On the Douglas–Rachford algorithm

Heinz H. Bauschke* and Walaa M. Moursi†

October 21, 2016

Abstract

The Douglas–Rachford algorithm is a very popular splitting technique for finding a zero of the sum of two maximally monotone operators. The behaviour of the algorithm remains mysterious in the general inconsistent case, i.e., when the sum problem has no zeros. However, more than a decade ago, it was shown that in the (possibly inconsistent) convex feasibility setting, the shadow sequence remains bounded and its weak cluster points solve a best approximation problem.

In this paper, we advance the understanding of the inconsistent case significantly by providing a complete proof of the full weak convergence in the convex feasibility setting. In fact, a more general sufficient condition for the weak convergence in the general case is presented. Our proof relies on a new convergence principle for Fejér monotone sequences. Numerous examples illustrate our results.

2010 Mathematics Subject Classification: Primary 47H05, 47H09, 49M27; Secondary 49M29, 49N15, 90C25.

Keywords: Attouch–Théra duality, Douglas–Rachford algorithm, inconsistent case, maximally monotone operator, nonexpansive mapping, paramonotone operator, sum problem, weak convergence.

*Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. E-mail: heinz.bauschke@ubc.ca.

†Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada, and Mansoura University, Faculty of Science, Mathematics Department, Mansoura 35516, Egypt. E-mail: walaa.moursi@ubc.ca.

1 Introduction

In this paper, we assume that

$$X \text{ is a real Hilbert space,}$$

with inner product $\langle\cdot,\cdot\rangle$ and induced norm $\|\cdot\|$. A classical problem in optimization is to find a minimizer of the sum of two proper lower semicontinuous convex functions. This problem can be modelled as

$$\text{find } x \in X \text{ such that } 0 \in (A+B)x, \tag{1}$$

where A and B are maximally monotone operators on X, namely the subdifferential operators of the functions under consideration. For detailed discussions on problem (1) and the connection to optimization problems, we refer the reader to [10], [19], [23], [25], [27], [52], [53], [51], [56], [57], and the references therein.

Due to its general convergence results, the Douglas–Rachford algorithm has become a very popular splitting technique to solve the sum problem (1) provided that the solution set is nonempty. Indeed, the method has found many applications; see, e.g., [30], [31] and [32]. The algorithm was first introduced in [35] to numerically solve certain types of heat equations. Let $x \in X$, let $T = T_{(A,B)}$ be the Douglas–Rachford operator associated with the ordered pair (A, B) (see (5)) and let J_A be the resolvent of A (see Fact 2.3). In their masterpiece [47] (see also [46]), Lions and Mercier extended the algorithm to be able to find a zero of the sum of two, not necessarily linear and possibly set-valued, maximally monotone operators. They proved that the "governing sequence" $(T^n x)_{n\in\mathbb{N}}$ converges weakly to a fixed point of T, and that if $A+B$ is maximally monotone, then the weak cluster points of the "shadow sequence" $(J_A T^n x)_{n\in\mathbb{N}}$ are solutions of (1). In [54], weak convergence of the shadow sequence[1], regardless of the maximal monotonicity of $A+B$, was established.

Nonetheless, very little is known about the behaviour of the algorithm in the inconsistent setting, i.e., when the set of zeros of the sum is empty. In [11] (see Remark 4.6), the authors showed that if A and B are normal cone operators of two nonempty closed convex subsets of X, and $P_{\overline{\operatorname{ran}}(\operatorname{Id}-T)}0 \in \operatorname{ran}(\operatorname{Id}-T)$ (see Fact 4.1), then the shadow sequence $(J_A T^n x)_{n\in\mathbb{N}}$ is bounded and its weak cluster points solve a certain best approximation problem. This is an important case of the Douglas–Rachford algorithm since it has often been applied in the context of feasibility problems (both in convex and nonconvex settings); see, e.g., [1], [2], [20], [21], [27], [39], [40], [42], [43], and [45].

[1] To the best of our knowledge, the Douglas–Rachford algorithm has never been applied in infinite-dimensional spaces; nonetheless, Svaiter's result is a milestone in abstract infinite-dimensional optimization.

In this paper, we derive some new and useful identities for the Douglas–Rachford operator. The main contribution of the paper is generalizing the results in [11] by proving the *full weak convergence of the shadow sequence* in the convex feasibility setting (see Theorem 4.5). As seen in the previous paragraph, this is a problem of actual importance in applications even in the inconsistent case (see [11], [26] and [34]). Moreover, this convergence result is new *even when the Hilbert space is finite-dimensional!* Our proof crucially relies on a *new convergence principle for Fejér monotone sequences* (see Lemma 4.3). While the general case remains open (see Example 4.9, Remark 4.10, and Example 4.11), we do provide some sufficient conditions for the convergence of the shadow sequence in some special cases (see Theorem 4.4 and Corollary 4.8).

As a byproduct of our analysis, we are able to present *a new proof* for Svaiter's result [54] concerning the weak convergence of the shadow sequence (see Theorem 6.2) in the consistent case. Our proof is not only in the spirit of the techniques used in the original paper by Lions and Mercier [47] but it is also significantly shorter and more elegant than Svaiter's proof. The notation used in the paper is standard and follows largely [10].

2 Useful identities for the Douglas–Rachford operator

We start with two elementary identities which are easily verified directly.

Lemma 2.1. *Let* $(a, b, z) \in X^3$*. Then the following hold:*

(i) $\langle z, z - a + b\rangle = \|z - a + b\|^2 + \langle a, z - a\rangle + \langle b, 2a - z - b\rangle$.
(ii) $\langle z, a - b\rangle = \|a - b\|^2 + \langle a, z - a\rangle + \langle b, 2a - z - b\rangle$.
(iii) $\|z\|^2 = \|z - a + b\|^2 + \|b - a\|^2 + 2\langle a, z - a\rangle + 2\langle b, 2a - z - b\rangle$.

Lemma 2.2. *Let* $(a, b, x, y, a^*, b^*, u, v) \in X^8$*. Then*

$$\begin{aligned}\langle (a, b) - (x, y), (a^*, b^*) - (u, v)\rangle &= \langle a - b, a^*\rangle + \langle x, u\rangle - \langle x, a^*\rangle - \langle a - b, u\rangle \\ &\quad + \langle b, a^* + b^*\rangle + \langle y, v\rangle - \langle y, b^*\rangle - \langle b, u + v\rangle. \qquad (2)\end{aligned}$$

Unless stated otherwise, we assume from now on that

$$A : X \rightrightarrows X \text{ and } B : X \rightrightarrows X \text{ are maximally monotone operators.}$$

The following result concerning the *resolvent* $J_A := (\mathrm{Id} + A)^{-1}$ and the *reflected resolvent* $R_A := 2J_A - \mathrm{Id}$ is well known; see, e.g., [10, Corollary 23.10(i)&(ii)].

Fact 2.3. $J_A : X \to X$ *is firmly nonexpansive and* $R_A : X \to X$ *is nonexpansive.*

Let us recall the well-known inverse resolvent identity (see [51, Lemma 12.14])

$$J_A + J_{A^{-1}} = \mathrm{Id} \tag{3}$$

and the following useful description of the graph of A.

Fact 2.4 (Minty parametrization). (See [49].) *$M\colon X \to \operatorname{gra} A : x \mapsto (J_A x, J_{A^{-1}} x)$ is a continuous bijection, with continuous inverse $M^{-1}\colon \operatorname{gra} A \to X : (x, u) \to x + u$; consequently,*

$$\operatorname{gra} A = M(X) = \big\{ (J_A x, x - J_A x) \;\big|\; x \in X \big\}. \tag{4}$$

Definition 2.5. *The* Douglas–Rachford splitting operator *associated with (A, B) is*

$$T := T_{(A,B)} = \tfrac{1}{2}(\mathrm{Id} + R_B R_A) := \mathrm{Id} - J_A + J_B R_A. \tag{5}$$

We will simply use T instead of $T_{(A,B)}$ provided there is no cause for confusion.

The following result will be useful.

Lemma 2.6. *Let $x \in X$. Then the following hold:*

(i) $x - Tx = J_A x - J_B R_A x = J_{A^{-1}} x + J_{B^{-1}} R_A x$.
(ii) $(J_A x, J_B R_A x, J_{A^{-1}} x, J_{B^{-1}} R_A x)$ *lies in* $\operatorname{gra}(A \times B)$.

Proof. (i): The first identity is a direct consequence of (5). In view of (3) $J_A x - J_B R_A x - J_{A^{-1}} x - J_{B^{-1}} R_A x = J_A x - (x - J_A x) - (J_B + J_{B^{-1}}) R_A x = R_A x - R_A x = 0$, which proves the second identity. (ii): Use (32) and Fact 2.4 applied to $A \times B$ at $(x, R_A x) \in X \times X$. ■

The next theorem is a direct consequence of the key identities presented in Lemma 2.1.

Theorem 2.7. *Let $x \in X$ and let $y \in X$. Then the following hold:*

(i) $\langle Tx - Ty, x - y \rangle = \|Tx - Ty\|^2 + \langle J_A x - J_A y, J_{A^{-1}} x - J_{A^{-1}} y \rangle$
$\qquad + \langle J_B R_A x - J_B R_A y, J_{B^{-1}} R_A x - J_{B^{-1}} R_A y \rangle$.

(ii) $\langle (\mathrm{Id} - T)x - (\mathrm{Id} - T)y, x - y \rangle = \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 + \langle J_A x - J_A y, J_{A^{-1}} x - J_{A^{-1}} y \rangle$
$\qquad + \langle J_B R_A x - J_B R_A y, J_{B^{-1}} R_A x - J_{B^{-1}} R_A y \rangle$.

(iii) $\|x - y\|^2 = \|Tx - Ty\|^2 + \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 + 2\langle J_A x - J_A y, J_{A^{-1}} x - J_{A^{-1}} y \rangle$
$\qquad + 2\langle J_B R_A x - J_B R_A y, J_{B^{-1}} R_A x - J_{B^{-1}} R_A y \rangle$.

(iv) $\|J_A x - J_A y\|^2 + \|J_{A^{-1}} x - J_{A^{-1}} y\|^2 - \|J_A Tx - J_A Ty\|^2 - \|J_{A^{-1}} Tx - J_{A^{-1}} Ty\|^2$
$\quad = \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 + 2\langle J_A Tx - J_A Ty, J_{A^{-1}} Tx - J_{A^{-1}} Ty \rangle$
$\qquad + 2\langle J_B R_A x - J_B R_A y, J_{B^{-1}} R_A x - J_{B^{-1}} R_A y \rangle$.

(v) $\|J_A Tx - J_A Ty\|^2 + \|J_{A^{-1}} Tx - J_{A^{-1}} Ty\|^2 \le \|J_A x - J_A y\|^2 + \|J_{A^{-1}} x - J_{A^{-1}} y\|^2$.

Proof. (i)–(iii): Use Lemma 2.1(i)–(iii) respectively, with $z = x - y$, $a = J_A x - J_A y$ and $b = J_B R_A x - J_B R_A y$, (4) and (5). (iv): It follows from (3) that

$$\|x - y\|^2 = \|J_A x - J_A y + J_{A^{-1}} x - J_{A^{-1}} y\|^2 \tag{6a}$$

$$= \|J_A x - J_A y\|^2 + \|J_{A^{-1}} x - J_{A^{-1}} y\|^2 + 2\langle J_A x - J_A y, J_{A^{-1}} x - J_{A^{-1}} y\rangle. \tag{6b}$$

Applying (6) to (Tx, Ty) instead of (x, y) yields

$$\begin{aligned}\|Tx - Ty\|^2 = \|J_A Tx - J_A Ty\|^2 + \|J_{A^{-1}} Tx - J_{A^{-1}} Ty\|^2 \\ + 2\langle J_A Tx - J_A Ty, J_{A^{-1}} Tx - J_{A^{-1}} Ty\rangle.\end{aligned} \tag{7}$$

Now combine (6), (7), and (iii) to obtain (iv). (v): In view of (4), the monotonicity of A and B implies $\langle J_A Tx - J_A Ty, J_{A^{-1}} Tx - J_{A^{-1}} Ty\rangle \geq 0$ and $\langle J_B R_A x - J_B R_A y, J_{B^{-1}} R_A x - J_{B^{-1}} R_A y\rangle \geq 0$. Now use (iv). ■

3 The Douglas–Rachford operator and duality

The *Attouch–Théra dual pair* (see [3] and [48, page 40]) of (A, B) is $(A, B)^* := (A^{-1}, B^{-\circledcirc})$, where

$$B^{\circledcirc} := (-\operatorname{Id}) \circ B \circ (-\operatorname{Id}) \quad \text{and} \quad B^{-\circledcirc} := (B^{-1})^{\circledcirc} = (B^{\circledcirc})^{-1}. \tag{8}$$

We use

$$Z := Z_{(A,B)} = (A + B)^{-1}(0) \qquad \text{and} \qquad K := K_{(A,B)} = (A^{-1} + B^{-\circledcirc})^{-1}(0) \tag{9}$$

to denote the primal and dual solutions, respectively (see, e.g., [9]).

Let us record some useful properties of $T_{(A,B)}$.

Fact 3.1. *The following hold:*

(i) **(Lions and Mercier).** $T_{(A,B)}$ *is firmly nonexpansive.*
(ii) **(Eckstein).** $T_{(A,B)} = T_{(A^{-1}, B^{-\circledcirc})}$.
(iii) **(Combettes).** $Z = J_A(\operatorname{Fix} T)$.
(iv) $K = J_{A^{-1}}(\operatorname{Fix} T)$.

Proof. (i): See, e.g., [47, Lemma 1], [36, Corollary 4.2.1 on page 139], [37, Corollary 4.1], or Theorem 2.7(i). (ii): See e.g., [36, Lemma 3.6 on page 133] or [9, Corollary 4.3]. (iii): See [28, Lemma 2.6(iii)]. (iv): See [9, Corollary 4.9]. ■

The following notion, coined by Iusem [44], is very useful. We say that $C : X \rightrightarrows X$ is *paramonotone* if it is monotone and we have the implication

$$\left.\begin{array}{c}(x,u) \in \operatorname{gra} C\\ (y,v) \in \operatorname{gra} C\\ \langle x-y, u-v\rangle = 0\end{array}\right\} \quad \Rightarrow \quad \big\{(x,v),(y,u)\big\} \subseteq \operatorname{gra} C. \tag{10}$$

Example 3.2. *Let $f : X \to \left]-\infty,+\infty\right]$ be convex, lower semicontinuous, and proper. Then ∂f is paramonotone by [44, Proposition 2.2] (or by [10, Example 22.3(i)]).*

We now recall that the so-called "extended solution set" (see [38, Section 2.1] and also [9, Section 3]) is defined by

$$\mathcal{S} := \mathcal{S}_{(A,B)} := \{(z,k) \in X \times X \mid -k \in Bz, k \in Az\} \subseteq Z \times K. \tag{11}$$

Fact 3.3. *Recalling Fact 2.4, we have the following:*

(i) $\mathcal{S} = M(\operatorname{Fix} T) = \big\{(J_A \times J_{A^{-1}})(y,y) \,\big|\, y \in \operatorname{Fix} T\big\}$.
(ii) $\operatorname{Fix} T = M^{-1}(\mathcal{S}) = \big\{z + k \,\big|\, (z,k) \in \mathcal{S}\big\}$.
(iii) *(**Eckstein and Svaiter).** $\mathcal{S}$ is closed and convex.*

If A and B are paramonotone, then we additionally have:

(iv) $\mathcal{S} = Z \times K$.
(v) $\operatorname{Fix} T = Z + K$.

Proof. (i)&(ii): This is [9, Theorem 4.5]. (iii): See [38, Lemma 2]. Alternatively, since $\operatorname{Fix} T$ is closed, and M and M^{-1} are continuous, we deduce the closedness from (i). The convexity was proved in [9, Corollary 3.7]. (iv)&(v): See [9, Corollary 5.5(ii)&(iii)]. ■

4 Main results

In this section, we consider the case when the set Z is possibly empty.

We recall the following important fact.

Fact 4.1 (minimal displacement vector). (See, e.g., [4],[24], and [50].) *Let $T : X \to X$ be nonexpansive. Then $\overline{\operatorname{ran}}(\operatorname{Id} - T)$ is convex; consequently, the minimal displacement vector*

$$v := P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)} 0 \tag{12}$$

is the unique and well-defined element in $\overline{\operatorname{ran}}(\operatorname{Id} - T)$ such that $\|v\| = \inf_{x \in X} \|x - Tx\|$.

Following [8], the *normal problem* associated with the ordered pair (A, B) is to[2]

$$\text{find } x \in X \text{ such that } 0 \in {}_vAx + B_vx = Ax - v + B(x - v). \tag{13}$$

We shall use

$$Z_v := Z_{({}_vA, B_v)} \qquad \text{and} \qquad K_v := K_{(({}_vA)^{-1}, (B_v)^{-\circledcirc})}, \tag{14}$$

to denote the *primal normal* and *dual normal solutions*, respectively. It follows from [8, Proposition 3.3] that

$$Z_v \neq \varnothing \Leftrightarrow v \in \operatorname{ran}(\operatorname{Id} - T). \tag{15}$$

Corollary 4.2. *Let $x \in X$ and let $y \in X$. Then the following hold:*

$$\sum_{n=0}^{\infty} \|(\operatorname{Id} - T)T^n x - (\operatorname{Id} - T)T^n y\|^2 < +\infty, \tag{16a}$$

$$\sum_{n=0}^{\infty} \underbrace{\langle J_A T^n x - J_A T^n y, J_{A^{-1}} T^n x - J_{A^{-1}} T^n y \rangle}_{\geq 0} < +\infty, \tag{16b}$$

$$\sum_{n=0}^{\infty} \underbrace{\langle J_B R_A T^n x - J_B R_A T^n y, J_{B^{-1}} R_A T^n x - J_{B^{-1}} R_A T^n y \rangle}_{\geq 0} < +\infty. \tag{16c}$$

Consequently,

$$(\operatorname{Id} - T)T^n x - (\operatorname{Id} - T)T^n y \to 0, \tag{17a}$$

$$\langle J_A T^n x - J_A T^n y, J_{A^{-1}} T^n x - J_{A^{-1}} T^n y \rangle \to 0, \tag{17b}$$

$$\langle J_B R_A T^n x - J_B R_A T^n y, J_{B^{-1}} R_A T^n x - J_{B^{-1}} R_A T^n y \rangle \to 0. \tag{17c}$$

Proof. Let $n \in \mathbb{N}$. Applying (4), to the points $T^n x$ and $T^n y$, we learn that $\{(J_A T^n x, J_{A^{-1}} T^n x), (J_A T^n y, J_{A^{-1}} T^n y)\} \subseteq \operatorname{gra} A$, hence, by monotonicity of A, we have $\langle J_A T^n x - J_A T^n y, J_{A^{-1}} T^n x - J_{A^{-1}} T^n y \rangle \geq 0$. Similarly, $\langle J_B R_A T^n x - J_B R_A T^n y, J_{B^{-1}} R_A T^n x - J_{B^{-1}} R_A T^n y \rangle \geq 0$. Now (16) and (17) follow from Theorem 2.7(iii) by telescoping. ■

The next result on Fejér monotone sequences is of critical importance in our analysis. (When $(u_n)_{n \in \mathbb{N}} = (x_n)_{n \in \mathbb{N}}$ one obtains a well-known result; see, e.g., [10, Theorem 5.5].)

Lemma 4.3 (new Fejér monotonicity principle). *Suppose that E is a nonempty closed convex subset of X, that $(x_n)_{n \in \mathbb{N}}$ is a sequence in X that is* Fejér monotone with respect to E, *i.e.,*

$$(\forall e \in E)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - e\| \leq \|x_n - e\|, \tag{18}$$

[2]Let $w \in X$ be fixed. The *inner and outer shifts* associated with A are defined by $A_w\colon X \rightrightarrows X\colon x \mapsto A(x - w)$ and ${}_wA\colon X \rightrightarrows X\colon x \mapsto Ax - w$. Note that A_w and ${}_wA$ are maximally monotone because A is.

that $(u_n)_{n\in\mathbb{N}}$ is a bounded sequence in X such that its weak cluster points lie in E, and that

$$(\forall e \in E)\ \ \langle u_n - e, u_n - x_n\rangle \to 0. \tag{19}$$

Then $(u_n)_{n\in\mathbb{N}}$ converges weakly to some point in E.

Proof. It follows from (19) that

$$(\forall (e_1, e_2) \in E \times E) \quad \langle e_2 - e_1, u_n - x_n\rangle = \langle u_n - e_1, u_n - x_n\rangle - \langle u_n - e_2, u_n - x_n\rangle \to 0. \tag{20}$$

Now obtain four subsequences $(x_{k_n})_{n\in\mathbb{N}}$, $(x_{l_n})_{n\in\mathbb{N}}$, $(u_{k_n})_{n\in\mathbb{N}}$ and $(u_{l_n})_{n\in\mathbb{N}}$ such that $x_{k_n} \rightharpoonup y_1$, $x_{l_n} \rightharpoonup y_2$, $u_{k_n} \rightharpoonup e_1$ and $u_{l_n} \rightharpoonup e_2$. Taking the limit in (20) along these subsequences, we have $\langle e_2 - e_1, e_1 - y_1\rangle = 0 = \langle e_2 - e_1, e_2 - y_2\rangle$; hence,

$$\|e_2 - e_1\|^2 = \langle e_2 - e_1, y_2 - y_1\rangle. \tag{21}$$

Since $\{e_1, e_2\} \subseteq E$, we conclude, in view of [5, Theorem 6.2.2(ii)] or [13, Lemma 2.2], that $\langle e_2 - e_1, y_2 - y_1\rangle = 0$. By (20), $e_1 = e_2$. ■

We are now ready for our main result.

Theorem 4.4 (shadow convergence). *Suppose that $x \in X$, that the sequence $(J_A T^n x)_{n\in\mathbb{N}}$ is bounded and its weak cluster points lie in Z_v, that $Z_v \subseteq \operatorname{Fix}(v+T)$ and that $(\forall n \in \mathbb{N})$ $(\forall y \in Z_v)$ $J_A T^n y = y$. Then the "shadow" sequence $(J_A T^n x)_{n\in\mathbb{N}}$ converges weakly to some point in Z_v.*

Proof. Let $y \in Z_v \subseteq \operatorname{Fix}(v+T)$. Using (17b) and [16, Proposition 2.4(iv)], we have

$$\begin{aligned}\langle J_A T^n x - y, T^n x + nv - J_A T^n x\rangle &= \langle J_A T^n x - y, T^n x - J_A T^n x - (y - nv - y)\rangle && \text{(22a)}\\ &= \langle J_A T^n x - J_A T^n y, (\operatorname{Id} - J_A) T^n x - (\operatorname{Id} - J_A) T^n y\rangle && \text{(22b)}\\ &\to 0. && \text{(22c)}\end{aligned}$$

Note that [16, Proposition 2.4(vi)] implies that $(T^n x + nv)_{n\in\mathbb{N}}$ is Fejér monotone with respect to $\operatorname{Fix}(v+T)$ and consequently with respect to Z_v. Now apply Lemma 4.3 with E replaced by Z_v, $(u_n)_{n\in\mathbb{N}}$ replaced by $(J_A T^n x)_{n\in\mathbb{N}}$, and $(x_n)_{n\in\mathbb{N}}$ replaced by $(T^n x + nv)_{n\in\mathbb{N}}$. ■

As a powerful application of Theorem 4.4, we obtain the following striking result on the convergence of the Douglas–Rachford algorithm for feasibility problems.

Theorem 4.5 (possibly inconsistent feasibility problem). *Suppose that U and V are nonempty closed convex subsets of X, that $A = N_U$, that $B = N_V$, that $v = P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)} 0$ and that $U \cap (v + V) \neq \varnothing$. Let $x \in X$. Then $(P_U T^n x)_{n\in\mathbb{N}}$ converges weakly to some point in $Z_v = U \cap (v + V)$.*

Proof. It follows from [11, Theorem 3.13(iii)(b)] that $(P_U T^n x)_{n\in\mathbb{N}}$ is bounded and its weak cluster points lie in $U \cap (v+V)$. Moreover, [11, Theorem 3.5] implies that $Z_v = U \cap (v+V) \subseteq U \cap (v+V) + N_{\overline{U-V}}(v) \subseteq \operatorname{Fix}(v+T)$. Finally, [11, Lemma 3.12 & Proposition 2.4(ii)] imply that $(\forall y \in \operatorname{Fix}(v+T))(\forall n \in \mathbb{N})\ P_U T^n y = P_U(y - nv) = y$. Hence all the assumptions of Theorem 4.4 are satisfied and the result follows. ■

In Figure 1 below, we provide a visualization of the conclusion of Theorem 4.5 in $\mathbb{R}^2$.

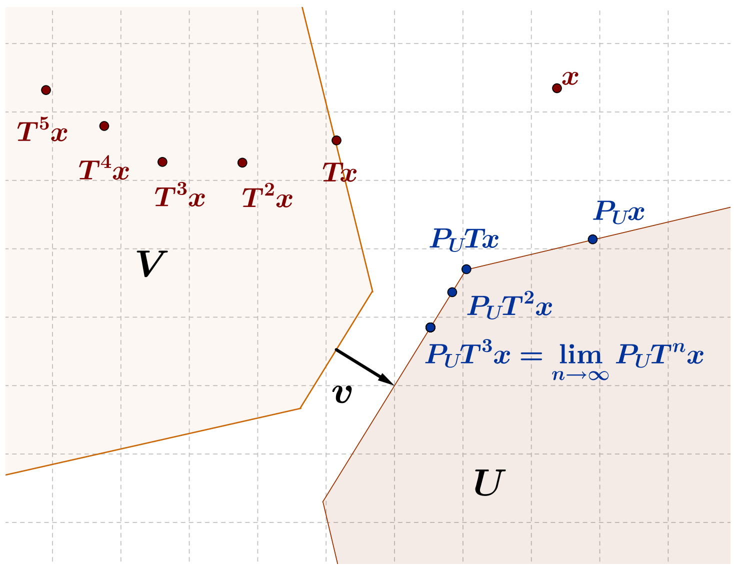

Figure 1: A `GeoGebra` [41] snapshot that illustrates Theorem 4.5. Two nonintersecting polyhedral sets in $\mathbb{R}^2$, U and V, are shown along with the minimal displacement vector v. The first few iterates of the sequence $(T^n x)_{n\in\mathbb{N}}$ (red points) and the sequence $(P_U T^n x)_{n\in\mathbb{N}}$ (blue points) are also depicted.

Remark 4.6. *Suppose that $v \in \operatorname{ran}(\operatorname{Id} - T)$. More than a decade ago, it was shown in [11] that if $A = N_U$ and $B = N_V$, where U and V are nonempty closed convex subsets of X, then $(P_U T^n x)_{n\in\mathbb{N}}$ is bounded and its weak cluster points lie in $U \cap (v+V)$. Theorem 4.5 yields the much stronger result that $(P_U T^n x)_{n\in\mathbb{N}}$ actually converges weakly to a point in $U \cap (v+V)$. As seen in the introduction, the feasibility setting is of practical importance and occurs in real-world applications. We also stress the fact that Theorem 4.5 is* entirely new even in finite-dimensional spaces!

In proving Corollary 4.8, which is another variant of Theorem 4.4, we require a couple of auxiliary results. First, if U is a closed affine subspace of X and[3] $w \in (\operatorname{par} U)^\perp$, then

[3]Suppose that U is a closed affine subspace of X. We use $\operatorname{par} U$ to denote the *parallel space* of U which is defined by $\operatorname{par} U = U - U$.

(by, e.g., [16, Lemma 4.2(iv)])

$$(\forall \alpha \in \mathbb{R})(\forall x \in X) \quad P_U(x + \alpha w) = P_U x. \tag{23}$$

Proposition 4.7. *Suppose that U is a closed affine subspace of X, that $A = N_U$, that B is paramonotone, that $v = P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)} 0 \in \operatorname{ran}(\operatorname{Id} - T)$, and that $\operatorname{zer}({}_vA) \cap \operatorname{zer}(B_v) \neq \varnothing$. Let $x \in X$. Then the following hold:*

(i) $Z_v = \operatorname{zer}({}_vA) \cap \operatorname{zer}(B_v) = \operatorname{zer}(N_U - v) \cap \operatorname{zer}(B(\cdot - v))$.
(ii) $0 \in K_v$.
(iii) $v \in (\operatorname{par} U)^{\perp}$.
(iv) $Z_v \subseteq U$.
(v) $Z_v \subseteq Z_v + K_v = \operatorname{Fix}(T_{-v}) \subseteq \operatorname{Fix}(v + T)$.
(vi) $(\forall n \in \mathbb{N})(\forall y \in Z_v)\ J_A T^n y = P_U T^n y = y$.
(vii) $P_U T^n x - J_B R_U T^n x = T^n x - T^{n+1} x \to v$.
(viii) $(J_A T^n x)_{n \in \mathbb{N}} = (P_U T^n x)_{n \in \mathbb{N}}$ *is bounded.*
(ix) $(J_B R_U T^n x)_{n \in \mathbb{N}}$ *is bounded.*

Proof. (i)&(ii): Apply [14, Theorem 4.7(i)] to $({}_vA, B_v) = (-v + N_U, B(\cdot - v))$. (iii): In view of (15) we learn that $Z_v \neq \varnothing$. Therefore, (i) implies that $\operatorname{zer}({}_vA) = \operatorname{zer}(-v + N_U) \neq \varnothing$. Hence $(\exists x \in X)\ 0 \in -v + N_U x = -v + (\operatorname{par} U)^{\perp}$; equivalently, $v \in (\operatorname{par} U)^{\perp}$. (iv): Using (i), we see that $Z_v \subseteq \operatorname{zer}(N_U - v)$. We claim that $\operatorname{zer}(N_U - v) = U$. Indeed, let $y \in X$. In view of (iii) we have $0 \in N_U y - v \Leftrightarrow [y \in U \text{ and } 0 \in (\operatorname{par} U)^{\perp} - v = (\operatorname{par} U)^{\perp}]$. (v): The left-hand inclusion follows from (ii). Note that, as a subdifferential operator, A is paramonotone. To prove the identity, apply Fact 3.3(v) to $({}_vA, B_v)$ and use [8, Proposition 2.24]. The right-hand inclusion follows from [16, Lemma 2.4(v)]. (vi): Let $y \in Z_v$ and note that $y \in U \cap \operatorname{Fix}(T_{-v})$ by (iv) and (v). Therefore, by [16, Proposition 2.5(vii)], $T^n y = y - nv$. Now use (23) and (iii). (vii): By definition of T, we have $P_U T^n x - J_B R_U T^n x = (\operatorname{Id} - T) T^n x$, which proves the first identity. The limit follows from e.g., [24, Corollary 1.5] and [4, Corollary 2.3]. (viii): It follows from [16, Lemma 2.4(vi)] that the sequence $(T^n x + nv)_{n \in \mathbb{N}}$ is Fejér monotone with respect to $\operatorname{Fix}(T_{-v})$, hence bounded. In view of (23) and (iii), we learn that $(\forall n \in \mathbb{N})\ J_A T^n x = P_U T^n x = P_U(T^n x + nv)$ and therefore $(J_A T^n x)_{n \in \mathbb{N}} = (P_U T^n x)_{n \in \mathbb{N}}$ is bounded. (ix): Combine (vii) and (viii). ∎

Here is another instance of Theorem 4.4.

Corollary 4.8. *Let $x \in X$. Suppose that U is a closed affine subspace of X, that $A = N_U$, that B is paramonotone, that $v = P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)} 0 \in \operatorname{ran}(\operatorname{Id} - T)$, that $\operatorname{zer}({}_vA) \cap \operatorname{zer}(B_v) \neq \varnothing$, and that all weak cluster points of $(J_A T^n x)_{n \in \mathbb{N}} = (P_U T^n x)_{n \in \mathbb{N}}$ lie in Z_v. Then $(J_A T^n x)_{n \in \mathbb{N}} = (P_U T^n x)_{n \in \mathbb{N}}$ converges weakly to some point in Z_v.*

Proof. Combine Theorem 4.4 and Proposition 4.7(v), (vi), and (viii). ∎

Before we proceed, we need to recall that, by [17, Corollary 5.3(ii)],

$$\overline{\operatorname{ran}}(\operatorname{Id} - T) = \overline{(\operatorname{dom} A - \operatorname{dom} B)} \cap \overline{(\operatorname{ran} A + \operatorname{ran} B)}. \tag{24}$$

The following provides an example of Corollary 4.8.

Example 4.9. *Suppose that U is a closed linear subspace of X, that $b \in U^\perp \smallsetminus \{0\}$, that $A = N_U$, and that $B = \operatorname{Id} + N_{(-b+U)}$. Then $Z = \varnothing$, $v = b \in \operatorname{ran}(\operatorname{Id} - T)$, $Z_v = \{0\}$, and $K_v = U^\perp$. Moreover, $(\forall x \in X)$ $(\forall n \in \mathbb{N})$ $P_U T^n x = \frac{1}{2^n} P_U x \to 0$, $\|P_{U^\perp} T^n x\| \to \infty$, and $0 \in \operatorname{zer}({}_vA) \cap \operatorname{zer}(B_v)$.*

Proof. By the Brezis-Haraux theorem (see [22, Theorems 3 & 4] or [10, Theorem 24.20]), we have $X = \operatorname{int} X \subseteq \operatorname{int}\operatorname{ran} B = \operatorname{int}(\operatorname{ran}\operatorname{Id} + \operatorname{ran} N_{(-b+U)}) \subseteq X$, hence $\operatorname{ran} B = X$. By (24), $\overline{\operatorname{ran}}(\operatorname{Id} - T) = (U + b - U) \cap (U^\perp + X) = b + U$. Consequently, using (12) and [10, Proposition 3.17], we have

$$v = P_{\overline{\operatorname{ran}}(\operatorname{Id} - T)} 0 = P_{b+U} 0 = b + P_U(-b) = b \in U^\perp \smallsetminus \{0\}. \tag{25}$$

Note that $\operatorname{dom} {}_vA = \operatorname{dom} A = U$ and $\operatorname{dom} B_v = v + \operatorname{dom} B = b - b + U = U$, hence $\operatorname{dom}({}_vA + B_v) = U \cap U = U$. Let $x \in U$. Using (25), we have

$$\begin{aligned} x \in Z_v &\Leftrightarrow 0 \in N_U x - b + x - b + N_{-b+U}(x - b) = N_U x - b + x - b + N_U x && \text{(26a)} \\ &\Leftrightarrow 0 \in U^\perp - b + x - b + U^\perp = x + U^\perp \Leftrightarrow x \in U^\perp, && \text{(26b)} \end{aligned}$$

hence $Z_v = \{0\}$, as claimed. As subdifferential operators, both A and B are paramonotone, and so are the translated operators ${}_vA$ and B_v. Since $Z_v = \{0\}$, in view of [9, Remark 5.4] and (25), we learn that

$$K_v = (N_U 0 - b) \cap (0 - b + N_{-b+U}(0 - b)) = (U^\perp - b) \cap (-b + U^\perp) = U^\perp. \tag{27}$$

Next, we claim that

$$(\forall x \in X) \quad P_U T x = \tfrac{1}{2} P_U x. \tag{28}$$

Indeed, note that $J_B = (\operatorname{Id} + B)^{-1} = (2\operatorname{Id} + N_{-b+U})^{-1} = (2\operatorname{Id} + 2N_{-b+U})^{-1} = (\operatorname{Id} + N_{-b+U})^{-1} \circ (\frac{1}{2}\operatorname{Id}) = P_{-b+U} \circ (\frac{1}{2}\operatorname{Id}) = -b + \frac{1}{2}P_U$, where the last identity follows from [10, Proposition 3.17)] and (25). Now, using that[4] $P_U R_U = P_U$ and (25) we have $P_U T = P_U(P_{U^\perp} + J_B R_U) = P_U J_B R_U = P_U(-b + \frac{1}{2}P_U)R_U = P_U(-b + \frac{1}{2}P_U) = \frac{1}{2}P_U$. To show that $(\forall x \in X)$ $(\forall n \in \mathbb{N})$ $P_U T^n x = \frac{1}{2^n} P_U x$, we use induction. Let $x \in X$. Clearly, when $n = 0$, the base case holds. Now suppose that for some $n \in \mathbb{N}$, we have, for every $x \in X$, $P_U T^n x = \frac{1}{2^n} P_U x$. Applying the inductive hypothesis with x replaced by Tx,

[4] It follows from [10, Corollary 3.20] that P_U is linear, hence, $P_U R_U = P_U(2P_U - \operatorname{Id}) = 2P_U - P_U = P_U$.

and using (28), we have $P_U T^{n+1}x = P_U T^n(Tx) = \frac{1}{2^n} P_U Tx = \frac{1}{2^n} P_U(\frac{1}{2} P_U x) = \frac{1}{2^{n+1}} P_U x$, as claimed. Finally, using (25) and [50, Corollary 6(a)], we have $\|T^n x\| \to +\infty$, hence

$$\|P_{U^\perp} T^n x\|^2 = \|T^n x\|^2 - \|P_U T^n x\|^2 = \|T^n x\|^2 - \tfrac{1}{4^n}\|P_U x\|^2 \to +\infty. \tag{29}$$

■

In fact, as we shall now see, the shadow sequence may be unbounded in the general case, even when one of the operators is a normal cone operator.

Remark 4.10. (shadows in the presence of normal solutions)

(i) *Example 4.9 illustrates that even when normal solutions exists, the shadows need not converge. Indeed, we have $K_v = U^\perp \neq \varnothing$ but the* dual *shadows satisfy $\|P_{U^\perp} T^n x\| \to +\infty$.*

(ii) *Suppose that A and B are defined as in Example 4.9. Set $\widetilde{A} = A^{-1}$, $\widetilde{B} = B^{-\circledcirc}$ and $\widetilde{Z} = Z_{(\widetilde{A},\widetilde{B})}$. By [9, Proposition 2.4(v)] $\widetilde{Z} \neq \varnothing \Leftrightarrow K_{(\widetilde{A},\widetilde{B})} = Z_{(A,B)} \neq \varnothing$, hence $\widetilde{Z} = \varnothing$. Moreover, [8, Remarks 3.13 & 3.5] imply that $v = b \in \operatorname{ran}(\operatorname{Id} - T)$ and $\widetilde{Z}_v = U^\perp + b = U^\perp \neq \varnothing$. However, in the light of (i), the* primal *shadows satisfy $\|J_{\widetilde{A}} T^n x\| = \|J_{A^{-1}} T^n x\| = \|P_{U^\perp} T^n x\| \to +\infty$.*

(iii) *Concerning Theorem 4.4, it would be interesting to find other conditions sufficient for weak convergence of the shadow sequence or to even characterize this behaviour.*

The assumption that $\operatorname{zer}({}_vA) \cap \operatorname{zer}(B_v) \neq \varnothing$ is critical in the conclusions of Proposition 4.7 and Corollary 4.8 as we illustrate in Example 4.11. It also provides another example where the shadow sequence diverges to infinity.

Example 4.11. *Suppose that U is a closed linear subspace of X, that $w \in X \smallsetminus U^\perp$, that $A = N_U$, and that $(\forall x \in X)\ Bx = w$. Let $x \in X$. Then $Z = \varnothing$, $v = P_U w \notin U^\perp$, $Z_v = U$, $T = P_U - w$, and $(\forall n \in \mathbb{N})\ P_U T^n x = P_U x - nP_U w$, hence $\|P_U T^n x\| \to +\infty$. Note that $\operatorname{zer}({}_vA) = \operatorname{zer}(B_v) = \operatorname{zer}({}_vA) \cap \operatorname{zer}(B_v) = \varnothing$.*

Proof. One can readily verify that $(\forall y \in X)\ N_U y + By \subseteq U^\perp + w = U^\perp + P_U w \not\ni 0$ and therefore $Z = \varnothing$. Using (24), we have $\overline{\operatorname{ran}}(\operatorname{Id} - T) = (U - X) \cap (U^\perp + w) = w + U^\perp$. Therefore, using (12) and [10, Proposition 3.17], we have $v = P_{w+U^\perp} 0 = w - P_{U^\perp} w = P_U w \notin U^\perp$. Now let $y \in X$. Then $y \in Z_v \Leftrightarrow [y \in U \text{ and } 0 \in -v + N_U y + B(y - v) = U^\perp - P_U w + w = U^\perp + P_{U^\perp} w = U^\perp]$. Hence $Z_v = U$. Clearly $J_B x = (\operatorname{Id} + B)^{-1} x = x - w$, therefore $T = \operatorname{Id} - P_U + J_B(2P_U - \operatorname{Id}) = \operatorname{Id} - P_U + 2P_U - \operatorname{Id} - w = P_U - w$. Using simple induction, we get $(\forall n \in \mathbb{N})\ T^n x = P_U - (n-1)P_U w - w$. Now apply P_U and notice that P_U is linear. ■

We point out that the assumption that $Z_v \subseteq \operatorname{Fix}(v + T)$ in Theorem 4.4 does not hold true in general, even in the consistent case when $v = 0$ as we illustrate in the next two examples.

Example 4.12 ($v \neq 0$ **and** $Z_v \not\subseteq \operatorname{Fix}(v+T)$**).** *Suppose that A and B are defined as in Example 4.11, and that $w \notin U$. Then*

$$U = Z_v \not\subseteq \operatorname{Fix}(v+T) = -w + U. \tag{30}$$

Proof. Let $x \in X$. Then $x \in \operatorname{Fix}(v+T) = \operatorname{Fix}(P_U w + P_U - w) = \operatorname{Fix}(P_U - P_{U^\perp} w) \Leftrightarrow x = P_U x - P_{U^\perp} w \Leftrightarrow P_{U^\perp} x = -P_{U^\perp} w \Leftrightarrow P_{U^\perp}(x+w) = 0 \Leftrightarrow x + w \in U \Leftrightarrow x \in -w + U$. Using Example 4.11, we have $U = Z_v \subseteq \operatorname{Fix}(v+T) = -w + U \Leftrightarrow w \in U$ which is not true. ■

Example 4.13 ($v = 0$ **and** $Z \not\subseteq \operatorname{Fix} T$**).** *Suppose that $(a,b) \in X \times (X \smallsetminus \{0\})$, and that $(\forall x \in X)$ $Ax = x - a$ and $Bx = x - b$. Then*

$$\left\{\tfrac{1}{2}(a+b)\right\} = Z \not\subseteq \operatorname{Fix} T = \{b\} \tag{31}$$

whenever $a \neq b$.

Proof. Clearly $Z = \frac{1}{2}(a+b)$. One can readily verify that $(\forall x \in x)$ $J_A x = \frac{1}{2}(x+a)$, $J_B x = \frac{1}{2}(x+b)$, hence $R_A x = a$ and $R_B x = b$. Therefore, $R_B R_A x = b$. Using (5), we learn that $\operatorname{Fix} T = \operatorname{Fix}(R_B R_A) = \{b\}$ and the conclusion follows. ■

5 The Douglas–Rachford operator and the solution sets

Working in $X \times X$, we recall that (see, e.g., [10, Proposition 23.16])

$$A \times B \text{ is maximally monotone} \quad \text{and} \quad J_{A\times B} = J_A \times J_B. \tag{32}$$

Corollary 5.1. *Let $x \in X$ and let $(z,k) \in \mathcal{S}$. Then*

$$\begin{aligned} \|(J_A Tx, J_{A^{-1}} Tx) - (z,k)\|^2 &= \|J_A Tx - z\|^2 + \|J_{A^{-1}} Tx - k\|^2 && \text{(33a)} \\ &\leq \|J_A x - z\|^2 + \|J_{A^{-1}} x - k\|^2 && \text{(33b)} \\ &= \|(J_A x, J_{A^{-1}} x) - (z,k)\|^2. && \text{(33c)} \end{aligned}$$

Proof. It follows from [9, Theorem 4.5] that $z + k \in \operatorname{Fix} T$, $J_A(z+k) = z$ and $J_{A^{-1}}(z+k) = k$. Now combine with Theorem 2.7(v) with y replaced by $z+k$. ■

We recall, as consequence of [10, Corollary 22.19] and Example 3.2, that when $X = \mathbb{R}$, the operators A and B are paramonotone. In view of Fact 3.3(iv), we then have $\mathcal{S} = Z \times K$.

Lemma 5.2. *Suppose that $X = \mathbb{R}$. Let $x \in X$ and let $(z,k) \in Z \times K$. Then the following hold:*

(i) $|J_ATx - z|^2 \leq |J_Ax - z|^2$.
(ii) $|J_{A^{-1}}Tx - k|^2 \leq |J_{A^{-1}}x - k|^2$.

Proof. (i): Set
$$q(x,z) := |J_ATx - z|^2 - |J_Ax - z|^2. \tag{34}$$
If $x \in \operatorname{Fix} T$ we get $q(x,z) = 0$. Suppose that $x \in \mathbb{R} \smallsetminus \operatorname{Fix} T$. Since T is firmly nonexpansive we have that $\operatorname{Id} - T$ is firmly nonexpansive (see [10, Proposition 4.2]), hence monotone by [10, Example 20.27]. Therefore, $(\forall x \in \mathbb{R} \smallsetminus \operatorname{Fix} T)(\forall f \in \operatorname{Fix} T)$ we have

$$(x - Tx)(x - f) = \big((\operatorname{Id} - T)x - (\operatorname{Id} - T)f\big)\,(x - f) > 0. \tag{35}$$

Notice that (34) can be rewritten as

$$q(x,z) = (J_ATx - J_Ax)\big((J_ATx - z) + (J_Ax - z)\big). \tag{36}$$

We argue by cases.

Case 1: $x < Tx$.
It follows from (35) that
$$(\forall f \in \operatorname{Fix} T)\ x < f. \tag{37}$$
On the one hand, since J_A is firmly nonexpansive, we have J_A is monotone and, therefore, $J_ATx - J_Ax \geq 0$. On the other hand, it follows from Fact 3.1(iii) that $(\exists f \in \operatorname{Fix} T)\ z = J_Af = J_ATf$. Using (37) and the fact that J_A is monotone, we conclude that $J_Ax - z = J_Ax - J_Af \leq 0$. Moreover, since J_A and T are firmly nonexpansive operators on $\mathbb{R}$, we have $J_A \circ T$ is firmly nonexpansive, hence monotone. Therefore, (37) implies that $J_ATx - z = J_ATx - J_ATf \leq 0$. Combining with (36), we conclude that (i) holds.

Case 2: $x > Tx$.
The proof follows similar to *Case 1*.

(ii) Apply the results of (i) to A^{-1} and use (3). ■

In view of (11), one might conjecture that Corollary 5.1 holds when we replace $\mathcal{S}$ by $Z \times K$. The following example gives a negative answer to this conjecture. It also illustrates that when $X \neq \mathbb{R}$, the conclusion of Lemma 5.2 fails in general.

Example 5.3. *Suppose that $X = \mathbb{R}^2$, that A is the normal cone operator of $\mathbb{R}^2_+$, and that $B : X \to X : (x_1, x_2) \mapsto (-x_2, x_1)$ is the rotator by $\pi/2$. Then $\operatorname{Fix} T = \mathbb{R}_+ \cdot (1, -1)$, $Z = \mathbb{R}_+ \times \{0\}$ and $K = \{0\} \times \mathbb{R}_+$. Moreover, $(\exists x \in \mathbb{R}^2)$ $(\exists (z,k) \in Z \times K)$ such that $\|(J_ATx, J_{A^{-1}}Tx) - (z,k)\|^2 - \|(J_Ax, J_{A^{-1}}x) - (z,k)\|^2 > 0$ and $\|J_ATx - z\|^2 - \|J_Ax - z\|^2 > 0$.*

Proof. Let $(x_1, x_2) \in \mathbb{R}^2$. Using [8, Proposition 2.10], we have

$$J_B(x_1,x_2) = \big(\tfrac{1}{2}(x_1 + x_2), \tfrac{1}{2}(-x_1 + x_2)\big) \quad \text{and} \quad R_B(x_1,x_2) = (x_2, -x_1) = -B(x_1,x_2). \tag{38}$$

Hence, $R_B^{-1} = (-B)^{-1} = B$. Using (5), we conclude that $(x_1, x_2) \in \operatorname{Fix} T \Leftrightarrow (x_1, x_2) \in \operatorname{Fix} R_B R_A$. Hence

$$\begin{aligned}
\operatorname{Fix} T &= \big\{(x_1,x_2) \in \mathbb{R}^2 \bigm| (x_1,x_2) = R_B R_A(x_1,x_2)\big\} && \text{(39a)}\\
&= \big\{(x_1,x_2) \in \mathbb{R}^2 \bigm| R_B^{-1}(x_1,x_2) = 2J_A(x_1,x_2) - (x_1,x_2)\big\} && \text{(39b)}\\
&= \big\{(x_1,x_2) \in \mathbb{R}^2 \bigm| B(x_1,x_2) + (x_1,x_2) = 2P_{\mathbb{R}^2_+}(x_1,x_2)\big\} && \text{(39c)}\\
&= \big\{(x_1,x_2) \in \mathbb{R}^2 \bigm| (x_1 - x_2, x_1 + x_2) = 2P_{\mathbb{R}^2_+}(x_1,x_2)\big\}. && \text{(39d)}
\end{aligned}$$

We argue by cases.

Case 1: $x_1 \geq 0$ and $x_2 \geq 0$. Then $(x_1, x_2) \in \operatorname{Fix} T \Leftrightarrow (x_1 - x_2, x_1 + x_2) = 2P_{\mathbb{R}^2_+}(x_1, x_2) = (2x_1, 2x_2) \Leftrightarrow x_1 = -x_2$ and $x_1 = x_2 \Leftrightarrow x_1 = x_2 = 0$.

Case 2: $x_1 < 0$ and $x_2 < 0$. Then $(x_1, x_2) \in \operatorname{Fix} T \Leftrightarrow (x_1 - x_2, x_1 + x_2) = 2P_{\mathbb{R}^2_+}(x_1, x_2) = (0, 0) \Leftrightarrow x_1 = x_2$ and $x_1 = -x_2 \Leftrightarrow x_1 = x_2 = 0$, which contradicts that $x_1 < 0$ and $x_2 < 0$.

Case 3: $x_1 \geq 0$ and $x_2 < 0$. Then $(x_1, x_2) \in \operatorname{Fix} T \Leftrightarrow (x_1 - x_2, x_1 + x_2) = 2P_{\mathbb{R}^2_+}(x_1, x_2) = (2x_1, 0) \Leftrightarrow x_1 = -x_2$.

Case 4: $x_1 < 0$ and $x_2 \geq 0$. Then $(x_1, x_2) \in \operatorname{Fix} T \Leftrightarrow (x_1 - x_2, x_1 + x_2) = 2P_{\mathbb{R}^2_+}(x_1, x_2) = (0, 2x_2) \Leftrightarrow x_1 = x_2$, which never occurs since $x_1 < 0$ and $x_2 \geq 0$. Altogether, we conclude that $\operatorname{Fix} T = R_+ \cdot (1, -1)$, as claimed.

Using Fact 3.1(iii)&(iv), we have $Z = J_A(\operatorname{Fix} T) = \mathbb{R}_+ \times \{0\}$, and $K = J_{A^{-1}}(\operatorname{Fix} T) = (\operatorname{Id} - J_A)(\operatorname{Fix} T) = \{0\} \times \mathbb{R}_-$.

Now let $a > 0$, let $x = (a, 0)$, set $z := (2a, 0) \in Z$, and set $k := (0, -a) \in K$. Notice that $Tx = x - P_{\mathbb{R}^2_+} + \frac{1}{2}(\operatorname{Id} - B)x = (a, 0) - (a, 0) + \frac{1}{2}((a, 0) - (0, a)) = (\frac{1}{2}a, -\frac{1}{2}a)$. Hence, $J_A x = P_{\mathbb{R}^2_+}(a, 0) = (a, 0)$, $J_{A^{-1}}x = P_{\mathbb{R}^2_-}(a, 0) = (0, 0)$, $J_A Tx = P_{\mathbb{R}^2_+}(\frac{1}{2}a, -\frac{1}{2}a) = (\frac{1}{2}a, 0)$, and $J_{A^{-1}}x = P_{\mathbb{R}^2_-}(\frac{1}{2}a, -\frac{1}{2}a) = (0, -\frac{1}{2}a)$. Therefore

$$\begin{aligned}
&\|(J_A Tx, J_{A^{-1}}Tx) - (z,k)\|^2 - \|(J_A x, J_{A^{-1}}x) - (z,k)\|^2\\
&= \|J_A Tx - z\|^2 + \|J_{A^{-1}}Tx - k\|^2 - \|J_A x - z\|^2 - \|J_{A^{-1}}x - k\|^2\\
&= \|(\tfrac{1}{2}a, 0) - (2a, 0)\|^2 + \|(0, -\tfrac{1}{2}a) - (0, -a)\|^2 - \|(a, 0) - (2a, 0)\|^2 - \|(0, 0) - (0, -a)\|^2\\
&= \tfrac{9}{4}a^2 + \tfrac{1}{4}a^2 - a^2 - a^2 = \tfrac{1}{2}a^2 > 0.
\end{aligned}$$

Similarly, one can verify that $\|J_A Tx - z\|^2 - \|J_A x - z\|^2 = \frac{5}{4}a^2 > 0$. ■

Throughout the rest of this section, we assume that[5]

$$A : X \rightrightarrows X \text{ and } B : X \rightrightarrows X \text{ are maximally monotone linear relations;}$$

equivalently, by [15, Theorem 2.1(xviii)], that

$$J_A \text{ and } J_B \text{ are linear operators from } X \text{ to } X. \tag{40}$$

This additional assumption leads to stronger conclusions.

Lemma 5.4. $\mathrm{Id} - T = J_A - 2J_BJ_A + J_B$.

Proof. Let $x \in X$. Then indeed, $x - Tx = J_Ax - J_BR_Ax = J_Ax - J_B(2J_Ax - x) = J_Ax - 2J_BJ_Ax + J_Bx$. ■

Lemma 5.5. *Suppose that U is a linear subspace of X and that $A = P_U$. Then A is maximally monotone,*

$$J_A = J_{P_U} = \tfrac{1}{2}(\mathrm{Id} + P_{U^\perp}), \quad \textit{and} \quad R_A = P_{U^\perp} = \mathrm{Id} - A. \tag{41}$$

Proof. Let $(x, y) \in X \times X$. Then

$$y = J_Ax \Leftrightarrow x = y + P_Uy. \tag{42}$$

Now assume that $y = J_Ax$. Since P_U is linear, (42) implies that $P_{U^\perp}x = P_{U^\perp}y$. Moreover, $y = x - P_Uy = \frac{1}{2}(x + x - 2P_Uy) = \frac{1}{2}(x + y + P_Uy - 2P_Uy) = \frac{1}{2}(x + y - P_Uy) = \frac{1}{2}(x + P_{U^\perp}y) = \frac{1}{2}(x + P_{U^\perp}x)$. Finally, $R_A = 2J_A - \mathrm{Id} = (\mathrm{Id} + P_{U^\perp}) - \mathrm{Id} = P_{U^\perp}$. ■

We say that a linear relation A is *skew* (see, e.g., [18]), if $(\forall (a, a^*) \in \operatorname{gra} A)\ \langle a, a^* \rangle = 0$.

Lemma 5.6. *Suppose that $A : X \to X$ and $B : X \to X$ are both skew, and $A^2 = B^2 = -\mathrm{Id}$. Then $\mathrm{Id} - T = \frac{1}{2}(\mathrm{Id} - BA)$.*

Proof. It follows from [8, Proposition 2.10] that $R_A = A$ and $R_B = B$. Therefore, (5) implies that $\mathrm{Id} - T = \frac{1}{2}(\mathrm{Id} - R_BR_A) = \frac{1}{2}(\mathrm{Id} - BA)$. ■

Example 5.7. *Suppose that A and B are skew. Let $x \in X$, and let $y \in X$. Then the following hold:*

(i) $\langle Tx - Ty, x - y \rangle = \|Tx - Ty\|^2$.
(ii) $\langle (\mathrm{Id} - T)x - (\mathrm{Id} - T)y, x - y \rangle = \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2$.
(iii) $\|x - y\|^2 = \|Tx - Ty\|^2 + \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2$.

[5] $A\colon X \rightrightarrows X$ is a *linear relation* if $\operatorname{gra} A$ is a linear subspace of $X \times X$.

(iv) $\|J_A x - J_A y\|^2 + \|J_{A^{-1}} x - J_{A^{-1}} y\|^2 - \|J_A Tx - J_A Ty\|^2 - \|J_{A^{-1}} Tx - J_{A^{-1}} Ty\|^2$
$= \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2$.
(v) $\|x\|^2 = \|Tx\|^2 + \|x - Tx\|^2$.
(vi) $\langle Tx, x - Tx\rangle = 0$.

Proof. (i)–(iv): Apply Theorem 2.7, and use (4) as well as the skewness of A and B. (v): Apply (iii) with $y = 0$. (vi): We have $2\langle Tx, x - Tx\rangle = \|x\|^2 - \|Tx\|^2 - \|x - Tx\|^2$. Now apply (v). ■

Suppose that U is a closed affine subspace of X. One can easily verify that

$$(\forall x \in X)(\forall y \in X) \quad \langle P_U x - P_U y, (\mathrm{Id} - P_U)x - (\mathrm{Id} - P_U)y\rangle = 0. \tag{43}$$

Example 5.8. *Suppose that U and V are closed affine subspaces of X such that $U \cap V \neq \varnothing$, that $A = N_U$, and that $B = N_V$. Let $x \in X$, and let $(z, k) \in Z \times K$. Then*

$$\|(P_U x, (\mathrm{Id} - P_U)x) - (z,k)\|^2 - \|(P_U Tx, (\mathrm{Id} - P_U)Tx) - (z,k)\|^2 \tag{44a}$$
$$= \|x - (z+k)\|^2 - \|Tx - (z+k)\|^2 \tag{44b}$$
$$= \|x - Tx\|^2 \tag{44c}$$
$$= \|P_U x - P_V x\|^2. \tag{44d}$$

Proof. As subdifferential operators, A and B are paramonotone (by Example 3.2). It follows from Fact 3.3(v), and [9, Theorem 4.5] that

$$z + k \in \operatorname{Fix} T, \quad P_U(z+k) = z, \quad \text{and} \quad (\mathrm{Id} - P_U)(z+k) = k. \tag{45}$$

Hence, in view of (43), we have

$$\|(P_U x, (\mathrm{Id} - P_U)x) - (z,k)\|^2 \tag{46a}$$
$$= \|P_U x - z\|^2 + \|(\mathrm{Id} - P_U)x - k\|^2 \tag{46b}$$
$$= \|P_U x - P_U(z+k)\|^2 + \|(\mathrm{Id} - P_U)x - (\mathrm{Id} - P_U)(z+k)\|^2 \tag{46c}$$
$$+ 2\langle P_U x - P_U(z+k), (\mathrm{Id} - P_U)x - (\mathrm{Id} - P_U)(z+k)\rangle \tag{46d}$$
$$= \|P_U x - P_U(z+k) + (\mathrm{Id} - P_U)x - (\mathrm{Id} - P_U)(z+k)\|^2 \tag{46e}$$
$$= \|x - (z+k)\|^2. \tag{46f}$$

Applying (46), with x replaced by Tx, yields

$$\|(P_U Tx, (\mathrm{Id} - P_U)Tx) - (z,k)\|^2 = \|Tx - (z+k)\|^2. \tag{47}$$

Combining (46) and (47) yields (44b). It follows from (43), and Theorem 2.7(iii) applied with (A, B, y) replaced by $(N_U, N_V, z+k)$, that $\|x - (z+k)\|^2 - \|Tx - T(z+k)\|^2 = \|x - Tx - ((z+k) - T(z+k))\|^2$. In view of (45), this proves (44c).

Now we turn to (44d). Let $w \in U \cap V$. Then $U = w + \operatorname{par} U$ and $V = w + \operatorname{par} V$. Suppose momentarily that $w = 0$. In this case, $\operatorname{par} U = U$ and $\operatorname{par} V = V$. Using [7, Proposition 3.4(i)], we have

$$T = T_{(U,V)} = P_V P_U + P_{V^\perp} P_{U^\perp}. \tag{48}$$

Therefore

$$\begin{aligned} x - Tx &= P_U x + P_{U^\perp} x - P_V P_U x - P_{V^\perp} P_{U^\perp} x = (\mathrm{Id} - P_V) P_U x + (\mathrm{Id} - P_{V^\perp}) P_{U^\perp} x && \text{(49a)} \\ &= P_{V^\perp} P_U x + P_V P_{U^\perp} x. && \text{(49b)} \end{aligned}$$

Using (49b), we have

$$\begin{aligned} \|x - Tx\|^2 &= \|P_{V^\perp} P_U x + P_V P_{U^\perp} x\|^2 = \|P_U x - P_V P_U x + P_V x - P_V P_U x\|^2 && \text{(50a)} \\ &= \|P_U x - 2 P_V P_U x + P_V x\|^2 && \text{(50b)} \\ &= \|P_U x\|^2 + \|P_V x\|^2 + 4\|P_V P_U x\|^2 && \text{(50c)} \\ &\quad + 2\langle P_U x, P_V x\rangle - 4\langle P_U x, P_V P_U x\rangle - 4\langle P_V x, P_V P_U x\rangle && \text{(50d)} \\ &= \|P_U x\|^2 + \|P_V x\|^2 - 2\langle P_U x, P_V x\rangle = \|P_U x - P_V x\|^2. && \text{(50e)} \end{aligned}$$

Now, if $w \neq 0$, then by [14, Proposition 5.3] we have $Tx = T_{(\operatorname{par} U, \operatorname{par} V)}(x - w) + w$. Therefore, (50) yields $\|x - Tx\|^2 = \|(x - w) - T_{(\operatorname{par} U, \operatorname{par} V)}(x - w)\|^2 = \|P_{\operatorname{par} U}(x - w) - P_{\operatorname{par} V}(x - w)\|^2 = \|w + P_{\operatorname{par} U}(x - w) - (w + P_{\operatorname{par} V}(x - w))\|^2 = \|P_U x - P_V x\|^2$, where the last equality follows from [10, Proposition 3.17]. ■

6 A new proof of the Lions-Mercier-Svaiter theorem

In this section, we work under the assumptions that

$$Z \neq \varnothing \quad \text{and} \quad \operatorname{Fix} T \neq \varnothing.$$

Parts of the following two results are implicit in [54]; however, our proofs are different.

Proposition 6.1. *Let $x \in X$. Then the following hold:*

(i) $T^n x - T^{n+1} x = J_A T^n x - J_B R_A T^n x = J_{A^{-1}} T^n x + J_{B^{-1}} R_A T^n x \to 0$.
(ii) *The sequence* $(J_A T^n x, J_B R_A T^n x, J_{A^{-1}} T^n x, J_{B^{-1}} R_A T^n x)_{n \in \mathbb{N}}$ *is bounded and lies in* $\operatorname{gra}(A \times B)$.

Suppose that (a, b, a^*, b^*) is a weak cluster point of $(J_A T^n x, J_B R_A T^n x, J_{A^{-1}} T^n x, J_{B^{-1}} R_A T^n x)_{n \in \mathbb{N}}$. Then:

(iii) $a - b = a^* + b^* = 0$.
(iv) $\langle a, a^*\rangle + \langle b, b^*\rangle = 0$.
(v) $(a, a^*) \in \operatorname{gra} A$ *and* $(b, b^*) \in \operatorname{gra} B$.
(vi) *For every* $x \in X$, *the sequence* $(J_A T^n x, J_{A^{-1}} T^n x)_{n\in\mathbb{N}}$ *is bounded and its weak cluster points lie in* $\mathcal{S}$.

Proof. (i): Apply Lemma 2.6(i) with x replaced by $T^n x$. The claim of the strong limit follows from combining Fact 3.1(i) and [4, Corollary 2.3] or [10, Theorem 5.14(ii)]. (ii): The boundedness of the sequence follows from the weak convergence of $(T^n x)_{n\in\mathbb{N}}$ (see, e.g.,[10, Theorem 5.14(iii)]), and the nonexpansiveness of the resolvents and reflected resolvents of monotone operators (see, e.g., [10, Corollary 23.10(i) and (ii)]). Now apply Lemma 2.6(ii) with x replaced by $T^n x$. (iii): This follows from taking the weak limit along the subsequences in (i). (iv): In view of (iii) we have $\langle a, a^*\rangle + \langle b, b^*\rangle = \langle a, a^* + b^*\rangle = \langle a, 0\rangle = 0$. (v): Let $((x, y),(u, v)) \in \operatorname{gra}(A \times B)$ and set

$$a_n := J_A T^n x, a_n^* := J_{A^{-1}} T^n x, b_n := J_B R_A T^n x, b_n^* := J_{B^{-1}} R_A T^n x. \tag{51}$$

Applying Lemma 2.2 with (a, b, a^*, b^*) replaced by (a_n, b_n, a_n^*, b_n^*) yields

$$\begin{aligned}\langle (a_n, b_n) - (x, y),(a_n^*, b_n^*) - (u, v)\rangle &= \langle a_n - b_n, a_n^*\rangle + \langle x, u\rangle - \langle x, a_n^*\rangle - \langle a_n - b_n, u\rangle \\ &\quad + \langle b_n, a_n^* + b_n^*\rangle + \langle y, v\rangle - \langle y, b_n^*\rangle - \langle b_n, u + v\rangle.\end{aligned} \tag{52}$$

By (32), $A \times B$ is monotone. In view of (51), (52) and Proposition 6.1(ii), we deduce that

$$\begin{aligned}&\langle a_n - b_n, a_n^*\rangle + \langle x, u\rangle - \langle x, a_n^*\rangle - \langle a_n - b_n, u\rangle \\ &+ \langle b_n, a_n^* + b_n^*\rangle + \langle y, v\rangle - \langle y, b_n^*\rangle - \langle b_n, u + v\rangle \geq 0.\end{aligned} \tag{53}$$

Taking the limit in (53) along a subsequence and using (51), Proposition 6.1(i), (iii) and (iv), we obtain

$$\begin{aligned}0 &\leq \langle x, u\rangle - \langle x, a^*\rangle + \langle y, v\rangle - \langle y, b^*\rangle - \langle b, u + v\rangle \\ &= \langle x, u\rangle - \langle x, a^*\rangle + \langle y, v\rangle - \langle y, b^*\rangle - \langle a, u\rangle - \langle b, v\rangle + \langle a, a^*\rangle + \langle b, b^*\rangle \\ &= \langle a - x, a^* - u\rangle + \langle b - y, b^* - v\rangle = \langle (a, b) - (x, y),(a^*, b^*) - (u, v)\rangle.\end{aligned} \tag{54}$$

By maximality of $A \times B$ (see (32)), we deduce that $((a, b),(a^*, b^*)) \in \operatorname{gra}(A \times B)$. Therefore, $(a, a^*) \in \operatorname{gra} A$ and $(b, b^*) \in \operatorname{gra} B$. (vi): The boundedness of the sequence follows from (ii). Now let (a, b, a^*, b^*) be a weak cluster point of $(J_A T^n x, J_B R_A T^n x, J_{A^{-1}} T^n x, J_{B^{-1}} R_A T^n x)_{n\in\mathbb{N}}$. By (v), we know that $(a, a^*) \in \operatorname{gra} A$ and $(b, b^*) = (a, b^*) \in \operatorname{gra} B$, which, in view of (iv), implies $a^* \in Aa$ and $-a^* = b^* \in Bb = Ba$. Hence, $(a, a^*) \in \mathcal{S}$, as claimed (see (11)). ■

Theorem 6.2. *Let* $x \in X$ *and let* $(z, k) \in \mathcal{S}$. *Then the following hold:*

(i) *For every* $n \in \mathbb{N}$,

$$\|(J_A T^{n+1}x, J_{A^{-1}} T^{n+1}x) - (z,k)\|^2 = \|J_A T^{n+1}x - z\|^2 + \|J_{A^{-1}} T^{n+1}x - k\|^2 \tag{55a}$$

$$\leq \|J_A T^n x - z\|^2 + \|J_{A^{-1}} T^n x - k\|^2 \tag{55b}$$

$$= \|(J_A T^n x, J_{A^{-1}} T^n x) - (z,k)\|^2. \tag{55c}$$

(ii) *The sequence* $(J_A T^n x, J_{A^{-1}} T^n x)_{n\in\mathbb{N}}$ *is Fejér monotone with respect to* $\mathcal{S}$.

(iii) *The sequence* $(J_A T^n x, J_{A^{-1}} T^n x)_{n\in\mathbb{N}}$ *converges weakly to some point in* $\mathcal{S}$.

Proof. (i): Apply Corollary 5.1 with x replaced by $T^n x$. (ii): This follows directly from (i). (iii): Combine Proposition 6.1(vi), (ii), Fact 3.3(iii) and [10, Theorem 5.5]. ■

We are now ready for the main result of this section, i.e., an alternative shorter proof of the Lions–Mercier–Svaiter result.

Corollary 6.3. (Lions–Mercier–Svaiter). $(J_A T^n x)_{n\in\mathbb{N}}$ *converges weakly to some point in* Z.

Proof. This follows from Theorem 6.2(iii) and (11); see also Lions and Mercier's [47, Theorem 1] and Svaiter's [54, Theorem 1]. ■

Remark 6.4 (brief history). *The Douglas–Rachford algorithm has its roots in the 1956 paper [35] as a method for solving a system of linear equations, where the matrices are symmetric and positive semi-definite. In 1969, Lieutaud (see [46]) extended their method to deal with (possibly nonlinear) maximally monotone operators that are defined everywhere. Lions and Mercier, in their paper [47] from 1979, presented a broad and powerful generalization to its current form, i.e., to handle the sum of any two maximally monotone operators that are possibly nonlinear, possibly set-valued and not necessarily defined everywhere. (For details on this connection, we refer the reader to [46] and [28].) In their seminal work, they showed that, for every* $x \in X$, $(T^n x)_{n\in\mathbb{N}}$ *converges weakly to a point in* $\operatorname{Fix} T$ *and that the bounded shadow sequence* $(J_A T^n x)_{n\in\mathbb{N}}$ *has all its weak cluster points in* $\operatorname{zer}(A+B)$, *provided that* $A+B$ *was maximally monotone. (Note that resolvents are* not *weakly continuous in general; see, e.g., [55] or [10, Example 4.12].) In their joint work from 1992, Eckstein and Bertsekas proved that* $J_A(\operatorname{Fix} T) \subseteq \operatorname{zer}(A+B)$ *(see [37, Theorem 5]). Later on, in 2006, Combettes refined the results by Eckstein and Bertsekas by providing the first characterization of the set of zeros of the sum to be precisely the shadows of the fixed points of* T, *namely* $J_A(\operatorname{Fix} T)$ *(see [28, Lemma 2.6(iii)]). In the finite-dimensional setting, together with the earlier results by Lions and Mercier [47], the work by Eckstein and Bertsekas and later by Combettes, asserts the convergence of the shadow sequence to a solution of the sum (without assuming maximality). Working with the shadow sequence, we point out that the explicit proof of weak convergence of the shadow sequence under additional assumptions and strong convergence results in special cases can be found in [29]. The first proof that the weak cluster points of the shadow sequence are zeros of* $A+B$ *(without assuming the maximality of the sum) in the convex feasibility setting appeared in [12, Fact 5.9] in 2002. Building on [6] and*

[38], Svaiter provided a complete answer in 2011 (see [54]) demonstrating that $A + B$ does not have to be maximally monotone and that the shadow sequence $(J_A T^n x)_{n\in\mathbb{N}}$ in fact does converge weakly to a point in $\operatorname{zer}(A + B)$. (He used Theorem 6.2; however, his proof differs from ours which is more in the style of the original paper by Lions and Mercier [47].) Nonetheless, when $Z = \varnothing$, the complete understanding of $(J_A T^n x)_{n\in\mathbb{N}}$ remains open — to the best of our knowledge, Theorem 4.4 is currently the most powerful result available.

In our final result, we show that when $X = \mathbb{R}$, the Fejér monotonicity of the sequence $(J_A T^n x, J_{A^{-1}} T^n x)_{n\in\mathbb{N}}$ with respect to S can be decoupled to yield Fejér monotonicity of $(J_A T^n x)_{n\in\mathbb{N}}$ and $(J_{A^{-1}} T^n x)_{n\in\mathbb{N}}$ with respect to Z and K, respectively.

Lemma 6.5. *Suppose that $X = \mathbb{R}$. Let $x \in X$ and let $(z, k) \in Z \times K$. Then the following hold:*

(i) *The sequence $(J_A T^n x)_{n\in\mathbb{N}}$ is Fejér monotone with respect to Z.*
(ii) *The sequence $(J_{A^{-1}} T^n x)_{n\in\mathbb{N}}$ is Fejér monotone with respect to K.*

Proof. Apply Lemma 5.2 with x replaced by $T^n x$. ■

We point out that the conclusion of Lemma 6.5 does not hold when $\dim X \geq 2$, see, Example 5.3 or [7, Section 5 & Figure 1].

Acknowledgments

We are grateful to Patrick Combettes for helpful comments and pointing out additional references. We also would like to thank the anonymous referees for comments that were constructive. HHB was partially supported by the Natural Sciences and Engineering Research Council of Canada and by the Canada Research Chair Program.

References

[1] F.J. Aragón Artacho and J.M. Borwein, Global convergence of a non-convex Douglas–Rachford iteration, *Journal of Global Optimization* 57 (2013), 753–769.
[2] F.J. Aragón Artacho, J.M. Borwein, and M.K. Tam, Recent results on Douglas–Rachford methods for combinatorial optimization problems, *Journal of Global Optimization* 163 (2014), 1–30.
[3] H. Attouch and M. Théra, A general duality principle for the sum of two operators, *Journal of Convex Analysis* 3 (1996), 1–24.
[4] J.B. Baillon, R.E. Bruck, and S. Reich, On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces, *Houston Journal of Mathematics* 4 (1978), 1–9.

[5] H.H. Bauschke, *Projection Algorithms and Monotone Operators*, PhD thesis, Simon Fraser University, Burnaby, B.C., Canada, August 1996.

[6] H.H. Bauschke, A note on the paper by Eckstein and Svaiter on "General projective splitting for sums of maximal monotone operators", *SIAM Journal on Control and Optimization* 48 (2009), 2513–2515.

[7] H.H. Bauschke, J.Y. Bello Cruz, T.T.A. Nghia, H.M. Phan, and X. Wang, The rate of linear convergence of the Douglas–Rachford algorithm for subspaces is the cosine of the Friedrich's angle, *Journal of Approximation Theory* 185 (2014), 63–79.

[8] H.H. Bauschke, W.L. Hare, and W.M. Moursi, Generalized solutions for the sum of two maximally monotone operators, *SIAM Journal on Control and Optimization* 52 (2014), 1034–1047.

[9] H.H. Bauschke, R.I. Boţ, W.L. Hare, and W.M. Moursi, Attouch–Théra duality revisited: paramonotonicity and operator splitting, *Journal of Approximation Theory* 164 (2012), 1065–1084.

[10] H.H. Bauschke and P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, 2011.

[11] H.H. Bauschke, P.L. Combettes, and D.R. Luke, Finding best approximation pairs relative to two closed convex sets in Hilbert spaces, *Journal of Approximation Theory* 127 (2004), 178–192.

[12] H.H. Bauschke, P.L. Combettes, and D.R. Luke, Phase retrieval, error reduction algorithm, and Fienup variants: a view from convex optimization, *Journal of the Optical Society of America* 19 (2002), 1334–1345.

[13] H.H. Bauschke, M.N. Dao, and W.M. Moursi, On Fejér monotone sequences and nonexpansive mappings, *Linear and Nonlinear Analysis* 1 (2015), 287–295.

[14] H.H. Bauschke, B. Lukens, and W.M. Moursi, Affine nonexpansive operators, Attouch–Théra duality and the Douglas–Rachford algorithm, `arXiv:1603.09418 [math.OC]`.

[15] H.H. Bauschke, S.M. Moffat, and X. Wang, Firmly nonexpansive mappings and maximally monotone operators: correspondence and duality, *Set-Valued and Variational Analysis* 20 (2012), 131–153.

[16] H.H. Bauschke and W.M. Moursi, The Douglas–Rachford algorithm for two (not necessarily intersecting) affine subspaces, to appear in *SIAM Journal on Optimization* 26 (2016), 968–985.

[17] H.H. Bauschke, W.L. Hare, and W.M. Moursi, On the range of the Douglas–Rachford operator, *Mathematics of Operation Research* 41 (2016), 884–897.

[18] H.H. Bauschke, X. Wang, and L. Yao, Examples of discontinuous maximal monotone linear operators and the solution to a recent problem posed by B.F. Svaiter, *Journal of Mathematical Analysis and Applications* 370 (2010), 224–241.

[19] J.M. Borwein, Fifty years of maximal monotonicity, *Optimization Letters* 4 (2010), 473–490.

[20] J.M. Borwein and B. Sims, The Douglas–Rachford algorithm in the absence of convexity, in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, Optimization and Applications 49, Springer, New York, pp. 93–109, 2011.

[21] J.M. Borwein and M.K. Tam, The cyclic Douglas–Rachford method for inconsistent feasibility problems. *Journal of Nonlinear Convex Analysis* 16 (2015), 537–584.

[22] H. Brezis and A. Haraux, Image d'une Somme d'opérateurs Monotones et Applications, *Israel Journal of Mathematics* 23 (1976), 165–186.

[23] H. Brezis, *Operateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert*, North-Holland/Elsevier, 1973.

[24] R.E. Bruck and S. Reich, Nonexpansive projections and resolvents of accretive operators in Banach spaces, *Houston Journal of Mathematics* 3 (1977), 459–470.
[25] R.S. Burachik and A.N. Iusem, *Set-Valued Mappings and Enlargements of Monotone Operators*, Springer-Verlag, 2008.
[26] P.L. Combettes, Inconsistent signal feasibility problems: Least-squares solutions in a product space, *IEEE Transactions on Signal Processing* 42 (1994), 2955–2966.
[27] P.L. Combettes, The convex feasibility problem in image recovery, *Advances in Imaging and Electron Physics* 25 (1995), 155–270.
[28] P.L. Combettes, Solving monotone inclusions via compositions of nonexpansive averaged operators, *Optimization* 53 (2004), 475–504.
[29] P.L. Combettes, Iterative construction of the resolvent of a sum of maximal monotone operators, *Journal of Convex Analysis* 16 (2009), 727–748.
[30] P.L. Combettes and J.-C. Pesquet, A Douglas–Rachford splitting approach to nonsmooth convex variational signal recovery, *IEEE Journal of Selected Topics in Signal Processing* 1 (2007), 564–574.
[31] P.L. Combettes and J.-C. Pesquet, A proximal decomposition method for solving convex variational inverse problems, *Inverse Problems* 24, article 065014, (2008).
[32] P.L. Combettes and J.-C. Pesquet, Stochastic quasi-Fejér block-coordinate fixed point iterations with random sweeping, *SIAM Journal on Optimization* 25 (2015), 1221–1248.
[33] L. Demanet and X. Zhang, Eventual linear convergence of the Douglas–Rachford iteration for basis pursuit, *Mathematics of Computation* 85 (2016), 209–238.
[34] D. Drusvyatskiy, C.-K. Li, D.C. Pelejo, Y.-L. Voronin, and H. Wolkowicz, Projection Methods for Quantum Channel Construction, *Quantum Information Processing* 14 (2015), 3075–3095.
[35] J. Douglas and H.H. Rachford, On the numerical solution of the heat conduction problem in 2 and 3 space variables, *Transactions of the AMS* 82 (1956), 421–439.
[36] J. Eckstein, *Splitting Methods for Monotone Operators with Applications to Parallel Optimization*, Ph.D. thesis, MIT, 1989.
[37] J. Eckstein and D.P. Bertsekas, On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators, *Mathematical Programming* 55 (1992), 293–318.
[38] J. Eckstein and B.F. Svaiter, A family of projective splitting methods for the sum of two maximal monotone operators, *Mathematical Programming (Series B)* 111 (2008), 173–199.
[39] V. Elser and I. Rankenburg, Deconstructing the energy landscape: constraint-based algorithm for folding heteropolymers, *Physics Reviews E* 73 (2006), 026702.
[40] V. Elser, I. Rankenburg and P. Thibault, Searching with iterated maps, *Proceedings of the National Academy of Sciences of the United States of America* 102 (2007), 418–423.
[41] `GeoGebra`, `http://www.geogebra.org`.
[42] R. Hesse and D.R. Luke, Nonconvex notions of regularity and convergence of fundamental algorithms for feasibility problems, *SIAM Journal on Optimization* 23 (2013), 2397–2419.
[43] R. Hesse, D.R. Luke, and P. Neumann, Alternating projections and Douglas–Rachford for sparse affine feasibility, *IEEE Transactions on Signal Processing* 62 (2014).
[44] A.N. Iusem, On some properties of paramonotone operators, *Journal of Convex Analysis* 5 (1998), 269–278.
[45] G. Li and T.K. Pong, Douglas–Rachford splitting for nonconvex optimization with application to nonconvex feasibility problems, *Mathematical Programming (Series A)* 159 (2016), 371–401.

[46] J. Lieutaud, *Approximations d'opérateurs monotones par des méthodes de splitting*, doctoral thesis, University of Paris, 1969.

[47] P.L. Lions and B. Mercier, Splitting algorithms for the sum of two nonlinear operators, *SIAM Journal on Numerical Analysis* 16 (1979), 964–979.

[48] B. Mercier, *Inéquations Variationnelles de la Mécanique* (Publications Mathématiques d'Orsay, no. 80.01), Orsay, France: Université de Paris-XI, 1980. `http://portail.mathdoc.fr/PMO/PDF/M_MERCIER-87.pdf`

[49] G.J. Minty, Monotone (nonlinear) operators in Hilbert spaces, *Duke Mathematical Journal* 29 (1962), 341–346.

[50] A. Pazy, Asymptotic behavior of contractions in Hilbert space, *Israel Journal of Mathematics* 9 (1971), 235–240.

[51] R.T. Rockafellar and R. J-B Wets, *Variational Analysis*, Springer-Verlag, corrected 3rd printing, 2009.

[52] S. Simons, *Minimax and Monotonicity*, Springer-Verlag, 1998.

[53] S. Simons, *From Hahn-Banach to Monotonicity*, Springer-Verlag, 2008.

[54] B.F. Svaiter, On weak convergence of the Douglas–Rachford method, *SIAM Journal on Control and Optimization* 49 (2011), 280–287.

[55] E.H. Zarantonello, Projections on convex sets in Hilbert space and spectral theory, in: E.H. Zarantonello (Ed.), *Contributions to Nonlinear Functional Analysis*, Academic Press, New York, (1971), 237–424.

[56] E. Zeidler, *Nonlinear Functional Analysis and Its Applications II/A: Linear Monotone Operators*, Springer-Verlag, 1990.

[57] E. Zeidler, *Nonlinear Functional Analysis and Its Applications II/B: Nonlinear Monotone Operators*, Springer-Verlag, 1990.