &= P_{X_3}(x^n_3 - \gamma_n e(z^n_1) + \gamma_n e(x^n_1)) \\
  z^n_{n+1} &= (z^n_{1+1}, z^n_{2+1}, z^n_{3+1})
\end{align*}
\]

(4.14)

where \( \gamma_n \) is the largest \( \gamma \in \{\rho\sigma, \rho\sigma^2, \rho\sigma^3, \ldots\} \) satisfying

\[
\gamma^2 \left( \left\| \sum_{i=1}^p z^n_{3,i} \nabla e_i(z^n_1) - \Phi^3_{z^n,i}(y) \nabla e(\Phi^1_{z^n}(y)) \right\|_2^2 + \left\| e(z^n_1) - e(\Phi^1_{z^n}(y)) \right\|_2^2 \right) \leq \theta^2 \|z^n - \Phi_{z^n}(y)\|_2^2.
\]

(4.15)

Moreover, assume that at least one of the following additional statements holds:

(i) \( \lim\inf_{n \to \infty} \gamma_n = \delta > 0 \).

(ii) For every \( i \in \{1, \ldots, p\}, \nabla e_i \) is bounded and uniformly continuous in every weakly compact subset of \( \text{dom } \partial f \).

Then, \( (z^n)_{n \in \mathbb{N}} \) converges weakly to a solution to Problem 4.1.

**Proof** Let \( \mathcal{H} = H \times G \times \mathbb{R}^p \), define

\[
\begin{align*}
  A: \mathcal{H} \to 2\mathcal{H}: (x, u, v) \mapsto \partial f(x) \times \partial g^*(u) \times N_{[0, +\infty]^p}(v), \\
  B_1: \mathcal{H} \to \mathcal{H}: (x, u, v) \mapsto (\nabla h(x), 0, 0), \\
  B_2: \mathcal{H} \to \mathcal{H}: (x, u, v) \mapsto (M^*u, -Mx, 0),
\end{align*}
\]

(4.16)

and consider the operator \( B_3 \) defined in (4.6). Note that \( A \) is maximally monotone [2, Proposition 20.23 & Proposition 20.25], \( B_1 \) is \( \beta \)-cocoercive [2, Corollary 18.17], \( B_2 \) is \( \|M\| \)-Lipschitzian [10, Proposition 2.7(ii)] & [2, Fact 2.20], and the operator \( B_3 \) is maximally monotone by Proposition 4.2. Furthermore, note that \( \text{dom } A = \text{dom } (\partial f) \times \text{dom } (\partial g^*) \times [0, +\infty]^p \) and \( \text{dom } B_3 = \bigcap_{i=1}^p \text{intdom } e_i \times G \times [0, +\infty]^p \).
Hence, it follows from \( \text{dom}(\partial f) \subset \cap_{i=1}^{n} \text{intdom} \, e_i \); that \( \text{dom} \, A \cup X \subset \text{dom} \, B_3 \) and \( 0 \in \text{int}(\text{dom} \, A \rangle \text{dom} \, B_3) \). Therefore, \( A + B_3 \) is maximally monotone \([2, \text{Corollary 25.5(ii)}]\). Altogether, the inclusion in (4.5) is a particular instance of Problem 3.1. Define, for every \( n \in \mathbb{N} \), \( x_n = (x_n^1, x_n^2, x_n^3) \). Hence, (4.13), (4.6), and (4.16) yield

\[
(\forall n \in \mathbb{N}) \quad \begin{aligned}
 x_n &= J_{\gamma A}(z_n - \gamma_n(B_1 + B_2 + B_3)z_n) \\
 z_{n+1} &= P_X(x_n + \gamma_n(B_2 + B_3)z_n - \gamma_n(B_2 + B_3)x_n),
\end{aligned}
\]

(4.17)

where \( \gamma_n \), by (4.15) and (4.16), satisfies

\[
\gamma \| B_3z_n - B_3J_{\gamma A}(z_n - \gamma(B_1 + B_2 + B_3)z_n) \| \leq \theta \| z_n - J_{\gamma A}(z_n - \gamma(B_1 + B_2 + B_3)z_n) \|.
\]

(4.18)

Note that if we assume (ii), by Proposition 4.2, \( B_3 \) is uniformly continuous in every weak compact subset of \( \overline{\text{dom}} \, \partial f \times \overline{\text{dom}} \, \partial g^* \times [0, +\infty] = \text{conv} \left( \text{dom} \, A \right) \leq \overline{\text{conv}} \left( \text{dom} \, A \cup X \right) [2, \text{Corollary 21.14 & Exercise 3.2}]. \) In view of \( 0 \in \text{sr}(\text{dom} \, g - M(\text{dom} \, f)) \), (4.1), and [2, Proposition 27.21], there exists a solution \( \tilde{x} = (\hat{x}, \hat{u}, \hat{v}) \in H \times \mathcal{G} \times \mathbb{R}^p \) to (4.5) such that \( \hat{x} \) is a solution to Problem 4.1. Altogether, since \( X \subset \text{dom} \, A \), Theorem 3.3 implies that \( z_n \to \tilde{x} \) and the result follows. \( \square \)

5 Numerical Experiments

In this section, we consider the following optimization problem

\[
\min_{\begin{subarray}{l}
y^0 \leq x \leq y^1, \\
i \in \{1, \ldots, n\}
\end{subarray}} \frac{\alpha \| Mx \|_1 + \frac{1}{2} \| Ax - z \|^2}{
\begin{aligned}
x_i &\left( \ln(x_i/a_i) - 1 \right) - r_i \leq 0,
\end{aligned}
\]

(5.1)

where \( M \in \mathbb{R}^{r \times n}, A \in \mathbb{R}^{m \times n}, z \in \mathbb{R}^m, y^0 = (\eta^0_i)_{1 \leq i \leq n} \in \mathbb{R}^n, y^1 = (\eta^1_i)_{1 \leq i \leq N} \in \mathbb{R}^N, \) and, for every \( i \in \{1, \ldots, p\}, r_i \in ] - a_i, 0[ \) and \( a_i \in ]0, +\infty[ \). Set

\[
\begin{aligned}
 C &= \times_{i=1}^n [\eta^0_i, \eta^1_i], \\
f &= 1_c, \\
g &= \alpha \| \cdot \|_1, \\
h &= \| A \cdot -z \|^2/2, \\
e &= (e_i(\cdot))^n_i,
\end{aligned}
\]

(5.2)

where

\[
(\forall i = 1, \ldots, n) \quad e_i: \mathbb{R}^n \to ] - \infty, +\infty[; x \mapsto \begin{cases} x_i(\ln(x_i/a_i) - 1) - r_i, & \text{if } x_i > 0; \\
-r_i, & \text{if } x_i = 0; \\
+\infty, & \text{otherwise.}
\end{cases}
\]

(5.3)

\( \copyright \) Springer
Then, we have $f \in \Gamma_0(\mathbb{R}^n)$, $g \in \Gamma_0(\mathbb{R}^r)$, and $\nabla h$ is $\|A\|^2$-Lipschitzian. Additionally, for every $i \in \{1, \ldots, n\}$, $e_i$ is Gâteaux differentiable in $[0, +\infty)^n$,

\[
(\nabla e_i(x))_k = \begin{cases} 
\ln x_k, & \text{if } k = i; \\
0, & \text{otherwise}, 
\end{cases} \tag{5.4}
\]

dom $e_i$ is closed, $\cap_{i=1}^n \text{dom } e_i = [0, +\infty)^n$, (4.1) holds, and

\[
0 \in \text{int} \left( \text{dom} \ (\partial f) - \cap_{i=1}^n \text{dom } e_i \right) = x^N = -\infty, 1^1.
\]

Hence, the optimization problem in (5.1) is a particular instance of Problem 4.1. In this setting, since $g^* = \ell[-\alpha, \alpha]^r$ [2, Example 13.32(v) & Proposition 13.23(i)], we consider $X_1 \times X_2 \times X_3 = C \times [-\alpha, \alpha]^r \times [0, +\infty]^p$ in order to write the recurrence in (4.14) as Algorithm 1.

Algorithm 1

1: Fix $z_0 = (z_0^1, z_0^2, z_0^3) \in \mathbb{R}^n \times \mathbb{R}^r \times \mathbb{R}^r$. Let $\sigma \in [0, 1]$, let $\varepsilon = \|A\|^4 \sqrt{1 + 16|\nabla h|^2/\|A\|^4 - 1}$, let $\theta = 2\varepsilon \|M\|(1 - \sigma)/\|A\|^2$, and let $\varepsilon > 0$.
2: while $r_n > \varepsilon$ do
3: \quad $\gamma = 2\varepsilon \|M\|$
4: \quad $\nu = 0$
5: \quad while $\nu = 0$ do
6: \quad \quad $\gamma \rightarrow \gamma \cdot \sigma$
7: \quad \quad $\Phi^1(\gamma) = P_{[0, \rho]}(z^1_n - \gamma(A^*A_z^1_n - \nu + M^*z_n^1 + \sum_{i=1}^p z^1_n,z^i_n, \ln(z^1_n,z))$\)
8: \quad \quad $\Phi^2(\gamma) = \gamma(1 - \text{prox}_{\sigma \|\cdot\|_1/\gamma})(z^2_n/\gamma + Mz_n^1)$
9: \quad \quad $\Phi^3(\gamma) = P_{[0, +\infty]}(z^2_n/\gamma + \gamma e(z_n^3))$
10: \quad \quad $\Phi(\gamma) = (\Phi^1(\gamma), \Phi^2(\gamma), \Phi^3(\gamma))$
11: \quad \quad if $\sum_{i=1}^p |z^3_n,z|_i - \Phi^3(\gamma) \ln(\Phi^1(\gamma))_i^2 + \|e(z^1_n) - e(\Phi^1(\gamma))\|^2 \leq \theta^2 \varepsilon \|z_n - \Phi(\gamma)\|^2$
then
12: \quad \quad \quad $\nu = 1$
13: \quad \quad end if
14: \quad end while
15: \quad $y_n = \gamma$
16: \quad $(x_n^1, x_n^2, x_n^3) = (\Phi^1(y_n), \Phi^2(y_n), \Phi^3(y_n))$
17: \quad $z_n^1 + 1 = P_{[0, \rho]}(x_n^1 + y_n(M^*x_n^2 + \sum_{i=1}^p z^1_n,z^i_n, \ln(z^1_n,z)) - y_n(M^*x_n^2 + \sum_{i=1}^p z^1_n,z^i_n, \ln(z^1_n,z))$\)
18: \quad $z_n^2 + 1 = P_{[-\alpha, \alpha]^r}(x_n^2 - y_n Mz_n^1 + y_n Mx_n^1)$
19: \quad $z_n^3 + 1 = P_{[0, +\infty]}(x_n^3 - y_n e(z_n^3) + y_n e(x_n^1))$
20: \quad $z_n + 1 = (z_n^1, z_n^2, z_n^3)$
21: \quad $r_n = R(z_n, z)$
22: \quad $n \rightarrow n + 1$
23: end while
24: return $z_{n+1}$

We compare Algorithm 1 with the algorithm propose in [12] called FBHF and with the MATLAB’s fmincon (interior point).
Table 1  Average time and average number of iterations of 20 random realizations of problem in (5.1) for Algorithm 1, FBHF, and fmincon

| $N_1$ | $\epsilon = 10^{-6}$ | $N_2 = N_1/3$ | $N_2 = N_1/2$ | $N_2 = 2N_1/3$ |
|-------|-------------------|----------------|----------------|----------------|
|       | Algorithm | Av. Time (s) | Av. Iter | Av. Time (s) | Av. Iter | A. Time (s) | A. Iter |
| 600   | Alg. 1 | 6.82 | 7845 | 9.61 | 10384 | 15.79 | 16136 |
|      | FBHF | 10.47 | 8280 | 14.22 | 10885 | 23.28 | 16772 |
|      | fmincon | 52.52 | 238 | 66.25 | 276 | 69.78 | 251 |
| 900   | Alg. 1 | 19.26 | 8185 | 28.89 | 11932 | 52.69 | 20653 |
|      | FBHF | 31.28 | 8568 | 46.06 | 12375 | 84.17 | 21757 |
|      | fmincon | 256.71 | 350 | 309.33 | 408 | 292.21 | 368 |
| 1200  | Alg. 1 | 36.01 | 8809 | 62.82 | 14490 | 110.76 | 24778 |
|      | FBHF | 59.41 | 9231 | 98.60 | 14783 | 174.08 | 25633 |
|      | fmincon | 694.86 | 457 | 839.06 | 528 | 790.66 | 462 |

To solve problem in (4.2) with FBHF algorithm, we consider $X = X_1 \times X_2 \times X_3$ and the operators (see (4.5) and [12, Theorem 2.3])

$$A = \begin{pmatrix} \frac{\partial f(\hat{x})}{\partial g^*(\hat{u})} \\ N_{[0, +\infty]^p}(\hat{v}) \end{pmatrix}, \quad B_1 = \begin{pmatrix} \nabla h(\hat{x}) \\ 0 \end{pmatrix}, \quad B_2 + B_3 = \begin{pmatrix} M^*\hat{u} + \sum_{i=1}^p \hat{v}_i \nabla e_i(\hat{x}) \\ -e(\hat{x}) \end{pmatrix}.$$  (5.5)

In our numerical experiments, we generate 20 random realizations of $A$, $M$, $z$, and $r_1, \ldots, r_n$ for dimensions $n = m \in \{600, 900, 1200\}$ and $r \in \{n/3, n/2, 2n/3\}$. In each realization, we define $a_i = 9$ for $i \in \{1, \ldots, n\}$, $\alpha = 0.05$, and $y^0 = \hat{y}_0$ and $y^1 = \hat{y}_1 + \text{rand}(n)$, where $\hat{y}_1$ and $\hat{y}_2$ satisfy $e(\hat{y}_0) = e(\hat{y}_1) = 0$. For Algorithm 1, we consider $\sigma = 0.99$. For FBHF, we consider $\epsilon = 0.8$, $\theta = \sqrt{1-\epsilon}/2$, $\sigma = 0.99$, the maximally monotone operator $A$, the cocoercive operator $B_1$, and the continuous operator $B_1 + B_2$ on (5.5) (see [12, Theorem 2.3]).

In Table 1, we provide the average time and iterations to achieve a tolerance $\epsilon = 10^{-6}$ for the instances mentioned above. We can observe that, for each instance, Algorithm 1 is more efficient than the method FBHF and fmincon. Algorithm 1 and FBHF are similar in number of iterations, but each iteration of FBHF is more expensive in time than Algorithm 1. This is because FBHF needs, additionally, to activate the operators $M^*$ and $M$ in each line search. This difference is larger as the dimension of the problem increases. Although fmincon needs less iterations than Algorithm 1 and FBHF to reach the stop criterion, each iteration is very expensive in CPU time. Indeed, Algorithm 1 reaches the stop criterion in 20% of the time that fmincon takes.

6 Conclusions

We provide an algorithm for finding a zero of the sum of a maximally monotone, a cocoercive, a monotone-Lipschitzian, and a monotone-continuous operators in a...
Hilbertian setting which splits their influence in its implementation. The proposed method exploits the intrinsic properties of each operator activating implicitly the maximally monotone operator via its resolvent, explicitly the cocoercive and the monotone-Lipschitzian operator, and for the continuous operator activation, a line search procedure is needed. This method generalizes previous results in [12, 29, 38], and it is more efficient than other algorithms in the literature when applied to nonlinearly constrained convex optimization problems.

**Acknowledgements** The first author thanks the support of Centro de Modelamiento Matemático (CMM), ACE210010 and FB210005, BASAL funds for centers of excellence, grant Redes 180032, and grant FONDECYT 1190871 from ANID-Chile. The second author thanks the support of ANID-Subdirección de Capital Humano/Doctorado Nacional/2018-21181024 and by the Dirección de Postgrado y Programas from UTFSM through Programa de Incentivos a la Iniciación Científica (PIIC).

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

**Declarations**

**Conflict of interest** The authors have no conflicts of interest to declare that are relevant to the content of this article.

**References**

1. Aubin, J.P., Frankowska, H.: Set-Valued Analysis. Modern Birkhäuser Classics, Birkhäuser Boston, Inc., Boston, MA. [https://doi.org/10.1007/978-0-8176-4848-0](https://doi.org/10.1007/978-0-8176-4848-0) (2009)
2. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Springer, New York (2007). [https://doi.org/10.1007/978-3-319-48311-5](https://doi.org/10.1007/978-3-319-48311-5)
3. Boţ, R.I., Csetnek, E.R.: An inertial forward-backward-forward primal-dual splitting algorithm for solving monotone inclusion problems. Numer. Algorithms 71, 519–540 (2016). [https://doi.org/10.1007/s11075-015-0007-5](https://doi.org/10.1007/s11075-015-0007-5)
4. Boţ, R.I., Csetnek, E.R.: ADMM for monotone operators: convergence analysis and rates. Adv. Comput. Math. 45, 327–359 (2019). [https://doi.org/10.1007/s10444-018-9619-3](https://doi.org/10.1007/s10444-018-9619-3)
5. Boţ, R.I., Csetnek, E., Heinrich, A.: A primal-dual splitting algorithm for finding zeros of sums of maximal monotone operators. SIAM J. Optim. 23, 2011–2036 (2013). [https://doi.org/10.1137/12088255X](https://doi.org/10.1137/12088255X)
6. Boţ, R.I., Csetnek, E.R., Hendrich, C.: Inertial Douglas-Rachford splitting for monotone inclusion problems. Appl. Math. Comput. 256, 472–487 (2015). [https://doi.org/10.1016/j.amc.2015.01.017](https://doi.org/10.1016/j.amc.2015.01.017)
7. 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). [https://doi.org/10.1137/120901106](https://doi.org/10.1137/120901106)
8. Briceño-Arias, L.M.: Forward-Douglas-Rachford splitting and forward-partial inverse method for solving monotone inclusions. Optimization 64, 1239–1261 (2015). [https://doi.org/10.1080/02331934.2013.855210](https://doi.org/10.1080/02331934.2013.855210)
9. Briceño-Arias, L.M.: Forward-partial inverse-forward splitting for solving monotone inclusions. J. Optim. Theory Appl. 166, 391–413 (2015). [https://doi.org/10.1007/s10957-015-0703-2](https://doi.org/10.1007/s10957-015-0703-2)
10. 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). [https://doi.org/10.1137/10081602X](https://doi.org/10.1137/10081602X)
11. Briceño-Arias, L.M., Combettes, P.L.: Monotone Operator Methods for Nash Equilibria in Non-Potential Games. In: Bailey D H, Bauschke H H, Borwein P, Garvan F, Théra M, Vanderwerff J, Wolkowicz H (eds) Computational and Analytical Mathematics. Springer Proceedings in Mathematics & Statistics 50. Springer, New York, NY, 143–159. [https://doi.org/10.1007/978-1-4614-7621-4_9](https://doi.org/10.1007/978-1-4614-7621-4_9) (2013)
12. Briceño-Arias, L.M., Davis, D.: Forward-backward-half forward algorithm for solving monotone inclusions. SIAM J. Optim. 28, 2839–2871 (2018). [https://doi.org/10.1137/17M1120099](https://doi.org/10.1137/17M1120099)
13. Bùi, M.N., Combettes, P.L.: Multivariate monotone inclusions in saddle form. Math. Operat. Res. 47, 1082–1109 (2021). https://doi.org/10.1287/moor.2021.1161
14. Cevher, V., Vü, B.C.: A reflected forward-backward splitting method for monotone inclusions involving lipschitzian operators. Set-Valued Var. Anal. 29, 163–174 (2021). https://doi.org/10.1007/s11228-020-00542-4
15. Combettes, P.L.: Quasi-Fejérian Analysis of Some Optimization Algorithms. In: Butnariu D, Censor Y, Reich S (eds) Inherently parallel algorithms in feasibility and optimization and their applications. Stud Comput Math 8. North-Holland, Amsterdam, 115–152. (2001) https://doi.org/10.1016/S1570-579X(01)80010-0
16. Combettes, P.L., Eckstein, J.: Asynchronous block-iterative primal-dual decomposition methods for monotone inclusions. Math. Program. 168, 645–672 (2018). https://doi.org/10.1007/s10107-016-1044-4
17. 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). https://doi.org/10.1007/s11228-011-0191-y
18. Combettes, P.L., Vü, B.C.: variable metric forward-backward splitting with applications to monotone inclusions in duality. Optimization 63, 1289–1318 (2014). https://doi.org/10.1080/02331934.2012.733883
19. Csetnek, E., Malitsky, Y., Tam M.: Shadow Douglas-Rachford splitting for monotone inclusions. Appl. Math. Optim. 80, 665–678 (2019). https://doi.org/10.1007/s00245-019-09597-8
20. Davis, D., Yin, W.: A three-operator scheme and its optimization applications. Set-Valued Var. Anal. 25, 829–858 (2017). https://doi.org/10.1007/s11228-017-0421-z
21. Dong, Y.: Weak convergence of an extended splitting method for monotone inclusions. J. Global. Optim. 79, 257–277 (2021). https://doi.org/10.1007/s10898-020-00940-w
22. Dung, D.: Vü B.C.: A splitting algorithm for system of composite monotone inclusions. Vietnam J. Math. 43, 323–341 (2015). https://doi.org/10.1007/s10013-015-0121-7
23. Eckstein, J.: A simplified form of block-iterative operator splitting and an asynchronous algorithm resembling the multi-block alternating direction method of multipliers. J. Optim. Theory Appl. 173, 155–182 (2017). https://doi.org/10.1007/s10957-017-1074-7
24. Gabay, D.: Chapter IX Applications of the method of multipliers to variational inequalities. In: Fortin M, Glowinski R (eds.) Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems. Studies in Mathematics and Its Applications, 15. Elsevier, 299–331. (1983) https://doi.org/10.1016/S0168-2024(08)70034-1
25. Glowinski, R., Marrocco, A.: Sur l’approximation, Par Éléments Finis d’ordre un, et la Résolution, par Pénalisation-dualité, d’une Classe de Problèmes de Dirichlet non Linéaires. Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge Anal Numér 9, 41–76 (1975)
26. Goldstein, A.A.: Convex programming in Hilbert space. Bull. Amer. Math. Soc. 70, 709–710 (1964). https://doi.org/10.1090/S0002-9904-1964-11178-2
27. Johnstone, P.R., Eckstein, J.: Projective splitting with forward steps only requires continuity. Optim. Lett. 14, 229–247 (2020). https://doi.org/10.1007/s11590-019-01509-7
28. Johnstone, P.R., Eckstein, J.: Single-forward-step projective splitting: exploiting cocoercivity. Comput. Optim. Appl. 78, 125–166 (2021). https://doi.org/10.1007/s10589-020-00238-3
29. Lions, P., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. 16, 964–979 (1979). https://doi.org/10.1137/0716071
30. Malitsky, Y., Tam, M.K.: A forward-backward splitting method for monotone inclusions without cocoercivity. SIAM J. Optim. 30, 1451–1472 (2020). https://doi.org/10.1137/18M1207260
31. Malitsky, Y., Tam, M.K.: A forward-backward splitting method for monotone inclusions without cocoercivity. SIAM J. Imag. Sci. 6, 1199–1226 (2013). https://doi.org/10.1137/120872802
32. Rieger, J., Tam, M.K.: Backward-forward-reflected-backward splitting for three operator monotone inclusions. Appl. Math. Comput. 381, 125248 2020).https://doi.org/10.1016/j.amc.2020.125248
33. Rockafellar, R.T.: Monotone operators associated with saddle-functions and mininax problems. In: Browder F E (ed) Nonlinear Functional Analysis (Proc. Sympos. Pure Math., Vol. XVIII, Part 1, Chicago, Ill., 1968). Amer Math Soc. Providence, R.I., 241–250 (1970)
34. Ryu, E.K., Vü, B.C.: Finding the forward-Douglas-Rachford-forward method. J. Optim. Theory Appl. 184, 858–876 (2020). https://doi.org/10.1007/s10957-019-01601-z
35. Salzo, S.: The variable metric forward-backward splitting algorithm under mild differentiability assumptions. SIAM J. Optim. 27, 2153–2181 (2017). https://doi.org/10.1137/16M1073741
36. Showalter, R.E.: Monotone Operators in Banach Space and Nonlinear Partial Differential Equations. Mathematical surveys and monographs 49. Amer. Math. Soc. Providence, RI. (1997) https://doi.org/10.1090/surv/049
37. Spingarn, J.E.: Partial inverse of a monotone operator. Appl. Math. Optim. 10, 247–265 (1983). https://doi.org/10.1007/BF01448388
38. Tseng, P.: A modified forward-backward splitting method for maximal monotone mappings. SIAM J. Control Optim. 38, 431–446 (2000). https://doi.org/10.1137/S0363012998338806
39. Vău, B.C.: A splitting algorithm for dual monotone inclusions involving cocoercive operators. Adv. Comput. Math. 38, 667–681 (2013). https://doi.org/10.1007/s10444-011-9254-8

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.