SIAM J. OPTIM.
Vol. 30, No. 3, pp. 2559–2576

ON THE BEHAVIOR OF THE DOUGLAS–RACHFORD ALGORITHM FOR MINIMIZING A CONVEX FUNCTION SUBJECT TO A LINEAR CONSTRAINT*

HEINZ H. BAUSCHKE† AND WALAA M. MOURSI‡

Abstract. The Douglas–Rachford algorithm (DRA) is a powerful optimization method for minimizing the sum of two convex (not necessarily smooth) functions. The vast majority of previous research dealt with the case when the sum has at least one minimizer. In the absence of minimizers, it was recently shown that for the case of two indicator functions, the DRA converges to a best approximation solution. In this paper, we present a new convergence result on the DRA applied to the problem of minimizing a convex function subject to a linear constraint. Indeed, a normal solution may be found even when the domain of the objective function and the linear subspace constraint have no point in common. As an important application, a new parallel splitting result is provided. We also illustrate our results through various examples.

Key words. convex optimization problem, Douglas–Rachford splitting, inconsistent constrained optimization, least squares solution, normal problem, parallel splitting method, projection operator, proximal mapping

AMS subject classifications. Primary, 49M27, 65K10, 90C25; Secondary, 47H14, 49M29

DOI. 10.1137/19M1281538

1. Introduction.

Throughout, we assume that

$$X \text{ is a real Hilbert space} \tag{1}$$

with inner product $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{R}$ and induced norm $\|\cdot\|$. We further assume that

$$U \text{ is a closed linear subspace of } X \tag{2}$$

and that

$$g\colon X \to \left]-\infty, +\infty\right] \text{ is convex, lower semicontinuous, and proper.} \tag{3}$$

Our aim is to discuss the behavior of the Douglas–Rachford algorithm [15] applied to solving the optimization problem[1]

$$\underset{x \in X}{\text{minimize}} \ \ \iota_U(x) + g(x), \tag{4}$$

*Received by the editors August 15, 2019; accepted for publication (in revised form) July 9, 2020; published electronically September 23, 2020.
https://doi.org/10.1137/19M1281538
Funding: The research of the first author was partially supported by a Discovery Grant of the Natural Sciences and Engineering Research Council of Canada. The research of the second author was partially supported by the Natural Sciences and Engineering Research Council of Canada Postdoctoral Fellowship.

†Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada (heinz.bauschke@ubc.ca).
‡Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada (walaa.moursi@uwaterloo.ca).

[1] Let us point out that if $\widetilde{U} = \widetilde{u} + U$ is an *affine* subspace and $\widetilde{g}$ is convex, lower semicontinuous, and proper, then all our results are applicable by working with U and $g = \widetilde{g}(\cdot - \widetilde{u})$ instead.

where $\iota_U(x) = 0$ if $x \in U$ and $\iota_U(x) = +\infty$ if $x \notin U$. Note that we do *not* assume a priori that (4) has a solution. Given any starting point $x_0 \in X$, the Douglas–Rachford algorithm generates the so-called *governing sequence*

$$(T^n x_0)_{n\in\mathbb{N}}, \tag{5}$$

where

$$T = \operatorname{Id} - \mathrm{P}_U + \mathrm{P}_g \mathrm{R}_U \tag{6}$$

is the Douglas–Rachford operator, P_U is the projector of U, P_g is the *proximal mapping* of the function g, and $\mathrm{R}_U = 2\mathrm{P}_U - \operatorname{Id} = \mathrm{P}_U - \mathrm{P}_{U^\perp}$ is the reflector of U. The basic convergence result (see [20], [16], and [25]), guarantees that the *shadow sequence*

$$(\mathrm{P}_U T^n x_0)_{n\in\mathbb{N}} \tag{7}$$

converges weakly to a solution of (4) provided that $(\mathrm{N}_U + \partial g)^{-1}(0) \neq \varnothing$.

To deal with the potential lack of solutions of (4), we define the *minimal displacement vector*

$$v = \mathrm{P}_{\overline{\operatorname{ran}}\,(\operatorname{Id} - T)}(0). \tag{8}$$

This vector is well defined because $\overline{\operatorname{ran}}\,(\operatorname{Id} - T)$ is convex, closed, and trivially nonempty. We now assume that the so-called *normal problem* corresponding to (4), which asks to find a zero of the operator $-v + \mathrm{N}_U + \partial g(\cdot - v)$, admits at least one *normal solution*[2] (see [8, Definition 3.7]):

$$Z = \big\{x \in X \;\big|\; v \in \mathrm{N}_U(x) + \partial g(x - v)\big\} \neq \varnothing. \tag{9}$$

We also assume throughout that

$$\mathrm{P}_Z \text{ is weak-to-weak continuous,} \tag{10}$$

which is automatically the case when X is finite dimensional, and that

$$0 \in U^\perp + \operatorname{dom} g^*, \tag{11}$$

which is a rather mild constraint qualification that is satisfied, for instance, if g has minimizers.[3] Note that if (4) has a solution and $\partial(\iota_U + g) = \mathrm{N}_U + \partial g$ (this sum formula is typically guaranteed through a regularity condition), then $v = 0$ and $Z = \operatorname{argmin}(\iota_U + g)$. Our main result (see Theorem 5.1 below) can now be concisely stated as follows: *Under the above assumptions, which we assume for the rest of the paper, we have*

$$\mathrm{P}_U T^n x_0 \rightharpoonup \textit{some minimizer of } \iota_U + g(\cdot - v). \tag{12}$$

This is a completely new (and very beautiful) variant of the classical result which is proven with a careful function value analysis in section 4! *It reveals the Douglas–Rachford algorithm to be a method for solving the following bilevel optimization problem:* First, obtain the gap vector between $U = \operatorname{dom} \iota_U$ and $\operatorname{dom} g$. This level is

[2]Note that it is possible that Z is empty: Indeed, consider the case when $X = \mathbb{R} = U$ and $g = \exp$. In this case, $|T^n x| \to +\infty$ for every $x \in \mathbb{R}$.

[3]Also note that (11) implies that the Fenchel dual of (4) is feasible and hence that (4) is implicitly assumed to be bounded below.

purely geometrical, depending on the sets U and $\operatorname{dom} g$, and revealing the minimal displacement vector v. Second, if $v \neq 0$, rather than minimizing the original $\iota_U + g$ which would have the optimal value $+\infty$, we then instead minimize the minimal perturbation function $\iota_U + g(\cdot - v)$. This has consequences for minimizing the "sum of convex functions" by using a product space technique; in fact, real world applications inspired this research (see the last section).

Let us now comment on related previous works which will illustrate the complementary nature of the present work. To the best of our knowledge, none of these works contains the result (12) in the generality of the setting of Theorem 5.1. The paper [1] by Banjac et al. applies the Douglas–Rachford algorithm with the function f being the sum of a quadratic function and the indicator function of an affine subspace rather than ι_U and with g being the indicator function of a nonempty closed convex set. The Douglas–Rachford method (equivalent to alternating direction method of multipliers (ADMM) in this setting) is shown to be useful in providing certificates of infeasibility. The paper [7] concerns the more restrictive case when g is the indicator function of a nonempty closed convex set; however, the underlying assumptions there do not require (10). The paper [8] introduces the normal problem but it does not contain any algorithmic/dynamic results. Similarly to [7], the paper [11] deals with the case when g is assumed to be an indicator function of a closed affine subspace. Under suitable assumptions, the shadow sequence $(\mathrm{P}_U T^n x_0)_{n\in\mathbb{N}}$ is shown to converge strongly. The paper [12] considers an infinite-dimensional setting that encompasses two indicator functions; however, our present main result is not covered by these results (see Remark 5.4 below). In the paper [21] by Liu, Ryu, and Yin, the authors study the behavior of the Douglas–Rachford algorithm applied to conic programming, where g is the indicator function of a nonempty closed convex cone while ι_U is replaced by the sum of a linear function and the indicator function of an affine subspace. The Douglas–Rachford method is shown to reveal information on the type of pathologies the conic program may exhibit. Finally, the paper [24] by Ryu, Liu, and Yin is the first to provide a comprehensive function-value analysis in pathological cases. It differs from the present work in that Ryu, Liu, and Yin allow for a general function f rather than the indicator function ι_U considered here. However, our main result Theorem 5.1 gives information on the iterates and the function values that are not covered by the results in [24] when strong duality fails.

The remainder of this paper is organized as follows. In section 2 we review known facts and present new auxiliary results that are needed in the main analysis. Section 3 presents new descriptions of the minimal displacement vector and the set of minimizers which are crucial in the convergence proofs. The building blocks of our analysis and the main result are presented in sections 4 and 5, respectively. In the final section 6, we provide a useful application of our theory to describe the behavior of a parallel splitting method.

We employ standard notation from convex analysis and optimization as can be found, e.g., in [5] and [23].

2. Known and new auxiliary results.

Because $Z \neq \varnothing$ (see (9)), the generalized fixed point set introduced in [8] is very well behaved in the sense that

$$F := \operatorname{Fix} T(\cdot + v) = \big\{x \in X \;\big|\; x = T(x+v)\big\} \text{ is convex, closed, and nonempty.} \tag{13}$$

The Douglas–Rachford operator T defined in (6) enjoys the following nice properties which also underline the importance of F for understanding the Douglas–Rachford algorithm.

FACT 2.1. *Let $x \in X$ and $y \in F$. Then*[4]

$$(\forall n \in \mathbb{N}) \quad T^n y = y - nv; \tag{14}$$

the sequence $(nv + T^n x)_{n\in\mathbb{N}}$ *is* Fejér monotone *with respect to F, i.e.,*

$$(\forall n \in \mathbb{N}) \quad \|(n+1)v + T^{n+1}x - y\| \le \|nv + T^n x - y\|; \tag{15}$$

$$\sum_{n=0}^{+\infty} \|T^{n+1}x - T^n x - v\|^2 < +\infty, \tag{16}$$

$$T^n x - T^{n+1}x \to v; \tag{17}$$

and the limit

$$\lim_{n\to+\infty} \mathrm{P}_F(nv + T^n x) \in F \tag{18}$$

exists.

Proof. See [12, Corollary 4.2], [11, Proposition 2.5(vi)], and [5, Proposition 5.7] for the proof. ∎

Before we proceed, we recall the following useful fact that will be used in the proofs of Propositions 2.3 and 3.1.

FACT 2.2. *Let C be a nonempty closed convex subset of X. Set $w = \mathrm{P}_{\overline{U-C}}(0)$ and let $x \in X$. Then $w = \lim_{n\to\infty}(\mathrm{P}_U - \mathrm{Id})(\mathrm{P}_C\mathrm{P}_U)^n x \in \overline{\operatorname{ran}}\,(\mathrm{P}_U - \mathrm{Id}) = -U^\perp = U^\perp$.*

Proof. See [2, Corollary 4.6] for the proof. ∎

The next result will also be used in the proof of Proposition 3.1.

PROPOSITION 2.3. *Let C_1 and C_2 be nonempty closed convex subsets of X, and set $S_1 := U - C_1$ and $S_2 := U^\perp - C_2$. Define*

$$v_D := \mathrm{P}_{\overline{S_1}}(0), \quad v_R := \mathrm{P}_{\overline{S_2}}(0), \quad v := \mathrm{P}_{\overline{S_1}\cap\overline{S_2}}(0). \tag{19}$$

Then the following hold:

(i) $(v_D, v_R) \in U^\perp \times U$.
(ii) $\mathrm{P}_{U^\perp}(\overline{S_1}) \subseteq \overline{S_1}$.
(iii) $\mathrm{P}_U(\overline{S_2}) \subseteq \overline{S_2}$.
(iv) $v_D + v_R \in \overline{S_1} \cap \overline{S_2}$.
(v) $v = v_D + v_R$.

Proof. (i): Apply Fact 2.2 with (C, w) replaced by (C_1, v_D) (respectively, (C, w) replaced by (C_2, v_R)). (ii): Let $y \in \overline{S_1}$. Then there exist $(u_n)_{n\in\mathbb{N}}$ in U and $(c_{1,n})_{n\in\mathbb{N}}$ is C_1 such that $u_n - c_{1,n} \to y$. Now, $\mathrm{P}_{U^\perp} y \leftarrow \mathrm{P}_{U^\perp}(u_n - c_{1,n}) = -\mathrm{P}_{U^\perp} c_{1,n} = \mathrm{P}_U c_{1,n} - c_{1,n} \in U - C_1$. Hence, $\mathrm{P}_{U^\perp} y \in \overline{U - C_1} = \overline{S_1}$ and the claim follows. (iii): Proceed similarly to the proof of (ii). (iv): Indeed, note that by (i) we have $v_R \in U$, hence $v_D + v_R \in \overline{S_1} + v_R = \overline{U - C_1} + v_R = \overline{U - C_1 + v_R} = \overline{U - C_1} = \overline{S_1}$. Similarly, we show that $v_D + v_R \in \overline{S_2}$ and the conclusion follows. (v): Note that (ii) and (iii) imply that $(\mathrm{P}_U v, \mathrm{P}_{U^\perp} v) \in \overline{S_2} \times \overline{S_1}$. Consequently, $\|v_R\| \le \|\mathrm{P}_U v\|$ and $\|v_D\| \le \|\mathrm{P}_{U^\perp} v\|$. Altogether, in view of (i), we learn that $\|v_D + v_R\|^2 = \|v_D\|^2 + \|v_R\|^2 \le \|\mathrm{P}_U v\|^2 + \|\mathrm{P}_{U^\perp} v\|^2 = \|v\|^2$. Combining this with (iv), and the definition of v, we obtain the result. ∎

[4]We point out that Fact 2.1 holds in the more general setting when T is any firmly nonexpansive mapping.

The following simple result, which relies on the assumption that U is a closed linear subspace, will be used in the proof of Theorem 5.1.

LEMMA 2.4. *Let C be a nonempty closed convex subset of U. Then*

$$\mathrm{P}_C = \mathrm{P}_C \circ \mathrm{P}_U. \tag{20}$$

Proof. Let $x \in X$ and let $c \in C \subseteq U$. Then $\mathrm{P}_C\mathrm{P}_U x \in C$ and

$$\langle c - \mathrm{P}_C\mathrm{P}_U x, x - \mathrm{P}_C\mathrm{P}_U x\rangle = \big\langle \underbrace{c - \mathrm{P}_C\mathrm{P}_U x}_{\in U}, \underbrace{x - \mathrm{P}_U x}_{\in U^\perp}\big\rangle + \langle c - \mathrm{P}_C\mathrm{P}_U x, \mathrm{P}_U x - \mathrm{P}_C\mathrm{P}_U x\rangle \tag{21a}$$

$$= \langle c - \mathrm{P}_C\mathrm{P}_U x, \mathrm{P}_U x - \mathrm{P}_C\mathrm{P}_U x\rangle \tag{21b}$$

$$\leq 0, \tag{21c}$$

and we are done. ☐

We now turn to the minimization of a convex function subject to a linear constraint. The following result will be used in the proof of Theorem 3.4.

LEMMA 2.5. *Let $h\colon X \to \left]-\infty, +\infty\right]$ be a proper lower semicontinuous convex function. Furthermore, let x and y be points in U, and let $x^* \in X$. Then the following hold:*

(i) *If $U^\perp \cap \partial h(x) \neq \varnothing$, then x is a minimizer of $\iota_U + h$.*

(ii) *If $x^* \in U^\perp \cap \partial h(x)$ and y is a minimizer of $\iota_U + h$, then $x^* \in U^\perp \cap \partial h(y)$.*

Proof. (i): Suppose that $U^\perp \cap \partial h(x) \neq \varnothing$. Then, since $U^\perp$ is a subspace, $(-U^\perp) \cap \partial h(x) \neq \varnothing$. Suppose that $x^* \in \partial h(x)$. Then $-x^* \in U^\perp = \mathrm{N}_U(x)$. It follows that $0 = (-x^*) + x^* \in \mathrm{N}_U(x) + \partial h(x) = \partial \iota_U(x) + \partial h(x) \subseteq \partial(\iota_U + h)(x)$. By Fermat's rule, x is a minimizer of $\iota_U + h$.

(ii): Suppose that $x^* \in U^\perp \cap \partial h(x) \neq \varnothing$. Then

$$(\forall z \in X) \quad h(z) \geq h(x) + \langle z - x, x^*\rangle \tag{22}$$

and

$$\langle y - x, x^*\rangle = 0. \tag{23}$$

On the other hand, because y is a minimizer of $\iota_U + h$, we learn from (i) that

$$h(x) = h(y). \tag{24}$$

Altogether,

$$(\forall z \in X) \quad h(z) \geq h(x) + \langle z - x, x^*\rangle \tag{25a}$$

$$= h(y) + \langle z - y, x^*\rangle + \langle y - x, x^*\rangle \tag{25b}$$

$$= h(y) + \langle z - y, x^*\rangle\,. \tag{25c}$$

Therefore, $x^* \in \partial h(y)$. ☐

The assumption that $U^\perp \cap \partial h(x) \neq \varnothing$ in Lemma 2.5(ii) is critical.

Example 2.6. Suppose that $X = \mathbb{R}$, that $U = \{0\}$, and that $h(\xi) = -\sqrt{\xi}$, if $\xi \geq 0$ and $h(\xi) = +\infty$ if $\xi < 0$. Then 0 minimizes $\iota_U + h = \iota_U$ yet $U^\perp \cap \partial h(0) = \partial h(0) = \varnothing$.

Remark 2.7. Let $h\colon X \to \left]-\infty, +\infty\right]$ be a proper lower semicontinuous convex function. Then Lemma 2.5 implies that the set-valued operator

$$\operatorname{argmin}(\iota_U + h) \rightrightarrows X\colon x \mapsto U^\perp \cap \partial h(x) \tag{26}$$

is constant.

3. New static results. We start with the following useful result for the minimal displacement vector v from (8).

PROPOSITION 3.1. *Set* $w = \mathrm{P}_{\overline{U - \operatorname{dom} g}}(0)$. *Then the following hold:*

(i) $w \in U^\perp$.

(ii) *If X is finite dimensional, then* $v = w = \mathrm{P}_{\overline{U - \operatorname{dom} g}}(0) \in U^\perp$.

Proof. Clearly $\overline{U - \operatorname{dom} g} = \overline{U - \overline{\operatorname{dom} g}}$ and $\overline{U^\perp + \operatorname{dom} g^*} = \overline{U^\perp + \overline{\operatorname{dom} g^*}}$. (i): Apply Fact 2.2 with C replaced by $\overline{\operatorname{dom} g}$. (ii): Note that $\iota_U^* = \iota_{U^\perp}$ and thus $\operatorname{dom} \iota_U^* = U^\perp$. Hence (11) states exactly that $0 \in \operatorname{dom} \iota_U^* + \operatorname{dom} g^*$. It follows from [9, Proposition 6.1(ii) and Corollary 6.5(i)] that $v = \mathrm{P}_{\overline{(U - \operatorname{dom} g)} \cap \overline{(U^\perp + \operatorname{dom} g^*)}}(0)$. By Proposition 2.3 applied with (C_1, C_2) replaced by $(\overline{\operatorname{dom} g}, -\overline{\operatorname{dom} g^*})$ we have

$$v = \mathrm{P}_{\overline{U - \operatorname{dom} g}}(0). \tag{27}$$

Now combine with (i). ∎

The result in Proposition 3.1(ii) was first proved—in an even more general form—by Ryu, Liu, and Yin with a different argument relying on recession functions (see [24, Lemma 3]). From now on, we assume

$$v = \mathrm{P}_{\overline{U - \operatorname{dom} g}}(0). \tag{28}$$

Note that (28) holds if X is finite dimensional by Proposition 3.1(ii). In view of Proposition 3.1(i), we have

$$v \in U^\perp. \tag{29}$$

The fact that v belongs to $U^\perp$ is new and crucial to our analysis.

We now turn towards alternative descriptions of the set Z of normal solutions, defined in (9). In passing, we mention that the next result is true even if $Z = \varnothing$.

PROPOSITION 3.2. *We have*

$$Z = \big\{ x \in U \;\big|\; U^\perp \cap \partial g(x - v) \neq \varnothing \big\} \tag{30}$$

and

$$
\begin{aligned}
&U \cap (v + \operatorname{argmin} g) \subseteq \big\{ x \in U \;\big|\; U^\perp \cap \partial g(x - v) \neq \varnothing \big\} && \text{(31a)}\\
&\qquad = \operatorname{zer}\big(\mathrm{N}_U + \partial g(\cdot - v)\big) && \text{(31b)}\\
&\qquad \subseteq U \cap (v + \operatorname{dom} \partial g) \cap \operatorname{argmin}(\iota_U + g(\cdot - v)) && \text{(31c)}\\
&\qquad \subseteq \operatorname{argmin}(\iota_U + g(\cdot - v)) && \text{(31d)}\\
&\qquad \subseteq U \cap (v + \operatorname{dom} g). && \text{(31e)}
\end{aligned}
$$

Proof. Recall that $v \in U^\perp$ by (29). Hence $\mathrm{N}_U = -v + \mathrm{N}_U$. Now let $x \in X$. Then

$$
\begin{aligned}
&x \in U \cap (v + \operatorname{argmin} g) \Leftrightarrow \big[x \in U \text{ and } x - v \in \operatorname{argmin} g\big] && \text{(32a)}\\
&\qquad \Leftrightarrow \big[x \in \operatorname{zer} \mathrm{N}_U \text{ and } 0 \in \partial g(x - v)\big] && \text{(32b)}\\
&\qquad \Leftrightarrow \big[x \in \operatorname{zer}(-v + \mathrm{N}_U) \text{ and } 0 \in \partial g(x - v)\big] && \text{(32c)}\\
&\qquad \Leftrightarrow \big[0 \in -v + \mathrm{N}_U(x) \text{ and } 0 \in \partial g(x - v)\big] && \text{(32d)}\\
&\qquad \Rightarrow 0 \in -v + \mathrm{N}_U(x) + \partial g(x - v) && \text{(32e)}
\end{aligned}
$$

$$\Leftrightarrow x \in Z \tag{32f}$$
$$\Leftrightarrow v \in \mathrm{N}_U(x) + \partial g(x-v) \tag{32g}$$
$$\Leftrightarrow \big[x \in U \text{ and } v \in U^\perp + \partial g(x-v)\big] \tag{32h}$$
$$\Leftrightarrow \big[x \in U \text{ and } 0 \in U^\perp + \partial g(x-v)\big] \tag{32i}$$
$$\Leftrightarrow \big[x \in U \text{ and } U^\perp \cap \partial g(x-v) \neq \varnothing\big] \tag{32j}$$
$$\Leftrightarrow x \in \operatorname{zer}\big(\mathrm{N}_U + \partial g(\cdot - v)\big) \tag{32k}$$
$$\Leftrightarrow 0 \in \big(\mathrm{N}_U + \partial g(\cdot - v)(x), \tag{32l}$$

which proves (30), (31a), and (31b). Turning to (31c), let $x \in \operatorname{zer}(\mathrm{N}_U + \partial g(\cdot - v))$. On the one hand, $x \in \operatorname{dom}(\mathrm{N}_U + \partial g(\cdot - v))$ and thus $\mathrm{N}_U(x) \neq \varnothing$ and $\partial g(x-v) \neq \varnothing$. Hence $x \in U$ and $x - v \in \operatorname{dom} \partial g$, i.e., $x \in U \cap (v + \operatorname{dom} \partial g)$. On the other hand, $\operatorname{zer}(\mathrm{N}_U + \partial g(\cdot - v)) = \operatorname{zer}(\partial \iota_U + \partial g(\cdot - v))$. Hence $0 \in \partial \iota_U(x) + \partial g(\cdot - v)(x) \subseteq \partial(\iota_U + g(\cdot - v))(x)$ and therefore x minimizes $\iota_U + g(\cdot - v)$. Finally, (31d) and (31e) are obvious. ∎

Example 3.3 (linear-convex feasibility). Suppose that $g = \iota_W$, where W is a nonempty closed convex subset of X. Then $v = \mathrm{P}_{\overline{U-W}}(0)$, $\operatorname{argmin} g = \operatorname{dom} \partial g = W$, and $v + \operatorname{argmin} g = v + W = v + \operatorname{dom} g$. Thus Proposition 3.2 yields

$$Z = U \cap (v + V), \tag{33}$$

a result that is well known (see [6]).

We are now ready for our first main result which provides a useful description of Z.

THEOREM 3.4. *Because Z is nonempty, we have*

$$Z = U \cap (v + \operatorname{dom} \partial g) \cap \operatorname{argmin}\big(\iota_U + g(\cdot - v)\big) = \operatorname{argmin}\big(\iota_U + g(\cdot - v)\big). \tag{34}$$

Proof. Proposition 3.2 yields the inclusions $Z \subseteq U \cap (v + \operatorname{dom} \partial g) \cap (\iota_U + g(\cdot - v)) \subseteq \operatorname{argmin}(\iota_U + g(\cdot - v))$. Because $Z \neq \varnothing$, we let $x \in Z$, and also let $y \in \operatorname{argmin}(\iota_U + g(\cdot - v)) \subseteq U$. First, by (30), $x \in U$ and $U^\perp \cap \partial g(x - v) \neq \varnothing$. Second, it follows from Lemma 2.5 (applied with $h = g(\cdot - v)$) that $U^\perp \cap \partial g(y - v) \neq \varnothing$. Therefore, by again using (30), we obtain $y \in Z$. ∎

Here is an example of a case where $Z \neq \varnothing$.

Example 3.5. Suppose that g is polyhedral. Then [3, Theorem 5.6.1] implies that $U \cap (v + \operatorname{dom} g) = U \cap \operatorname{dom} g(\cdot - v) \neq \varnothing$. Hence, by [5, Corollary 27.3(c)] we have $Z = \operatorname{argmin}(\iota_U + g(\cdot - v))$.

The underlying assumption that Z be nonempty (see (9)) in Theorem 3.4 is critical.

Example 3.6. Suppose that $X = \mathbb{R}^2$, that $U = \{0\} \times \mathbb{R}$, and that g is the Rockafellar function defined by

$$g(\xi_1, \xi_2) = \begin{cases} \max\{1 - \sqrt{\xi_1}, |\xi_2|\} & \text{if } \xi_1 \geq 0, \\ +\infty & \text{otherwise} \end{cases} \tag{35}$$

(see [23, Example on p. 218]). Then $v = 0$ and it follows from [22, Example 7.5] that $Z = \varnothing$, $\operatorname{argmin}(\iota_U + g(\cdot - v)) = \{0\} \times [-1, 1]$, and $U \cap (v + \operatorname{dom} \partial g) \cap \operatorname{argmin}(\iota_U + g(\cdot - v)) = \{0\} \times \{-1, 1\}$.

Proof. Clearly we have $U^\perp = \mathbb{R}\times\{0\}$ and $\operatorname{dom} g = \mathbb{R}_+\times\mathbb{R}$. Moreover, [22, Example 6.5] implies that $\operatorname{dom}\partial g = \big\{(\xi_1,\xi_2)\;\big|\;\xi_1>0, \xi_2\in\mathbb{R}\big\}\cup\big\{(0,\xi_2)\;\big|\;\xi_2\geq 1\big\}$ and $\operatorname{dom}\partial g^* = \operatorname{dom} g^* = \big\{(\xi_1,\xi_2)\;\big|\;\xi_1\leq 0, |\xi_2|\leq 1\big\}$. Therefore, using [9, Corollary 6.5(i)] we learn that $v = \mathrm{P}_{(\overline{U-\operatorname{dom} g})\cap(\overline{U^\perp+\operatorname{dom} g^*})}(0)=0$. It follows from Proposition 3.2 that $Z = \big\{(0,\xi_2)\;\big|\;U^\perp\cap\partial g((0,\xi_2))\neq\varnothing\big\}$. Now let $(0,\xi_2)\in U\cap\operatorname{dom} g$ and note that [22, Example 6.5] implies that

$$\partial g(0,\xi_2) = \begin{cases}\varnothing & \text{if } |\xi_2|<1,\\ \mathbb{R}_-\times\{1\} & \text{if } |\xi_2|\geq 1,\\ \mathbb{R}_-\times\{-1\} & \text{if } |\xi_2|\leq -1,\end{cases} \tag{36}$$

which proves the claim that $Z=\varnothing$. Finally, using (35), we see that $\operatorname{argmin}(\iota_U + g(\cdot - v)) = \operatorname{argmin}(\iota_U+g) = \{0\}\times[-1,1]$ and the conclusion follows. ∎

When $X=\mathbb{R}$, then we obtain the following positive result, which holds even when $Z=\varnothing$.

PROPOSITION 3.7. *Suppose that $X=\mathbb{R}$. Then*

$$Z = U\cap(v+\operatorname{dom}\partial g)\cap\operatorname{argmin}\big(\iota_U + g(\cdot - v)\big). \tag{37}$$

More precisely, exactly one of the following cases holds:

(i) $U=\{0\}$, $v=\mathrm{P}_{\overline{-\operatorname{dom} g}}(0)$, $Z = 0\cdot\partial g(-v)$, *and either* $\iota_U + g(\cdot - v) = \iota_{\{0\}}$ *if* $-v\in\operatorname{dom} g$ *or* $\iota_U+g(\cdot-v)=\iota_\varnothing$ *if* $-v\notin\operatorname{dom} g$.

(ii) $U=\mathbb{R}$, $v=0$, *and* $Z = \operatorname{dom}\partial g\cap\operatorname{argmin} g = \operatorname{argmin} g$.

Proof. Denote the right side of (37) by R. It is clear from Proposition 3.2 that $Z\subseteq R$. Now let $x\in R$. On the one hand,

$$0\in\partial(\iota_U+g(\cdot-v))(x). \tag{38}$$

On the other hand, $x\in\operatorname{dom}\partial\iota_U\cap\operatorname{dom}\partial g(\cdot-v)$. By the sum rule for the real line, we have

$$\partial\iota_U(x)+\partial g(x-v) = \partial\big(\iota_U+g(\cdot-v)\big)(x). \tag{39}$$

Altogether, $0\in\partial\iota_U(x)+\partial g(x-v)$ and thus $x\in Z$ by Proposition 3.2. The remaining statements follow readily. ∎

The previous results make it tempting to conjecture that when $X=\mathbb{R}$ and $Z=\varnothing$, then we have $\operatorname{argmin}(\iota_U+g(\cdot-v))=\varnothing$. Unfortunately, this conjecture is false.

Example 3.8. Suppose that $X=\mathbb{R}$, that $U=\{0\}$, and that $g(x)z-\sqrt{x}$ with $\operatorname{dom} g = \mathbb{R}_+$. Then $v=\mathrm{P}_{\overline{-\operatorname{dom} g}}(0)=0$. Hence $Z=\{0\}\cdot\partial g(0)=\varnothing$ by Proposition 3.7 while $\operatorname{argmin}(\iota_U+g(\cdot-v))=\{0\}$ because $\iota_U+g(\cdot-v)=\iota_U+g=\iota_U=\iota_{\{0\}}$.

We conclude this section with another useful consequence of (29).

PROPOSITION 3.9. *We have $Z=\mathrm{P}_U(F)$ and*

$$\mathrm{P}_U\circ\mathrm{P}_F = \mathrm{P}_Z. \tag{40}$$

Proof. Set $A=-v+\mathrm{N}_U$ and $B=\partial g(\cdot-v)$, and note that by (29) $A=\mathrm{N}_U$. Then the Douglas–Rachford operator corresponding to (A,B) is [8, Proposition 3.2]

$$T(\cdot+v). \tag{41}$$

Moreover $\mathrm{J}_A := (\mathrm{Id}+A)^{-1} = \mathrm{P}_U$. Note that A and B are subdifferential operators, hence paramonotone by [17, Theorem 2.2]. So [4, Corollary 5.6] yields $F = Z + K$, $Z = \mathrm{J}_A(F) = \mathrm{P}_U(F)$, where $K := (\mathrm{Id} - \mathrm{J}_{A^{-1}})(F) = \mathrm{P}_{U^\perp}(F) \subseteq U^\perp$. Moreover, because $Z - Z \subseteq U$ and so $Z - Z \perp K$, we have $\mathrm{J}_A \mathrm{P}_{Z+K} = \mathrm{P}_Z$, equivalently, $\mathrm{P}_U \mathrm{P}_F = \mathrm{P}_Z$, by [4, Theorem 6.7(ii)]. ∎

4. New dynamic results. Recall that

$$T = \mathrm{Id} - \mathrm{P}_U + \mathrm{P}_g \mathrm{R}_U. \tag{42}$$

We start with a result that provides some information on the shadow sequence $(\mathrm{P}_U T^n x)_{n\in\mathbb{N}}$. (In passing, we note that only item (v) requires that Z be nonempty.)

LEMMA 4.1. *Let $x \in X$. Then the following hold:*

(i) $\mathrm{P}_U T^n x - \mathrm{P}_g \mathrm{R}_U T^n x = T^n x - T^{n+1} x \to v \in U^\perp$.

(ii) $\mathrm{P}_U T^n x - \mathrm{P}_U \mathrm{P}_g \mathrm{R}_U T^n x = \mathrm{P}_U T^n x - \mathrm{P}_U T^{n+1} x \to 0$.

(iii) $-\mathrm{P}_{U^\perp} \mathrm{P}_g \mathrm{R}_U T^n x = \mathrm{P}_{U^\perp} T^n x - \mathrm{P}_{U^\perp} T^{n+1} x \to v$.

(iv) *All weak cluster points of $(\mathrm{P}_U T^n x)_{n\in\mathbb{N}}$ lie in $U \cap (v + \overline{\mathrm{dom}}\, g)$.*

(v) *The sequences $(nv + T^n x)_{n\in\mathbb{N}}$, $(\mathrm{P}_U T^n x)_{n\in\mathbb{N}}$, and $(\mathrm{P}_g \mathrm{R}_U T^n x)_{n\in\mathbb{N}}$ are bounded.*

Proof. (i): This is clear from the definition of T, (17), and (29). (ii): Apply P_U to (i). (iii): Apply $\mathrm{P}_{U^\perp}$ to (i). (iv): On the one hand, $(T^n x - T^{n+1} x) + \mathrm{P}_g \mathrm{R}_U T^n x = \mathrm{P}_U T^n x \in U$. On the other hand, $\mathrm{P}_g \mathrm{R}_U T^n x \in \mathrm{dom}\, \partial g \subseteq \overline{\mathrm{dom}}\, g$. Altogether, combined with (i), we obtained the desired result. (v): By Fact 2.1 and (13), the sequence $(nv + T^n x)_{n\in\mathbb{N}}$ is Fejér monotone with respect to $F \neq \varnothing$, hence it is bounded. Therefore, $(\mathrm{P}_U T^n x)_{n\in\mathbb{N}} = (\mathrm{P}_U(nv + T^n x))_{n\in\mathbb{N}}$ is also bounded. The boundedness of $(\mathrm{P}_g \mathrm{R}_U T^n x)_{n\in\mathbb{N}}$ follows from (i). ∎

Note that Proposition 3.2 yields that $Z - v \subseteq (U - v) \cap \mathrm{dom}\, g$, and thus $(U - v) \cap \mathrm{dom}\, g$ is nonempty. The next result provides information on function values of g of a sequence occurring in the Douglas–Rachford algorithm.

LEMMA 4.2. *Let $x \in X$, let $y \in (U - v) \cap \mathrm{dom}\, g$, and let $n \in \mathbb{N}$. Then*

$$\begin{aligned}
g(y) \geq{}& g\big(\mathrm{P}_g(\mathrm{R}_U T^n x)\big) && \text{(43a)}\\
&+ \langle y - \mathrm{P}_g(\mathrm{R}_U T^n x), (\mathrm{P}_U T^n x - v) - \mathrm{P}_g(\mathrm{R}_U T^n x)\rangle && \text{(43b)}\\
&- \big\langle \mathrm{P}_{U^\perp} T^n x - \mathrm{P}_{U^\perp} T^{n+1} x - v, \mathrm{P}_{U^\perp}(nv + T^n x)\big\rangle && \text{(43c)}\\
&- (n+1)\langle (\mathrm{Id} - T) T^n x - v, 0 - v\rangle && \text{(43d)}\\
\geq{}& g\big(\mathrm{P}_g(\mathrm{R}_U T^n x)\big) && \text{(43e)}\\
&+ \langle y - \mathrm{P}_g(\mathrm{R}_U T^n x), (\mathrm{P}_U T^n x - v) - \mathrm{P}_g(\mathrm{R}_U T^n x)\rangle && \text{(43f)}\\
&- \big\langle \mathrm{P}_{U^\perp} T^n x - \mathrm{P}_{U^\perp} T^{n+1} x - v, \mathrm{P}_{U^\perp}(nv + T^n x)\big\rangle. && \text{(43g)}
\end{aligned}$$

Proof. The characterization of the prox operator P_g gives

$$g(y) \geq g\big(\mathrm{P}_g(\mathrm{R}_U T^n x)\big) + \langle y - \mathrm{P}_g(\mathrm{R}_U T^n x), \mathrm{R}_U T^n x - \mathrm{P}_g(\mathrm{R}_U T^n x)\rangle. \tag{44}$$

We also have

$$\begin{aligned}
&\langle y - \mathrm{P}_g(\mathrm{R}_U T^n x), \mathrm{R}_U T^n x - \mathrm{P}_g(\mathrm{R}_U T^n x)\rangle && \text{(45a)}\\
&\quad = \langle y - \mathrm{P}_g(\mathrm{R}_U T^n x), \mathrm{R}_U T^n x - (\mathrm{P}_U T^n x - v)\rangle && \text{(45b)}\\
&\qquad + \langle y - \mathrm{P}_g(\mathrm{R}_U T^n x), (\mathrm{P}_U T^n x - v) - \mathrm{P}_g(\mathrm{R}_U T^n x)\rangle && \text{(45c)}\\
&\quad = \langle y - \mathrm{P}_g(\mathrm{R}_U T^n x), -\mathrm{P}_{U^\perp} T^n x + v\rangle && \text{(45d)}\\
&\qquad + \langle y - \mathrm{P}_g(\mathrm{R}_U T^n x), (\mathrm{P}_U T^n x - v) - \mathrm{P}_g(\mathrm{R}_U T^n x)\rangle. && \text{(45e)}
\end{aligned}$$

Now write $y = u - v$, where $u \in U$. Then, using also the identity in Lemma 4.1(iii) to derive (46e), we have

$$
\begin{aligned}
&\text{(46a)} && \langle y - \mathrm{P}_g(\mathrm{R}_U T^n x), -\mathrm{P}_{U^\perp} T^n x + v\rangle \\
&\text{(46b)} && = \langle (u - v) - \mathrm{P}_g(\mathrm{R}_U T^n x), -\mathrm{P}_{U^\perp} T^n x + v\rangle \\
&\text{(46c)} && = \Big\langle \underbrace{(u - \mathrm{P}_U \mathrm{P}_g(\mathrm{R}_U T^n x)}_{\in U} - \underbrace{(v + \mathrm{P}_{U^\perp}\mathrm{P}_g(\mathrm{R}_U T^n x))}_{\in U^\perp}, \underbrace{-\mathrm{P}_{U^\perp} T^n x + v}_{\in U^\perp}\Big\rangle \\
&\text{(46d)} && = \big\langle -v - \mathrm{P}_{U^\perp}\mathrm{P}_g(\mathrm{R}_U T^n x), -\mathrm{P}_{U^\perp} T^n x + v\big\rangle \\
&\text{(46e)} && = \big\langle -v + \mathrm{P}_{U^\perp} T^n x - \mathrm{P}_{U^\perp} T^{n+1} x, -\mathrm{P}_{U^\perp} T^n x + v\big\rangle \\
&\text{(46f)} && = -\big\langle \mathrm{P}_{U^\perp} T^n x - \mathrm{P}_{U^\perp} T^{n+1} x - v, \mathrm{P}_{U^\perp} T^n x - v\big\rangle \\
&\text{(46g)} && = -\big\langle \mathrm{P}_{U^\perp} T^n x - \mathrm{P}_{U^\perp} T^{n+1} x - v, \mathrm{P}_{U^\perp}(nv + T^n x) - (n+1)v\big\rangle \\
&\text{(46h)} && = -\big\langle \mathrm{P}_{U^\perp} T^n x - \mathrm{P}_{U^\perp} T^{n+1} x - v, \mathrm{P}_{U^\perp}(nv + T^n x)\big\rangle \\
&\text{(46i)} && \quad - (n+1)\big\langle \mathrm{P}_{U^\perp} T^n x - \mathrm{P}_{U^\perp} T^{n+1} x - v, -v\big\rangle \\
&\text{(46j)} && = -\big\langle \mathrm{P}_{U^\perp} T^n x - \mathrm{P}_{U^\perp} T^{n+1} x - v, \mathrm{P}_{U^\perp}(nv + T^n x)\big\rangle \\
&\text{(46k)} && \quad - (n+1)\,\big\langle T^n x - T^{n+1} x - v, -v\big\rangle \\
&\text{(46l)} && = -\big\langle \mathrm{P}_{U^\perp} T^n x - \mathrm{P}_{U^\perp} T^{n+1} x - v, \mathrm{P}_{U^\perp}(nv + T^n x)\big\rangle \\
&\text{(46m)} && \quad - (n+1)\,\underbrace{\langle (\mathrm{Id} - T) T^n x - v, 0 - v\rangle}_{\le 0 \text{ by } (8)} \\
&\text{(46n)} && \ge -\big\langle \mathrm{P}_{U^\perp} T^n x - \mathrm{P}_{U^\perp} T^{n+1} x - v, \mathrm{P}_{U^\perp}(nv + T^n x)\big\rangle.
\end{aligned}
$$

Therefore, substituting (45) and (46) into (44), we obtain

$$
\begin{aligned}
&\text{(47a)} && g(y) \ge g\big(\mathrm{P}_g(\mathrm{R}_U T^n x)\big) \\
&\text{(47b)} && \qquad + \langle y - \mathrm{P}_g(\mathrm{R}_U T^n x), (\mathrm{P}_U T^n x - v) - \mathrm{P}_g(\mathrm{R}_U T^n x)\rangle \\
&\text{(47c)} && \qquad - \big\langle \mathrm{P}_{U^\perp} T^n x - \mathrm{P}_{U^\perp} T^{n+1} x - v, \mathrm{P}_{U^\perp}(nv + T^n x)\big\rangle \\
&\text{(47d)} && \qquad - (n+1)\,\underbrace{\langle (\mathrm{Id} - T) T^n x - v, 0 - v\rangle}_{\le 0} \\
&\text{(47e)} && \quad \ge g\big(\mathrm{P}_g(\mathrm{R}_U T^n x)\big) \\
&\text{(47f)} && \qquad + \langle y - \mathrm{P}_g(\mathrm{R}_U T^n x), (\mathrm{P}_U T^n x - v) - \mathrm{P}_g(\mathrm{R}_U T^n x)\rangle \\
&\text{(47g)} && \qquad - \big\langle \mathrm{P}_{U^\perp} T^n x - \mathrm{P}_{U^\perp} T^{n+1} x - v, \mathrm{P}_{U^\perp}(nv + T^n x)\big\rangle,
\end{aligned}
$$

which completes the proof. ∎

We are now able to locate weak cluster points of the shadow sequence $(\mathrm{P}_U T^n x)_{n\in\mathbb{N}}$.

LEMMA 4.3. *Let $x \in X$ and let $y \in (U - v) \cap \operatorname{dom} g$. Then there exists a sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ such that*

$$\varepsilon_n \to 0 \tag{48}$$

and for every $n \in \mathbb{N}$, we have

$$
\begin{aligned}
&\text{(49a)} && g(y) \ge g\big(\mathrm{P}_g(\mathrm{R}_U T^n x)\big) + \varepsilon_n + (n+1)\big\langle T^n x - T^{n+1} x - v, v\big\rangle \\
&\text{(49b)} && \quad \ge g\big(\mathrm{P}_g(\mathrm{R}_U T^n x)\big) + \varepsilon_n.
\end{aligned}
$$

Moreover, the sequence

$$(\mathrm{P}_g(\mathrm{R}_U T^n x))_{n\in\mathbb{N}} \text{ is bounded, all its weak cluster points are minimizers of } \iota_{U-v} + g, \tag{50}$$

$$g\big(\mathrm{P}_g(\mathrm{R}_U T^n x)\big) \to \inf g(U - v), \tag{51}$$

and

$$(n+1)\big\langle T^n x - T^{n+1}x - v, v\big\rangle \to 0. \tag{52}$$

Finally, the sequence

$$(\mathrm{P}_U T^n x)_{n\in\mathbb{N}} \text{ is bounded and all its weak cluster points are minimizers of } \iota_U + g(\cdot - v). \tag{53}$$

Proof. Lemmas 4.1(v) and (i) yield that $(y - \mathrm{P}_g\mathrm{R}_U T^n x)_{n\in\mathbb{N}}$ is bounded and that $\mathrm{P}_U T^n x - v - \mathrm{P}_g \mathrm{R}_U T^n x \to 0$. Thus

$$\langle y - \mathrm{P}_g\mathrm{R}_U T^n x, (\mathrm{P}_U T^n x - v) - \mathrm{P}_g(\mathrm{R}_U T^n x)\rangle \to 0. \tag{54}$$

Lemmas 4.1(iii) and (i) yield that $\mathrm{P}_{U^\perp} T^n x - \mathrm{P}_{U^\perp} T^{n+1} x - v \to 0$ and that $(\mathrm{P}_{U^\perp}(nv + T^n x))_{n\in\mathbb{N}}$ is bounded. Hence

$$-\big\langle \mathrm{P}_{U^\perp} T^n x - \mathrm{P}_{U^\perp} T^{n+1} x - v, \mathrm{P}_{U^\perp}(nv + T^n x)\big\rangle \to 0. \tag{55}$$

Setting

$$\varepsilon_n = \langle y - \mathrm{P}_g\mathrm{R}_U T^n x), (\mathrm{P}_U T^n x - v) - \mathrm{P}_g(\mathrm{R}_U T^n x)\rangle \tag{56}$$

$$\qquad - \big\langle \mathrm{P}_{U^\perp} T^n x - \mathrm{P}_{U^\perp} T^{n+1} x - v, \mathrm{P}_{U^\perp}(nv + T^n x)\big\rangle, \tag{57}$$

we see that (49) is a consequence of Lemma 4.2, (54), and (55).

By Lemma 4.1(v), $(\mathrm{P}_g\mathrm{R}_U T^n x)_{n\in\mathbb{N}}$ is bounded. Let c be a weak cluster point of $(\mathrm{P}_g\mathrm{R}_U T^n x)_{n\in\mathbb{N}}$, say $\mathrm{P}_g\mathrm{R}_U T^{k_n} x \rightharpoonup c$. Lemma 4.1(i) implies that

$$\mathrm{P}_g\mathrm{R}_U T^{k_n} x \rightharpoonup c \in U - v. \tag{58}$$

Now abbreviate $\alpha_n = (n+1)\big\langle T^n x - T^{n+1} x - v, v\big\rangle$. Then (49) yields

$$g(y) \geq g\big(\mathrm{P}_g(\mathrm{R}_U T^n x)\big) + \varepsilon_n + \alpha_n \geq g\big(\mathrm{P}_g(\mathrm{R}_U T^n x)\big) + \varepsilon_n. \tag{59}$$

The weak lower semicontinuity of g now yields

$$g(y) \geq \overline{\lim}\, g(\mathrm{P}_g(\mathrm{R}_U T^{k_n} x)) \geq \underline{\lim}\, g(\mathrm{P}_g(\mathrm{R}_U T^{k_n} x)) \geq g(c). \tag{60}$$

Combining with (58), we deduce that

$$c \in (U - v) \cap \operatorname{dom} g. \tag{61}$$

Set $\mu = \inf g(U - v)$. Choosing $y = c$ in (60) yields

$$g(\mathrm{P}_g(\mathrm{R}_U T^{k_n} x)) \to g(c) \geq \mu. \tag{62}$$

Now choosing y so that $g(y)$ is as close to μ as we like, we deduce from (60) and (62) that

$$g(\mathrm{P}_g(\mathrm{R}_U T^{k_n} x)) \to g(c) = \mu. \tag{63}$$

Hence c is a minimizer of $\iota_{U-v} + g$. Because c was an *arbitrary* weak cluster point of $(\mathrm{P}_g \mathrm{R}_U T^n x)_{n\in\mathbb{N}}$, we obtain through a simple proof by contradiction that

$$g(\mathrm{P}_g(\mathrm{R}_U T^n x)) \to \mu, \tag{64}$$

i.e., (51) holds.

Next, (59) with $y = c$ yields $\mu = g(c) \ge \mu + \overline{\lim}\, \alpha_n \ge \mu + \underline{\lim}\, \alpha_n \ge \mu$. Thus $\alpha_n \to 0$ and (52) follows.

Finally, (53) follows from (50) and Lemma 4.1(i). ∎

Remark 4.4. Note that (52) is equivalent to $n \cdot \langle T^n x - T^{n+1}x - v, v\rangle \to 0$. On the other hand, (15) and (16) combined with [19, Chapter III, section 14, Theorem on p. 124] (or [18, Problem 3.2.35]) yields $n \cdot \|T^n x - T^{n+1}x - v\|^2 \to 0$. We do not know whether $n \cdot \|T^n x - T^{n+1}x - v\| \to 0$.

5. The main result. We are now ready for the main result. In the following we set

$$y\colon X \to X\colon x \mapsto \lim_{n\to\infty} \mathrm{P}_F(nv + T^n x), \tag{65}$$

which is well defined by Fact 2.1.

Theorem 5.1 (main result). *Let $x \in X$. Then*

$$\mathrm{P}_U T^n x \rightharpoonup \mathrm{P}_U y(x) \in \operatorname{argmin}(\iota_U + g(\cdot - v)), \tag{66}$$

$$T^{n+1}x - T^n x + \mathrm{P}_U T^n x = \mathrm{P}_g(\mathrm{R}_U T^n x) \rightharpoonup -v + \mathrm{P}_U y(x), \text{ and}$$

$$g(\mathrm{P}_g \mathrm{R}_U T^n x) \to \min(\iota_U + g(\cdot - v)). \tag{67}$$

Proof. For brevity, we write $y = y(x)$. Because P_U is continuous, we have

$$\mathrm{P}_U \mathrm{P}_F(nv + T^n x) \to \mathrm{P}_U y. \tag{68}$$

On the other hand, $\mathrm{P}_U \mathrm{P}_F = \mathrm{P}_Z = \mathrm{P}_Z \mathrm{P}_U$ by (40) and (20). Invoking the fact that $v \in U^\perp$ (see (29)), we conclude altogether that

$$\mathrm{P}_Z \mathrm{P}_U T^n x = \mathrm{P}_Z \mathrm{P}_U(nv + T^n x) \to \mathrm{P}_U y. \tag{69}$$

Recall from (53) and (34) that $(\mathrm{P}_U T^n x)_{n\in\mathbb{N}}$ is bounded and that all its cluster points lie in $\operatorname{argmin}(\iota_U + g(\cdot - v)) = Z$. Now let z be an arbitrary weak cluster point of $(\mathrm{P}_U T^n x)_{n\in\mathbb{N}}$, say $\mathrm{P}_U T^{k_n} x \rightharpoonup z \in Z \subseteq U$. Then $\mathrm{P}_Z \mathrm{P}_U T^{k_n} x \rightharpoonup \mathrm{P}_Z z = z$ using (10). Combining with (69), we deduce that $z = \mathrm{P}_U y$. Hence *every* weak cluster point of $(\mathrm{P}_U T^n x)_{n\in\mathbb{N}}$ coincides with $\mathrm{P}_U y$. In view of the boundedness of $(\mathrm{P}_U T^n x)_{n\in\mathbb{N}}$, we obtain (66). The remainder follows from Lemma 4.1(i) and (51). ∎

Example 5.2 (linear-convex feasibility). Suppose that $g = \iota_W$, where W is a nonempty closed convex subset of X such that $U \cap (v + W) \neq \varnothing$. Then, $0 \in \operatorname{dom} g^*$ which implies that $0 \in U^\perp + \operatorname{dom} g^*$, hence (11) is verified. Moreover, $v = \mathrm{P}_{\overline{U-W}}(0)$ by [8, Proposition 3.16] and $(\forall x \in X)$ $\mathrm{P}_U T^n x \rightharpoonup \mathrm{P}_U y \in U \cap (v + W)$, where $y = \lim_{n\to\infty} \mathrm{P}_F(nv + T^n x)$ by Theorem 5.1.

Example 5.3. Suppose that W is a linear subspace of X such that $\{0\} \subsetneqq W \subsetneqq U^\perp$. Let $w \in W \smallsetminus \{0\}$, let $b \in (U^\perp \cap W^\perp) \smallsetminus \{0\}$, and suppose that $g = \frac{1}{2}\|\cdot\|^2 + \langle w, \cdot\rangle + \iota_{-b+W}$. Let $x \in X$. Then the following hold:

(i) $\partial g = w + \mathrm{Id} + \mathrm{N}_{-b+W}$.
(ii) $U \cap W = \{0\}$.
(iii) $\operatorname{dom} g = \operatorname{dom} \partial g = -b + W$, $\operatorname{dom} g^* = X$, and $0 \in U^\perp + \operatorname{dom} g^* = X$.
(iv) $v = b \in U^\perp \cap W^\perp$.
(v) $-v + \mathrm{N}_U = \mathrm{N}_U$.
(vi) $Z = \{0\}$.
(vii) $\mathrm{P}_g = -b - \frac{1}{2}w + \frac{1}{2}\mathrm{P}_W$.
(viii) $T = -b - \frac{1}{2}w + \mathrm{Id} - \mathrm{P}_U - \frac{1}{2}\mathrm{P}_W$.
(ix) $F = U^\perp \cap (-w + W^\perp)$.
(x) $0 \notin F$.
(xi) $(\forall n \geq 1)\ T^n x = (\mathrm{P}_{U^\perp} - (1 - \frac{1}{2^n})\mathrm{P}_W)x - nb - (1 - \frac{1}{2^n})w$.
(xii) $(\forall n \geq 1)\ \mathrm{P}_U T^n x = 0$.

Proof. Note that $U + W \subsetneqq U + U^\perp = X$ and thus $U^\perp \cap W^\perp = (U+W)^\perp \supsetneqq \{0\}$. Hence the choice of b is possible. (i): Proof is clear. (ii): Indeed, $\{0\} \subseteq U \cap W \subseteq U \cap U^\perp = \{0\}$. (iii): It is clear that $\operatorname{dom} g = \operatorname{dom} \partial g = -b + W$. Because $\lim_{\|x\| \to +\infty} g(x)/\|x\| = +\infty$, it follows that $\operatorname{dom} g^* = \operatorname{dom} \partial g^* = X$ by, e.g., [5, Propositions 14.15 and 16.27]. (iv): Using (29) and (iii), we obtain $v = \mathrm{P}_{\overline{U - \operatorname{dom} g}}(0) = \mathrm{P}_{b+U+W}(0) = b + \mathrm{P}_{U+W}(0 - b) = \mathrm{P}_{(U+W)^\perp}(b) = \mathrm{P}_{U^\perp \cap W^\perp}(b) = b$. (v): Proof is clear from (iv). (vi): This follows from (9), (i), (ii), and (iii). (vii): Set $y = -b - \frac{1}{2}w + \frac{1}{2}\mathrm{P}_W x$. Then $y \in -b + W$. Thus, $\mathrm{P}_{W^\perp} x \in -2b + W^\perp \Leftrightarrow x \in 2(-b - \frac{1}{2}w + \frac{1}{2}\mathrm{P}_W x) + w + W^\perp = 2y + w + W^\perp = y + w + y + \mathrm{N}_{-b+W}(y) = (\mathrm{Id} + \partial g)(y) \Leftrightarrow y = \mathrm{P}_g(x)$. (viii): This follows from (6) and (vii). (ix): Using (13) and (viii), we obtain $x \in F \Leftrightarrow x = T(x+v) = T(x+b) \Leftrightarrow x = -b - \frac{1}{2}w + x + b - \mathrm{P}_U(x+b) - \frac{1}{2}\mathrm{P}_W(x+b) \Leftrightarrow 0 = \frac{1}{2}w + \frac{1}{2}\mathrm{P}_U x + \frac{1}{2}\mathrm{P}_W x \Leftrightarrow [x \in U^\perp \text{ and } x \in -w + W^\perp]$. (x): We have the equivalences $0 \in F \Leftrightarrow 0 = T(0+v) \Leftrightarrow 0 = T(b) \Leftrightarrow 0 = -b - \frac{1}{2}w + b - \mathrm{P}_U b - \frac{1}{2}\mathrm{P}_W b \Leftrightarrow 0 = -\frac{1}{2}w$, which is absurd. (xi): This follows from (ix) and induction. (xii): Proof is clear from (xi). ☐

Remark 5.4. We point out that in [12, Theorem 4.4] the authors provide an instance where the shadow sequence converges. The proof in [12] critically relies on the assumption that $Z \subseteq F$. Our new result does not require this assumption. Indeed, by Example 5.3(vi) and (x), $Z = \{0\}$ and $Z \cap F = \varnothing$.

Example 5.5. Suppose that X is finite dimensional,[5] that $U \neq \{0\}$, let $u^* \in U \smallsetminus \{0\}$, suppose that[6] $g = \frac{1}{2}\mathrm{dist}_U^2 + \langle u^*, \cdot \rangle$, and let $x \in X$. Then the following hold:
(i) $\partial g = \nabla g = u^* + \mathrm{P}_{U^\perp}$.
(ii) $U - \operatorname{dom} \nabla g = U - \operatorname{dom} g = X$.
(iii) $\operatorname{ran} \mathrm{N}_U + \operatorname{ran} \partial g = U^\perp + \operatorname{dom} g^* = U^\perp + \operatorname{dom} \partial g^* = u^* + U^\perp$ is closed.
(iv) $0 \notin \overline{U^\perp + \operatorname{dom} g^*} = \overline{\operatorname{ran} \mathrm{N}_U + \operatorname{ran} \partial g}$.
(v) $v = u^* \in U \smallsetminus \{0\}$.
(vi) $Z = U$.
(vii) $\mathrm{P}_g = -u^* + \mathrm{Id} - \frac{1}{2}\mathrm{P}_{U^\perp}$.
(viii) $T = \mathrm{P}_g = -u^* + \mathrm{Id} - \frac{1}{2}\mathrm{P}_{U^\perp}$.
(ix) $F = U$.
(x) $(\forall n \in \mathbb{N})\ T^n x = -nu^* + \mathrm{P}_U x + \frac{1}{2^n}\mathrm{P}_{U^\perp} x$.
(xi) $(\forall n \in \mathbb{N})\ \mathrm{P}_U T^n x = -nu^* + \mathrm{P}_U x$.
(xii) $(\forall n \in \mathbb{N})\ \|T^n x\| \geq \|\mathrm{P}_U T^n x\| \geq n\|u^*\| - \|\mathrm{P}_U x\| \to +\infty$.

[5]We require this assumption in the proof of item (v) which relies on [9].

[6]Given a nonempty closed convex subset C of X, the associated distance function to the set C is denoted by dist_C.

Proof. (i): Proof is clear since $\nabla \frac{1}{2}\operatorname{dist}_U^2 = \operatorname{Id} - \mathrm{P}_U = \mathrm{P}_{U^\perp}$. Note that $\nabla g = u^* + \operatorname{Id} - \mathrm{P}_U = u^* + \mathrm{P}_{U^\perp}$. (ii): $U - \operatorname{dom}\partial g = U - X = X$. (iii): $\operatorname{dom}\partial g^* = \operatorname{ran}\nabla g = u^* + U^\perp$ is closed. On the other hand, $\operatorname{dom}\partial g^*$ is a dense subset of $\overline{\operatorname{dom}}\, g^*$. Hence $\operatorname{dom}\partial g^* = \operatorname{dom} g^* = u^* + U^\perp$ and thus $\operatorname{ran} \mathrm{N}_U + \operatorname{ran}\partial g = U^\perp + (u^* + U^\perp) = u^* + U^\perp$. (iv): Proof is clear from (iii) and the assumption that $u^* \neq 0$. (v): By [9, Proposition 6.1], (ii), and (iii), we have $v = \mathrm{P}_{\overline{U - \operatorname{dom} g}\, \cap\, \overline{U^\perp + \operatorname{dom} g^*}}(0) = \mathrm{P}_{u^* + U^\perp}(0) = u^* + \mathrm{P}_{U^\perp}(0 - u^*) = \mathrm{P}_U(u^*) = u^*$. (vi): Using (9), (i), and (v), we have $x \in Z \Leftrightarrow v \in \mathrm{N}_U(x) + \partial g(x - v) \Leftrightarrow [x \in U$ and $u^* \in U^\perp + u^* + \mathrm{P}_{U^\perp}(x - u^*)] \Leftrightarrow x \in U$. (vii): Set $y = -u^* + x - \frac{1}{2}\mathrm{P}_{U^\perp}x$. By (i) and (v), $y + \nabla g(y) = (-u^* + x - \frac{1}{2}\mathrm{P}_{U^\perp}x) + (u^* + \mathrm{P}_{U^\perp}(-u^* + x - \frac{1}{2}\mathrm{P}_{U^\perp}x)) = x$. Thus $y = \mathrm{P}_g(x)$ as claimed. (viii): Using (6) and (vii), we obtain $T = \operatorname{Id} - \mathrm{P}_U + \mathrm{P}_g \mathrm{R}_U = \mathrm{P}_{U^\perp} + \mathrm{P}_g(\mathrm{P}_U - \mathrm{P}_{U^\perp}) = \mathrm{P}_{U^\perp} - u^* + (\operatorname{Id} - \frac{1}{2}\mathrm{P}_{U^\perp})(\mathrm{P}_U - \mathrm{P}_{U^\perp}) = -u^* + \mathrm{P}_U + \frac{1}{2}\mathrm{P}_{U^\perp} = -u^* + \mathrm{P}_U + \mathrm{P}_{U^\perp} - \frac{1}{2}\mathrm{P}_{U^\perp} = -u^* + \operatorname{Id} - \frac{1}{2}\mathrm{P}_{U^\perp} = \mathrm{P}_g$. (ix): Using (13), (v), and (viii), we have $x \in F \Leftrightarrow x = T(x + v) \Leftrightarrow x = -u^* + \mathrm{P}_U x + \frac{1}{2}\mathrm{P}_{U^\perp}(x + v) \Leftrightarrow x = \mathrm{P}_U x + \frac{1}{2}\mathrm{P}_{U^\perp}x \Leftrightarrow x \in U$. (x): This follows from (viii) and (v) by a straightforward induction. (xi): Apply P_U to (x) and use (v). (xii): This follows from (xi). ∎

Remark 5.6. Example 5.5 illustrates the importance of the constraint qualification (11); indeed, it provides a scenario where (11) fails (see item (iv)) and the shadow sequence never converges (see item (xii)).

Remark 5.7. While Theorem 5.1 guarantees that $(\mathrm{P}_U T^n x)_{n\in\mathbb{N}}$ converges weakly to a minimizer of $\iota_U + g(\cdot - v)$, we leave numerical experiments and the development of meaningful termination criteria as topics for future research. A promising starting point appears to be the analysis in [1, section 5].

The remaining results in this section were inspired by a referee's question.

THEOREM 5.8 (switching the order of the operators). *Set* $\widetilde{T} = \operatorname{Id} - \mathrm{P}_g + \mathrm{P}_U \mathrm{R}_g = \operatorname{Id} - \mathrm{P}_g + \mathrm{P}_U(2\mathrm{P}_g - \operatorname{Id})$. *Suppose that*[7] $\mathrm{P}_{\overline{\operatorname{ran}}\,(\operatorname{Id} - \widetilde{T})}(0) = -v$. *Let* $x \in X$. *Then the following hold:*

(i) $(\forall n \in \mathbb{N})\ \mathrm{P}_U \widetilde{T}^n = \mathrm{P}_U T^n \mathrm{R}_U$.

(ii) $\widetilde{T}^n x - \widetilde{T}^{n+1}x = \mathrm{P}_g \widetilde{T}^n x - 2\mathrm{P}_U \mathrm{P}_g \widetilde{T}^n x + \mathrm{P}_U \widetilde{T}^n x = \mathrm{P}_U \widetilde{T}^n x - \mathrm{R}_U \mathrm{P}_g \widetilde{T}^n x \to -v$.

(iii) $\mathrm{P}_U \widetilde{T}^n x - \mathrm{P}_U \mathrm{P}_g \widetilde{T}^n x \to \mathrm{P}_U(-v) = 0$.

(iv) $\mathrm{P}_U T^n x \rightharpoonup \mathrm{P}_U y(x) \in \operatorname{argmin}(\iota_U + g(\cdot - v))$.

(v) $\mathrm{P}_U \widetilde{T}^n x \rightharpoonup \mathrm{P}_U y(\mathrm{R}_U x) \in \operatorname{argmin}(\iota_U + g(\cdot - v))$.

(vi) $\mathrm{P}_g \widetilde{T}^n x \rightharpoonup \mathrm{P}_U y(\mathrm{R}_U x) - v \in \operatorname{dom} g$.

Proof. Observe that $\mathrm{P}_U \mathrm{R}_U = \mathrm{P}_U$ and $\mathrm{R}_U^2 = \operatorname{Id}$. (i): Using [13, Theorem 2.7(i)] we learn that $(\forall n \in \mathbb{N})\ \mathrm{P}_U \widetilde{T}^n = \mathrm{P}_U \mathrm{R}_U \widetilde{T}^n \mathrm{R}_U \mathrm{R}_U = \mathrm{P}_U T^n \mathrm{R}_U$. (ii): $\widetilde{T}^n - \widetilde{T}^{n+1} = \mathrm{P}_g \widetilde{T}^n - \mathrm{P}_U \mathrm{R}_g \widetilde{T}^n = \mathrm{P}_g \widetilde{T}^n - 2\mathrm{P}_U \mathrm{P}_g \widetilde{T}^n + \mathrm{P}_U \widetilde{T}^n = \mathrm{P}_U \widetilde{T}^n - \mathrm{R}_U \mathrm{P}_g \widetilde{T}^n$. Now combine with (17). (iii): Recall that $-v \in U^\perp$ by (29). Now combine with (ii). (iv): This is Theorem 5.1. (v): Combine (i) and (iv) with x replaced by $\mathrm{R}_U x$. (vi): It follows from (iii) and (v) that $\mathrm{P}_U \mathrm{P}_g \widetilde{T}^n x \rightharpoonup \mathrm{P}_U y(\mathrm{R}_U x)$. Now combine with (ii). ∎

In the setting of Theorem 5.1, we point out that no general conclusion can be drawn about the sequence $(\mathrm{P}_g T^n x)_{n\in\mathbb{N}}$ as we illustrate below.

Example 5.9 ($(\mathrm{P}_g T^n x)_{n\in\mathbb{N}}$ may converge). Suppose that $(U, g) = (X, \iota_X)$. Then $\mathrm{P}_U = \mathrm{P}_g = T = \widetilde{T} = \operatorname{Id}$. Hence, $\operatorname{ran}(\operatorname{Id} - T) = \operatorname{ran}(\operatorname{Id} - \widetilde{T}) = \{0\}$. Consequently, $v = -v = 0$ and $(\forall n \in \mathbb{N})\ (\forall x \in X)\ \mathrm{P}_g T^n x = x = \lim_{n\to\infty} \mathrm{P}_g T^n x$.

[7] This assumption is satisfied if, for instance, X is finite dimensional. To see this, proceed as in the proof of Proposition 3.1(ii), with the roles of ι_U and g switched.

Example 5.10 (($(\mathrm{P}_g T^n x)_{n\in\mathbb{N}}$ may have no cluster points). Suppose that $X = \mathbb{R}^2$, that $U = \mathbb{R} \times \{0\}$, that $C = \operatorname{epi}(|\cdot| + 1)$, and that $g = \iota_C$. Let $x \in [-1, 1] \times \{0\}$. Using induction (see also [7, Example 2.6]), one can show that $(\forall n \in \{1, 2, \ldots\})$ $T^n x = (0, n) \in C$. Consequently, $\|\mathrm{P}_g T^n x\| = \|\mathrm{P}_C T^n x\| = n \to +\infty$.

6. Minimizing the sum of finitely many functions. In this section we assume for simplicity that

$$X \text{ is finite dimensional,} \tag{70}$$

that $m \in \{2, 3, \ldots\}$, that $I = \{1, 2, \ldots, m\}$, and that

$$g_i \colon X \to \left]-\infty, +\infty\right] \text{ is convex, lower semicontinuous, and proper} \tag{71}$$

for every $i \in I$. Furthermore, we set (see also [5] and [14])

$$\begin{cases} \mathbf{X} = \bigoplus_{i\in I} X, \\ \mathbf{g} = \bigoplus_{i\in I} g_i, \\ \boldsymbol{\Delta} = \big\{(x, x, \ldots, x) \in \mathbf{X} \;\big|\; x \in X\big\}, \\ \mathbf{Z} = \big\{\mathbf{x} \in \mathbf{X} \;\big|\; \mathbf{v} \in N_{\boldsymbol{\Delta}}(\mathbf{x}) + \partial\mathbf{g}(\mathbf{x} - \mathbf{v})\big\}, \\ (\forall i \in I) \quad D_i = \overline{\operatorname{dom}} g_i, \\ \mathbf{D} = \bigtimes_{i\in I} D_i, \\ \mathbf{v} = (v_i)_{i\in I} = \mathrm{P}_{\overline{\operatorname{ran}}\,(\mathrm{Id} - \mathbf{T})}(\mathbf{0}), \\ \mathbf{T} = \mathrm{Id} - \mathrm{P}_{\boldsymbol{\Delta}} + \mathrm{P}_{\mathbf{g}} \mathrm{R}_{\boldsymbol{\Delta}}, \\ \mathbf{j} \colon X \to \boldsymbol{\Delta} \colon x \mapsto (x, x, \ldots, x), \\ e \colon \mathbf{X} \to X \colon (x_i)_{i\in I} \mapsto \frac{1}{m}\big(\sum_{i\in I} x_i\big). \end{cases} \tag{72}$$

Remark 6.1. In passing we point out that, by [10, Theorem 2.16], we have $(\forall i \in I)$ $D_i = \overline{\operatorname{dom}}\, \partial g_i = \overline{\operatorname{dom}}\, g_i$.

FACT 6.2. *Write* $\mathbf{x} = (x_i)_{i\in I} \in \mathbf{X}$. *Then the following hold:*

(i) $\mathbf{g} \colon \mathbf{X} \to \left]-\infty, +\infty\right]$ *is convex, lower semicontinuous, and proper.*

(ii) $\mathbf{g}^* = \bigoplus_{i\in I} g_i^*$.

(iii) $\partial\mathbf{g} = \bigtimes_{i\in I} \partial g_i$.

(iv) $\mathrm{P}_{\boldsymbol{\Delta}}\mathbf{x} = \mathbf{j}\big(\frac{1}{m}\sum_{i\in I} x_i\big)$.

(v) $\mathrm{P}_{\mathbf{g}} = \bigtimes_{i\in I} \mathrm{P}_{g_i}$.

(vi) $\boldsymbol{\Delta}^\perp = \big\{\mathbf{u} \in \mathbf{X} \;\big|\; \sum_{i\in I} u_i = 0\big\}$.

Proof. (i): Proof is clear. (ii): This is [5, Proposition 13.30]. (iii): This is [5, Proposition 16.9]. (iv): This is [5, Proposition 26.4(ii)]. (v): This is [5, Proposition 24.11]. (vi): This is [5, Proposition 26.4(i)]. ∎

Next we define the set of least squares solutions of $(D_i)_{i\in I}$

$$L = \operatorname{argmin} \sum_{i\in I} \operatorname{dist}^2_{D_i}. \tag{73}$$

Finally, throughout the remainder of this section, we assume that

$$\mathbf{0} \in \boldsymbol{\Delta}^\perp + \operatorname{dom}\mathbf{g}^* \text{ and } \mathbf{Z} \neq \varnothing. \tag{74}$$

Remark 6.3. In many applications, the individual functions g_i have minimizers. In such cases, $(\forall i \in I)$ $0 \in \operatorname{dom}\partial g_i^* \subseteq \operatorname{dom} g_i^*$ and, therefore, $\mathbf{0} \in \operatorname{dom}\mathbf{g}^* \subseteq \boldsymbol{\Delta}^\perp + \operatorname{dom}\mathbf{g}^*$.

PROPOSITION 6.4. *The following hold:*

(i) $\mathbf{v} = \mathrm{P}_{\overline{\mathbf{\Delta}-\operatorname{dom}\mathbf{g}}}(\mathbf{0}) = \mathrm{P}_{\overline{\mathbf{\Delta}-\mathbf{D}}}(\mathbf{0}) \in \mathbf{\Delta}^{\perp}$.

(ii) $\operatorname{Fix}\mathrm{P}_{\mathbf{\Delta}}\mathrm{P}_{\mathbf{D}} = \mathbf{\Delta}\cap(\mathbf{v}+\mathbf{D}) \neq \varnothing$.

(iii) $(\forall y \in \operatorname{Fix}\mathrm{P}_{\mathbf{\Delta}}\mathrm{P}_{\mathbf{D}})\ \mathbf{v} = \mathbf{y} - \mathrm{P}_{\mathbf{D}}(\mathbf{y})$.

(iv) $\mathbf{Z} = \big\{\mathbf{x}\in\mathbf{\Delta} \ \big|\ \mathbf{\Delta}^{\perp}\cap\partial\mathbf{g}(\mathbf{x}-\mathbf{v}) \neq \varnothing\big\} = \mathbf{j}\big(\operatorname{zer}\sum_{i\in I}\partial g_i(\cdot - v_i)\big)$.

(v) $\operatorname{zer}\Big(\sum_{i\in I}\partial g_i(\cdot - v_i)\Big) \neq \varnothing$.

(vi) $L = \operatorname{Fix}\Big(\frac{1}{m}\sum_{i\in I}\mathrm{P}_{D_i}\Big) = \bigcap_{i\in I}(v_i + D_i)$.

(vii) $e(\mathbf{Z}) = \operatorname{zer}\big(\sum_{i\in I}\partial g_i(\cdot - v_i)\big) \subseteq \cap_{i\in I}(\operatorname{dom}\partial g_i(\cdot - v_i)) \subseteq \cap_{i\in I}(v_i + D_i) = L$.

Proof. (i): Observe that $\overline{\mathbf{\Delta} - \operatorname{dom}\mathbf{g}} = \overline{\mathbf{\Delta} - \overline{\operatorname{dom}}\mathbf{g}} = \overline{\mathbf{\Delta} - \mathbf{D}}$. Now combine this with (74) and Proposition 3.1(ii) applied with (X, U, g) replaced by $(\mathbf{X}, \mathbf{\Delta}, \mathbf{g})$. (ii) and (iii): Combine [2, Lemmas 2.2(i) and (iv)] and (34) applied with (X, U, g) replaced by $(\mathbf{X}, \mathbf{\Delta}, \mathbf{g})$. (iv): The first identity follows from applying (30) with (X, U, g) replaced by $(\mathbf{X}, \mathbf{\Delta}, \mathbf{g})$. The second identity follows from [5, Propositions 26.4(vii) and (viii)]. (v): This is a direct consequence of item (iv). (vi): Combine item (i), [2, Lemma 2.2(i)] and [7, Corollary 3.1]. (vii): This is a direct consequence of (iv) and (vi). ∎

PROPOSITION 6.5. *Suppose that* $j \in I$ *satisfies that* $\operatorname{dom} g_j = X$. *Then* $v_j = 0$.

Proof. Set $\mathbf{A} = \operatorname{argmin}(\iota_{\mathbf{\Delta}} + \mathbf{g}(\cdot - \mathbf{v}))$ and observe that Propositions 6.4(i) and (ii) imply that $\mathbf{A} \subseteq \mathbf{\Delta}\cap(\mathbf{v}+\operatorname{dom}\mathbf{g}) \subseteq \mathbf{\Delta}\cap(\mathbf{v}+\mathbf{D}) = \operatorname{Fix}\mathrm{P}_{\mathbf{\Delta}}\mathrm{P}_{\mathbf{D}}$. Note that (74) and Theorem 3.4 (applied with (U, g) replaced by $(\mathbf{\Delta}, \mathbf{g})$) imply that $\mathbf{A} = \mathbf{Z}$. Hence, $e(\mathbf{A}) = e(\mathbf{Z}) \subseteq L$, by Proposition 6.4(vii). Now, let $\mathbf{y} \in \operatorname{Fix}\mathrm{P}_{\mathbf{\Delta}}\mathrm{P}_{\mathbf{D}}$. Then Proposition 6.4(iii) implies that $\mathbf{v} = \mathbf{y} - \mathrm{P}_{\mathbf{D}}(\mathbf{y}) = (y_1, \ldots, y_m) - (\mathrm{P}_{D_1}y_1, \ldots, \mathrm{P}_{D_m}y_m)$. Consequently, if $D_j = X$ then $v_j = y_j - \mathrm{P}_{D_j}y_j = 0$. ∎

THEOREM 6.6. *Let* $\mathbf{x} = (x_i)_{i\in I} \in \mathbf{X}$ *and set* $\mathbf{y} = \lim_{n\to\infty}\mathrm{P}_{\operatorname{Fix}\mathbf{T}}(n\mathbf{v} + \mathbf{T}^n\mathbf{x})$. *Then*

$$\mathrm{P}_{\mathbf{\Delta}}\mathbf{T}^n\mathbf{x} \to \mathrm{P}_{\mathbf{\Delta}}\mathbf{y} \in \operatorname{argmin}(\iota_{\mathbf{\Delta}} + \mathbf{g}(\cdot - \mathbf{v})), \tag{75}$$

$$\begin{aligned} \mathbf{T}^{n+1}\mathbf{x} - \mathbf{T}^n\mathbf{x} + \mathrm{P}_{\mathbf{\Delta}}\mathbf{T}^n\mathbf{x} = \mathrm{P}_{\mathbf{g}}(\mathrm{R}_{\mathbf{\Delta}}\mathbf{T}^n\mathbf{x}) &\to -\mathbf{v} + \mathrm{P}_{\mathbf{\Delta}}\mathbf{y}, \\ \textit{and } \mathbf{g}(\mathrm{P}_{\mathbf{g}}\mathrm{R}_{\mathbf{\Delta}}\mathbf{T}^n\mathbf{x}) &\to \min(\iota_{\mathbf{\Delta}} + \mathbf{g}(\cdot - \mathbf{v})). \end{aligned} \tag{76}$$

Furthermore,

$$e(\mathrm{P}_{\mathbf{\Delta}}\mathbf{y}) \in \operatorname{argmin}\left(\sum_{i\in I} g_i(\cdot - v_i)\right). \tag{77}$$

Proof. (75) and (76) follow from applying Theorem 5.1 with (X, U, g) replaced by $(\mathbf{X}, \mathbf{\Delta}, \mathbf{g})$. It follows from combining (75) and Theorem 3.4 (applied with (U, g) replaced by $(\mathbf{\Delta}, \mathbf{g})$) that $\mathrm{P}_{\mathbf{\Delta}}\mathbf{y} \in \operatorname{argmin}(\iota_{\mathbf{\Delta}} + \mathbf{g}(\cdot - \mathbf{v})) = \mathbf{Z}$. Now combine with Proposition 6.4(vii). ∎

COROLLARY 6.7. *Let* $x_0 \in X$ *and set* $\overline{x}_0 = x_{0,1} = \cdots = x_{0,m} = x_0$. *Update via* $(\forall n \in \mathbb{N})$

$$(\forall i \in I) \quad x_{n+1,i} = x_{n,i} - \overline{x}_n + \mathrm{P}_{g_i}(2\overline{x}_n - x_{n,i}), \tag{78a}$$

$$\overline{x}_{n+1} = \tfrac{1}{m}\sum_{i\in I} x_{n+1,i}. \tag{78b}$$

Then $\overline{x}_n \to \overline{x} \in \operatorname{argmin}(\sum_{i\in I} g_i(\cdot - v_i))$.

Proof. Combine Theorem 6.6 and Propositions 6.4(v), (iv), and (v) in view of (74). ∎

COROLLARY 6.8. *Suppose that* $J \subseteq I$, *that for every* $i \in I \smallsetminus J$, $f_i\colon X \to \mathbb{R}$ *is convex and satisfies* $\operatorname{dom} f_i = X$ *and* $\operatorname{argmin} f_i \neq \varnothing$, *and that for every* $i \in J$, $C_i \neq X$ *is convex, closed, and nonempty. Set* $L_C = \operatorname{argmin} \sum_{i\in J} \operatorname{dist}^2_{C_i}$. *Consider the problem*

$$\text{minimize} \sum_{i\in I\smallsetminus J} f_i(x) \ \text{subject to}\ x \in \bigcap_{i\in J} C_i. \tag{79}$$

Suppose that $\operatorname{zer}(\sum_{i\in I\smallsetminus J} \partial f_i + \sum_{i\in J} \mathrm{N}_{C_i}(\cdot - v_i)) \neq \varnothing$. *Let* $x_0 \in X$, *and set* $\overline{x}_0 = x_{0,1} = \cdots = x_{0,m} = x_0$. *Update via* $(\forall n \in \mathbb{N})$

$$(\forall i \in I \smallsetminus J) \quad x_{n+1,i} = x_{n,i} - \overline{x}_n + \mathrm{P}_{g_i}(2\overline{x}_n - x_{n,i}), \tag{80a}$$

$$(\forall i \in J) \quad x_{n+1,i} = x_{n,i} - \overline{x}_n + \mathrm{P}_{C_i}(2\overline{x}_n - x_{n,i}), \tag{80b}$$

$$\overline{x}_{n+1} = \tfrac{1}{m}\sum_{i\in I} x_{n+1,i}. \tag{80c}$$

Then $\overline{x}_n \to \overline{x} \in X$, *and* $\overline{x}$ *is a solution of*

$$\text{minimize} \sum_{i\in I\smallsetminus J} f_i(x) \ \text{subject to}\ x \in L_C. \tag{81}$$

In particular, if $\cap_{i\in J} C_i \neq \varnothing$, *then* $L_C = \cap_{i\in J} C_i \neq \varnothing$ *and* $\overline{x}$ *is a solution of* (79).

Proof. Suppose that $g_i = f_i$ if $i \in I \smallsetminus J$ and $g_i = \iota_{C_i}$ if $i \in J$, and observe that (79) reduces to

$$\text{minimize} \sum_{i\in I} g_i(x). \tag{82}$$

Note that combining (78) and [5, Example 23.4] yields (80). It follows from Proposition 6.5 that $(\forall i \in I \smallsetminus J)$ $v_i = 0$. Consequently, $\operatorname{zer}(\sum_{i\in I} \partial g_i(\cdot - v_i)) = \operatorname{zer}(\sum_{i\in I\smallsetminus J} \partial f_i + \sum_{i\in J} \mathrm{N}_{C_i}(\cdot - v_i)) \neq \varnothing$, and by Corollary 6.7 we have $\overline{x}_n \to \overline{x} \in X$, and $\overline{x} \in \operatorname{zer}(\sum_{i\in I\smallsetminus J} \partial f_i + \sum_{i\in J} \mathrm{N}_{C_i}(\cdot - v_i))$. Finally, using Proposition 6.4(vi), $(\exists u \in X)$ $-u \in \sum_{i\in I\smallsetminus J} \partial f_i(\overline{x}) = \partial(\sum_{i\in I\smallsetminus J} f_i)(\overline{x})$ and $u \in \sum_{i\in J} \mathrm{N}_{C_i}(\overline{x} - v_i) \subseteq \mathrm{N}_{\cap_{i\in J}(v_i + C_i)}(\overline{x}) = \mathrm{N}_{L_C}(\overline{x})$. Therefore, $\overline{x}$ solves (81). ∎

Acknowledgments. The authors thank the editor and three anonymous referees for insightful comments that led to a substantially improved manuscript.

REFERENCES

[1] G. Banjac, P. Goulart, B. Stellato, and S. Boyd, *Infeasibility detection in the alternating direction method of multipliers for convex optimization*, J. Optim. Theory Appl., 183 (2019), pp. 490–519.

[2] H.H. Bauschke and J.M. Borwein, *Dykstra's alternating projection algorithm for two sets*, J. Approx. Theory, 79 (1994), pp. 418–443.

[3] H.H. Bauschke, J.M. Borwein, and A.S. Lewis, *The method of cyclic projections for closed convex sets in Hilbert space*, in Recent Developments in Optimization Theory and Nonlinear Analysis (Jerusalem 1995), Contemp. Math. 204, American Mathematical Society, Providence, RI, 1997, pp. 1–38.

[4] H.H. Bauschke, R.I. Boţ, W.L. Hare, and W.M. Moursi, *Attouch-Théra duality revisited: Paramonotonicity and operator splitting*, J. Approx. Theory, 164 (2012), pp. 1065–1084.

[5] H.H. Bauschke and P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed., Springer, Cham, Switzerland, 2017.

[6] H.H. Bauschke, P.L. Combettes, and D.R. Luke, *Finding best approximation pairs relative to two closed convex sets in Hilbert spaces*, J. Approx. Theory, 127 (2004), pp. 178–192.
[7] H.H. Bauschke, M.N. Dao, and W.M. Moursi, *The Douglas–Rachford algorithm in the affine-convex case*, Oper. Res. Lett., 44 (2016), pp. 379–382.
[8] H.H. Bauschke, W.L. Hare, and W.M. Moursi, *Generalized solutions for the sum of two maximally monotone operators*, SIAM J. Control Optim., 52 (2014), pp. 1034–1047.
[9] H.H. Bauschke, W.L. Hare, and W.M. Moursi, *On the range of the Douglas–Rachford operator*, Math. Oper. Res., 41 (2016), pp. 884–897.
[10] H.H. Bauschke, S.M. Moffat, and X. Wang, *Near equality, near convexity, sums of maximally monotone operators, and averages of firmly nonexpansive mappings*, Math. Program. Ser. B, 139 (2013), pp. 55–70.
[11] H.H. Bauschke and W.M. Moursi, *The Douglas–Rachford algorithm for two (not necessarily intersecting) affine subspaces*, SIAM J. Optim., 26 (2016), pp. 968–985.
[12] H.H. Bauschke and W.M. Moursi, *On the Douglas–Rachford algorithm*, Math. Program. Ser. A, 164 (2017), pp. 263–284.
[13] H.H. Bauschke and W.M. Moursi, *On the order of the operators in the Douglas–Rachford algorithm*, Optim. Lett., 10 (2016), pp. 447–455.
[14] P.L. Combettes, *Iterative construction of the resolvent of a sum of maximal monotone operators*, J. Convex Anal., 16 (2009), pp. 727–748.
[15] J. Douglas and H.H. Rachford, *On the numerical soluion of heat conduction problems in two and three variables*, Trans. Amer. Math. Soc., 82 (1956), pp. 421–439.
[16] J. Eckstein and D.P. Bertsekas, *On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone opeators*, Math. Program. (Ser. A), 55 (1992), pp. 293–318.
[17] A.N. Iusem, *On some properties of paramonotone operators*, J. Convex Anal., 5 (1998), pp. 269–278.
[18] W.J. Kaczor and M.T. Nowak, *Problems in Mathematical Analysis* I, American Mathematical Society, Providence, RI, 2000.
[19] K. Knopp, *Infinite Sequences and Series*, Dover, New York, 1956.
[20] P.L. Lions and B. Mercier, *Splitting algorithms for the sum of two nonlinear operators*, SIAM J. Numer. Anal., 16 (1979), pp. 964–979.
[21] Y. Liu, E.K. Ryu, and W. Yin, *A new use of Douglas-Rachford splitting for identifying infeasible, unbounded, and pathological conic programs*, Math. Program. Ser. A, 177 (2019), pp. 225–253.
[22] S.M. Moffat, W.M. Moursi, and S. Wang, *Nearly convex sets: Fine properties and domains or ranges of subdifferentials of convex functions*, Math. Program. Ser. A, 126 (2016), pp. 193–223.
[23] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
[24] E.K. Ryu, Y. Liu, and W. Yin, *Douglas–Rachford splitting and ADMM for pathological convex optimization*, Comput. Optim. Appl., 74 (2019), pp. 747–778, https://doi.org/10.1007/s10589-019-00130-9.
[25] B.F. Svaiter, *On weak convergence of the Douglas–Rachford method*, SIAM J. Control Optim., 49 (2011), pp. 280–287.