Primal-dual splittings as fixed point iterations in the range of linear operators

Luis Briceño-Arias · Fernando Roldán

Received: 25 January 2022 / Accepted: 14 September 2022 / Published online: 3 October 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
In this paper we study the convergence of the relaxed primal-dual algorithm with critical preconditioners for solving composite monotone inclusions in real Hilbert spaces. We prove that this algorithm define Krasnosel’skii-Mann (KM) iterations in the range of a particular monotone self-adjoint linear operator with non-trivial kernel. Our convergence result generalizes (Condat in J Optim Theory Appl 158: 460–479, 2013, Theorem 3.3) and follows from that of KM iterations defined in the range of linear operators, which is a real Hilbert subspace under suitable conditions. The Douglas–Rachford splitting (DRS) with a non-standard metric is written as a particular instance of the primal-dual algorithm with critical preconditioners and we recover classical results from this new perspective. We implement the algorithm in total variation reconstruction, verifying the advantages of using critical preconditioners and relaxation steps.

Keywords Convex optimization · Douglas–Rachford splitting · Krasnosel’skii-Mann iterations · Monotone operator theory · Primal-dual algorithm · Quasinonexpansive operators

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

1 Introduction
In this paper we provide a theoretical study of the relaxed primal-dual splitting [45] for solving the following composite monotone inclusion.

Proposition 1.1 Let \( \mathcal{H} \) and \( \mathcal{G} \) be real Hilbert spaces, let \( A : \mathcal{H} \to 2^{\mathcal{H}} \) and \( B : \mathcal{G} \to 2^{\mathcal{G}} \) be maximally monotone operators, and let \( L : \mathcal{H} \to \mathcal{G} \) be a linear bounded operator. The problem is to find \((\hat{x}, \hat{u}) \in \mathcal{Z}\), where
\[ Z = \{(x, u) \in \mathcal{H} \times \mathcal{G} \mid 0 \in Ax + L^*u, \ 0 \in B^{-1}u - Lx\} \] (1.1)
is assumed to be non-empty.

It follows from [8, Proposition 2.8] that any solution \((\hat{x}, \hat{u})\) to Problem 1.1 satisfies that \(\hat{x}\) is a solution to the primal inclusion
\[
\text{find } \ x \in \mathcal{H} \quad \text{such that} \quad 0 \in Ax + L^*BLx
\] (1.2)
and \(\hat{u}\) is a solution to the dual inclusion
\[
\text{find } \ u \in \mathcal{G} \quad \text{such that} \quad 0 \in B^{-1}u - LA^{-1}(-L^*u).
\] (1.3)

Conversely, if \(\hat{x}\) is a solution to (1.2) then there exists \(\bar{u}\) solution to (1.3) such that \((\hat{x}, \bar{u}) \in Z\) and the dual argument also holds. In the particular case when \(A = \partial f\) and \(B = \partial g\), where \(f : \mathcal{H} \to ]-\infty, +\infty]\) and \(g : \mathcal{G} \to ]-\infty, +\infty]\) are proper, convex, and lower semicontinuous, we have that \(Z \subset \mathcal{P} \times \mathcal{D}\), where \(\mathcal{P}\) is the set of solutions to the convex optimization problem
\[
\min_{x \in \mathcal{H}} f(x) + g(Lx)
\] (1.4)
and \(\mathcal{D}\) is the set of solutions to its Fenchel-Rockafellar dual
\[
\min_{u \in \mathcal{G}} g^*(u) + f^*(-L^*u).
\] (1.5)

Problem 1.1 and (1.4) model a wide class of problems in engineering including mechanical problems [29, 31, 32], differential inclusions [1, 43], game theory [9], image processing problems as image restoration and denoising [16, 18, 25], traffic theory [7, 28, 30], among other disciplines.

In the last years, several algorithms have been proposed in order to solve Problem 1.1 and some generalizations involving single-valued monotone operators (see, e.g., [8, 10–12, 21, 34, 45]). One of the most used, is the algorithm proposed in [23] (see also [5, 6, 45] and [17, 40] in the context of (1.4)), which iterates
\[
(\forall n \in \mathbb{N}) \quad \begin{cases} 
p_{n+1} = J_{\gamma A}(x_n - \gamma L^*u_n) 
q_{n+1} = J_{\lambda B^{-1}}(u_n + \lambda L(2p_{n+1} - x_n)) 
(x_{n+1}, y_{n+1}) = (1 - \lambda_n)(x_n, y_n) + \lambda_n(p_{n+1}, q_{n+1}),
\end{cases}
\] (1.6)
where \((x_0, u_0) \in \mathcal{H} \times \mathcal{G}, (\lambda_n)_{n \in \mathbb{N}} \subset ]0, 2[, Y : \mathcal{H} \to \mathcal{H}\) and \(\Sigma : \mathcal{G} \to \mathcal{G}\) are strongly monotone self-adjoint linear bounded preconditioners satisfying \(\|\sqrt{\Sigma L}\sqrt{Y}\| < 1\), and \(J_A = (\text{Id} + A)^{-1}\) stands for the resolvent of \(A\). It turns out that (1.6) corresponds to the relaxed proximal-point algorithm [37, 41] associated to the operator \(V^{-1}M\), where
\[
M : \mathcal{H} \to 2^{\mathcal{H}} : (x, u) \mapsto \{(y, v) \in \mathcal{H} \mid y \in Ax + L^*u, \ v \in B^{-1}u - Lx\}
\] (1.7)
is maximally monotone in \(\mathcal{H} \oplus \mathcal{G}\) [8, Proposition 2.7(iii)] and the self-adjoint linear bounded operator
\[
V : \mathcal{H} \to \mathcal{H} : (x, u) \mapsto (\gamma^{-1}x - L^*u, \Sigma^{-1}u - Lx)
\] (1.8)
is strongly monotone because \(\|\sqrt{\Sigma L}\sqrt{Y}\| < 1\). Hence, \(V^{-1}M\) is maximally monotone in the real Hilbert space \((\mathcal{H} \times \mathcal{G}, \langle \cdot | \cdot \rangle_Y)\) and the convergence is a consequence of [37, 41]. Note that \(J_M\) is also firmly nonexpansive in \(\mathcal{H} \oplus \mathcal{G}\), however it has no explicit computation. Non-standard metrics as \(\langle \cdot | \cdot \rangle_Y\) are widely used not only to obtain explicit resolvent computations but also to accelerate algorithms [10, 22, 24, 26, 34, 45]. In the presence of critical preconditioners, i.e., \(\|\sqrt{\Sigma L}\sqrt{Y}\| = 1\), the non-standard metric approach fails since
ker $V \neq \{0\}$, and hence $\langle \cdot | V \cdot \rangle$ is not an inner product in $\mathcal{H} \times \mathcal{G}$. The convergence of (1.6) when $\|\sqrt{\Sigma} L \sqrt{\Upsilon}\| \leq 1$ in the context of Problem 1.1 is an open question, which is partially answered in [24, Theorem 3.3] for solving (1.4) in a finite dimensional setting when $\Sigma = \sigma I$ and $\Upsilon = \tau I$. In the particular case when $\lambda_n \equiv 1$, the weak convergence of (1.6) when $\|\sqrt{\Sigma} L \sqrt{\Upsilon}\| \leq 1$ is deduced in [13] from an alternative formulation of (1.6). This formulation in the case when $L = I$, generates primal-dual iterates in the graph of $A$, and therefore the argument does not hold when $\lambda_n \neq 1$.

In this paper we generalize [24, Theorem 3.3] to the monotone inclusion in Problem 1.1 in the infinite dimensional setting with critical preconditioners. Our approach is based on a fixed point theory restricted to $(\text{ran } V, \langle \cdot | V \cdot \rangle)$, which is a real Hilbert space under the condition ran $V$ closed. We obtain the weak convergence of Krasnosel’skii-Mann (KM) iterations governed by firmly quasinonexpansive and averaged operators in $(\text{ran } V, \langle \cdot | V \cdot \rangle)$, which generalizes [19, Theorem 5.2(i)] and [3, Proposition 5.16]. Our approach gives new insights on primal-dual algorithms: the convergence of primal-dual iterates in $\mathcal{H}$ follows from the convergence of their shadows in $\text{ran } V$.

We also provide a detailed analysis of the case $L = I$ and relations of primal-dual algorithms with the relaxed DRS [27, 35]. We give a primal-dual version of DRS derived from (1.6) when $L = I$ and we recover the weak convergence of an auxiliary sequence whose primal-dual shadow is a solution to Problem 1.1, as in [27, 35].

We finish this paper by providing a numerical experiment on total variation image reconstruction, in which the advantages of using critical preconditioners and relaxation steps are illustrated.

The paper is organized as follows. In Sect. 2.1 we set our notation and some preliminaries. In Sect. 2.2 we study the fixed point problem on the range of linear operators and we provide conditions for the convergence of KM iterations governed by firmly quasinonexpansive or averaged nonexpansive operators. In Sect. 3 we apply fixed point results to the particular case of primal-dual algorithms for monotone inclusions and we provide several connections with other results in the literature. In Sect. 3.1, we study in detail the particular case when $L = I$, which is connected with DRS splitting. Finally, in Section 4, we provide numerical experiments in image processing showing the advantages of our approach.

## 2 Notation and preliminaries

In this section we first provide our notation. Next, we study the convergence of Krasnosel’skii-Mann iterations defined in the range of a monotone self-adjoint linear operator, which is crucial for the convergence of the relaxed primal-dual algorithm (1.6) with critical preconditioners.

### 2.1 Notation

Throughout this paper $\mathcal{H}$ and $\mathcal{G}$ are real Hilbert spaces. We denote the scalar product by $\langle \cdot | \cdot \rangle$ and the associated norm by $\| \cdot \|$. The symbols $\rightharpoonup$ and $\to$ denotes the weak and strong convergence, respectively. Given a linear bounded operator $L : \mathcal{H} \to \mathcal{G}$, we denote its adjoint by $L^* : \mathcal{G} \to \mathcal{H}$, its kernel by $\ker L$, and its range by $\text{ran} L$. $I$ denotes the identity operator on $\mathcal{H}$. Let $D \subset \mathcal{H}$ be non-empty and let $T : D \to \mathcal{H}$. The set of fixed points of $T$ is given

---

**Journal of Global Optimization (2023) 85:847–866 849**
by Fix \( T = \{ x \in D \mid x = Tx \} \). Let \( \beta \in [0, +\infty[ \). The operator \( T \) is \( \beta \)-cocoercive if
\[
(\forall x \in D)(\forall y \in D) \quad (x - y \mid Tx - Ty) \geq \beta \|Tx - Ty\|^2,
\]
(2.1)
it is \( \beta \)-strongly monotone if
\[
(\forall x \in D)(\forall y \in D) \quad (x - y \mid Tx - Ty) \geq \beta \|x - y\|^2,
\]
(2.2)
it is nonexpansive if
\[
(\forall x \in D)(\forall y \in D) \quad \|Tx - Ty\| \leq \|x - y\|,
\]
(2.3)
it is quasinonexpansive if
\[
(\forall x \in D)(\forall y \in Fix T) \quad \|Tx - y\| \leq \|x - y\|,
\]
(2.4)
and it is firmly quasinonexpansive if
\[
(\forall x \in D)(\forall y \in Fix T) \quad \|Tx - y\|^2 \leq \|x - y\|^2 - \|Tx - x\|^2.
\]
(2.5)
Let \( \alpha \in [0, 1[ \). The operator \( T \) is \( \alpha \)-averaged nonexpansive if \( T = (1 - \alpha)I + \alpha R \) for some nonexpansive operator \( R : \mathcal{H} \to \mathcal{H} \), and \( T \) is firmly nonexpansive if it is \( \frac{1}{2} \)-averaged nonexpansive.

Given a self-adjoint monotone linear bounded operator \( V : \mathcal{H} \to \mathcal{H} \), we denote \( \langle \cdot \mid \cdot \rangle_{V} = \langle \cdot \mid V \cdot \rangle \), which is bilinear, positive semi-definite, and symmetric. Moreover, there exists a self-adjoint monotone linear bounded operator \( \sqrt{V} : \mathcal{H} \to \mathcal{H} \) such that
\[
V = \sqrt{V} \sqrt{V}, \quad (\forall x \in \mathcal{H}) \quad \langle x \mid Vx \rangle = \|\sqrt{V}x\|^2,
\]
(2.6)
and \( \text{ran} \ V = \text{ran} \sqrt{V} \). In addition, if \( V \) is strongly monotone, \( \langle \cdot \mid \cdot \rangle_{V} \) defines an inner product on \( \mathcal{H} \) and we denote by \( \| \cdot \|_V = \sqrt{\langle \cdot \mid \cdot \rangle_V} \) the induced norm.

Let \( A : \mathcal{H} \to 2^{\mathcal{H}} \) be a set-valued operator. The domain, range, and graph of \( A \) are \( \text{dom} \ A = \{ x \in \mathcal{H} \mid Ax \neq \emptyset \} \), \( \text{ran} \ A = \{ u \in \mathcal{H} \mid (\exists x \in \mathcal{H}) \ u \in Ax \} \), and \( \text{gra} \ A = \{ (x, u) \in \mathcal{H} \times \mathcal{H} \mid u \in Ax \} \), respectively. The set of zeros of \( A \) is \( \text{zer} \ A = \{ x \in \mathcal{H} \mid 0 \in Ax \} \), the inverse of \( A \) is \( A^{-1} : \mathcal{H} \to 2^{\mathcal{H}} : u \mapsto \{ x \in \mathcal{H} \mid u \in Ax \} \), and the resolvent of \( A \) is \( J_A = (I + A)^{-1} \). We have \( \text{zer} \ A = \text{Fix} \ J_A \). The operator \( A \) is monotone if
\[
(\forall (x, u) \in \text{gra} \ A)(\forall (y, v) \in \text{gra} \ A) \quad \langle x - y \mid u - v \rangle \geq 0
\]
(2.7)
and it is maximally monotone if it is monotone and there exists no monotone operator \( B : \mathcal{H} \to 2^{\mathcal{H}} \) such that \( \text{gra} \ B \) properly contains \( \text{gra} \ A \), i.e., for every \( (x, u) \in \mathcal{H} \times \mathcal{H} \),
\[
(x, u) \in \text{gra} \ A \iff (\forall (y, v) \in \text{gra} \ A) \quad \langle x - y \mid u - v \rangle \geq 0.
\]
(2.8)
Let \( C \) be a non-empty subset of \( \mathcal{H} \) and let \( (x_n)_{n \in \mathbb{N}} \) be a sequence in \( \mathcal{H} \). Then \( (x_n)_{n \in \mathbb{N}} \) is Fejér monotone with respect to \( C \) if
\[
(\forall x \in C)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - x\| \leq \|x_n - x\|.
\]
(2.9)
Let \( D \) be a non-empty weakly sequentially closed subset of \( \mathcal{H} \), i.e., the weak limit of every weakly convergent sequence in \( D \) is also in \( D \) [3, page 35], let \( T : D \to \mathcal{H} \), and let \( u \in \mathcal{H} \). Then \( T \) is demiclosed at \( u \) in \( (\mathcal{H}, (\cdot \mid \cdot)) \) if, for every sequence \( (x_n)_{n \in \mathbb{N}} \) in \( D \) and every \( x \in D \) such that \( x_n \to x \) and \( Tx_n \to u \) in \( (\mathcal{H}, (\cdot \mid \cdot)) \), we have \( Tx = u \). In addition, \( T \) is demiclosed if it is demiclosed at every point in \( D \).

We denote by \( \Gamma_0(\mathcal{H}) \) the class of proper lower semicontinuous convex functions \( f : \mathcal{H} \to ]-\infty, +\infty[ \). 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 \( I_0(\mathcal{H}) \). The subdifferential of \( f \) is the maximally monotone operator
\[ \partial f : x \mapsto \{ u \in \mathcal{H} \mid (\forall y \in \mathcal{H}) \ f(x) + \langle y - x | u \rangle \leq f(y) \}, \]
we have that \((\partial f)^{-1} = \partial f^*\), and that zero \( \partial f \) is the set of minimizers of \( f \), which is denoted by \( \arg \min_{x \in \mathcal{H}} f \). Given a strongly monotone self-adjoint linear operator \( \Upsilon : \mathcal{H} \to \mathcal{H} \), we denote by
\[ \text{prox}_f : x \mapsto \arg \min_{y \in \mathcal{H}} (f(y) + \frac{1}{2} \| x - y \|^2), \quad (2.10) \]
and by \( \text{prox}_f = \text{prox}_{\text{Id}} \). We have \( \text{prox}_f = J_{\Upsilon^{-1}} \partial f \) [3, Proposition 24.24(i)] and it is single valued since the objective function in (2.10) is strongly convex. Moreover, it follows from [3, Proposition 24.24] that
\[ \text{prox}_f = \text{Id} - \Upsilon^{-1} \text{prox}_{f^*} = \Upsilon^{-1} (\text{Id} - \text{prox}_{f^*}) \Upsilon. \quad (2.11) \]
Given a non-empty 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 [3].

### 2.2 Preliminaries: fixed points in the range of linear operators

The following result allows us to define algorithms in a real Hilbert space defined by the range of non-invertible self-adjoint monotone linear bounded operators. The result is a direct consequence of [3, Fact 2.26] and (2.6).

**Proposition 2.1** Let \( V : \mathcal{H} \to \mathcal{H} \) be a monotone self-adjoint linear bounded operator. The following statements are equivalent.

1. \( \text{ran} \ V \) is closed.
2. \((\exists \alpha > 0)(\forall x \in \text{ran} \ V) \ \langle Vx | x \rangle \geq \alpha \| x \|^2. \quad (2.12)\)

Moreover, if 1. or 2. holds, then \( \text{ran} \ V, \langle \cdot | \cdot \rangle_V \) is a real Hilbert space.

The following example exhibits a monotone self-adjoint linear bounded operator whose range is not closed.

**Example 2.2** Let \( \ell^2(\mathbb{R}) \) be the real Hilbert space defined by square summable sequences in \( \mathbb{R} \) endowed by the inner product \( \langle \cdot | \cdot \rangle : (x, y) \mapsto \sum_{j \geq 1} x_j y_j \) and consider the monotone self-adjoint bounded linear operator
\[ V : \ell^2(\mathbb{R}) \to \ell^2(\mathbb{R}) : (x_n)_{n \in \mathbb{N} \setminus \{0\}} \mapsto \left( x_1, \frac{x_2}{2}, \frac{x_3}{3}, \ldots \right). \]
By considering the sequence \((x^n)_{n \in \mathbb{N}} \subset \ell^2(\mathbb{R})\) defined by \( x^n_j = 1 \) for \( j \leq n \) and \( x^n_j = 0 \) for \( j > n \)... we have \( Vx^n \to y = (1/j)_{j \in \mathbb{N} \setminus \{0\}} \in \ell^2(\mathbb{R}) \) as \( n \to +\infty \), and \( y \notin \text{ran} \ V \), which implies that \( \text{ran} \ V \) is not closed.

The following fixed point problem is the basis for the analysis of primal-dual algorithms.
**Proposition 2.5** In the context of Problem 2.3 via their shadows, we first need the following technical lemma.

**Proof** Now we prove that Krasnosel’skii–Mann iterations defined by $S$ approximate the solutions in $\text{ran} V$. 

**Lemma 2.4** Let $Q : \mathcal{H} \to \mathcal{H}$ and let $S : \mathcal{H} \to \mathcal{H}$ be such that $\text{Fix} S \neq \emptyset$ and $S = S \circ Q$. Then $S(\text{Fix} (Q \circ S)) = \text{Fix} S$ and, in particular, $\text{Fix} (Q \circ S) \neq \emptyset$.

**Proof** See the appendix. □

Now we prove that Krasnosel’skii–Mann iterations defined by $S$ approximate the solutions to Problem 2.3 via their shadows in $\text{ran} V$.

**Proposition 2.5** In the context of Problem 2.3, define

$$T : \text{ran} V \to \text{ran} V : x \mapsto P_{\text{ran} V} \circ Sx$$

and consider the sequence $(x_n)_{n \in \mathbb{N}}$ defined by the recurrence

$$x_0 \in \mathcal{H}, \quad (\forall n \in \mathbb{N}) \quad x_{n+1} = (1 - \lambda_n) x_n + \lambda_n S x_n.$$  

(2.14)

Moreover, suppose that one of the following holds:

(i) $T$ is firmly quasinonexpansive, $\text{Id} - T$ is demiclosed at $0$ in $(\text{ran} V, \langle \cdot \mid \cdot \rangle_V)$, and $(\lambda_n)_{n \in \mathbb{N}}$ is a sequence in $[\varepsilon, 2 - \varepsilon]$ for some $\varepsilon \in ]0, 1[$.

(ii) $T$ is $\alpha$-averaged nonexpansive in $(\text{ran} V, \langle \cdot \mid \cdot \rangle_V)$ for some $\alpha \in ]0, 1[$ and $(\lambda_n)_{n \in \mathbb{N}}$ is a sequence in $[0, 1/\alpha]$ such that $\sum_{n \in \mathbb{N}} \lambda_n (1 - \alpha \lambda_n) = +\infty$.

Then the following hold:

1. $(P_{\text{ran} V} x_n)_{n \in \mathbb{N}}$ is Féjer monotone in $(\text{ran} V, \langle \cdot \mid \cdot \rangle_V)$ with respect to $\text{Fix} T$.
2. $(P_{\text{ran} V} (S x_n - x_n))_{n \in \mathbb{N}}$ converges strongly to $0$ in $(\text{ran} V, \langle \cdot \mid \cdot \rangle_V)$.
3. $(P_{\text{ran} V} x_n)_{n \in \mathbb{N}}$ converges weakly in $(\text{ran} V, \langle \cdot \mid \cdot \rangle_V)$ to some $\hat{x} \in \text{Fix} T$ and $S \hat{x}$ is a solution to Problem 2.3.
Remark 2.6

1. Previous results does not include summable errors for ease of the presentation, but they can be included effortlessly.

2. In the case when $V$ is strongly monotone, we have $\text{ran} V = \mathcal{H}$, $P_{\text{ran}} V = \text{Id}$, and Propositions 2.5(i) and 2.5(ii) are equivalent to [19, Theorem 5.2(i)] and [3, Proposition 5.16], respectively.

3. In [22, 26], a version of [3, Proposition 5.16] allowing for operators $S_k \in \mathbb{N}$ and $V_k \in \mathbb{N}$ varying among iterations is proposed. This modification allows to include variable step-sizes in primal-dual algorithms. In our context, the difficulty of including such generalization lies on the variation of the real Hilbert spaces $\text{ran} V_k, \langle \cdot | \cdot \rangle_{V_k} \in \mathbb{N}$, which complicates the asymptotic analysis.

3 Application to primal-dual algorithms for monotone inclusions

Now we focus on the asymptotic analysis of the relaxed primal-dual algorithm in (1.6) for solving Problem 1.1. First, note that $\mathcal{Z} = \text{zer} M$, where $M$, defined in (1.7), is maximally monotone in $(\mathcal{H}, \langle \cdot | \cdot \rangle)$, [8, Proposition 2.7(iii)], where

$$\mathcal{H} = \mathcal{H} \oplus \mathcal{G} \quad \text{and} \quad \langle \cdot | \cdot \rangle : ((x, u), (y, v)) \mapsto \langle x | y \rangle + \langle u | v \rangle.$$  

(3.1)

Consider the operator $V : \mathcal{H} \to \mathcal{H}$ defined in (1.8), where $\Sigma : \mathcal{G} \to \mathcal{G}$ and $\Upsilon : \mathcal{H} \to \mathcal{H}$ are strongly monotone self-adjoint linear operators such that $\| \sqrt{\Sigma} L \sqrt{\Upsilon} \| \leq 1$. In the case when $\| \sqrt{\Sigma} L \sqrt{\Upsilon} \| < 1$, $V$ is strongly monotone [23, eq. (6.15)] and the primal-dual algorithm is obtained by applying the proximal point algorithm (PPA) to the maximally monotone operator $V^{-1} M$ in the space $(\mathcal{H} \times \mathcal{G}, \langle \cdot | \cdot \rangle_{\mathcal{V}})$ [6, 22, 24, 34, 40, 45]. In the case when $\| \sqrt{\Sigma} L \sqrt{\Upsilon} \| = 1$, $V$ is no longer strongly monotone and $\langle \cdot | \cdot \rangle_{\mathcal{V}}$ does not define an inner product. However, if $\text{ran} V$ is closed, $(\text{ran} V, \langle \cdot | \cdot \rangle_{\mathcal{V}})$ is a real Hilbert space in view of Proposition 2.1, and we obtain the convergence of the primal-dual algorithm when $\| \sqrt{\Sigma} L \sqrt{\Upsilon} \| \leq 1$ in this Hilbert space using Proposition 2.5. The following result provides conditions on Problem 1.1 guaranteeing that $\text{ran} V$ is closed.

Proposition 3.1 In the context of Problem 1.1, set $\mathcal{H} = \mathcal{H} \oplus \mathcal{G}$, let $\Sigma : \mathcal{G} \to \mathcal{G}$ and $\Upsilon : \mathcal{H} \to \mathcal{H}$ be strongly monotone self-adjoint linear bounded operators such that $\| \sqrt{\Sigma} L \sqrt{\Upsilon} \| \leq 1$, and let $V$ be the operator defined in (1.8). Then, the following hold:
1. $V$ is linear, bounded, self-adjoint, and $\frac{\tau \sigma}{\tau + \sigma}$-cocoercive, where $\sigma > 0$ and $\tau > 0$ are the strongly monotone constants of $\Sigma$ and $\Upsilon$, respectively.

2. The followings statements are equivalent.

   (a) $\text{ran} V$ is closed in $\mathcal{H}$.
   (b) $\text{ran}(\Sigma^{-1} - L \Upsilon L^*)$ is closed in $G$.
   (c) $\text{ran}(\Upsilon^{-1} - L^* \Sigma L)$ is closed in $\mathcal{H}$.

**Proof** See the appendix. \qed

**Remark 3.2** 1. In the case when $\|\sqrt{\Sigma L \sqrt{\Upsilon}}\| < 1$, we have that $\Upsilon^{-1} - L \Sigma L^*$ is strongly monotone, and thus it is invertible. This is indeed an equivalence which follows from [13, eq. (2.7)]. Therefore, $\text{ran}(\Upsilon^{-1} - L \Sigma L^*) = G$ and $\text{ran} V$ is closed in view of Proposition 3.1.

2. Assume that $\text{ran} L = G$. Note that, for every $u \in G$, $(L \Upsilon L^* u | u) \geq \tau \|L^* u\|^2 \geq \tau \alpha^2 \|u\|^2$, where $\alpha > 0$ is the strong monotonicity parameter of $\Upsilon$ and the existence of $\Upsilon^{-1}$ is guaranteed by [3, Fact 2.26]. Hence, by setting $\Sigma = (L \Upsilon L^*)^{-1}$, we have $\Sigma^{-1} - L \Upsilon L^* = 0$. Hence, $\text{ran}(\Sigma^{-1} - L \Upsilon L^*) = \{0\}$ which is closed and Proposition 3.1 implies that $\text{ran} V$ is closed. This case arises in wavelets transformations in image and signal processing (see, e.g., [36]).

The next theorem is the main result of this section, in which we interpret the primal-dual splitting as a KM iteration on the real Hilbert space $(\text{ran} V, \langle \cdot | \cdot \rangle_V)$ when $\text{ran} V$ is closed. This iteration involves the resolvent of the primal-dual operator

$$
W : \mathcal{H} \to 2^\mathcal{H} : (x, u) \mapsto \{ (y, v) \in \mathcal{H} \mid V(y, v) \in M(x, u) \},
$$

(3.2)

where $M$ and $V$ are defined in (1.7) and (1.8), respectively. Note that, in the case when $\|\sqrt{\Sigma L \sqrt{\Upsilon}}\| < 1$, $V$ is invertible and $W = V^{-1} M$, which is maximally monotone in $(\mathcal{H}, \langle \cdot | \cdot \rangle_V)$, in view of [3, Proposition 20.24]. This implies that $J_W$ is firmly nonexpansive under the same metric [3, Proposition 23.8(iii)]. These properties do not hold when $\|\sqrt{\Sigma L \sqrt{\Upsilon}}\| = 1$, but $P_{\text{ran} V} \circ J_W$ is firmly nonexpansive in the real Hilbert space $(\text{ran} V, \langle \cdot | \cdot \rangle_V)$, from which the weak convergence of the primal-dual algorithm is obtained.

**Theorem 3.3** In the context of Problem 1.1, let $V$ be the operator defined in (1.8), where $\Sigma : \mathcal{G} \to \mathcal{G}$ and $\Upsilon : \mathcal{H} \to \mathcal{H}$ are self-adjoint linear strongly monotone operators such that $\|\sqrt{\Sigma L \sqrt{\Upsilon}}\| \leq 1$, and suppose that $\text{ran} V$ is closed. Moreover, let $\{(\lambda_n)_{n \in \mathbb{N}} \}$ be a sequence in $[0, 2]$ satisfying $\sum_{n \in \mathbb{N}} \lambda_n (2 - \lambda_n) = +\infty$, and consider the sequence $\{(x_n, u_n)\}_{n \in \mathbb{N}}$ defined by the recurrence

$$
\begin{align*}
(\forall n \in \mathbb{N}) & \quad p_{n+1} = J_{\Upsilon A}(x_n - \Upsilon L^* u_n) \\
q_{n+1} = J_{\Sigma B^{-1}}(u_n + \Sigma L(2p_{n+1} - x_n)) \\
(x_{n+1}, u_{n+1}) = (1 - \lambda_n)(x_n, u_n) + \lambda_n (p_{n+1}, q_{n+1}),
\end{align*}
$$

(3.3)

where $(x_0, u_0) \in \mathcal{H} \times \mathcal{G}$. Then $(P_{\text{ran} V}(x_n, u_n))_{n \in \mathbb{N}}$ converges weakly in $(\text{ran} V, \langle \cdot | \cdot \rangle_V)$ to some $(\hat{y}, \hat{v}) \in \text{Fix}(P_{\text{ran} V} \circ J_W)$, where $W$ is defined in (3.2). Moreover,

$$
(J_{\Upsilon A}(\hat{v} - \Upsilon L^* \hat{v}), J_{\Sigma B^{-1}}(\hat{v} + \Sigma L(2J_{\Upsilon A}(\hat{v} - \Upsilon L^* \hat{v}) - \hat{v})))
$$

(3.4)

is a solution to Problem 1.1.

**Proof** First, it follows from Proposition 3.1(1.) that $V$ is a monotone self-adjoint linear bounded operator. Note that

$$
\text{Fix} J_W = \text{zer} W = \text{zer} M = Z \neq \emptyset
$$

(3.5)
and, for every \((x, u)\) and \((p, q)\) in \(\mathcal{H}\),
\[
(p, q) \in J_W(x, u) \iff (x - p, u - q) \in W(p, q) \\
\iff V(x - p, u - q) \in M(p, q) \\
\iff \begin{cases} 
\gamma^{-1}(x - p) - L^*(u - q) \in Ap + L^*q, \\
\Sigma^{-1}(u - q) - L(x - p) \in B^{-1}q - Lp.
\end{cases}
\]
\quad \quad \quad \quad \quad \quad \quad \quad \quad (3.6)

Hence, \(J_W\) is single valued and, for every \((x, u)\) in \(\mathcal{H}\),
\[
J_W(x, u) = (J_{\gamma A}(x - \gamma L^*u), J_{\Sigma B^{-1}}(u - \Sigma L(x - p)))
\]
\[
= R(\gamma^{-1}x - L^*u, \Sigma^{-1}u - Lx)
\]
where \(R: (x, u) \mapsto (J_{\gamma A}(\gamma x), J_{\Sigma B^{-1}}(\Sigma u + 2\Sigma LJ_{\gamma A}(\gamma x)))\). Therefore
\[
J_W = R \circ V = R \circ V \circ P_{ran V} = J_W \circ P_{ran V}.
\] (3.7)

Moreover, define \(T = P_{ran V} \circ J_W\), let \(z\) and \(w\) in \(\text{ran} V\), and note that (3.2) yields \((J_Wz, V(z - J_Wz)) \in \text{gra} M\) and \((J_Ww, V(w - J_Ww)) \in \text{gra} M\). Then we deduce from \(V = V \circ P_{ran V}\), ker \(V \oplus \text{ran} V = \mathcal{H}\), and the monotonicity of \(M\) that
\[
\langle Tz - Tw, (I - T)z - (I - T)w \rangle_v
\]
\[
= \langle P_{ran V}(J_Wz - J_Ww) \mid V(z - P_{ran V}J_Wz - w + P_{ran V}J_Ww) \rangle_v
\]
\[
= \langle P_{ran V}(J_Wz - J_Ww) \mid V(z - J_Wz - w + J_Ww) \rangle_v
\]
\[
= \langle J_Wz - J_Ww \mid V(z - J_Wz - W(w - J_Ww)) \rangle_v
\]
\[
\geq 0,
\] (3.8)
which yields the firm nonexpansivity of \(T\) in \((\text{ran} V, \langle \cdot | \cdot \rangle_v)\). Therefore, it follows from (3.5) and (3.7) that Problem 1.1 is a particular instance of Problem 2.3 with \(S = J_W\). In addition, (3.3) and (3.6) yield
\[
(\forall n \in \mathbb{N}) \quad x_{n+1} = (1 - \lambda_n)x_n + \lambda_n J_Wx_n,
\] (3.9)
where, for every \(n \in \mathbb{N}\), \(x_n = (x_n, u_n)\). Altogether, we obtain the results by applying Theorem 2.5(ii) with \(\alpha = 1/2\) and \(S = J_W\).

**Remark 3.4** 1. Since \(J_W \circ P_{ran V} = J_W\), the sequence \(\{P_{ran V}(x_n, u_n)\}_{n \in \mathbb{N}}\) is not needed in practice. Indeed, since
\[
(\forall x \in \mathcal{H}) \quad \|x\|_V = \|P_{ran V}x\|_V,
\]
we can use a stopping criteria only involving \((x_n, u_n)\) and not \(P_{ran V}(x_n, u_n)\).

2. Suppose that \(G = \bigoplus_{i=1}^m G_i, B: (u_i)_{1 \leq i \leq m} \mapsto \times_{i=1}^m B_i u_i, \Sigma: (u_i)_{1 \leq i \leq m} \mapsto (\Sigma_i u_i)_{1 \leq i \leq m}\), and \(L: x \mapsto (L_i x)_{1 \leq i \leq m}\), where, for every \(i \in \{1, \ldots, m\}\), \(G_i\) is a real Hilbert space, \(B_i\) is maximally monotone, \(\Sigma_i: G_i \to G_i\) is a strongly monotone self-adjoint linear bounded operator, and \(L_i: \mathcal{H} \to \mathcal{G}_i\) is a linear bounded operator. In this context, the inclusion in (1.2) is equivalent to
\[
\text{find } x \in \mathcal{H} \text{ such that } 0 \in Ax + \sum_{i=1}^m L_i^* B_i L_i x.
\] (3.10)
Then, under the assumptions
\[
\sum_{i=1}^{m} \left\| \sqrt{\Sigma} L_i \sqrt{\Upsilon} \right\|^2 \leq 1 \quad \text{and} \quad \text{ran} \left( \Upsilon^{-1} - \sum_{i=1}^{m} L_i^* \Sigma L_i \right) \text{ is closed,} \quad (3.11)
\]

Proposition 3.1 and Theorem 3.3 ensures the convergence of (3.3), which reduces to
\[
(\forall n \in \mathbb{N}) \begin{cases}
    p_{n+1} = J_{T^A}(x_n - \mathcal{T} \sum_{i=1}^{m} L_i^* u_{i,n}) \\
    x_{n+1} = (1 - \lambda_n)x_n + \lambda_n p_{n+1} \\
    q_{i,n+1} = J_{\Sigma_i B_i^{-1}}(u_{i,n} + \Sigma_i L_i (2p_{n+1} - x_n)) \\
    u_{i,n+1} = (1 - \lambda_n)u_{i,n} + \lambda_n q_{i,n+1}.
\end{cases} \quad (3.12)
\]

Note that (3.12) has the same structure than the algorithm in [23, Corollary 6.2] without considering cocoercive operators and the convergence is guaranteed under the weaker assumption (3.11) in view of Remark 3.2(1).

3. In the context of the optimization problem in (1.4), (3.3) reduces to
\[
(\forall n \in \mathbb{N}) \begin{cases}
    p_{n+1} = \text{prox}_{f}^{\frac{1}{\mathcal{T}}}(x_n - \mathcal{T} L^* u_n) \\
    q_{n+1} = \text{prox}_{g^*}^{\frac{1}{\Sigma}}(u_n + \Sigma (2p_{n+1} - x_n)) \\
    (x_{n+1}, u_{n+1}) = (1 - \lambda_n)(x_n, u_n) + \lambda_n (p_{n+1}, q_{n+1}).
\end{cases} \quad (3.13)
\]

Under the additional condition \( \text{ran} V \) closed, Theorem 3.3 generalizes [24, Theorem 3.3] to infinite dimensional spaces and allowing preconditioners and a larger choice of parameters \((\lambda_n)_{n \in \mathbb{N}}\). Indeed, in finite dimensional spaces, \( \text{ran} V \) is closed, Theorem 3.3 implies that \( P_{\text{ran} V}(x_n, u_n) \rightarrow (\hat{y}, \hat{v}) \in \text{ran} V \) in \( (\mathcal{V}, \langle \cdot | \cdot \rangle_V) \) and, since \( J_W = J_{W} \circ P_{\text{ran} V} \) is continuous, we conclude \((p_{n+1}, q_{n+1}) = J_W(x_n, u_n) = J_W(P_{\text{ran} V}(x_n, u_n)) \) \( \rightarrow J_W(\hat{y}, \hat{v}) \in \mathcal{Z} \). In order to guarantee the convergence of the relaxed sequence \(((x_n, u_n))_{n \in \mathbb{N}}\), it is enough to suppose \((\lambda_n)_{n \in \mathbb{N}} \subset [\epsilon, 2 - \epsilon] \) for some \( \epsilon \in [0, 1] \), and use the argument in [24, p.473].

4. In the particular case when \( \| \sqrt{\Sigma} L \sqrt{\Upsilon} \| < 1 \), it follows from [23, eq. (6.15)] (see also [40, Lemma 1]) that \( V \) is strongly monotone, which yields which yields \( \text{ran} V = \mathcal{H} \) and \( P_{\text{ran} V} = \text{Id} \). Hence, we recover from Theorem 3.3 the weak convergence of \(((x_n, u_n))_{n \in \mathbb{N}}\) to a solution to Problem 1.1 proved in [6, 22, 24, 34, 40, 45].

5. In the particular instance when \( \lambda_n \equiv 1 \), the weak convergence of (3.3) is deduced in [13, Remark 3.4(4)] without any range closedness. The result is obtained from an alternative formulation of the algorithm and the extension to \( \lambda_n \neq 1 \) is not clear. As we will show in Sect. 4, the additional relaxation step is relevant in the efficiency of the algorithm.

### 3.1 Case \( L = \text{Id} \: \text{Douglas–Rachford splitting (DRS)} \)

In this section, we study the particular case of Problem 1.1 when \( L = \text{Id} \). In this context, the following result is a refinement of Theorem 3.3, which relates the primal-dual algorithm in (3.3) with DRS when
\[
\Upsilon = \Sigma^{-1} \text{ is strongly monotone.} \quad (3.14)
\]

When \( \Upsilon = \tau \text{Id} \) and \( \Sigma = \sigma \text{Id} \), (3.14) reads \( \sigma \tau = 1 \) and the connection of (3.3) with DRS is discovered in [17, Section 4.2] in the optimization context. However, the convergence is guaranteed only if \( \tau \sigma < 1 \), which is extended to the case \( \sigma \tau = 1 \) in [24, Section 3.1.3] in the
finite dimensional setting. Previous connection allows us to recover the classical convergence results in [27, 35] when \( \Upsilon = \tau \text{Id} \) with our approach. Define the operator

\[
G_{\Upsilon, B, A} = J_{TB} \circ (2J_{TA} - \text{Id}) + (\text{Id} - J_{TA}),
\]

and we recall that relaxed DRS iterations are defined by the recurrence

\[
z_0 \in \mathcal{H}, \quad (\forall n \in \mathbb{N}) \quad z_{n+1} = (1 - \lambda_n)z_n + \lambda_n G_{\Upsilon, B, A}z_n,
\]

where \((\lambda_n)_{n \in \mathbb{N}}\) is a sequence in \([0, 2]\).

**Proposition 3.5** In the context of Problem 1.1, set \( L = \text{Id} \), let \( \Upsilon \) be a strongly monotone self-adjoint linear operator, let \((\lambda_n)_{n \in \mathbb{N}}\) be a sequence in \([0, 2]\) satisfying \( \sum_{n \in \mathbb{N}} \lambda_n (2 - \lambda_n) = +\infty \), and consider the sequence \(( (x_n, u_n) )_{n \in \mathbb{N}} \) defined by the recurrence

\[
(\forall n \in \mathbb{N}) \quad \begin{cases}
p_{n+1} = J_{TA}(x_n - \Upsilon u_n) \\
q_{n+1} = J_{\Upsilon^{-1}B-1}(u_n + \Upsilon^{-1}(2p_{n+1} - x_n)) \\
x_{n+1}, u_{n+1} = (1 - \lambda_n)(x_n, u_n) + \lambda_n(p_{n+1}, q_{n+1}),
\end{cases}
\]

where \((x_0, u_0) \in \mathcal{H} \times \mathcal{H}\). Then, by setting, for every \( n \in \mathbb{N} \), \( z_n = x_n - \Upsilon y_n \), \((z_n)_{n \in \mathbb{N}}\) converges weakly in \( \mathcal{H} \) to some \( \hat{z} \in \text{Fix} G_{\Upsilon, B, A} \) and

\[
( J_{TA} \hat{z}, -\Upsilon^{-1}(\hat{z} - J_{TA} \hat{z}) )
\]

is a solution to Problem 1.1. Moreover, we have

\[
(\forall n \in \mathbb{N}) \quad z_{n+1} = (1 - \lambda_n)z_n + \lambda_n G_{\Upsilon, B, A}z_n.
\]

**Proof** Note that, since \( L = \text{Id} \), we have that \( \text{ran} L = G = \mathcal{H} \), and \( V : (x, u) \mapsto (\Upsilon^{-1}x - u, \Upsilon u - x) \) in view of (1.8) and (3.14). Therefore, Remark 3.2(2.) implies that \( \text{ran} V \) is closed. Hence, it follows from Theorem 3.3 and (3.6) in the case \( L = \text{Id} \) that \( (P_{\text{ran} V}(x_n, u_n))_{n \in \mathbb{N}} \) converges weakly in \( \text{ran} V \) to some \((\hat{\gamma}, \hat{v}) \in \text{Fix} (P_{\text{ran} V} \circ J_W) \)

\[
( \hat{\gamma}, \hat{v} ) = J_W (\hat{\gamma}, \hat{v}) = ( J_{TA}(\hat{\gamma} - \Upsilon \hat{v}), J_{\Upsilon^{-1}B-1}(\hat{v} + \Upsilon^{-1}(2\hat{\gamma} - \hat{\gamma})) ) \in \mathbb{Z}.
\]

Now set \( \Lambda : (x, u) \mapsto x - \Upsilon u \) and \( \Upsilon : (x, u) \mapsto (\Upsilon x, \Upsilon u) \). Note that \( \Lambda \) is surjective, that

\[
\Lambda^* \Lambda = V \circ \Upsilon, \quad \text{ran} V = \text{ran} \Lambda^*,
\]

and, in view of [3, Fact 2.25(iv)], that

\[
\mathcal{H} \times \mathcal{H} = \text{ran} \Lambda^* \oplus \ker \Lambda.
\]

Then, for every \((x, u) \in \mathcal{H} \times \mathcal{H}\), it follows from (3.2) and (3.6) in the case \( L = \text{Id} \), (3.14), [3, Proposition 23.34(iii)], and (3.15) that

\[
\Lambda (J_W(x, u)) = J_{TA} (x - \Upsilon u) - \Upsilon J_{\Upsilon^{-1}B-1}(\Upsilon u - x + 2J_{TA}(x - \Upsilon u))
\]

\[
= -J_{TA}(\Lambda(x, u)) + \Lambda(x, u) + J_{TB}(2J_{TA}(\Lambda(x, u)) - \Lambda(x, u))
\]

\[
= G_{\Upsilon, B, A}(\Lambda(x, u)).
\]

Moreover, since \((\hat{\gamma}, \hat{v}) \in \text{Fix} (P_{\text{ran} V} \circ J_W)\), by using (3.22), (3.21), and (3.20) we deduce

\[
G_{\Upsilon, B, A}(\Lambda(\hat{\gamma}, \hat{v})) = \Lambda(J_W(\hat{\gamma}, \hat{v}))
\]

\[
\Lambda \circ P_{\text{ran} \Lambda}^*(J_W(\hat{\gamma}, \hat{v}))
\]

\[
\Lambda(P_{\text{ran} V} \circ J_W(\hat{\gamma}, \hat{v}))
\]

\[
= \Lambda(\hat{\gamma}, \hat{v}).
\]
In turn, by setting $\hat{z} = \Lambda(\hat{y}, \hat{v})$, we obtain $\hat{z} \in \text{Fix } G_{\mathcal{Y}, B, A}$ and $\hat{x} = J_{\mathcal{T}A} \hat{z}$ in view of (3.19). In addition, since $(\hat{y}, \hat{v}) \in \text{Fix } (P_{\text{tan}} V \circ J_{W})$, we deduce from (3.21) and (3.6) that

$$\hat{z} = \Lambda(\hat{y}, \hat{v}) = \Lambda \left( P_{\text{tan}} V \circ J_{W}(\hat{y}, \hat{v}) \right) = \Lambda J_{W}(\hat{y}, \hat{v}) = \Lambda(\hat{x}, \hat{u}) = \hat{x} - \mathcal{T}\hat{u}, \quad (3.24)$$

which yields $\hat{u} = -\mathcal{T}^{-1}(\hat{x} - J_{\mathcal{T}A} \hat{z})$. Furthermore, noting that, for every $n \in \mathbb{N}$, $z_n = \Lambda(x_n, u_n)$, we deduce from (3.17), (3.6), and (3.22) that

$$\begin{align*}
(\forall n \in \mathbb{N}) & \quad z_{n+1} = \Lambda(x_{n+1}, u_{n+1}) \\
& = (1 - \lambda_n) \Lambda(x_n, u_n) + \lambda_n \Lambda(J_{W}(x_n, u_n)) \\
& = (1 - \lambda_n)z_n + \lambda_n G_{\mathcal{T}\Lambda, B, A}z_n. \quad (3.25)
\end{align*}$$

Finally, in order to prove the weak convergence of $(z_n)_{n \in \mathbb{N}}$ to $\hat{z}$, fix $w \in \mathcal{H}$ and set $(p, q) = ((\text{Id} + \mathcal{T}^2)^{-1} w, -\mathcal{T}(\text{Id} + \mathcal{T}^2)^{-1} w)$. We have $(p, q) \in \text{ran } \Lambda^*$, $\Lambda(p, q) = w$ and it follows from (3.21), (3.20), $(\hat{y}, \hat{v}) \in \text{ran } V = \text{ran } \Lambda^*$, and $(P_{\text{tan}} V(x_n, u_n))_{n \in \mathbb{N}} \rightharpoonup (\hat{y}, \hat{v})$ that

$$\begin{align*}
\langle z_n - \hat{z} | w \rangle &= \langle \Lambda(x_n - \hat{y}, u_n - \hat{v}) | \Lambda(p, q) \rangle \\
&= \langle \Lambda P_{\text{tan}} \Lambda^*(x_n - \hat{y}, u_n - \hat{v}) | \Lambda(p, q) \rangle \\
&= \langle P_{\text{tan}} \Lambda^*(x_n - \hat{y}, u_n - \hat{v}) | V(\mathcal{T}p, \mathcal{T}q) \rangle \\
&= \langle P_{\text{tan}} \Lambda^*(x_n - \hat{y}, u_n - \hat{v}) | (\mathcal{T}p, \mathcal{T}q) \rangle V \\
&\rightharpoonup \langle P_{\text{tan}} \Lambda^*(x_n, u_n) - (\hat{y}, \hat{v}) | (\mathcal{T}p, \mathcal{T}q) \rangle V \\n&\rightarrow 0. \quad (3.26)
\end{align*}$$

The proof is hence complete. $\square$

**Remark 3.1**

1. From the proof of Proposition 3.5, we deduce $\Lambda(\text{Fix } (P_{\text{tan}} V \circ J_{W})) \subset \text{Fix } G_{\mathcal{T}\Lambda, B, A}$. The converse inclusion is also true, as detailed in Proposition 6.1 in the Appendix.

2. Proposition 3.5 provides a connection between classical DRS scheme [27] and the primal-dual version in (3.17), and we obtain that the auxiliary sequence $(z_n)_{n \in \mathbb{N}}$ converges weakly to a $\hat{z}$ whose primal-dual shadow is a primal-dual solution. In [44] (see also [2, 4]) the weak convergence of the primal-dual shadow sequence is proved in the case $\lambda_n \equiv 1$, by reformulating DRS as an alternative algorithm with primal-dual iterates in gra $A$. This technique does not allow for relaxation steps, since after relaxation the iterates are no longer in gra $A$ unless it is affine linear.

### 4 Numerical experiments

A classical model in image processing is the total variation image restoration problem [42], which aims at recovering an image from a blurred and noisy observation under piecewise constant assumption on the solution. The model is formulated via the optimization problem

$$\min_{x \in [0, 255]^N} \frac{1}{2} \|Rx - b\|^2_2 + \alpha \|
abla x\|_1 =: F^T(x),$$

where $x \in [0, 255]^N$ is the image of $N = N_1 \times N_2$ pixels to be recovered from a blurred and noisy observation $b \in \mathbb{R}^m$, $R : \mathbb{R}^N \rightarrow \mathbb{R}^m$ is a linear operator representing a Gaussian blur, the discrete gradient $\nabla : x \mapsto (D_1 x, D_2 x)$ includes horizontal and vertical differences through linear operators $D_1 : \mathbb{R}^N \rightarrow \mathbb{R}^N$ and $D_2 : \mathbb{R}^N \rightarrow \mathbb{R}^N$, respectively, its adjoint $\nabla^*$ is the discrete divergence (see, e.g., [15]), and $\alpha \in [0, +\infty[$. A difficulty in this model is the
presence of the non-smooth $\ell^1$ norm composed with the discrete gradient operator $\nabla$, which is non-differentiable and its proximity operator has not a closed form.

Note that, by setting $f = \|R \cdot b\|^2/2$, $g_1 = \alpha \cdot 1 = g_2$, and $g_3 = \delta_{0,255}^\vee$, $L_1 = D_1$, $L_2 = D_2$, and $L_3 = \text{Id}$, (4.1) can be reformulated as $\min(f + \sum_{i=1}^3 g_i \circ L_i)$ or equivalently as (qualification condition holds)

$$ \text{find } x \in \mathbb{R}^N \text{ such that } 0 \in \partial f(x) + \sum_{i=1}^3 L_i^* g_i(L_i x), $$

(4.2)

which is a particular instance of (3.10), in view of [3, Theorem 20.25]. Moreover, for every $\tau > 0$, $J_{\tau} f = (\text{Id} + \tau R^* R)^{-1}(\text{Id} - \tau R^* b)$, for every $i \in \{1, 2, 3\}$, $J_{\tau} (\partial g_i)^{-1} = \tau (\text{Id} - \text{prox}_{g_i/\tau}(\text{Id} / \tau))$, $\text{prox}_{g_3/\tau} = P_{[0,255]^N}$, and, for $i \in \{1, 2\}$, $\text{prox}_{g_i/\tau} = \text{prox}_{\|\cdot\|_1/\tau}$ is the component-wise soft thresholder, computed in [3, Example 24.34]. Note that $(\text{Id} + \tau R^* R)^{-1}$ can be computed efficiently via a diagonalization of $R$ using the fast Fourier transform $F$ [33, Section 4.3]. Altogether, Remark 3.4.(2.) allows us to write algorithm in (3.12) as Algorithm 4.1 below, where we set $\gamma = \tau \text{Id}$, $\Sigma_1 = \sigma_1 \text{Id}$, $\Sigma_2 = \sigma_2 \text{Id}$, and $\Sigma_3 = \sigma_3 \text{Id}$, for $\tau > 0$, $\sigma_1 > 0$, $\sigma_2 > 0$, and $\sigma_3 > 0$. We denote by $\mathcal{R}$ the primal-dual error

$$ \mathcal{R}: (x_+, u_+, x, u) \mapsto \sqrt{\|x_+ - x\|^2 + \|u_+ - u\|^2 / \|x\|^2 + \|u\|^2} $$

(4.3)

and by $\varepsilon > 0$ the convergence tolerance.

**Algorithm 4.1:**

1: Fix $x_0, u_{1,0}, u_{2,0},$ and $u_{3,0}$ in $\mathbb{R}^N$, let $\tau, \sigma_1, \sigma_2,$ and $\sigma_3$ be in $[0, +\infty[$, let $(\lambda_n)_{n \in \mathbb{N}}$ in $[0, 2]$ such that $\sum_{n \in \mathbb{N}} \lambda_n (2 - \lambda_n) = +\infty$, and fix $r_0 > \varepsilon > 0$.

2: while $r_n > \varepsilon$ do

3: $p_{n+1} = (\text{Id} + \tau R^* R)^{-1}(x_n - \tau (D_1^* u_{1,n} + D_2^* u_{2,n} + u_{3,n} + R^* b))$

4: $x_{n+1} = (1 - \lambda_n)x_n + \lambda_n p_{n+1}$

5: $q_{1,n+1} = \sigma_1 (\text{Id} - \text{prox}_{\|\cdot\|_1/\sigma_1})(u_{1,n}/\sigma_1 + D_1(2p_{n+1} - x_n))$

6: $q_{2,n+1} = \sigma_2 (\text{Id} - \text{prox}_{\|\cdot\|_1/\sigma_2})(u_{2,n}/\sigma_2 + D_2(2p_{n+1} - x_n))$

7: $q_{3,n+1} = \sigma_3 (\text{Id} - P_{[0,255]^N})(u_{3,n}/\sigma_3 + 2p_{n+1} - x_n)$

for $i = 1, 2, 3$

8: $u_{i,n+1} = (1 - \lambda_n)u_{i,n} + \lambda_n q_{i,n+1}$

9: $r_n = \mathcal{R}(x_{n+1}, (u_{1,n+1}, u_{2,n+1}, u_{3,n+1}), x_n, (u_{1,n}, u_{2,n}, u_{3,n}))$

10: end while

11: return $(x_{n+1,1}, u_{1,n+1,1}, u_{2,n+1,1}, u_{3,n+1})$

In this case, (3.11) reduces to

$$ \tau (\sigma_1 \|D_1\|^2 + \sigma_2 \|D_2\|^2 + \sigma_3) \leq 1 $$

(4.4)

and the closed range condition is trivially satisfied. In view of [14, Theorem 3.1] we set $\|D_1\| = \|D_2\| = 2$

Observe that, when $\sigma_1 = \sigma_2 = \sigma_3 = \sigma$, Algorithm 4.1 reduces to the algorithm proposed in [17] (when $\sigma \tau (\|D_1\|^2 + \|D_2\|^2 + 1) < 1$) or [24, Theorem 3.3] (algorithm denoted by $\text{condat}$), where the case $\sigma \tau (\|D_1\|^2 + \|D_2\|^2 + 1) = 1$ is included.

Since in [13, Section 5.1], the critical step-sizes achieve the best performance, we provide a numerical experiment which compare the efficiency of Algorithm 4.1 for different values of the parameters $\tau, \sigma_1, \sigma_2,$ and $\sigma_3$ in the boundary of (4.4) and different relaxation parameters $\lambda_n$. In particular we compare with the case $\sigma_1 = \sigma_2 = \sigma_3 = \sigma$ ($\text{condat}$), which turns
Table 1  Averages number of iterations for Algorithm 4.1 with \( \tau(\sigma_1\|D_1\|^2 + \sigma_2\|D_1\|^2 + \sigma_3) = 1 \) and condat with tolerance \( 10^{-8} \)

| Algorithm | \( \tau \) | \( \sigma_1 \) | \( \sigma_2 \) | \( \lambda_n \) | \( \varepsilon = 10^{-8} \) | Av. Time(s) | Av. Iter. |
|-----------|-----------|-----------|-----------|-------------|----------------|-------------|----------|
| Alg. 4.1  | 0.2       | 0.7425    | 0.4950    | 1           | 82.6373        | 8844        |
|           | 0.2       | 0.7463    | 0.4975    | 1           | 82.2817        | 8827        |
|           | 0.2       | 0.7493    | 0.4995    | 1           | 82.5722        | 8833        |
|           | 0.2       | 0.7425    | 0.4950    | 1.5         | 63.1338        | 6766        |
|           | 0.2       | 0.7463    | 0.4975    | 1           | 63.0645        | 6754        |
|           | 0.2       | 0.7493    | 0.4995    | 1           | 63.0996        | 6758        |
|           | 0.2       | 0.8044    | 0.4331    | 1.9         | 53.9059        | 5770        |
|           | 0.2       | 0.8085    | 0.4353    | 1           | 53.8663        | 5767        |
|           | 0.2       | 0.8117    | 0.4371    | 1           | 53.8022        | 5761        |
| condat    | 0.2       | -         | -         | 1           | 92.7997        | 9326        |
|           | 0.2       | -         | -         | 1.5         | 67.0886        | 7131        |
|           | 0.2       | -         | -         | 1.9         | 57.8523        | 6121        |

Fig. 1  Comparison of Algorithm 4.1 with \( \tau(\sigma_1\|D_1\|^2 + \sigma_2\|D_1\|^2 + \sigma_3) = 1 \) and condat (observation \( b_4 \))

out to be more efficient than other methods as AFBS [39], MS [8], Condat-Vũ [24, 45] in this context [13, Section 5.1]. For these comparisons, we consider the test image \( \mathbf{x} \) shown in Fig. 2a of 256 \( \times \) 256 pixels \( (N_1 = N_2 = 256) \) inspired in [46, Section 5]. The blur operator \( R \) is set as a Gaussian blur of size 9 \( \times \) 9 and standard deviation 4 (applied by MATLAB function \( \text{fspecial} \)) and the observation \( b \) is obtained by \( b = R\mathbf{x} + e \in \mathbb{R}^{m_1 \times m_2} \), where \( m_1 = m_2 = 256 \) and \( e \) is an additive zero-mean white Gaussian noise with standard deviation \( 10^{-3} \) (using \( \text{imnoise} \) function in MATLAB). We generate 20 random realizations of the random variable \( e \) leading to 20 observations \( (b_i)_{1 \leq i \leq 20} \).

We study the efficiency of Algorithm 4.1 for different values of \( \tau, \sigma_1, \sigma_2, \sigma_3 \), and relaxation steps \( \lambda_n \equiv \lambda \in \{1, 1.5, 1.9\} \). In order to approximate the best performant step-sizes in the boundary of (4.4), we consider \( \tau \in \mathcal{C} := \{0.10 + 0.05 \cdot n\}_{n=0,10} \) and, for condat, we set \( \sigma_1 = \sigma_2 = \sigma_3 = \sigma = \tau/(1 + \|D_1\|^2 + \|D_2\|^2) \). In the case of Algorithm 4.1, we consider \( \sigma_1 = \gamma_1(1 - \gamma_2)/(\tau\|D_1\|^2), \sigma_2 = (1 - \gamma_1)(1 - \gamma_2)/(\tau\|D_2\|^2), \sigma_3 = \gamma_2/\tau \), where \( (\tau, \gamma_1, \gamma_2) \in \mathcal{C} \times [0.01, 0.005, 0.001] \times [0.5, 0.55, 0.6, 0.65] \).
In Table 1 we provide the average number of iterations obtained by applying Algorithm 4.1 for solving (4.1) considering the 20 observations \((b_i)_{1 \leq i \leq 20}\) and the best set of step-sizes found with the procedure described above. The tolerance is set as \(\varepsilon = 10^{-8}\). We observe that we can save up to 11, 3\%\(^1\) in computational time of Algorithm 4.1 if we allow for different parameters \(\sigma_1\), \(\sigma_2\), and \(\sigma_3\). Note that the set of step-sizes achieving the best performance gives more importance to horizontal differences, which is natural given the choice of the image (vertical black/white strips). In addition, Algorithm 4.1 becomes more efficient in iterations as long as the relaxation parameters are larger. The case \(\lambda = 1.9\) achieves the tolerance in approximately 35\% less iterations and computational time than the case \(\lambda = 1\). Therefore,

\(^{1}\) The computational saving time percentage of algorithm \(B\) with respect to algorithm \(A\) is given by \(100\times(\text{time}(A) - \text{time}(B))/\text{time}(A)\).
combining the use of critical preconditioners and relaxation steps, we can save up to 42% of computational time.

This conclusion is confirmed in Fig. 1, which shows the performance obtained with the observation $b_4$. This figure also shows that both algorithms achieve in less iterations the optimal objective value for higher relaxation parameters, with a slight advantage of Algorithm 4.1. Note that, since the algorithms under study has the same structure, the CPU time by iteration is very similar.

In Fig. 2 we provide the images reconstructed from observation $b_4$ by using condat and Algorithm 4.1 after 300 iterations. The best reconstruction, in terms of objective value $F^{TV}$ and PSNR (Peak signal-to-noise ratio) is obtained by Algorithm 4.1.

5 Conclusions

We study the primal-dual algorithm with relaxation steps and critical preconditioners for solving composite monotone inclusions. By writing this algorithm as the KM iterations in the range of a particular linear operator, we derive the weak convergence of the projection of the primal-dual sequence, from which we can recover a solution to the monotone primal-dual inclusion. Our results generalize [24, 40, 45] to infinite dimensional spaces including critical preconditioners. The DRS is interpreted as a particular instance of the primal-dual algorithm when the step-sizes are critical, deriving classical results from this new perspective. From the numerical experiments, we can note the benefit of including critical preconditioners and relaxation steps on the primal-dual algorithm applied to a total variation image restoration problem.

Acknowledgements The first author thanks the support of ANID under grants FONDECYT 1190871, Redes 180032, and Centro de Modelamiento Matemático (CMM), ACE210010 and FB210005, BASAL funds for centers of excellence. 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).

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

6 Appendix

Proof of Lemma 2.4: Let $x \in \text{Fix } S$. Since $S = S \circ Q$ we have

$$ Sx = x \iff S(Qx) = x \implies Q \circ S(Qx) = Qx, $$

which yields $Qx \in \text{Fix } (Q \circ S)$. Thus, $x = S(Qx) \in S(\text{Fix } (Q \circ S))$ and we conclude Fix $S \subseteq S(\text{Fix } (Q \circ S))$. Conversely, let $x \in \text{Fix } (Q \circ S)$. Since $S = S \circ Q$, we have

$$ Q(Sx) = x \implies S(Q(Sx)) = Sx \implies S(Sx) = Sx \implies Sx \in \text{Fix } S. $$

Thus $S(\text{Fix } (Q \circ S)) \subseteq \text{Fix } S$ and the result follows. \qed
Proof of Proposition 3.1: 1: It is a direct consequence of [13, Proposition 2.1]. 2.: (2a $\Rightarrow$ 2b). Let $(v_n)_{n \in \mathbb{N}}$ be sequence in ran($\Sigma^{-1} - L^*L$) such that $v_n \to v$. Therefore, for each $n \in \mathbb{N}$, there exists $u_n \in G$ such that $v_n = \Sigma^{-1}u_n - L^*L^*u_n$. Note that $V(\gamma L^*u_n, u_n) = (0, v_n) \to (0, v)$. Since ran $V$ is closed, there exists some $(x, u) \in \mathcal{H} \times G$ such that $V(x, u) = (0, v)$, i.e.,

$$V(x, u) = (0, v) \Leftrightarrow \begin{cases} \gamma^{-1}x - L^*u = 0 \\ \Sigma^{-1}u - Lx = v \end{cases} \implies \Sigma^{-1}u - L^*L^*u = v.$$  

Then $v \in$ ran($\Sigma^{-1} - L^*L$), and therefore ran($\Sigma^{-1} - L^*L^*$) is closed.

(2b $\Rightarrow$ 2a). Let $((y_n, v_n))_{n \in \mathbb{N}}$ be a sequence in ran $V$ such that $(y_n, v_n) \to (y, v)$. Then, for every $n \in \mathbb{N}$, there exists $(x_n, u_n)$ such that $(y_n, u_n) = V(x_n, u_n)$, or equivalently,

$$\begin{cases} y_n = \gamma^{-1}x_n - L^*u_n \\ v_n = \Sigma^{-1}u_n - Lx_n. \end{cases} \quad (6.1)$$

By applying $L^*$ to the first equation in $(6.1)$ and adding it to the second equation, by the continuity of $\gamma$ and $L$, we obtain

$$(\Sigma^{-1} - L^*L^*)u_n = L^*y_n + v_n \to L^*y + v. \quad (6.2)$$

Hence, since ran($\Sigma^{-1} - L^*L^*$) is closed, there exists $u \in G$ such that $L^*y + v = (\Sigma^{-1} - L^*L^*)u$. We deduce $V(\gamma(L^*u + y), u) = (y, v)$, and therefore ran $V$ is closed.

(2a $\Leftrightarrow$ 2c). Define $\tilde{V} : G \oplus \mathcal{H} \to G \oplus \mathcal{H}$ by $(u, x) \mapsto (\Sigma^{-1}u - Lx, \gamma^{-1}x - L^*u)$. By the equivalence $2a \Leftrightarrow 2b$ ran$\tilde{V}$ is closed if and only if ran$(\gamma^{-1} - L^*L)$ is closed. Consider the isometric map $\Lambda : \mathcal{H} \oplus G \to \mathcal{H} \oplus \mathcal{H}$ by $(x, u) \mapsto (u, x)$. Since $\Lambda \circ V = \tilde{V}$, ran $V$ is closed if and only if ran$\tilde{V}$ is closed and the result follows.

Proposition 6.1 In the context of Problem 1.1, set $L = \text{Id}$, let $\gamma : \mathcal{H} \to \mathcal{H}$ be a strongly monotone self adjoint linear bounded operator, set $\Lambda : \mathcal{H} \times \mathcal{H} \to \mathcal{H}$ by $\Lambda(x, u) = x - \gamma u$, let $V$, $W$, and $G_{\gamma, B, A}$ be the operators defined in (1.8), (3.2), and (3.15), respectively. Then, $\Lambda(\text{Fix}(P_{\text{ran}}V \circ J_W)) = \text{Fix} G_{\gamma, B, A}$.

Proof The inclusion $\subset$ is proved in (3.23). Conversely, since $\Lambda^* : z \mapsto (z, -\gamma z)$, we have $\Lambda \circ \Lambda^* = \text{Id} + \gamma^2$ and [3, Proposition 3.30 & Example 3.29] yields $P_{\text{ran}}V = P_{\text{ran}}\Lambda^* = \Lambda^*(\text{Id} + \gamma^2)^{-1} \Lambda$. Therefore, if $\hat{z} \in \text{Fix} G_{\gamma, B, A}$, by setting $(\hat{x}, \hat{u}) := \Lambda^*(\text{Id} + \gamma^2)^{-1} \hat{z}$, we have $\hat{z} = \Lambda(\hat{x}, \hat{u})$ and we deduce from (3.22) that

$$P_{\text{ran}}V \circ J_W(\hat{x}, \hat{u}) = \Lambda^*(\text{Id} + \gamma^2)^{-1}(\text{J}_W(\hat{x}, \hat{u}))$$

$$= \Lambda^*(\text{Id} + \gamma^2)^{-1}G_{\gamma, B, A}(\Lambda(\hat{x}, \hat{u}))$$

$$= \Lambda^*(\text{Id} + \gamma^2)^{-1}G_{\gamma, B, A} \hat{z}$$

$$= \Lambda^*(\text{Id} + \gamma^2)^{-1} \hat{z}$$

$$= (\hat{x}, \hat{u}). \quad (6.3)$$

Consequently, $(\hat{x}, \hat{u}) \in \text{Fix}(P_{\text{ran}}V \circ J_W)$ and $\hat{z} = \Lambda(\hat{x}, \hat{u}) \in \text{Fix}(P_{\text{ran}}V \circ J_W)$. \qed

References

1. Aubin, J.P., Frankowska, H.: Set-valued Analysis. Modern Birkhäuser Classics, Birkhäuser Boston, Inc., Boston, MA (2009) https://doi.org/10.1007/978-0-8176-4848-0

Springer
2. Bauschke, H.H.: New demiclosedness principles for (firmly) nonexpansive operators. In: Bailey, D.H., Bauschke, H.H., Borwein, P., Garvan, F., Théra, M., Vanderwerff, J.D., Wolkowicz, H. (eds.) Computational and Analytical Mathematics. Springer Proceedings in Mathematics & Statistics 50, pp. 19–28. Springer, New York (2013) https://doi.org/10.1007/978-1-4614-6414-4_2

3. Bauschke, H.H., Combettes, P.L.: Convex analysis and monotone operator theory in Hilbert spaces. Springer, New York (2017)

4. Bauschke, H.H., Moursi, W.M.: On the Douglas-Rachford algorithm. Math. Program. 164, 263–284 (2017). https://doi.org/10.1007/s10107-016-1086-3

5. Boţ, R.I., Csetnek, E.R., 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

6. 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

7. Briceño, L., Cominetti, R., Cortés, C.E., Martínez, F.: An integrated behavioral model of land use and transport system: a hyper-network equilibrium approach. Netw. Spat. Econ. 8, 201–224 (2008). https://doi.org/10.1007/s11067-007-9052-5

8. 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

9. Briceño-Arias, L.M., Combettes, P.L.: A primal-dual splitting method for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators in duality. SIAM J. Optim. 20, 307–330 (2010). https://doi.org/10.1137/090757291

10. 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

11. Briceño-Arias, L.M., Deride, J., Vega, C.: Random activations in primal-dual splittings for monotone inclusions with a priori information. J. Optim. Theory Appl. 192, 56–81 (2022). https://doi.org/10.1007/s10957-021-01944-6

12. Briceño-Arias, L.M., López Rivera, S.: A projected primal-dual method for solving constrained monotone inclusions. J. Optim. Theory Appl 180, 907–924 (2019). https://doi.org/10.1007/s10957-018-1430-2

13. Briceño-Arias, L.M., Roldán, F.: Split-Douglas-Rachford algorithm for composite monotone inclusions and split-ADMM. SIAM J. Optim. 31, 2987–3013 (2021). https://doi.org/10.1137/21M1395144

14. Chambolle, A.: An algorithm for total variation minimization and applications. J. Math. Imaging Vision 20, 89–97 (2004). https://doi.org/10.1023/B:JMIV.000001320.81911.38

15. Chambolle, A., Caselles, V., Cremers, D., Novaga, M., Pock, T.: An introduction to total variation image analysis. In: Fornasier, M. (ed.) Theoretical Foundations and Numerical Methods for Sparse Recovery. Radon Ser. Comput. Appl. Math. 9., pp. 263–340. Walter de Gruyter, Berlin (2010) https://doi.org/10.1515/9783110226157.263

16. Chambolle, A., Lions, P.L.: Image recovery via total variation minimization and related problems. Numer. Math. 76, 167–188 (1997). https://doi.org/10.1007/s002110050258

17. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. J. Math. Imaging Vision 40, 120–145 (2011). https://doi.org/10.1007/s10851-010-0251-1

18. Colas, J., Pustelnik, N., Oliver, C., Abry, P., Géminard, J.-C., Vidal, V.: Nonlinear denoising for characterization of solid friction under low confinement pressure. Phys. Rev. E 100, 032803 (2019). https://doi.org/10.1103/PhysRevE.100.032803

19. 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, pp. 115–152. North-Holland, Amsterdam (2001) https://doi.org/10.1016/S1570-579X(01)80010-0

20. 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

21. Combettes, P.L., Vu, B.C.: Variable metric quasi-Fejér monotonicity. Nonlinear Anal. 78, 17–31 (2013). https://doi.org/10.1016/j.na.2012.09.008

22. Combettes, P.L., Vu, 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
24. Condat, L.: A primal-dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms. J. Optim. Theory Appl. 158, 460–479 (2013). https://doi.org/10.1007/s10957-012-0245-9

25. Daubechies, I., Defrise, M., De Mol, C.: An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. Comm. Pure Appl. Math. 57, 1413–1457 (2004). https://doi.org/10.1002/cpa.20042

26. Davis, D.: Convergence rate analysis of primal-dual splitting schemes. SIAM J. Optim. 25, 1912–1943 (2015). https://doi.org/10.1137/151003076

27. Eckstein, J., Bertsekas, D.P.: On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. Math. Program. 55, 293–318 (1992). https://doi.org/10.1007/BF01581204

28. Fukushima, M.: The primal Douglas-Rachford splitting algorithm for a class of monotone mappings with application to the traffic equilibrium problem. Math. Program. 72, 1–15 (1996). https://doi.org/10.1016/0025-5610(95)00012-7

29. Gabay, D.: Chapter IX applications of the method of multipliers to variational inequalities. In: Fortin, M. and Glowinski, R. (eds.) Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems. Studies in Mathematics and Its Applications 15, pp. 299–331. Elsevier (1983) https://doi.org/10.1016/S0168-2024(08)70034-1

30. Gafni, E.M., Bertsekas, D.P.: Two-metric projection methods for constrained optimization. SIAM J. Control Optim. 22, 936–964 (1984). https://doi.org/10.1137/0322061

31. 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)

32. 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

33. Hansen, P.C., Nagy, J.G., O’Leary, D.P.: Deblurring Images: Matrices, Spectra, and Filtering. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2006) https://doi.org/10.1137/1.9780898718874

34. He, B., Yuan, X.: Convergence analysis of primal-dual algorithms for a saddle-point problem: from contraction perspective. SIAM J. Imaging Sci. 5, 119–149 (2012). https://doi.org/10.1137/100814494

35. Lions, P.L., 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

36. Mallat, S.: A wavelet tour of signal processing: the sparse way. Elsevier/Academic Press, Amsterdam (2009)

37. Martinet, B.: Brève communication. régularisation d’inéquations variationnelles par approximations successives. ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique 4, 154–158 (1970) http://www.numdam.org/item/M2AN_1970__4_3_154_0

38. Mises, R.V., Pollaczek-Geiringer, H.: Praktische verfahren der gleichungsauflosung. Z. Angew. Math. Mech. (1929). https://doi.org/10.1002/zamm.1929009090206

39. Molinari, C., Peypouquet, J., Roldan, F.: Alternating forward-backward splitting for linearly constrained optimization problems. Optim. Lett. 14, 1071–1088 (2020). https://doi.org/10.1007/s11590-019-01388-y

40. Pock, T., Chambolle, A.: Diagonal preconditioning for first order primal-dual algorithms in convex optimization. In: 2011 international conference on computer vision, pp. 1762–1769 (2011) https://doi.org/10.1109/ICCV.2011.6126441

41. Rockafellar, R.T.: Monotone operators and the proximal point algorithm. SIAM J. Control Optim. 14, 877–898 (1976). https://doi.org/10.1137/0314056

42. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Phys. D. 60, 259–268 (1992). https://doi.org/10.1016/0167-2789(92)90242-F

43. Showalter, R.E.: Monotone operators in banach space and nonlinear partial differential equations. Mathematical surveys and monographs 49, American mathematical society, Providence, RI (1997) https://doi.org/10.1090/surv/049

44. Svaiter, B.F.: On weak convergence of the Douglas-Rachford method. SIAM J. Control Optim. 49, 280–287 (2011). https://doi.org/10.1137/100788100

45. Vü, 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

46. Yang, Y., Tang, Y., Wen, M., Zeng, T.: Preconditioned Douglas-Rachford type primal-dual method for solving composite monotone inclusion problems with applications. Inverse Probl. Imaging 15, 787–825 (2021). https://doi.org/10.3934/ipi.2021014
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.