J Glob Optim (2016) 65:329–349
DOI 10.1007/s10898-015-0373-5

On Slater's condition and finite convergence of the Douglas–Rachford algorithm for solving convex feasibility problems in Euclidean spaces

Heinz H. Bauschke[1] · **Minh N. Dao**[1,2] · **Dominikus Noll**[3] · **Hung M. Phan**[4]

Received: 27 April 2015 / Accepted: 30 September 2015 / Published online: 13 October 2015

Abstract The Douglas–Rachford algorithm is a classical and very successful method for solving optimization and feasibility problems. In this paper, we provide novel conditions sufficient for finite convergence in the context of convex feasibility problems. Our analysis builds upon, and considerably extends, pioneering work by Spingarn. Specifically, we obtain finite convergence in the presence of Slater's condition in the affine-polyhedral and in a hyperplanar-epigraphical case. Various examples illustrate our results. Numerical experiments demonstrate the competitiveness of the Douglas–Rachford algorithm for solving linear equations with a positivity constraint when compared to the method of alternating projections and the method of reflection–projection.

Keywords Alternating projections · Convex feasibility problem · Convex set · Douglas–Rachford algorithm · Epigraph · Finite convergence · Method of reflection–projection · Monotone operator · Partial inverse · Polyhedral set · Projector · Slater's condition

✉ Heinz H. Bauschke
heinz.bauschke@ubc.ca

Minh N. Dao
minhdn@hnue.edu.vn

Dominikus Noll
noll@mip.ups-tlse.fr

Hung M. Phan
hung_phan@uml.edu

1 Department of Mathematics, University of British Columbia, Kelowna, BC V1V 1V7, Canada

2 Department of Mathematics and Informatics, Hanoi National University of Education, 136 Xuan Thuy, Hanoi, Vietnam

3 Institut de Mathématiques, Université de Toulouse, 118 route de Narbonne, 31062 Toulouse, France

4 Department of Mathematical Sciences, University of Massachusetts Lowell, 265 Riverside St., Olney Hall 428, Lowell, MA 01854, USA

Mathematics Subject Classification Primary 47H09 · 90C25; Secondary 47H05 · 49M27 · 65F10 · 65K05 · 65K10

1 Introduction

Throughout this paper, we assume that

$$X \text{ is a Euclidean space,} \tag{1}$$

i.e., a finite-dimensional real Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and induced norm $\|\cdot\|$, and

$$A \text{ and } B \text{ are closed convex subsets of } X \text{ such that } A \cap B \neq \varnothing. \tag{2}$$

Consider the *convex feasibility problem*

$$\text{find a point in } A \cap B \tag{3}$$

and assume that it is possible to evaluate the *projectors* (nearest point mappings) P_A and P_B corresponding to A and B, respectively. We denote the corresponding reflectors by $R_A := 2P_A - \mathrm{Id}$ and $R_B := 2P_B - \mathrm{Id}$, respectively. *Projection methods* combine the projectors and reflectors in a suitable way to generate a sequence converging to a solution of (3)—we refer the reader to [2,8], and [9] and the references therein for further information.

One celebrated algorithm for solving (3) is the so-called Douglas–Rachford algorithm (DRA) [11]. The adaption of this algorithm to optimization and feasibility is actually due to Lions and Mercier was laid out beautifully in their landmark paper [16] (see also [12]). The DRA is based on the Douglas–Rachford splitting operator,

$$T := \mathrm{Id} - P_A + P_B R_A, \tag{4}$$

which is used to generate a sequence $(z_n)_{n\in\mathbb{N}}$ with starting point $z_0 \in X$ via

$$(\forall n \in \mathbb{N}) \quad z_{n+1} := T z_n. \tag{5}$$

Then the "governing sequence" $(z_n)_{n\in\mathbb{N}}$ converges to a point $z \in \operatorname{Fix} T$, and, more importantly, the "shadow sequence" $(P_A z_n)_{n\in\mathbb{N}}$ converges to $P_A z$ which is a solution of (3).

An important question concerns the speed of convergence of the sequence $(P_A z_n)_{n\in\mathbb{N}}$. Linear convergence was more clearly understood recently, see [14,21], and [6].

The aim of this paper is to provide verifiable *conditions sufficient for* finite *convergence.*

Our two main results reveal that *Slater's condition*, i.e.,

$$A \cap \operatorname{int} B \neq \varnothing \tag{6}$$

plays a key role and guarantees finite convergence when **(MR1)** A is an affine subspace and B is a polyhedron (Theorem 3.7); or when **(MR2)** A is a certain hyperplane and B is an epigraph (Theorem 5.4). Examples illustrate that these results are applicable in situations where previously known conditions sufficient for finite convergence fail. When specialized to a product space setting, we derive a finite-convergence result due to Spingarn [26] for his method of partial inverses [25]. Indeed, the proof of Theorem 3.7 follows his pioneering work, but, at the same time, we simplify his proofs and strengthen the conclusions. These sharpenings allow us to obtain finite-convergence results for solving linear equations with a positivity constraint. Numerical experiments support the competitiveness of the DRA for solving (3).

It would be interesting to derive additional conditions guaranteeing finite convergence, not only in the present finite-dimensional setting but also in a more general infinite-dimensional setting. Moreover, how these results generalize (if at all) from convex feasibility to finding the zero of the sum of two maximally monotone operators is currently quite unclear.

1.1 Organization of the paper

The paper is organized as follows. In Sect. 2, we present several auxiliary results which make the eventual proofs of the main results more structured and transparent. Section 3 contains the first main result **(MR1)**. Applications using the product space set up, a comparison with Spingarn's work, and numerical experiments are provided in Sect. 4. The final Sect. 5 concerns the second main result **(MR2)**.

1.2 Notation

The notation employed is standard and follows largely [2]. The real numbers are $\mathbb{R}$, and the nonnegative integers are $\mathbb{N}$. Further, $\mathbb{R}_+ := \{x \in \mathbb{R} \mid x \geq 0\}$ and $\mathbb{R}_{++} := \{x \in \mathbb{R} \mid x > 0\}$. Let C be a subset of X. Then the closure of C is $\overline{C}$, the interior of C is $\operatorname{int} C$, the boundary of C is $\operatorname{bdry} C$, and the smallest affine and linear subspaces containing C are, respectively, $\operatorname{aff} C$ and $\operatorname{span} C$. The relative interior of C, $\operatorname{ri} C$, is the interior of C relative to $\operatorname{aff} C$. The orthogonal complement of C is $C^\perp := \{y \in X \mid (\forall x \in C)\ \langle x, y\rangle = 0\}$, and the dual cone of C is $C^\oplus := \{y \in X \mid (\forall x \in C)\ \langle x, y\rangle \geq 0\}$. The normal cone operator of C is denoted by N_C, i.e., $N_C(x) = \{y \in X \mid (\forall c \in C)\ \langle y, c - x\rangle \leq 0\}$ if $x \in C$, and $N_C(x) = \varnothing$ otherwise. If $x \in X$ and $\rho \in \mathbb{R}_{++}$, then $\operatorname{ball}(x; \rho) := \{y \in X \mid \|x - y\| \leq \rho\}$ is the closed ball centered at x with radius ρ.

2 Auxiliary results

In this section, we collect several auxiliary results that will be useful in the sequel.

2.1 Convex sets

Lemma 2.1 *Let C be a nonempty closed convex subset of X, let $x \in X$, and let $y \in C$. Then $x - P_C x \in N_C(y) \Leftrightarrow \langle x - P_C x, y - P_C x\rangle = 0$.*

Proof Because $y \in C$ and $x - P_C x \in N_C(P_C x)$, we have $\langle x - P_C x, y - P_C x\rangle \leq 0$. "$\Rightarrow$": From $x - P_C x \in N_C(y)$ and $P_C x \in C$, we have $\langle x - P_C x, P_C x - y\rangle \leq 0$. Thus $\langle x - P_C x, P_C x - y\rangle = 0$. "$\Leftarrow$": We have $(\forall c \in C)$ $\langle x - P_C x, c - y\rangle = \langle x - P_C x, c - P_C x\rangle + \langle x - P_C x, P_C x - y\rangle = \langle x - P_C x, c - P_C x\rangle \leq 0$. Hence $x - P_C x \in N_C(y)$. □

Lemma 2.2 *Let C be a nonempty convex subset of X. Then* $\operatorname{int} C \neq \varnothing \Leftrightarrow 0 \in \operatorname{int}(C - C)$.

Proof "$\Rightarrow$": Clear. "$\Leftarrow$": By [22, Theorem 6.2], $\operatorname{ri} C \neq 0$. After translating the set if necessary, we assume that $0 \in \operatorname{ri} C$. Then $0 \in C$, and so $Y := \operatorname{aff} C = \operatorname{span} C$. Since $0 \in \operatorname{int}(C - C) \subseteq \operatorname{int}(\operatorname{span} C) = \operatorname{int} Y$, this gives $\operatorname{int} Y \neq \varnothing$ and thus $Y = X$. In turn, $\operatorname{int} C = \operatorname{ri} C \neq \varnothing$. □

2.2 Cones

Lemma 2.3 *Let K be a nonempty convex cone in X. Then there exists $v \in \operatorname{ri} K \cap \operatorname{ri} K^{\oplus}$ such that*

$$(\forall x \in \overline{K} \smallsetminus (\overline{K} \cap (-\overline{K}))) \quad \langle v, x \rangle > 0. \tag{7}$$

Proof By [26, Lemma 2], there exists $v \in \operatorname{ri} K \cap \operatorname{ri} K^{\oplus}$. Then $v \in \operatorname{ri} K^{\oplus} = \operatorname{ri} \overline{K}^{\oplus}$. Noting that $\overline{K}$ is a closed convex cone and using [26, Lemma 3], we complete the proof. □

Lemma 2.4 *Let $(z_n)_{n \in \mathbb{N}}$ be a sequence in X, and let $f : X \to \mathbb{R}$ be linear. Assume that $z_n \to z \in X$, and that*

$$(\forall n \in \mathbb{N}) \quad f(z_n) > f(z_{n+1}). \tag{8}$$

Then there exist $n_0 \in \mathbb{N} \smallsetminus \{0\}$ and $(\mu_1, \ldots, \mu_{n_0}) \in \mathbb{R}_+^{n_0}$ such that $\sum_{k=1}^{n_0} \mu_k = 1$ and

$$\left\langle \sum_{k=1}^{n_0} \mu_k (z_{k-1} - z_k), \sum_{k=1}^{n_0} \mu_k (z_k - z) \right\rangle > 0. \tag{9}$$

Proof Introducing

$$(\forall n \in \mathbb{N} \smallsetminus \{0\}) \quad y_n := z_{n-1} - z_n, \tag{10}$$

we get

$$(\forall n \in \mathbb{N} \smallsetminus \{0\}) \quad f(y_n) > 0, \quad \text{and so} \quad y_n \neq 0. \tag{11}$$

Let K be the convex cone generated by $\{y_n\}_{n \in \mathbb{N} \smallsetminus \{0\}}$. We see that

$$(\forall x \in K \smallsetminus \{0\}) \quad f(x) > 0, \quad \text{and} \quad (\forall x \in -K \smallsetminus \{0\}) \quad f(x) < 0. \tag{12}$$

Therefore,

$$\overline{K} \cap (-\overline{K}) \subseteq \left\{ x \in X \,\middle|\, f(x) = 0 \right\}. \tag{13}$$

Setting

$$(\forall n \in \mathbb{N} \smallsetminus \{0\}) \quad w_n := z_n - z, \tag{14}$$

we immediately have $w_n \to 0$, and so

$$(\forall n \in \mathbb{N} \smallsetminus \{0\}) \quad w_n = z_n - z = \sum_{k=n+1}^{\infty} (z_{k-1} - z_k) = \sum_{k=n+1}^{\infty} y_k \in \overline{K}. \tag{15}$$

From (8) we get

$$(\forall n \in \mathbb{N} \smallsetminus \{0\}) \quad f(w_n) > f(w_{n+1}). \tag{16}$$

Moreover, $f(w_n) \to f(0) = 0$, hence

$$(\forall n \in \mathbb{N} \smallsetminus \{0\}) \quad f(w_n) > 0. \tag{17}$$

Together with (13) and (15), this gives

$$(\forall n \in \mathbb{N} \smallsetminus \{0\}) \quad w_n \in \overline{K} \smallsetminus (\overline{K} \cap (-\overline{K})). \tag{18}$$

By Lemma 2.3, there exists $v \in \operatorname{ri} K \cap \operatorname{ri} K^{\oplus}$ such that

$$(\forall n \in \mathbb{N} \smallsetminus \{0\}) \quad \langle v, w_n \rangle > 0. \tag{19}$$

Then we must have $v \neq 0$. Since $v \in K$, after scaling if necessary, there exist $n_0 \in \mathbb{N} \smallsetminus \{0\}$ and $(\mu_1, \ldots, \mu_{n_0}) \in \mathbb{R}_+^{n_0}$ such that

$$v = \sum_{k=1}^{n_0} \mu_k y_k, \quad \text{and} \quad \sum_{k=1}^{n_0} \mu_k = 1. \tag{20}$$

This combined with (19) implies

$$\left\langle \sum_{k=1}^{n_0} \mu_k y_k, \sum_{k=1}^{n_0} \mu_k w_k \right\rangle > 0, \tag{21}$$

and so (9) holds. □

Lemma 2.5 *Let K be a nonempty pointed*[1] *convex cone in X. Then the following hold:*

(i) Let $m \in \mathbb{N} \smallsetminus \{0\}$ and let $(x_1, \ldots, x_m) \in K^m$. Then $x_1 + \cdots + x_m = 0 \Leftrightarrow x_1 = \cdots = x_m = 0$.
(ii) If K is closed and $L\colon X \to X$ is linear such that

$$\ker L \cap K = \{0\}, \tag{22}$$

then $L(K)$ is a nonempty pointed closed convex cone.

Proof (i) Assume that $x_1 + \cdots + x_m = 0$. Then since K is a convex cone,

$$-x_1 = x_2 + \cdots + x_m \in K, \tag{23}$$

and so $x_1 \in K \cap (-K)$. Since K is pointed, we get $x_1 = 0$. Continuing in this fashion, we eventually conclude that $x_1 = \cdots = x_m = 0$. The converse is trivial.

(ii) Since K is a closed convex cone, so is $M := L(K)$ due to assumption (22) and [7, Proposition 3.4]. Now let $z \in M \cap (-M)$. Then $z = L(r) = -L(s)$ for some points r, s in K. Thus $L(r+s) = L(r) + L(s) = 0$, which gives $r + s \in \ker L$, and so $r + s \in \ker L \cap K$. By again (22), $r + s = 0$, and now (i) implies $r = s = 0$. Therefore, $z = 0$, and M is pointed.

□

Lemma 2.6 *Let $(a_n)_{n \in \mathbb{N}}$ be a sequence in X such that $a_n \to a \in X$, and K be a pointed closed convex cone of X. Assume that*

$$(\exists\, p \in \mathbb{N}) \quad a_p = a \quad \text{and} \quad (\forall n \geq p) \quad a_n - a_{n+1} \in K. \tag{24}$$

Then

$$(\forall n \geq p) \quad a_n = a. \tag{25}$$

Proof Since K is a closed convex cone and $a_n \to a$, it follows from (24) that

$$(\forall n \geq p) \quad a_n - a = \sum_{k=n}^{\infty} (a_k - a_{k+1}) \in K, \tag{26}$$

and so $a_{p+1} - a \in K$. Since $a_p = a$, (24) gives $a - a_{p+1} \in K$. Noting that K is pointed, this implies $a_{p+1} - a \in K \cap (-K) \subseteq \{0\}$, and hence $a_{p+1} = a$. Repeating this argument, we get the conclusion. □

[1] Recall that a cone K is pointed if $K \cap (-K) \subseteq \{0\}$.

2.3 Locally polyhedral sets

Definition 2.7 (*Local polyhedrality*) Let C be a subset of X. We say that C is *polyhedral at* $c \in C$ if there exist a polyhedral[2] set D and $\varepsilon \in \mathbb{R}_{++}$ such that $C \cap \mathrm{ball}(c;\varepsilon) = D \cap \mathrm{ball}(c;\varepsilon)$.

It is clear from the definition that every polyhedron is polyhedral at each of its points and that every subset C of X is polyhedral at each point in $\mathrm{int}\, C$.

Lemma 2.8 *Let C be a subset of X, and assume that C is polyhedral at $c \in C$. Then there exist $\varepsilon \in \mathbb{R}_{++}$, a finite set I, $(d_i)_{i\in I} \in (X \smallsetminus \{0\})^I$, $(\delta_i)_{i\in I} \in \mathbb{R}^I$ such that*

$$C \cap \mathrm{ball}(c;\varepsilon) = \left\{ x \in X \,\middle|\, \max_{i\in I} \big(\langle d_i, x\rangle - \delta_i \big) \le 0 \right\} \cap \mathrm{ball}(c;\varepsilon), \tag{27}$$

$(\forall i \in I)\ \langle d_i, c\rangle = \delta_i$, *and*

$$\big(\forall y \in C \cap \mathrm{ball}(c;\varepsilon)\big) \quad N_C(y) = \sum_{i\in I(y)} \mathbb{R}_+ d_i, \quad \textit{where } I(y) := \big\{ i \in I \,\big|\, \langle d_i, y\rangle = \delta_i \big\}. \tag{28}$$

Proof Combine Lemma 2.12 with [24, Theorem 6.46]. □

Lemma 2.9 *Let C be a nonempty closed convex subset of X that is polyhedral at $c \in C$. Then there exists $\varepsilon \in \mathbb{R}_{++}$ such that*

$$\big(\forall x \in P_C^{-1}(C \cap \mathrm{ball}(c;\varepsilon))\big) \quad \langle x - P_C x, c - P_C x\rangle = 0 \ \textit{ and } \ x - P_C x \in N_C(c). \tag{29}$$

Proof We adopt the notation of the conclusion of Lemma 2.8. Let $x \in X$ such that $y := P_C x \in \mathrm{ball}(c;\varepsilon)$. Then $x - y \in N_C(y)$ and Lemma 2.8 guarantees the existence of $(\lambda_i)_{i\in I(y)} \in \mathbb{R}_+^{I(y)}$ such that

$$x - y = \sum_{i\in I(y)} \lambda_i d_i. \tag{30}$$

Then

$$\langle x - y, c - y\rangle = \sum_{i\in I(y)} \lambda_i \langle d_i, c - y\rangle = \sum_{i\in I(y)} \lambda_i (\langle d_i, c\rangle - \langle d_i, y\rangle) = \sum_{i\in I(y)} \lambda_i (\delta_i - \delta_i) = 0 \tag{31}$$

and so $\langle x - P_C x, c - P_C x\rangle = 0$. Furthermore, by Lemma 2.1, $x - P_C x \in N_C(c)$. □

2.4 Two convex sets

Proposition 2.10 *Let A and B be closed convex subsets of X such that $A \cap B \neq \varnothing$. Then the following hold:*

(i) $(\forall a \in A)(\forall b \in B) \quad N_{A-B}(a - b) = N_A(a) \cap \big(- N_B(b)\big)$.
(ii) $0 \in \mathrm{int}(A - B) \quad \Leftrightarrow \quad N_{A-B}(0) = \{0\}$.
(iii) $\mathrm{int}\, B \neq \varnothing \quad \Leftrightarrow \quad (\forall x \in B)\ N_B(x) \cap (-N_B(x)) = \{0\}$.

Proof (i) Noting that $N_{-B}(-b) = -N_B(b)$ by definition, the conclusion follows from [20, Proposition 2.11(ii)] or from a direct computation. (ii) Clear from [2, Corollary 6.44]. (iii) Let $x \in B$. By (i), $N_B(x) \cap (-N_B(x)) = N_{B-B}(0)$. Now Lemma 2.2 and (ii) complete the proof. □

[2] Recall that a set is polyhedral if it is a finite intersection of halfspaces.

Corollary 2.11 *Let A be a linear subspace of X, and let B be a nonempty closed convex subset of X such that $A \cap B \neq \varnothing$. Then the following hold:*

(i) $0 \in \operatorname{int}(A - B) \quad \Leftrightarrow \quad \big[(\forall x \in A \cap B)\ A^{\perp} \cap N_B(x) = \{0\}\big]$.

(ii) $A \cap \operatorname{int} B \neq \varnothing \quad \Leftrightarrow \quad \big[0 \in \operatorname{int}(A - B)$ *and* $\operatorname{int} B \neq \varnothing\big]$.

Proof (i) Let $x \in A \cap B$. Since A is a linear subspace, Proposition 2.10(i) yields

$$A^{\perp} \cap N_B(x) = -(N_A(x) \cap (-N_B(x))) = -N_{A-B}(0). \tag{32}$$

Now apply Proposition 2.10(ii).

(ii) If $0 \in \operatorname{int}(A - B)$ and $\operatorname{int} B \neq \varnothing$, then $0 \in \operatorname{ri}(A - B)$ and $\operatorname{ri} B = \operatorname{int} B$. Since A is a linear subspace, we have $\operatorname{ri} A = A$, and using [22, Corollary 6.6.2], we get

$$0 \in \operatorname{ri}(A - B) = \operatorname{ri} A - \operatorname{ri} B = A - \operatorname{int} B, \tag{33}$$

which implies $A \cap \operatorname{int} B \neq \varnothing$. The converse is obvious.

□

Lemma 2.12 *Let A and B be closed convex subsets of X, and let $c \in A \cap B$ and $\varepsilon \in \mathbb{R}_{++}$ be such that $A \cap \operatorname{ball}(c; \varepsilon) = B \cap \operatorname{ball}(c; \varepsilon)$. Then $N_A(c) = N_B(c)$.*

Proof Let $u \in X$. Working with the directional derivative and using [2, Proposition 17.17(i)], we have $u \in N_A(c) = \partial \iota_A(c) \Leftrightarrow (\forall h \in X)\ \langle u, h\rangle \leq \iota'_A(c; h) \Leftrightarrow (\forall h \in X)\ \langle u, h\rangle \leq \iota'_B(c; h)$ $\Leftrightarrow u \in N_B(c) = \partial \iota_B(c)$. □

2.5 Monotone operators

Lemma 2.13 *Let $L\colon X \times X \to X \times X\colon (x, y) \mapsto (L_{11}x + L_{12}y, L_{21}x + L_{22}y)$, where each $L_{ij}\colon X \to X$ is linear. Assume that $L_{11}^{*}L_{22} + L_{21}^{*}L_{12} = \mathrm{Id}$ and that $L_{11}^{*}L_{21}$ and $L_{22}^{*}L_{12}$ are skew.[3] Then the following hold:*

(i) If $(x, y) \in X \times X$ and $(u, v) = L(x, y)$, then $\langle u, v\rangle = \langle x, y\rangle$.

(ii) Let $M\colon X \rightrightarrows X$ be a monotone operator, and define $M_L\colon X \rightrightarrows X$ via $\operatorname{gra} M_L = L(\operatorname{gra} M)$. Then for all ordered pairs (x, y) and (x', y') in $\operatorname{gra} M$ and $(u, v) = L(x, y)$, $(u', v') = L(x', y')$, we have $\big\langle u - v, u' - v'\big\rangle = \big\langle x - x', y - y'\big\rangle$; consequently, M_L is monotone.

Proof (i) The assumptions indeed imply

$$\langle u, v\rangle = \langle L_{11}x + L_{12}y, L_{21}x + L_{22}y\rangle \tag{34a}$$

$$= \big\langle x, (L_{11}^{*}L_{22} + L_{21}^{*}L_{12})y\big\rangle + \big\langle x, L_{11}^{*}L_{21}x\big\rangle + \big\langle y, L_{22}^{*}L_{12}y\big\rangle = \langle x, y\rangle\,. \tag{34b}$$

(ii) Since $(x - x', y - y') = (x, y) - (x', y')$ and L is linear, the result follows from (i).

□

Corollary 2.14 *Let A be a linear subspace of X, and let (x, y) and (x', y') be in $X \times X$. Then the following hold:*

(i) $\big\langle P_A y - P_{A^{\perp}} x, P_A x - P_{A^{\perp}} y\big\rangle = \langle y, x\rangle$

(ii) $\big\langle (P_A y - P_{A^{\perp}} x) - (P_A y' - P_{A^{\perp}} x'), (P_A x - P_{A^{\perp}} y) - (P_A x' - P_{A^{\perp}} y')\big\rangle$
$= \big\langle y - y', x - x'\big\rangle$.

[3] Recall that $S\colon X \to X$ is skew if $S^{*} = -S$.

Proof Set $L_{11} = L_{22} = P_A$ and $L_{12} = L_{21} = -P_{A^\perp}$ in Lemma 2.13. □

Remark 2.15 Spingarn's original *partial inverse* [25] arises from Lemma 2.13 by setting $L_{11} = L_{22} = P_A$ and $L_{12} = L_{21} = P_{A^\perp}$ while in Corollary 2.14 we used $L_{11} = L_{22} = P_A$ and $L_{12} = L_{21} = -P_{A^\perp}$. Other choices are possible: e.g., if $X = \mathbb{R}^2$ and R denotes the clockwise rotator by $\pi/4$, then a valid choice for Lemma 2.13 is $L_{11} = L_{22} = \frac{1}{2}R$ and $L_{12} = L_{21} = \frac{1}{2}R^*$.

2.6 Finite-convergence conditions for the proximal point algorithm

It is known (see, e.g., [12, Theorem 6]) that the DRA is a special case of the exact proximal point algorithm (with constant parameter 1). The latter generates, for a given maximally monotone operator $M\colon X \rightrightarrows X$ with resolvent $T := J_M = (\operatorname{Id} + M)^{-1}$, a sequence by

$$z_0 \in X, \quad (\forall n \in \mathbb{N}) \quad z_{n+1} := Tz_n = (\operatorname{Id} + M)^{-1} z_n, \tag{35}$$

in order to solve the problem

$$\text{find } z \in X \text{ such that } 0 \in Mz; \text{ equivalently, } z \in \operatorname{Fix} T. \tag{36}$$

A classical sufficient condition dates back to Rockafellar (see [23, Theorem 3]) who proved finite convergence when

$$(\exists\, \bar{z} \in X)(\exists\, \delta \in \mathbb{R}_{++}) \quad \operatorname{ball}(0; \delta) \subseteq M\bar{z}. \tag{37}$$

It is instructive to view this condition from the resolvent side:

$$(\exists\, \bar{z} \in X)(\exists\, \delta \in \mathbb{R}_{++})(\forall z \in X) \quad \|z - \bar{z}\| \leq \delta \;\Rightarrow\; Tz = \bar{z}. \tag{38}$$

Note that since $\operatorname{Fix} T$ is convex, this implies that

$$\operatorname{Fix} T = \{\bar{z}\} \tag{39}$$

is a singleton, which severely limits the applicability of this condition.

Later, Luque (see [17, Theorem 3.2]) proved finite convergence under the more general condition

$$M^{-1}0 \neq \varnothing \text{ and } (\exists\, \delta \in \mathbb{R}_{++}) \;\; M^{-1}\big(\operatorname{ball}(0; \delta)\big) \subseteq M^{-1}0. \tag{40}$$

On the resolvent side, his condition turns into

$$\operatorname{Fix} T \neq \varnothing \text{ and } (\exists \delta > 0) \quad \|z - Tz\| \leq \delta \quad \Rightarrow \quad Tz \in \operatorname{Fix} T. \tag{41}$$

However, when $M^{-1}0 = \operatorname{Fix} T \neq \varnothing$, it is well known that $z_n - Tz_n \to 0$; thus, the finite-convergence condition is essentially a tautology.

When illustrating our main results, we shall provide examples where both (37) and (40) fail while our results are applicable (see Remark 3.8 and Example 5.5 below).

2.7 Douglas–Rachford operator

For future use, we record some results on the DRA that are easily checked. Recall that, for two nonempty closed convex subsets A and B, the DRA operator is

$$T := \operatorname{Id} - P_A + P_B(2P_A - \operatorname{Id}) = \tfrac{1}{2}\left(\operatorname{Id} + R_B R_A\right). \tag{42}$$

The following result, the proof of which we omit since it is a direct verification, records properties in the presence of affinity/linearity.

Proposition 2.16 *Let A be an affine subspace of X and let B be a nonempty closed convex subset of X. Then the following hold:*

(i) *P_A is an affine operator and*

$$P_A R_A = P_A, \quad P_A T = P_A P_B R_A \quad \textit{and} \quad T = (P_A + P_B - \mathrm{Id}) R_A. \tag{43}$$

(ii) *If A is a linear subspace, then P_A is a symmetric linear operator and*

$$T = \big(P_A P_B - P_{A^\perp}(\mathrm{Id} - P_B)\big) R_A. \tag{44}$$

The next result will be used in Sect. 4.1 below to clarify the connection between the DRA and Spingarn's method.

Lemma 2.17 *Let A be a linear subspace of X, let B be a nonempty closed convex subset of X, let $(a, a^\perp) \in A \times A^\perp$, and set $(a_+, a_+^\perp) := (P_A P_B(a + a^\perp), P_{A^\perp}(\mathrm{Id} - P_B)(a + a^\perp)) \in A \times A^\perp$. Then $T(a - a^\perp) = a_+ - a_+^\perp$.*

Proof Clearly, $a_+ \in A$ and $a_+^\perp \in A^\perp$. Since $R_A(a - a^\perp) = (P_A - P_{A^\perp})(a - a^\perp) = a + a^\perp$, the conclusion follows from (44). □

3 The affine-polyhedral case with Slater's condition

In this section, we are able to state and prove finite convergence of the DRA in the case where A is an affine subspace and B is a polyhedral set such that Slater's condition, $A \cap \mathrm{int}\ B \neq \varnothing$, is satisfied. We start by recalling our standing assumptions. We assume that

$$X \text{ is a finite-dimensional real Hilbert space,} \tag{45}$$

and that

$$A \text{ and } B \text{ are closed convex subsets of } X \text{ such that } A \cap B \neq \varnothing. \tag{46}$$

The DRA is based on the operator

$$T := \mathrm{Id} - P_A + P_B R_A. \tag{47}$$

Given a starting point $z_0 \in X$, the DRA sequence $(z_n)_{n\in\mathbb{N}}$ is generated by

$$(\forall n \in \mathbb{N}) \quad z_{n+1} := T z_n. \tag{48a}$$

We also set

$$(\forall n \in \mathbb{N}) \quad a_n := P_A z_n, \quad r_n := R_A z_n = 2a_n - z_n. \tag{48b}$$

We now state the basic convergence result for the DRA.

Fact 3.1 *(Convergence of DRA) The DRA sequences (48) satisfy*

$$z_n \to z \in \mathrm{Fix}\, T = (A \cap B) + N_{A-B}(0) \quad \textit{and} \quad a_n \to P_A z \in A \cap B. \tag{49}$$

Proof Combine [3, Corollary 3.9 and Theorem 3.13]. □

Fact 3.1 can be strengthened when a constraint qualification is satisfied.

Lemma 3.2 *Suppose that $0 \in \mathrm{int}(A - B)$. Then there exists a point $c \in A \cap B$ such that the following hold for the DRA sequences (48):*

(i) $\mathrm{ri}(A) \cap \mathrm{ri}(B) \neq \varnothing$ *and hence* $(z_n)_{n\in\mathbb{N}}$, $(a_n)_{n\in\mathbb{N}}$ *and* $(r_n)_{n\in\mathbb{N}}$ *converge linearly to* c.
(ii) If $c \in \mathrm{int}\ B$, *then the convergence of* $(z_n)_{n\in\mathbb{N}}$, $(a_n)_{n\in\mathbb{N}}$ *and* $(r_n)_{n\in\mathbb{N}}$ *to* c *is finite.*

Proof (i) From $0 \in \mathrm{int}(A - B)$, we have $N_{A-B}(0) = \{0\}$ due to Proposition 2.10. This gives $\mathrm{Fix}\, T = A \cap B$, and Fact 3.1 implies $z_n \to c \in A \cap B$ and $a_n \to P_A c = c$. Then $r_n = 2a_n - z_n \to c$. Using again $0 \in \mathrm{int}(A - B)$, [22, Corollary 6.6.2] yields $0 \in \mathrm{ri}(A - B) = \mathrm{ri}\ A - \mathrm{ri}\ B$, and thus $\mathrm{ri}\ A \cap \mathrm{ri}\ B \neq \varnothing$. Now the linear convergence follows from [21, Theorem 4.14] or from [6, Theorem 8.5(i)].
(ii) Since $r_n \to c$ and $a_n \to c$ by (i), there exists $n \in \mathbb{N}$ such that $r_n \in B$ and $a_n \in B$. Then $P_B r_n = r_n$ and

$$z_{n+1} = z_n - a_n + P_B r_n = z_n - a_n + r_n = a_n \in A \cap B = \mathrm{Fix}\, T. \tag{50}$$

Hence $a_n = z_{n+1} = z_{n+2} = \cdots$ and we are done. □

Lemma 3.3 *Suppose that* A *is a linear subspace. Then the DRA sequences (48) satisfy*

$$a_n = P_A z_n = P_A r_n \quad \text{and} \quad a_{n+1} = P_A T z_n = P_A P_B R_A z_n = P_A P_B r_n, \tag{51}$$

and

$$(\forall n \in \mathbb{N}) \quad a_n - a_{n+1} = P_A(r_n - P_B r_n). \tag{52}$$

Proof (51): Clear from (i). (52): Use (51) and the linearity of P_A. □

Lemma 3.4 *Suppose that* A *is a linear subspace and that for the DRA sequences (48) there exists* $p \in \mathbb{N}$ *such that* $a_p = a_{p+1} = c \in A \cap B$, *and that there is a subset* N *of* X *such that* $r_p - P_B r_p \in N$ *and* $A^\perp \cap N = \{0\}$. *Then* $(\forall n \geq p + 1)\ z_n = c$.

Proof Since $a_p - a_{p+1} = 0$, (52) implies $r_p - P_B r_p \in A^\perp \cap N = \{0\}$. Thus $P_B r_p = r_p$ and therefore $z_{p+1} = z_p - a_p + r_p = a_p = c \in A \cap B \subseteq \mathrm{Fix}\, T$. □

Lemma 3.5 *Suppose that* A *is a linear subspace and let* $a \in A$. *Then the DRA sequence* (48a) *satisfies*

$$(\forall n \in \mathbb{N}) \quad \langle z_n - z_{n+1}, z_{n+1} - a\rangle = \langle r_n - P_B r_n, P_B r_n - a\rangle, \tag{53a}$$

and

$$(\forall n \in \mathbb{N})(\forall m \in \mathbb{N}) \quad \langle z_{n+1} - z_{m+1}, (z_n - z_{n+1}) - (z_m - z_{m+1})\rangle \geq 0. \tag{53b}$$

Proof Let $n \in \mathbb{N}$. Using (51), we find that

$$z_n - z_{n+1} = a_n - P_B r_n = P_A r_n - P_B r_n = P_A(r_n - P_B r_n) - P_{A^\perp}(P_B r_n - a). \tag{54}$$

Next, (44) yields

$$z_{n+1} - a = P_A P_B r_n - P_{A^\perp}(r_n - P_B r_n) - a = P_A(P_B r_n - a) - P_{A^\perp}(r_n - P_B r_n). \tag{55}$$

Moreover, $r_n - P_B r_n \in N_B(P_B r_n) = N_{B-a}(P_B r_n - a)$ and N_{B-a} is a monotone operator. The result thus follows from Corollary 2.14. □

Lemma 3.6 *Suppose that* A *is a linear subspace and that the DRA sequences (48) satisfy*

$$z_n \to a \in A \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad \langle r_n - P_B r_n, a - P_B r_n\rangle = 0. \tag{56}$$

Then there is no linear functional $f : X \to \mathbb{R}$ *such that*

$$(\forall n \in \mathbb{N}) \quad f(z_n) > f(z_{n+1}). \tag{57}$$

Proof Suppose to the contrary that there exists a linear function $f: X \to \mathbb{R}$ satisfying (57). Now set

$$(\forall n \in \mathbb{N}\smallsetminus\{0\}) \quad y_n := z_{n-1} - z_n \quad \text{and} \quad w_n := z_n - a. \tag{58}$$

On the one hand, Lemma 2.4 yields $n_0 \in \mathbb{N}\smallsetminus\{0\}$ and $(\mu_1, \ldots, \mu_{n_0}) \in \mathbb{R}_+^{n_0}$ such that

$$\sum_{k=1}^{n_0} \mu_k = 1 \quad \text{and} \quad \left\langle \sum_{k=1}^{n_0} \mu_k y_k, \sum_{k=1}^{n_0} \mu_k w_k \right\rangle > 0. \tag{59}$$

On the other hand, Lemma 3.5 and (56) yield

$$(\forall k \in \mathbb{N}\smallsetminus\{0\})(\forall j \in \mathbb{N}\smallsetminus\{0\}) \quad \langle y_k, w_k\rangle = 0 \quad \text{and} \quad \left\langle y_k - y_j, w_k - w_j\right\rangle \geq 0; \tag{60}$$

consequently, with the help of [2, Lemma 2.13(i)],

$$\left\langle \sum_{k=1}^{n_0} \mu_k y_k, \sum_{k=1}^{n_0} \mu_k w_k \right\rangle = \sum_{k=1}^{n_0} \mu_k \langle y_k, w_k\rangle - \tfrac{1}{2} \sum_{k=1}^{n_0} \sum_{j=1}^{n_0} \mu_k \mu_j \left\langle y_k - y_j, w_k - w_j\right\rangle \leq 0. \tag{61}$$

Comparing (59) with (61), we arrive at the desired contradiction. □

We are now ready for our first main result concerning the finite convergence of the DRA.

Theorem 3.7 (Finite convergence of DRA in the affine-polyhedral case) *Suppose that A is an affine subspace, that B is polyhedral at every point in $A \cap \operatorname{bdry} B$, and that* Slater's condition

$$A \cap \operatorname{int} B \neq \varnothing \tag{62}$$

holds. Then the DRA sequences (48) converge in finitely many steps *to a point in $A \cap B$.*

Proof After translating the sets if necessary, we can and do assume that A is a linear subspace of X. By Corollary 2.11(ii), (62) yields

$$0 \in \operatorname{int}(A - B) \quad \text{and} \quad \operatorname{int} B \neq \varnothing. \tag{63}$$

Lemma 3.2(i) thus implies that $(z_n)_{n\in\mathbb{N}}$, $(a_n)_{n\in\mathbb{N}}$ and $(r_n)_{n\in\mathbb{N}}$ converge linearly to a point $c \in A \cap B$. Since P_B is (firmly) nonexpansive, it also follows that $(P_B r_n)_{n\in\mathbb{N}}$ converges linearly to c. Since B is clearly polyhedral at every point in $\operatorname{int} B$ it follows from the hypothesis that B is polyhedral at c. Lemma 2.9 guarantees the existence of $n_0 \in \mathbb{N}$ such that

$$(\forall n \geq n_0) \quad \langle r_n - P_B r_n, c - P_B r_n\rangle = 0, \tag{64}$$

and

$$(\forall n \geq n_0) \quad r_n - P_B r_n \in N_B(c). \tag{65}$$

Because of $\operatorname{int} B \neq \varnothing$ and $c \in B$, Proposition 2.10(iii) yields $N_B(c) \cap (-N_B(c)) = \{0\}$. Hence $N_B(c)$ is a nonempty pointed closed convex cone. Using again $0 \in \operatorname{int}(A - B)$, Corollary 2.11(i) gives

$$\ker P_A \cap N_B(c) = A^\perp \cap N_B(c) = \{0\}. \tag{66}$$

In view of Lemma 2.5(ii),

$$K := P_A\big(N_B(c)\big) \text{ is a nonempty pointed closed convex cone.} \tag{67}$$

Combining (52) and (65), we obtain

$$(\forall n \geq n_0) \quad a_n - a_{n+1} \in K. \tag{68}$$

Since $a_n \to c$ and K is a closed convex cone, we have

$$(\forall n \geq n_0) \quad a_n - c = \sum_{k=n}^{\infty}(a_k - a_{k+1}) \in K. \tag{69}$$

We now consider two cases.

Case 1: $(\exists\, p \geq n_0)\ a_p = c$.

Using (67) and (68), we deduce from Lemma 2.6 that $a_p = a_{p+1} = c$. Now (65), (66), and Lemma 3.4 yield $(\forall n \geq p+1)\ z_n = c$ as required.

Case 2: $(\forall n \geq n_0)\ a_n \neq c$.

By (69), $(\forall n \geq n_0)\ a_n - c \in K \smallsetminus \{0\}$. Since K is pointed [see (67)], Lemma 2.3 yields $v \in \operatorname{ri} K \cap \operatorname{ri} K^{\oplus} \subseteq K$ such that

$$(\forall n \geq n_0) \quad \langle v, a_n - c\rangle > 0. \tag{70}$$

Recalling (67), we get $u \in N_B(c)$ such that $v = P_A u$. Clearly, $(\forall n \geq n_0)\ \langle u, a_n - c\rangle = \langle u, P_A(a_n - c)\rangle = \langle P_A u, a_n - c\rangle = \langle v, a_n - c\rangle$. It follows from (70) that

$$(\forall n \geq n_0) \quad \langle u, a_n - c\rangle > 0. \tag{71}$$

Since $u \in N_B(c)$ we also have

$$(\forall n \geq n_0) \quad \langle u, P_B r_n - c\rangle \leq 0. \tag{72}$$

Now define a linear functional on X by

$$f\colon X \to \mathbb{R}\colon x \mapsto \langle u, x\rangle\,. \tag{73}$$

In view of (71) with (72), we obtain $(\forall n \geq n_0)\ f(z_n - z_{n+1}) = f(a_n - P_B r_n) = f(a_n - c) - f(P_B r_n - c) = \langle u, a_n - c\rangle - \langle u, P_B r_n - c\rangle > 0$. Therefore,

$$(\forall n \geq n_0) \quad f(z_n) > f(z_{n+1}). \tag{74}$$

However, this and (56) together contradict Lemma 3.6 (applied to $(z_n)_{n \geq n_0}$). We deduce that *Case 2* never occurs which completes the proof of the theorem. □

Remark 3.8 Some comments on Theorem 3.7 are in order.

(i) Suppose we replace the Slater's condition "$A \cap \operatorname{int} B \neq \varnothing$" by "$A \cap \operatorname{ri} B \neq \varnothing$". Then one still obtains *linear convergence* (see [21, Theorem 4.14] or [6, Theorem 8.5(i)]); however, *finite convergence fails* in general: indeed, B can chosen to be an affine subspace (see [1, Section 5]).

(ii) Under the assumptions of Theorem 3.7, Rockafellar's condition (37) is only applicable when both $A = \{\bar{a}\}$ and $\bar{a} \in \operatorname{int} B$ hold. There are many examples where this condition is violated yet our condition is applicable (see, e.g., the scenario in the following item).

(iii) Suppose that $X = \mathbb{R}^2$, that $A = \mathbb{R} \times \{0\}$ and that $B = \operatorname{epi} f$, where $f\colon \mathbb{R} \to \mathbb{R}\colon x \mapsto |x| - 1$. It is clear that B is polyhedral and that $A \cap \operatorname{int} B \neq \varnothing$. Define $(\forall \varepsilon \in \mathbb{R}_{++})$ $z_\varepsilon := (1+\varepsilon, \varepsilon)$. Then

$$(\forall \varepsilon \in \mathbb{R}_{++}) \quad T z_\varepsilon = (1, \varepsilon) \notin \operatorname{Fix} T = [-1, 1] \times \{0\} \ \text{ yet } \ \|z_\varepsilon - T z_\varepsilon\| = \varepsilon. \tag{75}$$

We conclude that (the resolvent reformulation of) Luque's condition (41) fails.

4 Applications

4.1 Product space setup and Spingarn's method

Let us now consider a feasibility problem with possibly more than two sets, say

$$\text{find a point in } C := \bigcap_{j=1}^{M} C_j, \tag{76}$$

where

$$C_1, \dots, C_M \text{ are closed convex subsets of } X \text{ such that } C \neq \varnothing. \tag{77}$$

This problem is reduced to a two-set problem as follows. In the product Hilbert space $\mathbf{X} := X^M$, with the inner product defined by $((x_1, \dots, x_M), (y_1, \dots, y_M)) \mapsto \sum_{j=1}^{M} \langle x_j, y_j \rangle$, we set

$$\mathbf{A} := \big\{(x, \dots, x) \in \mathbf{X} \bigm| x \in X\big\} \quad \text{and} \quad \mathbf{B} := C_1 \times \cdots \times C_M. \tag{78}$$

Because of

$$x \in \bigcap_{j=1}^{M} C_j \quad \Leftrightarrow \quad (x, \dots, x) \in \mathbf{A} \cap \mathbf{B}, \tag{79}$$

the M-set problem (76), which is formulated in X, is equivalent to the two-set problem

$$\text{find a point in } \mathbf{A} \cap \mathbf{B}, \tag{80}$$

which is posed in $\mathbf{X}$. By, e.g., [2, Proposition 25.4(iii) and Proposition 28.3], the projections of $\mathbf{x} = (x_1, \dots, x_M) \in \mathbf{X}$ onto $\mathbf{A}$ and $\mathbf{B}$ are respectively given by

$$P_{\mathbf{A}}\mathbf{x} = \Big(\frac{1}{M}\sum_{j=1}^{M} x_j, \dots, \frac{1}{M}\sum_{j=1}^{M} x_j\Big) \quad \text{and} \quad P_{\mathbf{B}}\mathbf{x} = \big(P_{C_1}x_1, \dots, P_{C_M}x_M\big). \tag{81}$$

This opens the door of applying the DRA in $\mathbf{X}$: indeed, set

$$\mathbf{T} := \operatorname{Id} - P_{\mathbf{A}} + P_{\mathbf{B}} R_{\mathbf{A}}, \tag{82}$$

fix a starting point $\mathbf{z}_0 \in \mathbf{X}$, and generate the DRA sequence $(\mathbf{z}_n)_{n\in\mathbb{N}}$ via

$$(\forall n \in \mathbb{N}) \quad \mathbf{z}_{n+1} := \mathbf{T}\mathbf{z}_n. \tag{83}$$

We now obtain the following result as a consequence of Theorem 3.7.

Corollary 4.1 *Suppose that* $C_1, \dots, C_M$ *are polyhedral such that* $\operatorname{int} C \neq \varnothing$. *Then the DRA sequence defined by* (83) *converges finitely to* $\mathbf{z} = (z, \dots, z) \in \mathbf{A} \cap \mathbf{B}$ *with* $z \in C = \bigcap_{j=1}^{M} C_j$.

Proof Since $\operatorname{int} C \neq \varnothing$, there exists $c \in C$ and $\varepsilon \in \mathbb{R}_{++}$ such that $(\forall j \in \{1, \dots, M\})$ $\operatorname{ball}(c; \varepsilon) \subseteq C_j$. Then $(c, \dots, c) \in \mathbf{A} \cap \operatorname{int} \mathbf{B}$, and so $\mathbf{A} \cap \operatorname{int} \mathbf{B} \neq \varnothing$. Since $\mathbf{B} = C_1 \times \cdots \times C_M$ is a polyhedral subset of $\mathbf{X}$, the conclusion now follows from Theorem 3.7. □

Problem (76) was already considered by Spingarn [26] for the case where all sets $C_1, \dots, C_M$ are *halfspaces*. He cast the resulting problem into the form

$$\text{find } (\mathbf{a}, \mathbf{b}) \in \mathbf{A} \times \mathbf{A}^{\perp} \text{ such that } \mathbf{b} \in N_{\mathbf{B}}(\mathbf{a}), \tag{84}$$

and suggested solving it by a version of his *method of partial inverses* [25], which generates a sequence $(\mathbf{a}_n, \mathbf{b}_n)_{n\in\mathbb{N}}$ via

$$(\mathbf{a}_0, \mathbf{b}_0) \in \mathbf{A} \times \mathbf{A}^{\perp} \quad \text{and} \quad (\forall n \in \mathbb{N}) \quad \begin{cases} \mathbf{a}'_n := P_{\mathbf{B}}(\mathbf{a}_n + \mathbf{b}_n), & \mathbf{b}'_n := \mathbf{a}_n + \mathbf{b}_n - \mathbf{a}'_n, \\ \mathbf{a}_{n+1} := P_{\mathbf{A}}\mathbf{a}'_n, & \mathbf{b}_{n+1} := \mathbf{b}'_n - P_{\mathbf{A}}\mathbf{b}'_n. \end{cases} \tag{85}$$

It was pointed out in [15, Section 1], [12, Section 5], [19, Appendix] and [10, Remark 2.4] that this scheme is closely related to the DRA. (In fact, Lemma 2.17 above makes it completely clear why (85) is equivalent to applying the DRA to $\mathbf{A}$ and $\mathbf{B}$, with starting point $(\mathbf{a}_0 - \mathbf{b}_0)$.) However, Spingarn's proof of finite convergence in [26] requires

$$\mathbf{a}_n - \mathbf{a}_{n+1} \in N_C(c) \times \cdots \times N_C(c), \tag{86}$$

and he chooses a linear functional f based on the "diagonal" structure of $\mathbf{A}$—unfortunately, his proof does not work for problem (3). Our proof in the previous section at the same time simplifies and strengthens his proof technique to allow us to deal with polyhedral sets rather than just halfspaces. While every polyhedron is an intersection of halfspaces, the problems are theoretically equivalent—in practice, however, there can be *huge savings* as the requirement to work in Spingarn's setup might lead to much larger instances of the product space $\mathbf{X}$! It also liberates us from being forced to work in the product space. Our extension is also intrinsically more flexible as the following example illustrates.

Example 4.2 Suppose that $X = \mathbb{R}^2$, that $A = \{(x, x) \mid x \in \mathbb{R}\}$ is diagonal, and that $B = \{(x, y) \in \mathbb{R}^2 \mid -y \le x \le 2\}$. Clearly, B is polyhedral and $A \cap \operatorname{int} B \neq \varnothing$. Moreover, B is also *not* the Cartesian product of two polyhedral subsets of $\mathbb{R}$, i.e., of two intervals. Therefore, the proof of finite convergence of the DRA in [26] no longer applies. However, A and B satisfy all assumptions of Theorem 3.7, and thus the DRA finds a point in $A \cap B$ after a finite number of steps (regardless of the location of starting point). See Fig. 1 for an illustration, created with `GeoGebra` [13].

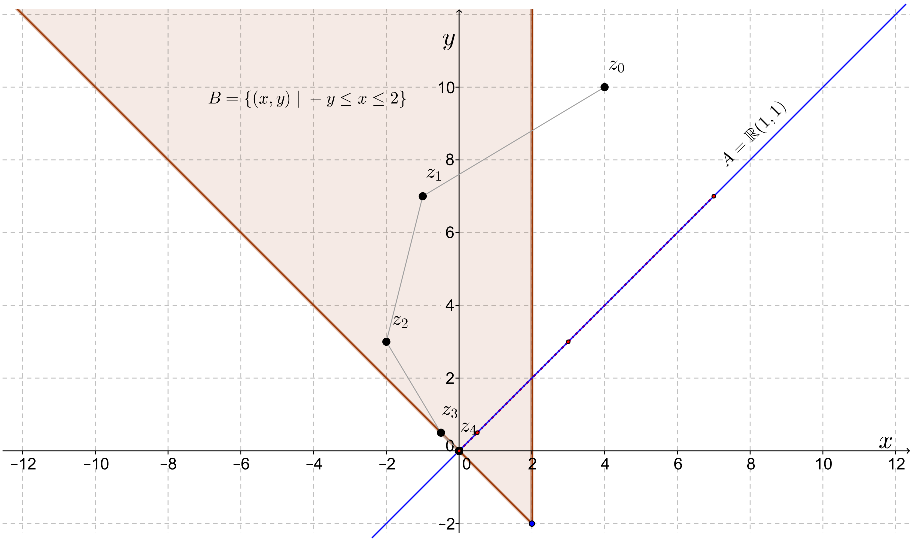

Fig. 1 The DRA for the case when B is not a Cartesian product

4.2 Solving linear equalities with a strict positivity constraint

In this subsection, we assume that

$$X = \mathbb{R}^N, \tag{87}$$

and that

$$A = \{x \in X \mid Lx = a\} \quad \text{and} \quad B = \mathbb{R}^N_+, \tag{88}$$

where $L \in \mathbb{R}^{M\times N}$ and $a \in \mathbb{R}^M$. Note that the set $A \cap B$ is polyhedral yet $A \cap B$ has empty interior (unless $A = X$ which is a case of little interest). Thus, Spingarn's finite-convergence result is *never applicable*. However, Theorem 3.7 guarantees finite convergence of the DRA provided that Slater's condition

$$\mathbb{R}^N_{++} \cap L^{-1}a \neq \varnothing \tag{89}$$

holds. Using [5, Lemma 4.1], we obtain

$$P_A : X \to X : x \mapsto x - L^\dagger(Lx - a), \tag{90}$$

where $L^\dagger$ denotes the Moore-Penrose inverse of L. We also have (see, e.g., [2, Example 6.28])

$$P_B : X \to X : x = (\xi_1, \dots, \xi_N) \mapsto x = (\xi_1^+, \dots, \xi_N^+), \tag{91}$$

where $\xi^+ = \max\{\xi, 0\}$ for every $\xi \in \mathbb{R}$. This implies $R_A : X \to X : x \mapsto x - 2L^\dagger(Lx - a)$ and $R_B : X \to X : x = (\xi_1, \dots, \xi_N) \mapsto (|\xi_1|, \dots, |\xi_N|)$. We will compare three algorithms, all of which generate a governing sequence with starting point $z_0 \in X$ via

$$(\forall n \in \mathbb{N}) \quad z_{n+1} = Tz_n. \tag{92}$$

The DRA uses, of course,

$$T = \operatorname{Id} - P_A + P_B R_A. \tag{93}$$

The second method is the classical *method of alternating projections (MAP)* where

$$T = P_A P_B; \tag{94}$$

while the third *method of reflection–projection (MRP)* employs

$$T = P_A R_B. \tag{95}$$

We now illustrate the performance of these three algorithms numerically. For the remainder of this section, we assume that that $X = \mathbb{R}^2$, that $A = \{(x, y) \in \mathbb{R}^2 \mid x + 5y = 6\}$, and that $B = \mathbb{R}^2_+$. Then A and B satisfy (89), and the sequence (93) generated by the DRA thus converges finitely to a point in $A \cap B$ regardless of the starting point. See Fig. 2 for an illustration, created with `GeoGebra` [13]. Note that the shadow sequence $(P_A z_n)_{n\in\mathbb{N}}$ for the DRA finds a point in $A \cap B$ even before a fixed point is reached. For each starting point $z_0 \in [-100, 100]^2 \subseteq \mathbb{R}^2$, we perform the DRA until $z_{n+1} = z_n$, and run the MAP and the MRP until $d_B(z_n) = \max\{d_A(z_n), d_B(z_n)\} < \varepsilon$, where we set the tolerance $\varepsilon = 10^{-4}$. Figure 3 compares the number of iterations needed to stop each algorithm. Note that even though we put the DRA at an "unfair disadvantage" (it must find a true fixed point while the MAP and the MRP will stop with ε-feasible solutions), it does extremely well. In Fig. 4, we level the playing field and compare the distance from $P_A z_n$ (for the DRA) or from z_n (for the MAP and the MRP) to B, where $n \in \{5, 10\}$.

Now we look at the process of reaching a solution for each algorithm. For the DRA, we monitor the shadow sequence $(P_A z_n)_{n\in\mathbb{N}}$ and for the MAP and the MRP, we monitor $(z_n)_{n\in\mathbb{N}}$,

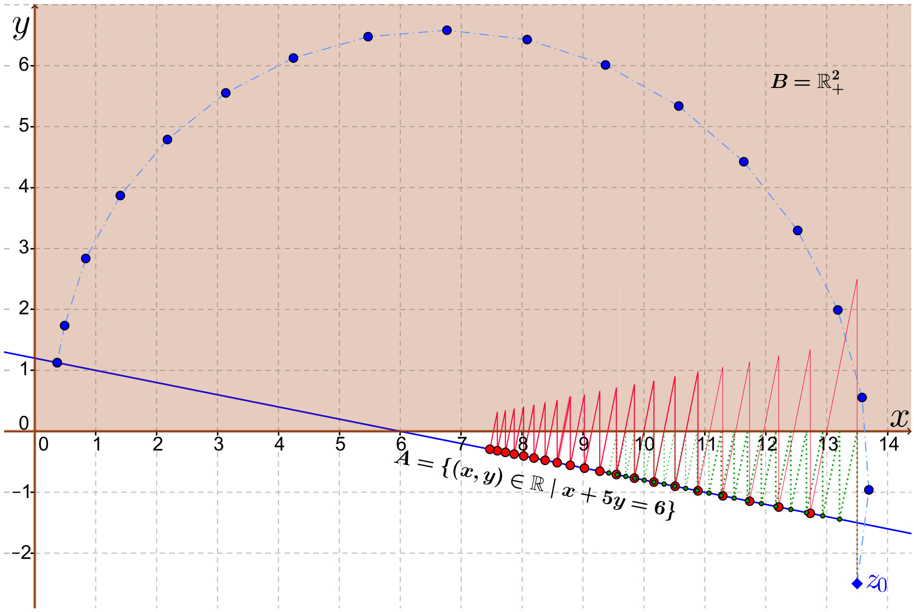

Fig. 2 The orbits of DRA (*dashed*), MAP (*dotted*) and MRP (*solid*)

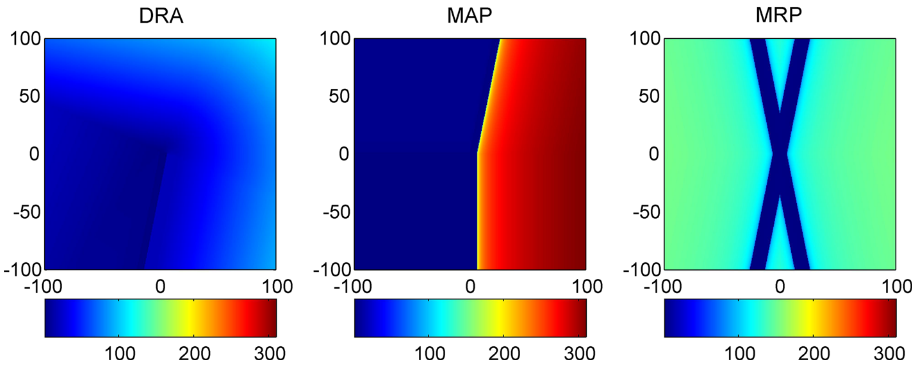

Fig. 3 Number of iterations needed to get the solution of DRA, MAP and MRP

which is actually the same as $(P_A z_n)_{n \in \mathbb{N}}$. Because all three monitored sequences lie in A, we are concerned about the distance to B. Our stopping criterion is therefore

$$d_B(P_A z_n) < \varepsilon \tag{96}$$

for all three algorithms. From top to bottom in Fig. 5, we check how many iterates are required to get to tolerance $\varepsilon = 10^{-m}$, where $m \in \{2, 4\}$, respectively. Computations were performed with MATLAB R2013b [18]. These experiments illustrate the superior convergence behaviour of the DRA compared to the MAP and the MRP.

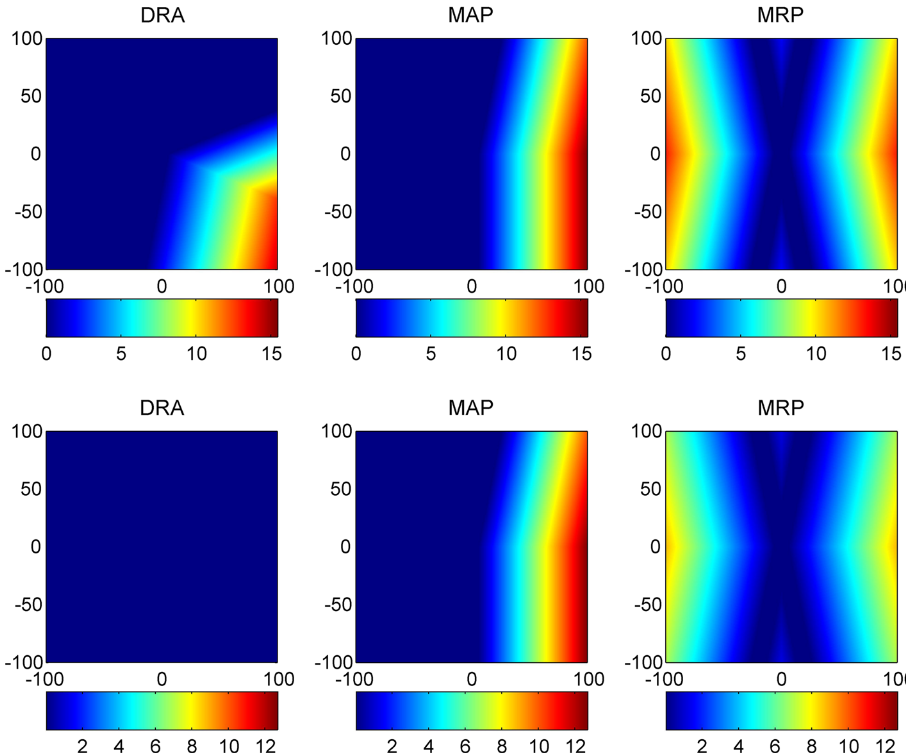

Fig. 4 The distance from the monitored iterate to B after 5 steps (*top*) and 10 steps (*bottom*)

5 The hyperplanar-epigraphical case with Slater's condition

In this section we assume that

$$f\colon X \to \mathbb{R} \text{ is convex and continuous.} \tag{97a}$$

We will work in $X \times \mathbb{R}$, where we set

$$A := X \times \{0\} \tag{97b}$$

and

$$B := \operatorname{epi} f := \big\{(x,\rho) \in X \times \mathbb{R} \;\big|\; f(x) \le \rho\big\}. \tag{97c}$$

Then

$$P_A\colon X \times \mathbb{R} \to X \times \mathbb{R}\colon (x,\rho) \mapsto (x,0), \tag{98}$$

and the projection onto B is described in the following result.

Lemma 5.1 *Let* $(x,\rho) \in (X \times \mathbb{R}) \smallsetminus B$. *Then there exists* $p \in X$ *such that* $P_B(x,\rho) = (p, f(p))$,

$$x \in p + \big(f(p) - \rho\big)\partial f(p) \text{ and } \rho < f(p) \le f(x) \tag{99}$$

and

$$(\forall y \in X) \quad \langle y - p, x - p\rangle \le \big(f(y) - f(p)\big)\big(f(p) - \rho\big). \tag{100}$$

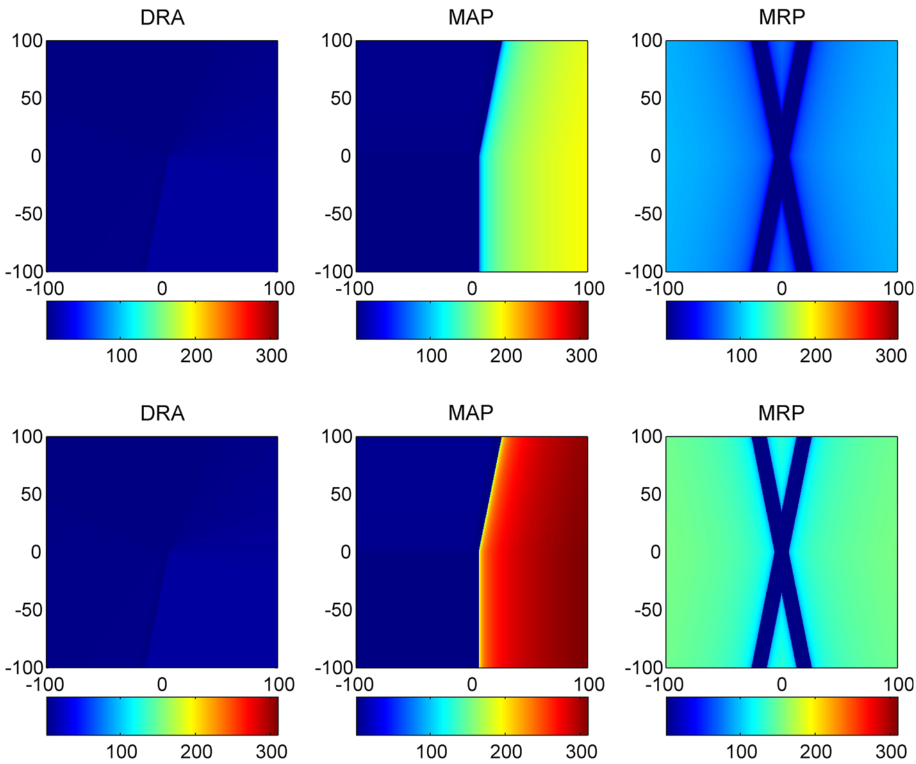

Fig. 5 Number of iterations needed to get the tolerance 10^{-2} (*top*) and 10^{-4} (*bottom*) of DRA, MAP and MRP

Moreover, the following hold:

(i) If u is a minimizer of f, then $\langle u - p, x - p\rangle \leq 0$.
(ii) If x is a minimizer of f, then $p = x$.
(iii) If x is not a minimizer of f, then $f(p) < f(x)$.

Proof According to [2, Proposition 28.28], there exists $(p, p^*) \in \operatorname{gra} \partial f$ such that $P_B(x, \rho) = (p, f(p))$ and $x = p + (f(p) - \rho)p^*$. Next, [2, Propositions 9.18] implies that $\rho < f(p)$ and (100). The subgradient inequality gives

$$f(x) - f(p) \geq \left\langle p^*, x - p\right\rangle = \left\langle p^*, (f(p) - \rho)p^*\right\rangle = (f(p) - \rho)\|p^*\|^2 \geq 0. \qquad (101)$$

Hence $f(p) \leq f(x)$, and this completes the proof of (99),

(i) It follows from (100) that $\langle u - p, x - p\rangle \leq (f(u) - f(p))(f(p) - \rho)$. Since $f(p) - \rho > 0$ and $f(u) \leq f(p)$, we have $\langle u - p, x - p\rangle \leq 0$.
(ii) Apply (i) with $u = x$.
(iii) By (99), $\rho < f(p) \leq f(x)$. We show the contrapositive and thus assume that $f(p) = f(x)$. Since $f(p) - \rho > 0$, (101) yields $p^* = 0$, and so $0 \in \partial f(p)$. It follows that p is a minimizer of f, and therefore x also minimizes f.

□

Remark 5.2 When $X = \mathbb{R}$ and u is a minimizer of f, then $(u - p)(x - p) \leq 0$ by Lemma 5.1(i). Therefore, p lies between x and u (see also [4, Corollary 4.2]).

Define the DRA operator by

$$T := \mathrm{Id} - P_A + P_B R_A. \tag{102}$$

It will be convenient to abbreviate

$$B' := R_A(B) = \big\{(x, \rho) \in X \times \mathbb{R} \;\big|\; \rho \le -f(x)\big\} \tag{103}$$

and to analyze the effect of performing one DRA step in the following result.

Corollary 5.3 *(One DRA step) Let $z = (x, \rho) \in X \times \mathbb{R}$, and set $z_+ := (x_+, \rho_+) = T(x, \rho)$. Then the following hold:*

(i) Suppose that $z \in B'$. Then $z_+ = (x, 0) \in A$. Moreover, either ($f(x) \le 0$ and $z_+ \in A \cap B$) or ($f(x) > 0$ and $z_+ \notin B \cup B'$).
(ii) Suppose that $z \notin B'$. Then there exists $x_+^* \in \partial f(x_+)$ such that

$$x_+ = x - \rho_+ x_+^*, \;\; f(x_+) \le f(x), \;\; \textit{and } \rho_+ = \rho + f(x_+) > 0. \tag{104}$$

Moreover, either ($\rho \ge 0$ and $z_+ \in B$) or ($\rho < 0$, $z_+ \notin B \cup B'$ and $Tz_+ \in B$).
(iii) Suppose that $z \in B \cap B'$. Then $z_+ \in A \cap B$.
(iv) Suppose that $z \in B \smallsetminus B'$. Then $z_+ \in B$.

Proof (i) We have $P_A z = (x, 0)$ and $R_A z = (x, -\rho) \in B$. Thus $P_B R_A z = P_B(x, -\rho) = (x, -\rho)$, which gives

$$z_+ = (\mathrm{Id} - P_A + P_B R_A)z = (x, \rho) - (x, 0) + (x, -\rho) = (x, 0) \in A. \tag{105}$$

If $f(x) \le 0$, then $z_+ = (x, 0) \in B$, and hence $z_+ \in A \cap B$. Otherwise, $f(x) > 0$ which implies $-f(x) < 0$ and further $z_+ = (x, 0) \notin B \cup B'$.

(ii) We have $R_A z = (x, -\rho) \notin B$, and by (99), $P_B(x, -\rho) = (p, f(p))$, where

$$p = x - (\rho + f(p))p^* \text{ for some } p^* \in \partial f(p), \quad \text{and} \quad -\rho < f(p) \le f(x). \tag{106}$$

We obtain

$$z_+ = (x_+, \rho_+) = (\mathrm{Id} - P_A + P_B R_A)z = (x, \rho) - (x, 0) + (p, f(p)) = (p, \rho + f(p)), \tag{107}$$

which gives $x_+ = p$ and $\rho_+ = \rho + f(p) > 0$, so (104) holds. If $\rho \ge 0$, then $\rho_+ = \rho + f(x_+) \ge f(x_+)$, and thus $z_+ \in B$. Otherwise, $\rho < 0$, so $\rho_+ < f(x_+)$ and also $f(x_+) > -\rho > 0$. Hence $\rho_+ > 0 > -f(x_+)$, which implies $z_+ \notin B \cup B'$, and then $Tz_+ \in B$ because $\rho_+ > 0$ and the previous case applies.

(iii) We have $f(x) \le \rho \le -f(x)$, and so $f(x) \le 0$. Now apply (i).

(iv) We have $\rho \ge f(x)$ and $\rho > -f(x)$. Then $\rho \ge |f(x)| \ge 0$, and (ii) gives $z_+ \in B$.

□

Theorem 5.4 *(Finite convergence of DRA in the hyperplanar-epigraphical case) Suppose that $\inf_X f < 0$, and, given a starting point $z_0 = (x_0, \rho_0) \in X \times \mathbb{R}$, generate the DRA sequence $(z_n)_{n \in \mathbb{N}}$ by*

$$(\forall n \in \mathbb{N}) \quad z_{n+1} = (x_{n+1}, \rho_{n+1}) = Tz_n. \tag{108}$$

Then $(z_n)_{n \in \mathbb{N}}$ converges finitely to a point $z \in A \cap B$.

Proof In view of Corollary 5.3(i)&(iii), we can and do assume that $z_0 \in B \smallsetminus B'$, where B' was defined in (103). It follows then from Corollary 5.3(iii)&(iv), that $(z_n)_{n\in\mathbb{N}}$ lies in B.

Case 1: $(\exists n \in \mathbb{N})\ z_n \in B \cap B'$.

By Corollary 5.3(iii), $z_{n+1} \in A \cap B$ and we are done.

Case 2: $(\forall n \in \mathbb{N})\ z_n \in B \smallsetminus B'$.

By Corollary 5.3(ii),

$$(\forall n \in \mathbb{N}) \quad f(x_{n+1}) \leq f(x_n) \ \text{ and } \ \rho_{n+1} = \rho_n + f(x_{n+1}) > 0. \tag{109}$$

Next, it follows from [22, Lemma 7.3] that $\operatorname{int} B = \{(x,\rho) \in X \times \mathbb{R} \mid f(x) < \rho\}$. Because $\inf_X f < 0$, we obtain $A \cap \operatorname{int} B \neq \varnothing$, which, due to Lemma 3.2(i) yields $z_n \to z = (x,\rho) \in A \cap B$. Since $z \in A$, we must have $\rho = 0$. If $(\forall n \in \mathbb{N})\ f(x_{n+1}) \geq 0$, then, by (109), $0 < \rho_1 \leq \rho_2 \leq \cdots \leq \rho_n \to \rho = 0$ which is absurd. Therefore,

$$(\exists\, n_0 \in \mathbb{N}) \quad f(x_{n_0+1}) < 0. \tag{110}$$

In view of (109), we see that $(\forall n \geq n_0+1)\ \rho_n \leq \rho_{n_0} + (n-n_0) f(x_{n_0+1})$. Since $f(x_{n_0+1}) < 0$, there exists $n_1 \in \mathbb{N}$, $n_1 \geq n_0 + 1$ such that

$$\rho_{n_0} + (n_1 - n_0) f(x_{n_0+1}) \leq -f(x_{n_0+1}). \tag{111}$$

Noting that $f(x_{n_1}) \leq f(x_{n_0+1})$, we then obtain

$$\rho_{n_1} \leq \rho_{n_0} + (n_1 - n_0) f(x_{n_0+1}) \leq -f(x_{n_0+1}) \leq -f(x_{n_1}). \tag{112}$$

Hence $z_{n_1} \in B'$, which contradicts the assumption of *Case 2*. Therefore, *Case 2* never occurs and the proof is complete. ◻

We conclude by illustrating that finite convergence may be deduced from Theorem 5.4 but not necessarily from the finite-convergence conditions of Sect. 2.6.

Example 5.5 Suppose that $X = \mathbb{R}$, that $A = \mathbb{R} \times \{0\}$, and that $B = \operatorname{epi} f$, where $f : \mathbb{R} \to \mathbb{R} : x \mapsto x^2 - 1$. Let $(\forall \varepsilon \in \mathbb{R}_{++})\ z_\varepsilon = (1+\varepsilon, -\varepsilon)$. Then $(\forall \varepsilon \in \mathbb{R}_{++})\ Tz_\varepsilon \notin \operatorname{Fix} T = [-1,1] \times \{0\}$, and $z_\varepsilon - Tz_\varepsilon \to 0$ as $\varepsilon \to 0^+$. Consequently, Luque's condition (41) fails.

Proof Let $\varepsilon \in \mathbb{R}_{++}$. Then $-f(1+\varepsilon) = -2\varepsilon - \varepsilon^2 < -\varepsilon$, and so $z_\varepsilon = (1+\varepsilon, -\varepsilon) \notin B'$. By Corollary 5.3(ii), there exists $x_\varepsilon \in \mathbb{R}$ such that

$$Tz_\varepsilon = (x_\varepsilon, -\varepsilon + f(x_\varepsilon)), \quad x_\varepsilon = 1+\varepsilon-(-\varepsilon+f(x_\varepsilon))2x_\varepsilon, \quad \text{and} \quad -\varepsilon + f(x_\varepsilon) > 0. \tag{113}$$

The last inequality shows that $Tz_\varepsilon \notin \operatorname{Fix} T = [-1,1] \times \{0\}$. It follows from the expression of x_ε that

$$2x_\varepsilon^3 - (1+2\varepsilon)x_\varepsilon - 1 - \varepsilon = 0. \tag{114}$$

Note that $(x_\varepsilon, f(x_\varepsilon)) + (0,-\varepsilon) = Tz_\varepsilon = z_\varepsilon - P_A z_\varepsilon + P_B R_A z_\varepsilon = (1+\varepsilon,-\varepsilon) - (1+\varepsilon,0) + P_B(1+\varepsilon,\varepsilon) = P_B(1+\varepsilon,\varepsilon) + (0,-\varepsilon)$ and hence $(x_\varepsilon, f(x_\varepsilon)) = P_B(1+\varepsilon,\varepsilon)$. Remark 5.2 gives $0 \leq x_\varepsilon \leq 1+\varepsilon$. If $0 \leq x_\varepsilon \leq 1$, then (114) yields $0 \leq 2x_\varepsilon - (1+2\varepsilon)x_\varepsilon - 1 - \varepsilon = (x_\varepsilon - 1) - 2\varepsilon x_\varepsilon - \varepsilon < 0$, which is absurd. Thus $1 < x_\varepsilon \leq 1+\varepsilon$. Now as $\varepsilon \to 0^+$, we have $x_\varepsilon \to 1$, $f(x_\varepsilon) \to f(1) = 0$, and at the same time, $z_\varepsilon - Tz_\varepsilon = (1+\varepsilon - x_\varepsilon, -f(x_\varepsilon)) \to 0$. ◻

Acknowledgments The authors thank an anonymous referee for careful reading and constructive comments. H.H.B. was partially supported by the Natural Sciences and Engineering Research Council of Canada and by the Canada Research Chair Program. M.N.D. was partially supported by an NSERC accelerator grant of H.H.B.

References

1. Bauschke, H.H., Bello Cruz, J.Y., Nghia, T.T.A., Phan, H.M., Wang, X.: The rate of linear convergence of the Douglas–Rachford algorithm for subspaces is the cosine of the Friedrichs angle. J. Approx. Theory **185**, 63–79 (2014)
2. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Springer, Berlin (2011)
3. Bauschke, H.H., Combettes, P.L., Luke, D.R.: Finding best approximation pairs relative to two closed convex sets in Hilbert spaces. J. Approx. Theory **127**, 178–192 (2004)
4. Bauschke, H.H., Dao, M.N., Noll, D., Phan, H.M.: Proximal point algorithm, Douglas–Rachford algorithm and alternating projections: a case study. J. Convex Anal. **23** (2016)
5. Bauschke, H.H., Kruk, S.G.: Reflection–projection method for convex feasibility problems with an obtuse cone. J. Optim. Theory Appl. **120**, 503–531 (2004)
6. Bauschke, H.H., Noll, D., Phan, H.M.: Linear and strong convergence of algorithms involving averaged nonexpansive operators. J. Math. Anal. Appl. **421**, 1–20 (2015)
7. Borwein, J.M., Moors, W.B.: Stability of closedness of convex cones under linear mappings. J. Convex Anal. **16**, 699–705 (2009)
8. Cegielski, A.: Iterative Methods for Fixed Point Problems in Hilbert Spaces. Springer, Berlin (2012)
9. Censor, Y., Zenios, S.A.: Parallel Optimization. Oxford University Press, Oxford (1997)
10. Combettes, P.L.: Iterative construction of the resolvent of a sum of maximal monotone operators. J. Convex Anal. **16**, 727–748 (2009)
11. Douglas, J., Rachford, H.H.: On the numerical solution of heat conduction problems in two and three space variables. Trans. AMS **82**, 421–439 (1956)
12. Eckstein, J., Bertsekas, D.P.: On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators. Math. Program. **55**, 293–318 (1992)
13. `GeoGebra` software. http://www.geogebra.org
14. Hesse, R., Luke, D.R.: Nonconvex notions of regularity and convergence of fundamental algorithms for feasibility problems. SIAM J. Optim. **23**, 2397–2419 (2013)
15. Lawrence, J., Spingarn, J.E.: On fixed points of nonexpansive piecewise isometric mappings. Proc. Lond. Math. Soc. Third Ser. **55**, 605–624 (1987)
16. Lions, P.-L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. **16**, 964–979 (1979)
17. Luque, F.J.: Asymptotic convergence analysis of the proximal point algorithm. SIAM J. Control Optim. **22**, 277–293 (1984)
18. `MATLAB` software. http://www.mathworks.com/products/matlab/
19. Mahey, P., Oualibouch, S., Tao, P.D.: Proximal decomposition on the graph of a maximal monotone operator. SIAM J. Optim. **5**, 454–466 (1995)
20. Mordukhovich, B.S., Nam, N.M.: An easy path to convex analysis and applications. Morgan & Claypool Publishers, San Rafael (2014)
21. Phan, H.M.: Linear convergence of the Douglas–Rachford method for two closed sets. Optimization (2015). doi:10.1080/02331934.2015.1051532
22. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)
23. Rockafellar, R.T.: Monotone operators and the proximal point algorithm. SIAM J. Control Optim. **14**(5), 877–898 (1976)
24. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Springer, Berlin (1998)
25. Spingarn, J.E.: Partial inverse of a monotone operator. Appl. Math. Optim. **10**, 247–265 (1983)
26. Spingarn, J.E.: A primal–dual projection method for solving systems of linear inequalities. Linear Algebra Appl. **65**, 45–62 (1985)