

Linear convergence of the generalized Douglas–Rachford algorithm for feasibility problems

Minh N. Dao* and Hung M. Phan†

April 16, 2018

Abstract

In this paper, we study the generalized Douglas–Rachford algorithm and its cyclic variants which include many projection-type methods such as the classical Douglas–Rachford algorithm and the alternating projection algorithm. Specifically, we establish several local linear convergence results for the algorithm in solving feasibility problems with finitely many closed possibly nonconvex sets under different assumptions. Our findings not only relax some regularity conditions but also improve linear convergence rates in the literature. In the presence of convexity, the linear convergence is global.

AMS Subject Classifications: Primary: 47H10, 49M27; Secondary: 41A25, 65K05, 65K10, 90C26.

Keywords: affine-hull regularity, cyclic algorithm, generalized Douglas–Rachford algorithm, linear convergence, linear regularity, strong regularity, superregularity, quasi Fejér monotonicity, quasi coercivity.

1. Introduction

The *feasibility problem* of finding a point in the intersection of closed constraint sets is of central importance in diverse areas of mathematics and engineering. Many methods have been proposed for this problem, and most of them naturally involve *nearest point projectors* and their variants with respect to underlying sets. We refer the readers to [3] and the references therein for more reviews and discussions on projection-type methods and applications.

Among methods for feasibility problems, the *Douglas–Rachford (DR) algorithm* has recently drawn much attention due to its interesting features which mysteriously allow for its success in both convex and nonconvex settings. The DR algorithm was first formulated in [21] for solving nonlinear heat flow problems numerically. Since then, it has been emerged in optimization theory and applications thanks to the seminal work [27]; more specifically, the DR algorithm was extended to the problem of finding a zero of the sum of two maximally monotone operators. When specialized to normal cone operators, the DR algorithm can be used for solving feasibility problems.

In the convex case, the convergence theory of the DR algorithm is well developed. In particular, weak convergence of the DR sequence to a fixed point was proved in [27] while that of the shadow sequence to a solution was proved in [37]. When the problem is *infeasible*, it was shown in [6] that the shadow sequence is bounded with cluster points solving a best approximation problem, in [8] that the

*CARMA, University of Newcastle, Callaghan, NSW 2308, Australia. E-mail: daonminh@gmail.com

†Department of Mathematical Sciences, Kennedy College of Sciences, University of Massachusetts Lowell, MA 01854, USA. E-mail: hung_phan@uml.edu.

entire shadow sequence is weakly convergent if one set is an affine subspace, and in [13] that the affine subspace condition can be removed. For nonconvex problems, the DR algorithm enjoys successful applications including, notably, phase retrieval, protein folding, and Sudoku, see, e.g., [23] (in which the DR algorithm is referred to as a special case of the *difference map*) and the references therein, even though the supporting theory is far from being complete.

Regarding convergence rate analysis of the DR algorithm, various results have been obtained, some of which even apply to certain nonconvex settings. For example, [1] proved linear convergence for two subspaces, [7, 9] proved several finite convergence results, in particular, when the sets involved are subspaces, halfspaces, epigraphs of convex functions, and finite sets, while [20, 24, 32] proved the linear convergence rate under some regularity conditions, see also [14]. However, it was observed in [10] that without regularity assumptions, linear convergence of the DR algorithm may fail even for simple cases in the Euclidean plane. Recently, linear convergence was studied in [19] for the so-called *generalized Douglas–Rachford algorithm*, which is a generalization of several projection-type methods including the DR algorithm and the *alternating projection algorithm*. Although the generalized DR algorithm in fact appeared earlier in [16, Example 3.5], little is known about its rate of convergence.

The goal of this paper is to provide linear convergence results for the generalized DR algorithm and its *cyclic* variants. To obtain linear convergence, the essential assumptions employed in recent works are *superregularity* of sets, *linear regularity*, *strong regularity*, and *affine-hull regularity* of the system of constraint sets. Indeed, these concepts were previously utilized in [11, 12, 15, 22, 24, 26, 29, 31] to prove linear convergence for the alternating projection algorithm. In the literature, linear regularity and strong regularity are also known as *subtransversality* and *transversality*, respectively. Our main contributions include various improvements in this direction. Precisely, in one of the new results, we obtain *better* linear convergence rate under the *same* assumptions used in recent studies (see Theorem 5.4 and Remark 5.5). In another new result, we *lessen* the assumptions required for linear convergence by exploiting the flexibility of generalized DR operators (see Theorem 5.8 and Corollary 5.9). To the best of our knowledge, the latter is the *first* result that puts together linear regularity and operator parameters to obtain linear convergence for variants of the DR algorithm. In the convex case, we show that the linear convergence is global (see Theorem 5.11 and Corollary 5.12). On the one hand, the linear convergence results in this work can be seen as a continuation of [19]. On the other hand, it is also worth noting that the generalized DR algorithm involving at most one reflection is convergent in convex case even when the problem is infeasible (see Theorem 2.14). Moreover, we present several new features of the generalized DR algorithm, for example, the parameters may decide whether the iterations converge to the set of fixed points or to the intersection. Our results provide another step toward understanding the behavior of the celebrated DR algorithm.

The remainder of the paper is organized as follows. Section 2 contains preliminary materials needed for our analysis. While part of Section 2 was previously presented in [19], we include it here for completeness. In Sections 3 and 4, we provide some improvements and new results on *quasi firm Fejér monotonicity* and *quasi coercivity* which are key ingredients in convergence rate analysis. Section 5 presents our main results on linear convergence for the generalized DR algorithm. Finally, the conclusion is given in Section 6.

2. Auxiliary results

Unless otherwise stated, X is a Euclidean space with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. The nonnegative integers are \mathbb{N} , the real numbers are \mathbb{R} , while $\mathbb{R}_+ := \{x \in \mathbb{R} \mid x \geq 0\}$ and $\mathbb{R}_{++} := \{x \in \mathbb{R} \mid x > 0\}$. If $w \in X$ and $\rho \in \mathbb{R}_+$, then $B(w; \rho) := \{x \in X \mid \|x - w\| \leq \rho\}$ is the closed ball centered at w with radius ρ , and B simply stands for the unit ball $B(0; 1)$. The set of fixed points of

a set-valued operator $T : X \rightrightarrows X$ is $\text{Fix} T := \{x \in X \mid x \in Tx\}$. Also, $\text{aff} A$ represents the affine hull of a set A and $L^\perp := \{u \in X \mid \forall x \in L, \langle u, x \rangle = 0\}$ is the orthogonal complementary subspace of L .

2.1. Distance function and relaxed projectors

Let C be a nonempty subset of X . The *distance function* to C is defined by

$$d_C : X \rightarrow \mathbb{R} : x \mapsto \inf_{c \in C} \|x - c\| \quad (1)$$

and the *projector* onto C is

$$P_C : X \rightrightarrows C : x \mapsto \underset{c \in C}{\text{argmin}} \|x - c\| = \{c \in C \mid \|x - c\| = d_C(x)\}. \quad (2)$$

The *relaxed projector* for C with parameter $\lambda \in \mathbb{R}_+$ is defined by

$$P_C^\lambda := (1 - \lambda)\text{Id} + \lambda P_C. \quad (3)$$

Note that $P_C^0 = \text{Id}$ is the *identity* operator, $P_C^1 = P_C$ is the projector onto C , and $P_C^2 = R_C := 2P_C - \text{Id}$ is the so-called *reflector* across C .

We now collect some useful properties of relaxed projectors.

Lemma 2.1. *Let L be an affine subspace of X . Then the following hold:*

- (i) P_L is an affine operator, and for every $x \in X$, $x - P_L x \in (L - L)^\perp$.
- (ii) $\forall x \in X, \forall z \in X, \|x - z\|^2 = \|P_L x - P_L z\|^2 + \|(x - P_L x) - (z - P_L z)\|^2$.
- (iii) $\forall x \in X, \forall w \in L, \|x - w\|^2 = \|P_L x - w\|^2 + \|x - P_L x\|^2$. Consequently, $\|P_L x - w\| \leq \|x - w\|$.

Proof. (i): This follows from [5, Corollary 3.22].

(ii): Let $x \in X$ and let $z \in X$. Then (i) implies that $\langle P_L x - P_L z, x - P_L x \rangle = 0$ and $\langle P_L x - P_L z, z - P_L z \rangle = 0$, which yields

$$\|x - z\|^2 = \|(P_L x - P_L z) + [(x - P_L x) - (z - P_L z)]\|^2 \quad (4a)$$

$$= \|P_L x - P_L z\|^2 + \|(x - P_L x) - (z - P_L z)\|^2. \quad (4b)$$

(iii): Apply (ii) to $z = w \in L$. ■

Lemma 2.2. *Let C be a nonempty closed subset of X , let L be an affine subspace of X containing C , and let $\lambda \in \mathbb{R}$. Then the following hold:*

- (i) $P_C^\lambda(L) \subseteq L$.
- (ii) $P_C P_L = P_C = P_L P_C$ and $P_L P_C^\lambda = P_C^\lambda P_L$.
- (iii) $(\text{Id} - P_L)P_C^\lambda = (1 - \lambda)(\text{Id} - P_L)$.
- (iv) $\forall x \in X, d_C^2(P_C^\lambda x) = d_C^2(P_L x) + (1 - \lambda)^2 d_L^2(x)$. In particular, $d_C^2(x) = d_C^2(P_L x) + d_L^2(x)$.
- (v) If Q is a nonempty closed subset of X such that $Q \subseteq (L - L)^\perp$, then

$$\forall y \in L, \forall q \in Q, P_{C+Q}(y + q) = q + P_{C+Q} y = q + P_C y, \quad (5a)$$

$$d_{C+Q}(y + q) = d_{C+Q}(y) = d_C(y). \quad (5b)$$

Proof. (i): Let $x \in L$. Since $P_C x \subseteq C \subseteq L$ and L is an affine subspace, we have that $P_C^\lambda x = (1 - \lambda)x + \lambda P_C x \subseteq (1 - \lambda)L + L \subseteq L$.

(ii): It follows from Fact 2.1(i) that P_L is an affine operator, and from [11, Lemma 3.3] that $P_C P_L = P_C = P_L P_C$. Thus,

$$P_L P_C^\lambda = P_L((1-\lambda)\text{Id} + \lambda P_C) = (1-\lambda)P_L + \lambda P_L P_C = (1-\lambda)P_L + \lambda P_C P_L = P_C^\lambda P_L. \quad (6)$$

(iii): Using (ii), we obtain that

$$(\text{Id} - P_L)P_C^\lambda = P_C^\lambda - P_L P_C^\lambda = P_C^\lambda - P_C^\lambda P_L \quad (7a)$$

$$= (1-\lambda)\text{Id} + \lambda P_C - (1-\lambda)P_L - \lambda P_C P_L = (1-\lambda)(\text{Id} - P_L). \quad (7b)$$

(iv): In Lemma 2.1(iii), substituting x by $P_L^\lambda x$ and using (iii), we get

$$\|P_L^\lambda x - w\|^2 = \|P_L x - w\|^2 + \|(\text{Id} - P_C)P_L^\lambda x\|^2 = \|P_L x - w\|^2 + (1-\lambda)^2 \|x - P_L x\|^2. \quad (8)$$

Taking the infimum over $w \in C$ yields

$$d_C^2(P_L^\lambda x) = d_C^2(P_L x) + (1-\lambda)^2 \|x - P_L x\|^2. \quad (9)$$

For the second conclusion, we apply (9) with $\lambda = 0$.

(v): Let $y \in L$ and let $q \in Q$. It is straightforward to see that $P_{C+Q}(y+q) = P_{C+Q+q}(y+q) = q + P_{C+Q}y$. Now for all $c \in C$, since $c - y \in L - L$, we have

$$\|(c+q) - y\|^2 = \|c - y\|^2 + 2\langle q, c - y \rangle + \|q\|^2 = \|c - y\|^2 + \|q\|^2 \quad (10a)$$

$$\geq \|c - y\|^2 \geq d_C^2(y) \geq d_{C+Q}^2(y). \quad (10b)$$

So $c + q \in P_{C+Q}y$ if and only if $q = 0$ and $c \in P_C y$. Therefore, $P_{C+Q}y = P_C y$, which completes the proof. \blacksquare

2.2. Normal cones and regularity of sets

Let C be a nonempty subset of X and let $x \in C$. The *proximal normal cone* to C at x (see [30, Section 2.5.2, D] and [35, Example 6.16]) is defined by

$$N_C^{\text{prox}}(x) := \{\lambda(z - x) \mid z \in P_C^{-1}(x), \lambda \in \mathbb{R}_+\}, \quad (11)$$

and the *limiting normal cone* to C at x (see [30, Definition 1.1(ii) and Theorem 1.6]) can be given by

$$N_C(x) := \{u \in X \mid \exists x_n \rightarrow x, u_n \rightarrow u \text{ with } x_n \in C, u_n \in N_C^{\text{prox}}(x_n)\}. \quad (12)$$

Let $w \in X$, $\varepsilon \in \mathbb{R}_+$, and $\delta \in \mathbb{R}_{++}$. We recall from [11, Definition 8.1] and [24, Definition 2.9] that C is (ε, δ) -regular at w if

$$\left. \begin{array}{l} x, y \in C \cap \mathbf{B}(w; \delta), \\ u \in N_C^{\text{prox}}(x) \end{array} \right\} \implies \langle u, x - y \rangle \geq -\varepsilon \|u\| \cdot \|x - y\| \quad (13)$$

and (ε, ∞) -regular at w if it is (ε, δ) -regular for all $\delta \in \mathbb{R}_{++}$. The set C is *superregular* at w if for all $\varepsilon \in \mathbb{R}_{++}$, there exists $\delta \in \mathbb{R}_{++}$ such that C is (ε, δ) -regular at w .

Fact 2.3. *Let C be a nonempty closed subset of X , $w \in C$, $\lambda \in \mathbb{R}_+$, and $\delta \in \mathbb{R}_{++}$. Then the following hold:*

(i) $P_C^\lambda(\mathbf{B}(w; \delta/2)) \subseteq \mathbf{B}(w; (1+\lambda)\delta/2)$. In particular, $P_C(\mathbf{B}(w; \delta/2)) \subseteq C \cap \mathbf{B}(w; \delta)$.

(ii) If $\lambda \in]0, 2]$ and C is (ε, δ) -regular at w with $\varepsilon \in [0, 1/3]$, then $P_C^\lambda(\mathbf{B}(w; \delta/2)) \subseteq \mathbf{B}(w; \delta/\sqrt{2})$.

Proof. This follows from [19, Lemma 3.4(i) and Proposition 3.5]. \blacksquare

2.3. Regularity of set systems

In this section, m is a positive integer, $I := \{1, \dots, m\}$, and $\{C_i\}_{i \in I}$ is a system of closed subsets of X . Recall that $\{C_i\}_{i \in I}$ is *linearly regular* with modulus $\kappa \in \mathbb{R}_{++}$ (or κ -*linearly regular*) on a subset U of X if

$$\forall x \in U, \quad d_C(x) \leq \kappa \max_{i \in I} d_{C_i}(x), \quad \text{where } C := \bigcap_{i \in I} C_i. \quad (14)$$

We say that $\{C_i\}_{i \in I}$ is *linearly regular* around $w \in X$ if there exist $\delta \in \mathbb{R}_{++}$ and $\kappa \in \mathbb{R}_{++}$ such that $\{C_i\}_{i \in I}$ is κ -linearly regular on $B(w; \delta)$. The system $\{C_i\}_{i \in I}$ is said to be *boundedly linearly regular* if for every bounded set S of X , there exists $\kappa_S \in \mathbb{R}_{++}$ such that $\{C_i\}_{i \in I}$ is κ_S -linearly regular on S . Interested readers can find more discussion on linear regularity in [3, 4, 14, 19].

The following is a generalization of strong regularity, see, e.g., [19, Definition 2.3], and affine-hull regularity [32, Definition 2.1].

Definition 2.4 (*L*-regularity of set systems). Let $w \in \bigcap_{i \in I} C_i$ and let L be an affine subspace of X that contains w . The system $\{C_i\}_{i \in I}$ is said to be *L-regular* at w if

$$\sum_{i \in I} u_i = 0 \quad \text{and} \quad u_i \in N_{C_i}(w) \cap (L - w) \implies \forall i \in I, u_i = 0. \quad (15)$$

We simply say that $\{C_i\}_{i \in I}$ is *strongly regular* at w when $L = X$, and say that $\{C_i\}_{i \in I}$ is *affine-hull regular* at w when $L = \text{aff} \bigcup_{i \in I} C_i$. In the case $I = \{1, 2\}$, condition (15) can be rewritten as $N_{C_1}(w) \cap (-N_{C_2}(w)) \cap (L - w) = \{0\}$.

By definition, if the system $\{C_i\}_{i \in I}$ is *L-regular* at w , then so are all the subsystems. As shown in [19], strong regularity in Definition 2.4 is equivalent to [25, Definition 1(vi)] and [24, Definition 3.2].

In what follows, $|I|$ denotes the number of elements in the set I .

Proposition 2.5 (*L*-regularity implies linear regularity). Let $w \in C := \bigcap_{i \in I} C_i$ and let L be an affine subspace of X containing $\bigcup_{i \in I} C_i$. Suppose that $\{C_i\}_{i \in I}$ is *L-regular* at w . Then the following hold:

- (i) $\{C_i\}_{i \in I}$ is linearly regular around w .
- (ii) If $m = |I| \geq 2$, then $L = \text{aff} \bigcup_{i \in J} C_i$ whenever $J \subseteq I$ and $|J| \geq 2$.
- (iii) Every subsystem of $\{C_i\}_{i \in I}$, including itself, is affine-hull regular at w .

Consequently, if $\{C_i\}_{i \in I}$ is strongly regular at w , then $X = \text{aff} \bigcup_{i \in I} C_i$.

Proof. (i): Consider the system $\{C_i\}_{i \in I}$ within L , then it is strongly regular at w within L . So we learn from [25, Theorem 1(ii)] that $\{C_i\}_{i \in I}$ is linearly regular around w within L , i.e., there exist $\delta, \kappa \in \mathbb{R}_{++}$ such that

$$\forall y \in B(w; \delta) \cap L, \quad d_C(y) \leq \kappa \max_{i \in I} d_{C_i}(y). \quad (16)$$

Let $x \in B(w; \delta)$ and $y = P_L x \in L$. By Lemma 2.1(iii), $y \in B(w; \delta) \cap L$. Now by Lemma 2.2(iv), $d_C^2(x) = d_C^2(y) + d_L^2(x)$ and $d_{C_i}^2(x) = d_{C_i}^2(y) + d_L^2(x)$ for every $i \in I$. Combining with (16), we obtain

$$d_C^2(x) \leq \kappa^2 \max_{i \in I} d_{C_i}^2(y) + d_L^2(x) \leq \max\{\kappa^2, 1\} \max_{i \in I} d_{C_i}^2(x), \quad (17)$$

and so $\{C_i\}_{i \in I}$ is $\max\{\kappa, 1\}$ -linearly regular at w on $B(w; \delta)$.

(ii): Take two distinct indices $i, j \in J$. By assumption, $\{C_i, C_j\}$ is *L-regular* at w . Suppose that L strictly contains $\text{aff}(C_i \cup C_j)$. Then $\text{aff}(C_i \cup C_j) - w$ is a proper subspace of $L - w$. So there exists

a unit vector $u \in L - w$ such that u is perpendicular to $\text{aff}(C_i \cup C_j) - w$. Now define $y_i = w + u$ and $y_j = w - u$. On the one hand, by [11, Lemma 3.2], $P_{C_i}y_i = P_{C_i}w = w$ and $P_{C_j}y_j = P_{C_j}w = w$, which yield $u \in N_{C_i}^{\text{prox}}(w) \cap (L - w)$ and $-u \in N_{C_j}^{\text{prox}}(w) \cap (L - w)$, respectively. On the other hand, $u + (-u) = 0$ while $u \neq 0$. This contradicts the L -regularity of $\{C_i, C_j\}$. We must therefore have $L = \text{aff}(C_i \cup C_j)$. Since $\text{aff}(C_i \cup C_j) \subseteq \text{aff} \bigcup_{i \in J} C_i \subseteq L$, it follows that $L = \text{aff} \bigcup_{i \in J} C_i$.

(iii): Let $\{C_i\}_{i \in J}$ be a subsystem of $\{C_i\}_{i \in I}$. If $|J| = 1$, then $\{C_i\}_{i \in J}$ is automatically affine-hull regular at w . If $|J| \geq 2$, then from (ii), we have $L = \text{aff} \bigcup_{i \in J} C_i$, which implies affine-hull regularity of $\{C_i\}_{i \in J}$ at w . ■

Remark 2.6 (linear regularity does not imply affine-hull regularity). Proposition 2.5 has shown that affine-hull regularity implies linear regularity. However, it is known that the reverse is not true, for example, in \mathbb{R}^2 , the system $\{\mathbb{R}_+^2, \mathbb{R}_-^2\}$ is linear regular, but not affine-hull regular at $(0, 0)$.

Remark 2.7 (affine-hull regularity of subsystems). Affine-hull regularity of every *proper* subsystem $\{C_i\}_{i \in J}$ with $J \subsetneq I$ and linear regularity of the entire system $\{C_i\}_{i \in I}$ do not imply affine-hull regularity of $\{C_i\}_{i \in I}$. For example, in \mathbb{R}^2 , consider $C_1 = \{(\xi, \zeta) \mid \xi + \zeta \leq 0\}$, $C_2 = \{(\xi, \zeta) \mid \xi - \zeta \leq 0\}$, $C_3 = \{(\xi, \zeta) \mid \xi \geq 0\}$, and $w = (0, 0) \in C_1 \cap C_2 \cap C_3$. Then one can check that $\{C_i\}_{i \in J}$ with $J \subsetneq \{1, 2, 3\}$ is affine-hull regular at w , and that $\{C_1, C_2, C_3\}$ is linearly regular around w , but $\{C_1, C_2, C_3\}$ is not affine-hull regular at w .

Let A and B be nonempty subsets of X and let L be an affine subspace of X containing $A \cup B$. We recall from [11, Definiton 6.1] that the *CQ-number* at a point $w \in X$ associated with (A, B, L) and $\delta \in \mathbb{R}_{++}$ is defined by

$$\theta_{A,B,L}(w, \delta) := \sup \{ \langle u, v \rangle \mid u \in N_A^{\text{prox}}(a) \cap (L - a) \cap B, a \in A \cap B(w; \delta), \\ v \in -N_B^{\text{prox}}(b) \cap (L - b) \cap B, b \in B \cap B(w; \delta) \} \quad (18)$$

and that the *limiting CQ-number* at w associated with (A, B, L) is defined by

$$\bar{\theta}_{A,B,L}(w) := \lim_{\delta \downarrow 0} \theta_{A,B,L}(w, \delta). \quad (19)$$

Clearly, $\theta_{A,B,L}(w, \delta) = \theta_{B,A,L}(w, \delta) \leq 1$. When $L = X$, we simply write (A, B) for (A, B, L) .

We end this section with a connection between the CQ-number and L -regularity for two sets.

Proposition 2.8 (L -regularity for two sets). *Let A and B be two nonempty subsets of X , let $w \in A \cap B$, and let L be an affine subspace of X containing $A \cup B$. Then the following are equivalent:*

- (i) $\{A, B\}$ is L -regular at w .
- (ii) The CQ-number $\theta_{A,B,L}(w, \delta) < 1$ for some $\delta \in \mathbb{R}_{++}$.
- (iii) The limiting CQ-number $\bar{\theta}_{A,B,L}(w) < 1$.

Proof. By [11, Example 7.2], it suffices to show the equivalence of (ii) and (iii). In fact, if (ii) holds, then by the definition of the CQ-number, $\theta_{A,B,L}(w, \delta') \leq \theta_{A,B,L}(w, \delta) < 1$ for all $\delta' \in]0, \delta]$, which implies that $\bar{\theta}_{A,B,L}(w) = \lim_{\delta \downarrow 0} \theta_{A,B,L}(w, \delta) < 1$. The converse is obvious. ■

2.4. Generalized Douglas–Rachford operator

Let A and B be nonempty closed subsets of X , let $\lambda, \mu \in]0, 2]$, and let $\alpha \in \mathbb{R}_{++}$. The *generalized Douglas–Rachford (gDR) operator* for the pair (A, B) with parameters (λ, μ, α) , which was also considered in [19], is defined by

$$T_{\lambda, \mu}^{\alpha} := (1 - \alpha) \text{Id} + \alpha P_B^{\mu} P_A^{\lambda}. \quad (20)$$

It is worth mentioning that $T_{1,1}^1 = P_B P_A$ is the *classical alternating projection operator* [18], that $T_{2,2}^{1/2} = \frac{1}{2}(\text{Id} + R_B R_A)$ is the *classical DR operator* [21, 27], and that

$$T_{2,2\alpha}^{1/2} = (1 - \alpha)P_A + \frac{\alpha}{2}(\text{Id} + R_B R_A) \quad (21)$$

is the *relaxed averaged alternating reflection operator* [28]. In addition, if B is an affine subspace of X , then by Lemma 2.1(i), P_B is an affine operator, and therefore

$$T_{1+\alpha,1+\alpha}^{1/(1+\alpha)} = (1 - \alpha)P_B P_A + \frac{\alpha}{2}(\text{Id} + R_B R_A) \quad (22)$$

is an affine combination of the classical alternating projection and DR operators.

It is interesting to see that the shadow of any *gDR step* on certain affine subspaces is again a *gDR step*. This phenomenon is referred to as *affine reduction* in [32, Section 3].

Lemma 2.9 (shadows of gDR steps). *Let L be an affine subspace of X containing $A \cup B$, $x \in X$, and $x_+ \in T_{\lambda,\mu}^\alpha x$. Define $y = P_L x$, $y_+ = P_L x_+$, and set $\eta := 1 - \alpha + \alpha(1 - \lambda)(1 - \mu)$. Then the following hold:*

- (i) $y_+ \in T_{\lambda,\mu}^\alpha y$.
- (ii) $x_+ - y_+ = \eta(x - y)$. Consequently, $d_L(x_+) = |\eta|d_L(x)$.
- (iii) $\|x - x_+\|^2 = \|y - y_+\|^2 + (1 - \eta)^2\|x - y\|^2$.

Proof. Since $x_+ \in T_{\lambda,\mu}^\alpha x$, there exist $r \in P_A^\lambda x$ and $s \in P_B^\mu r$ such that $x_+ = (1 - \alpha)x + \alpha s$.

(i): Using Lemma 2.2(ii), we have $P_L s \in P_L P_B^\mu P_A^\lambda x = P_B^\mu P_A^\lambda P_L x = P_B^\mu P_A^\lambda y$. By Lemma 2.1(i), P_L is an affine operator, and hence

$$y_+ = P_L x_+ = (1 - \alpha)P_L x + \alpha P_L s \in (1 - \alpha)y + \alpha P_B^\mu P_A^\lambda y = T_{\lambda,\mu}^\alpha y. \quad (23)$$

(ii): It follows from Lemma 2.2(iii) that

$$s - P_L s = (1 - \mu)(r - P_L r) = (1 - \mu)(1 - \lambda)(x - P_L x). \quad (24)$$

Combining with (23) yields

$$x_+ - y_+ = (1 - \alpha)(x - P_L x) + \alpha(s - P_L s) = \eta(x - P_L x) = \eta(x - y). \quad (25)$$

(iii): Apply Lemma 2.1(ii) to $z = x_+$ and take (ii) into account. ■

Next, we study the fixed points of *gDR operators*.

Lemma 2.10 (fixed points of gDR operators). *Suppose that $A \cap B \neq \emptyset$ and let L be an affine subspace of X containing $A \cup B$. Set $\eta := 1 - \alpha + \alpha(1 - \lambda)(1 - \mu)$. Then the following hold:*

- (i) $\forall x \in X, \forall u \in (L - L)^\perp, P_A^\lambda(x + u) = P_A^\lambda x + (1 - \lambda)u, P_B^\mu(x + u) = P_B^\mu x + (1 - \mu)u$, and $T_{\lambda,\mu}^\alpha(x + u) = T_{\lambda,\mu}^\alpha x + \eta u$.
- (ii) $\forall c \in A \cap B, \forall u \in (L - L)^\perp, P_A^\lambda c = P_B^\mu c = c, P_A(c + u) = P_B(c + u) = c \in A \cap B$, and $T_{\lambda,\mu}^\alpha(c + u) = c + \eta u$.
- (iii) $A \cap B \subseteq \text{Fix } T_{\lambda,\mu}^\alpha$ and, moreover, if $(1 - \lambda)(1 - \mu) = 1$, then $(A \cap B) + (L - L)^\perp \subseteq \text{Fix } T_{\lambda,\mu}^\alpha$. In particular, $(A \cap B) + (L - L)^\perp \subseteq \text{Fix } T_{2,2}^\alpha$.

Proof. (i): Let $x \in X$, $u \in (L-L)^\perp$, and $w \in A \cap B$. Then $(L-L)^\perp = (L-w)^\perp$, hence $u \in (\text{aff } A-w)^\perp$ and also $u \in (\text{aff } B-w)^\perp$. Applying [11, Lemma 3.2] implies that

$$P_A^\lambda(x+u) = (1-\lambda)(x+u) + \lambda P_A(x+u) = (1-\lambda)(x+u) + \lambda P_A x = P_A^\lambda x + (1-\lambda)u. \quad (26)$$

Similarly, $P_B^\mu(x+u) = P_B^\mu x + (1-\mu)u$ and the rest follows.

(ii): Let $c \in A \cap B$ and $u \in (L-L)^\perp$. Since $P_A c = P_B c = c$, we derive that $P_A^\lambda c = P_B^\mu c = c$, which yields $T_{\lambda,\mu}^\alpha c = c$. Combining with (i), $P_A(c+u) = P_B(c+u) = c \in A \cap B$ and $T_{\lambda,\mu}^\alpha(c+u) = T_{\lambda,\mu}^\alpha c + \eta u = c + \eta u$.

(iii): It follows from (ii) that $T_{\lambda,\mu}^\alpha c = c$ for all $c \in A \cap B$, which gives $A \cap B \subseteq \text{Fix } T_{\lambda,\mu}^\alpha$. Now if $(1-\lambda)(1-\mu) = 1$, then $\eta = 1$, which leads to $(A \cap B) + (L-L)^\perp \subseteq \text{Fix } T_{\lambda,\mu}^\alpha$ due to (i). ■

In the rest of this section, we assume that X is a real Hilbert space and that A and B are convex but need not intersect. Then $B-A$ is convex, hence we can take $g := P_{B-A} 0$ and set

$$E := A \cap (B-g) \quad \text{and} \quad F := (A+g) \cap B. \quad (27)$$

It is clear that if $A \cap B \neq \emptyset$, then $g = 0$ and $E = F = A \cap B$. We also note that

$$g \in B-A \iff E \neq \emptyset \iff F \neq \emptyset \quad (28)$$

and from [2, Lemma 2.2(i)&(iv)] that

$$a = P_A b \text{ and } b = P_B a \implies g = b - a. \quad (29)$$

Recall from [5, Definition 4.1 and 4.33] that a single-valued operator $T: X \rightarrow X$ is *nonexpansive* if it is Lipschitz continuous with constant 1, i.e.,

$$\forall x, y \in X, \quad \|Tx - Ty\| \leq \|x - y\|, \quad (30)$$

and is α -averaged if $\alpha \in]0, 1[$ and $T = (1-\alpha)\text{Id} + \alpha R$ for some nonexpansive operator $R: X \rightarrow X$.

Fact 2.11. *Let C be a nonempty closed convex subset of X , let $x \in X$ and $p = P_C x$. Then the following hold:*

- (i) *For every $\lambda \in \mathbb{R}_+$, $P_C(p + \lambda(x-p)) = P_C((1-\lambda)p + \lambda x) = p$.*
- (ii) *For every $\lambda \in]0, 2[$, P_C^λ is $\lambda/2$ -averaged. Consequently, for every $\lambda \in]0, 2[$, P_C^λ is nonexpansive.*

Proof. (i): See [5, Proposition 3.21]. (ii): Combine Proposition 4.16, Corollary 4.41, Remark 4.34(i), and Corollary 4.18 in [5]. ■

Hereafter, whenever dealing with the harmonic-like quantity $\beta := (\frac{1}{\beta_1} + \frac{1}{\beta_2} + \dots + \frac{1}{\beta_k})^{-1}$ of non-negative numbers $\beta_i \in \mathbb{R}_+$, we make a convention that $\beta = 0$ if at least one β_i equals 0.

Lemma 2.12. *Let $\lambda, \mu \in]0, 2[$. Then the following hold:*

- (i) *For every $\alpha \in \mathbb{R}_+$, $T_{\lambda,\mu}^\alpha$ is continuous and single-valued.*
- (ii) *For every $\alpha \in]0, 1 + \widehat{\beta}[$ where $\widehat{\beta} := (\frac{\lambda}{2-\lambda} + \frac{\mu}{2-\mu})^{-1}$, $T_{\lambda,\mu}^\alpha$ is $\alpha/(1 + \widehat{\beta})$ -averaged.*

Proof. By Fact 2.11(ii), P_A^λ and P_B^μ is nonexpansive, hence, continuous and single-valued. Thus, (i) follows. To prove (ii), we first have that $P_B^\mu P_A^\lambda$ is nonexpansive. If $\lambda = 2$ or $\mu = 2$, then $\widehat{\beta} = 0$ and $\alpha \in]0, 1[$, so $T_{\lambda,\mu}^\alpha = (1-\alpha)\text{Id} + \alpha P_B^\mu P_A^\lambda$ is α -averaged. If $\lambda, \mu < 2$, then P_A^λ and P_B^μ are respectively $\lambda/2$ - and $\mu/2$ -averaged due to Fact 2.11(ii). We derive from [5, Proposition 4.44] that $P_B^\mu P_A^\lambda$ is ξ -averaged with $\xi := 2(\lambda + \mu - \lambda\mu)/(4 - \lambda\mu) = 1/(1 + \widehat{\beta})$, and then from [5, Proposition 4.40] that $T_{\lambda,\mu}^\alpha$ is $\alpha/(1 + \widehat{\beta})$ -averaged. ■

Lemma 2.13 (fixed points of convex gDR operators). *Let $\lambda, \mu \in]0, 2]$ and $\alpha \in \mathbb{R}_{++}$. Then the following hold:*

(i) *If $E \neq \emptyset$ when $\min\{\lambda, \mu\} < 2$, and $A \cap B \neq \emptyset$ when $\lambda = \mu = 2$, then $P_A \text{Fix } T_{\lambda, \mu}^\alpha \subseteq E$ and*

$$\text{Fix } T_{\lambda, \mu}^\alpha = \begin{cases} E + \frac{\mu}{\lambda + \mu - \lambda\mu}g & \text{if } \min\{\lambda, \mu\} < 2, \\ A \cap B + N_{A-B}(0) & \text{if } \lambda = \mu = 2. \end{cases} \quad (31)$$

(ii) *If $A \cap B \neq \emptyset$ when $\min\{\lambda, \mu\} < 2$, and $0 \in \text{int}(B - A)$ when $\lambda = \mu = 2$, then $\text{Fix } T_{\lambda, \mu}^\alpha = A \cap B$.*
(iii) *If $\text{ri } A \cap \text{ri } B \neq \emptyset$, then $\text{Fix } T_{\lambda, \mu}^\alpha \cap \text{aff}(A \cup B) = A \cap B$.*

Proof. (i): First, it is straightforward to see that $\text{Fix } T_{\lambda, \mu}^\alpha = \text{Fix } P_B^\mu P_A^\lambda = \text{Fix } T_{\lambda, \mu}^{1/2}$. If $\lambda = \mu = 2$, then the conclusion follows from [6, Corollary 3.9].

Now assume that $\min\{\lambda, \mu\} < 2$. Since $\lambda, \mu \in]0, 2]$, we have $\lambda + \mu - \lambda\mu = 1 - (1 - \lambda)(1 - \mu) > 0$. Let $x \in \text{Fix } P_B^\mu P_A^\lambda$, that is, $x = P_B^\mu P_A^\lambda x$. Set $a = P_A x$, $r = (1 - \lambda)x + \lambda a$, and $b = P_B r$. It follows that $x = (1 - \mu)r + \mu b$ and hence

$$b = \frac{1}{\mu}x - \frac{1-\mu}{\mu}r = \frac{1}{\mu}x - \frac{1-\mu}{\mu}\left((1 - \lambda)x + \lambda a\right) = a + \frac{\lambda + \mu - \lambda\mu}{\mu}(x - a). \quad (32)$$

Since $\frac{\lambda + \mu - \lambda\mu}{\mu} > 0$, Fact 2.11(i) implies that $P_A b = a$. Similarly, $P_B a = b$. We then use (29) to deduce $g = b - a$. Combining with (32) yields

$$x = a + \frac{\mu}{\lambda + \mu - \lambda\mu}(b - a) \in E + \frac{\mu}{\lambda + \mu - \lambda\mu}g. \quad (33)$$

Conversely, take $x \in E + \frac{\mu}{\lambda + \mu - \lambda\mu}g$. It suffices to show that $x \in \text{Fix } P_B^\mu P_A^\lambda$ and $P_A x \in E$. By assumption, there exist $a \in A$ and $b \in B$ such that $a = b - g \in E$, $a = P_A b$, $b = P_B a$, and that $x = a + \frac{\mu}{\lambda + \mu - \lambda\mu}g = P_B b + \frac{\mu}{\lambda + \mu - \lambda\mu}(b - P_A b)$. Again, Fact 2.11(i) yields $P_A x = P_A b = a \in E$. In turn,

$$r := P_A^\lambda x = (1 - \lambda)\left(a + \frac{\mu}{\lambda + \mu - \lambda\mu}g\right) + \lambda a = b - \frac{\lambda}{\lambda + \mu - \lambda\mu}g = b + \frac{\lambda}{\lambda + \mu - \lambda\mu}(a - b) \quad (34)$$

and, by Fact 2.11(i), $P_B r = b$. It follows that

$$P_B^\mu P_A^\lambda x = P_B^\mu r = (1 - \mu)\left(b - \frac{\lambda}{\lambda + \mu - \lambda\mu}g\right) + \mu b = a + g - \frac{(1 - \mu)\lambda}{\lambda + \mu - \lambda\mu}g = x, \quad (35)$$

which completes the proof.

(ii): This follows from (i) by noting that if $A \cap B \neq \emptyset$, then $g = 0$ and $E = A \cap B$, and from, e.g., [9, Proposition 2.10(ii)] that if $0 \in \text{int}(B - A)$, then $N_{A-B}(0) = 0$.

(iii): If $\min\{\lambda, \mu\} < 2$, then the conclusion follows from (ii). If $\lambda = \mu = 2$, then the conclusion follows from [14, Proposition 4.1(iii)]. \blacksquare

Theorem 2.14 (convex possibly inconsistent case). *Let $\lambda, \mu \in]0, 2]$ and $\alpha \in]0, 1 + \widehat{\beta}[$ where $\widehat{\beta} := \left(\frac{\lambda}{2-\lambda} + \frac{\mu}{2-\mu}\right)^{-1}$. Suppose that $E \neq \emptyset$ when $\min\{\lambda, \mu\} < 2$, and $A \cap B \neq \emptyset$ when $\lambda = \mu = 2$. Then the gDR sequence $(x_n)_{n \in \mathbb{N}}$ generated by $T_{\lambda, \mu}^\alpha$ weakly converges to a point $\bar{x} \in \text{Fix } T_{\lambda, \mu}^\alpha$ with $P_A \bar{x} \in E$.*

Proof. According to Lemma 2.12(ii), $T_{\lambda, \mu}^\alpha = (1 - \alpha)\text{Id} + \alpha P_B^\mu P_A^\lambda$ is $\alpha/(1 + \widehat{\beta})$ -averaged. Now use Lemma 2.13(i) and apply [5, Proposition 5.16] with all $\lambda_n = 1$. \blacksquare

In the light of Theorem 2.14, if the gDR algorithm involves at most one reflection, then it is weakly convergent even in the inconsistent case.

3. Quasi firm Fejér monotonicity

In this section, we further refine some results on quasi firm Fejér monotonicity, which were partly developed in [19]. Let C and U be nonempty subsets of X , let $\gamma \in [1, +\infty[$, and let $\beta \in \mathbb{R}_+$. Recall from [19, Definition 3.1] that a set-valued operator $T: X \rightrightarrows X$ is (C, γ, β) -quasi firmly Fejér monotone on U if

$$\forall x \in U, \forall x_+ \in Tx, \forall \bar{x} \in C, \quad \|x_+ - \bar{x}\|^2 + \beta \|x - x_+\|^2 \leq \gamma \|x - \bar{x}\|^2. \quad (36)$$

Here, when $\beta = 0$, we simply say that T is (C, γ) -quasi Fejér monotone on U .

Lemma 3.1 (averaged quasi firmly Fejér monotone operators). *Let C and U be nonempty subsets of X , $\gamma \in [1, +\infty[$, $\beta \in \mathbb{R}_+$, $\lambda \in]0, 1 + \beta]$, and let $T: X \rightrightarrows X$ be a set-valued operator. Then the following are equivalent:*

- (i) T is (C, γ, β) -quasi firmly Fejér monotone on U .
- (ii) $T^\lambda := (1 - \lambda)\text{Id} + \lambda T$ is $(C, 1 - \lambda + \lambda\gamma, \frac{1 - \lambda + \beta}{\lambda})$ -quasi firmly Fejér monotone on U .
- (iii) $T = \frac{\beta}{1 + \beta}\text{Id} + \frac{1}{1 + \beta}T^{1 + \beta}$ with $T^{1 + \beta} := (1 + \beta)T - \beta\text{Id}$ being $(C, \gamma + \beta(\gamma - 1))$ -quasi Fejér monotone on U .

Proof. If (i) holds, then so does (ii) due to [19, Lemma 3.2]. Conversely, if (ii) holds, applying [19, Lemma 3.2] to $T = (1 - \frac{1}{\lambda})\text{Id} - \frac{1}{\lambda}T^\lambda$ and noting that $0 < \frac{1}{\lambda} \leq 1 + \frac{1 - \lambda + \beta}{\lambda}$, we get (i). So (i) and (ii) are equivalent, which implies the equivalence of (i) and (iii) by taking $\lambda = 1 + \beta$. \blacksquare

Lemma 3.2 (composition of quasi firmly Fejér monotone operators). *Let m be a positive integer, set $I := \{1, \dots, m\}$, and for every $i \in I$, let C_i and U_i be a nonempty subset of X , $\gamma_i \in [1, +\infty[$, $\beta_i \in \mathbb{R}_+$, and T_i a (C_i, γ_i, β_i) -quasi firmly Fejér monotone operator on U_i . Set $C := \bigcap_{i \in I} C_i$, $\gamma := \gamma_1 \cdots \gamma_m$,*

$$\beta' := \left(\frac{1}{\beta_1 \gamma_2 \cdots \gamma_m} + \frac{1}{\beta_2 \gamma_3 \cdots \gamma_m} + \cdots + \frac{1}{\beta_{m-1} \gamma_m} + \frac{1}{\gamma_m} \right)^{-1}, \quad \text{and } \beta := \left(\sum_{i \in I} \frac{1}{\beta_i} \right)^{-1}. \quad (37)$$

Let x_0, x_1, \dots, x_m be such that for every $i \in I$, $x_i \in T_i x_{i-1}$. Then $\beta' \geq \beta \geq 0$ and, if $x_{i-1} \in U_i$ for every $i \in I$, it holds that

$$\forall \bar{x} \in C, \quad \gamma \|x_0 - \bar{x}\|^2 \geq \|x_m - \bar{x}\|^2 + \beta' \left(\sum_{i \in I} \|x_{i-1} - x_i\| \right)^2 \quad (38a)$$

$$\geq \|x_m - \bar{x}\|^2 + \beta' \|x_0 - x_m\|^2. \quad (38b)$$

Consequently, if $T_i U_i \subseteq U_{i+1}$ for every $i \in I \setminus \{m\}$, then $T_m \cdots T_1$ is (C, γ, β') - and also (C, γ, β) -quasi firmly Fejér monotone on U_1 .

Proof. Because $\gamma_i \geq 1$ and $\beta_i \geq 0$ for every $i \in I$, we have $\beta' \geq \beta \geq 0$. Next, let $\bar{x} \in C$. For every $i \in I$, since $x_{i-1} \in U_i$ and T_i is (C, γ_i, β_i) -quasi firmly Fejér monotone on U_i , we derive that

$$\|x_1 - \bar{x}\|^2 + \beta_1 \|x_0 - x_1\|^2 \leq \gamma_1 \|x_0 - \bar{x}\|^2, \quad (39a)$$

$$\|x_2 - \bar{x}\|^2 + \beta_2 \|x_1 - x_2\|^2 \leq \gamma_2 \|x_1 - \bar{x}\|^2, \quad (39b)$$

$$\vdots \quad (39c)$$

$$\|x_m - \bar{x}\|^2 + \beta_m \|x_{m-1} - x_m\|^2 \leq \gamma_m \|x_{m-1} - \bar{x}\|^2. \quad (39d)$$

Using telescoping techniques yields

$$\begin{aligned} (\gamma_1 \cdots \gamma_m) \|x_0 - \bar{x}\|^2 &\geq \|x_m - \bar{x}\|^2 + (\beta_1 \gamma_2 \cdots \gamma_m \|x_0 - x_1\|^2 + \beta_2 \gamma_3 \cdots \gamma_m \|x_1 - x_2\|^2 \\ &\quad + \cdots + \beta_{m-1} \gamma_m \|x_{m-2} - x_{m-1}\|^2 + \beta_m \|x_{m-1} - x_m\|^2). \end{aligned} \quad (40)$$

By the coordinate version of Cauchy–Schwarz inequality,

$$\begin{aligned} \left(\frac{1}{\beta_1 \gamma_2 \cdots \gamma_m} + \cdots + \frac{1}{\beta_m} \right) (\beta_1 \gamma_2 \cdots \gamma_m \|x_0 - x_1\|^2 + \cdots + \beta_m \|x_{m-1} - x_m\|^2) \\ \geq \left(\sum_{i \in I} \|x_{i-1} - x_i\| \right)^2. \end{aligned} \quad (41)$$

Combining with (40), we obtain (38).

Now assume that $T_i U_i \subseteq U_{i+1}$ for every $i \in I \setminus \{m\}$. Let $x \in U_1$ and $x_+ \in (T_m \cdots T_1)x$. Then there exist x_0, x_1, \dots, x_m such that $x_0 = x$, $x_m = x_+$, and $x_i \in T x_{i-1}$ for every $i \in I$. We derive that $x_{i-1} \in U_i$ for every $i \in I$ and, by (38),

$$\forall \bar{x} \in C, \quad \gamma \|x - \bar{x}\|^2 \geq \|x_+ - \bar{x}\|^2 + \beta' \|x - x_+\|^2 \geq \|x_+ - \bar{x}\|^2 + \beta \|x - x_+\|^2. \quad (42)$$

The proof is complete. \blacksquare

Lemma 3.3 (quasi firm Fejér monotonicity of relaxed projectors). *Let C be a nonempty subset of X and L be an affine subspace of X containing C . Let also $w \in C$, $\varepsilon \in [0, 1[$, $\delta \in \mathbb{R}_{++}$, and $\lambda \in]0, 2]$. Set*

$$\Omega := C \cap \mathbf{B}(w; \delta), \quad \gamma := 1 + \frac{\lambda \varepsilon}{1 - \varepsilon}, \quad \text{and} \quad \beta := \frac{2 - \lambda}{\lambda}. \quad (43)$$

Suppose that C is (ε, δ) -regular at w . Then P_C^λ is $(\Omega + (L - L)^\perp, \gamma, \beta)$ -quasi firmly Fejér monotone and, in particular, R_C is $(\Omega + (L - L)^\perp, \frac{1 + \varepsilon}{1 - \varepsilon})$ -quasi Fejér monotone on $\mathbf{B}(w; \delta/2) \cap L$.

Proof. Let $x \in \mathbf{B}(w; \delta/2) \cap L \subseteq \mathbf{B}(w; \delta/2)$, let $p \in P_C x$, and let $\bar{x} \in \Omega + (L - L)^\perp$. By Fact 2.3(i), $p \in \Omega$. Writing $\bar{x} = u + v$ with $u \in \Omega$ and $v \in (L - L)^\perp$, we note that $x, p, u \in L$, so $\langle x - p, v \rangle = 0$ and $\|p - \bar{x}\|^2 = \|p - u\|^2 + \|v\|^2 \geq \|p - u\|^2$. Since C is (ε, δ) -regular at w and $x - p \in N_C^{\text{prox}}(p)$, we have

$$\langle x - p, p - \bar{x} \rangle = \langle x - p, p - u \rangle \geq -\varepsilon \|x - p\| \|p - u\| \quad (44a)$$

$$\geq -\frac{\varepsilon}{2} (\|x - p\|^2 + \|p - u\|^2) \geq -\frac{\varepsilon}{2} (\|x - p\|^2 + \|p - \bar{x}\|^2). \quad (44b)$$

Therefore,

$$\|x - \bar{x}\|^2 = \|x - p\|^2 + \|p - \bar{x}\|^2 + 2 \langle x - p, p - \bar{x} \rangle \quad (45a)$$

$$\geq \|x - p\|^2 + \|p - \bar{x}\|^2 - \varepsilon (\|x - p\|^2 + \|p - \bar{x}\|^2) \quad (45b)$$

$$= (1 - \varepsilon) (\|x - p\|^2 + \|p - \bar{x}\|^2), \quad (45c)$$

which yields

$$\frac{1}{1 - \varepsilon} \|x - \bar{x}\|^2 \geq \|x - p\|^2 + \|p - \bar{x}\|^2, \quad (46)$$

i.e., P_C is $(\Omega + (L - L)^\perp, \frac{1}{1 - \varepsilon}, 1)$ -quasi firmly Fejér monotone on $\mathbf{B}(w; \delta/2) \cap L$. In turn, Lemma 3.1 implies that P_C^λ is $(\Omega + (L - L)^\perp, \gamma, \beta)$ -quasi firmly Fejér monotone on $\mathbf{B}(w; \delta/2) \cap L$ with γ and β as in (43). In the case when $\lambda = 2$, since $\gamma = \frac{1 + \varepsilon}{1 - \varepsilon}$ and $\beta = 0$, the conclusion follows. \blacksquare

Proposition 3.4 (quasi firm Fejér monotonicity of gDR operators). *Let A and B be closed subsets of X such that $A \cap B \neq \emptyset$ and L be an affine subspace of X containing $A \cup B$. Let also $w \in A \cap B$, $\varepsilon_1 \in [0, 1/3]$, $\varepsilon_2 \in [0, 1[$, $\delta \in \mathbb{R}_{++}$, $\lambda, \mu \in]0, 2]$, and $\alpha \in]0, 1 + \widehat{\beta}]$ where $\widehat{\beta} := (\frac{\lambda}{2 - \lambda} + \frac{\mu}{2 - \mu})^{-1}$. Set $\Omega := A \cap B \cap \mathbf{B}(w; \delta)$,*

$$\gamma := 1 - \alpha + \alpha \left(1 + \frac{\lambda \varepsilon_1}{1 - \varepsilon_1} \right) \left(1 + \frac{\mu \varepsilon_2}{1 - \varepsilon_2} \right), \quad \text{and} \quad \beta := \frac{1 - \alpha + \widehat{\beta}}{\alpha}. \quad (47)$$

Suppose that A and B are (ε_1, δ) - and $(\varepsilon_2, \sqrt{2}\delta)$ -regular at w , respectively. Then the following hold:

- (i) $T_{\lambda,\mu}^\alpha$ is $(\Omega + (L - L)^\perp, \gamma, \beta)$ -quasi firmly Fejér monotone on $\mathbf{B}(w; \delta/2) \cap L$.
- (ii) $T_{\lambda,\mu}^\alpha$ is (Ω, γ, β) -quasi firmly Fejér monotone on $\mathbf{B}(w; \delta/2)$.
- (iii) $T_{2,2}^\alpha$ is $(\Omega + (L - L)^\perp, \gamma, \beta)$ -quasi firmly Fejér monotone on $\mathbf{B}(w; \delta/2)$.

Proof. (i): We first derive from Lemma 3.3 that P_A^λ is $(\Omega + (L - L)^\perp, \gamma_1, \beta_1)$ -quasi firmly Fejér monotone on $\mathbf{B}(w; \delta/2) \cap L$ and that P_B^μ is $(\Omega + (L - L)^\perp, \gamma_2, \beta_2)$ -quasi firmly Fejér monotone on $\mathbf{B}(w; \delta/\sqrt{2}) \cap L$, where

$$\gamma_1 := 1 + \frac{\lambda\varepsilon_1}{1-\varepsilon_1}, \quad \gamma_2 := 1 + \frac{\mu\varepsilon_2}{1-\varepsilon_2}, \quad \beta_1 := \frac{2-\lambda}{\lambda}, \quad \text{and} \quad \beta_2 := \frac{2-\mu}{\mu}. \quad (48)$$

Now let $x \in \mathbf{B}(w; \delta/2) \cap L$. On the one hand, Fact 2.3(ii) yields $P_A^\lambda x \subseteq \mathbf{B}(w; \delta/\sqrt{2})$. On the other hand, $P_A^\lambda x = (1 - \lambda)x + \lambda P_A x \subseteq L$ since $x \in L$, $P_A x \subseteq A \subseteq L$, and L is an affine subspace. We deduce that $P_A^\lambda(\mathbf{B}(w; \delta/2) \cap L) \subseteq \mathbf{B}(w; \delta/\sqrt{2}) \cap L$. Noting that $\widehat{\beta} = (\frac{1}{\beta_1} + \frac{1}{\beta_2})^{-1}$, we apply Lemma 3.2 to (P_A^λ, P_B^μ) to obtain the $(\Omega + (L - L)^\perp, \gamma_1\gamma_2, \widehat{\beta})$ -quasi firm Fejér monotonicity of $P_B^\mu P_A^\lambda$ on $\mathbf{B}(w; \delta/2) \cap L$. The conclusion follows from Lemma 3.1(i)–(ii) applied to the operators $P_B^\mu P_A^\lambda$ and $T_{\lambda,\mu}^\alpha = (1 - \alpha)\text{Id} + \alpha P_B^\mu P_A^\lambda$.

(ii): Apply (i) with $L = X$.

(iii): Let $x \in \mathbf{B}(w; \delta/2)$ and $x_+ \in T_{2,2}^\alpha x$. Define $y = P_L x$ and $y_+ = P_L x_+$. Then by Lemma 2.1(iii), $y \in \mathbf{B}(w; \delta/2) \cap L$ and by Lemma 2.9(i)–(ii), $y_+ \in T_{2,2}^\alpha y$ and $v := x_+ - y_+ = x - y$. Applying (i) to $T_{2,2}^\alpha$ yields

$$\forall \bar{y} \in \Omega + (L - L)^\perp, \quad \|y_+ - \bar{y}\|^2 + \beta \|y - y_+\|^2 \leq \gamma \|y - \bar{y}\|^2. \quad (49)$$

Now let $\bar{x} \in \Omega + (L - L)^\perp$. From Lemma 2.1(i), we observe that $v \in (L - L)^\perp$, hence $\bar{x} - v \in \Omega + (L - L)^\perp$. Substituting $\bar{y} = \bar{x} - v = \bar{x} - (x_+ - y_+) = \bar{x} - (x - y)$ into (49) completes the proof. \blacksquare

4. Quasi coercivity

Let C and U be nonempty subsets of X and let $\nu \in \mathbb{R}_{++}$. Recall from [19, Definition 3.3] that an operator $T: X \rightrightarrows X$ is (C, ν) -quasi coercive on U if

$$\forall x \in U, \forall x_+ \in Tx, \quad \|x - x_+\| \geq \nu d_C(x), \quad (50)$$

and C -quasi coercive around $w \in X$ if it is (C, ν) -quasi coercive on $\mathbf{B}(w; \delta)$ for some $\nu \in \mathbb{R}_{++}$ and $\delta \in \mathbb{R}_{++}$.

In this section, we show quasi coercivity of gDR operators under different assumptions on the system of sets. In particular, under *affine-hull regularity* assumption, Proposition 4.5 improves some existing results on quasi coercivity (see Remark 4.8), while under *linear regularity* assumption and some parameter restriction, Proposition 4.12 proves the quasi coercivity of gDR operators.

Lemma 4.1 (averaged quasi coercive operators). *Let C and U be nonempty subsets of X , $\nu \in \mathbb{R}_{++}$, $\lambda \in \mathbb{R}_{++}$, and let $T: X \rightrightarrows X$ be a set-valued operator. Then T is (C, ν) -quasi coercive on U if and only if $T^\lambda := (1 - \lambda)\text{Id} + \lambda T$ is $(C, \lambda\nu)$ -quasi coercive on U .*

Proof. Assume that T is (C, ν) -quasi coercive on U . Let $x \in U$ and let $x_+ \in T^\lambda x$. Then there exists $s \in Tx$ such that $x_+ = (1 - \lambda)x + \lambda s$. We obtain that $\|x - x_+\| = \lambda \|x - s\| \geq \lambda \nu d_C(x)$, and T^λ is thus $(C, \lambda\nu)$ -quasi coercive on U . Conversely, note that $T = (1 - \frac{1}{\lambda})\text{Id} + \frac{1}{\lambda}T^\lambda$. \blacksquare

Lemma 4.2 (global quasi coercivity). *Let C be a nonempty subset of X , L a closed subset of X , and $T: X \rightarrow X$ a continuous single-valued operator. Suppose that $\text{Fix} T \cap L \subseteq C$ and that, for every $w \in C$, T is C -quasi coercive on $\mathbf{B}(w; \delta) \cap L$ for some $\delta \in \mathbb{R}_{++}$. Then T is C -quasi coercive on $S \cap L$ for every bounded set S of X .*

Proof. Let S be a bounded set of X and suppose on the contrary that T is not C -quasi coercive on $S \cap L$. Then there exist sequences $\varepsilon_n \downarrow 0$ and $x_n \in S \cap L$ such that

$$\forall n \in \mathbb{N}, \quad 0 \leq \|x_n - Tx_n\| < \varepsilon_n d_C(x_n). \quad (51)$$

Since $(x_n)_{n \in \mathbb{N}}$ is bounded, so is $(d_C(x_n))_{n \in \mathbb{N}}$, and hence $x_n - Tx_n \rightarrow 0$. Extracting a convergent subsequence without relabeling, we can assume $x_n \rightarrow \bar{x}$. It follows from the continuity of T and the closedness of L that $\bar{x} \in \text{Fix } T \cap L \subseteq C$. In turn, there exist $\nu \in \mathbb{R}_{++}$ and $\delta \in \mathbb{R}_{++}$ such that T is (C, ν) -quasi coercive on $B(\bar{x}; \delta) \cap L$. Thus, for all n sufficiently large,

$$\varepsilon_n d_C(x_n) > \|x_n - Tx_n\| \geq \nu d_C(x_n), \quad (52)$$

which is a contradiction since $\varepsilon_n \downarrow 0$ and $d_C(x_n) > 0$. \blacksquare

4.1. In the presence of affine-hull regularity

In this section, we aim to improve the estimate for quasi coercivity constant previously obtained in [19]. To proceed, we need the following technical lemma.

Lemma 4.3. *Let $\theta \in [0, 1[$, $\mu \in]0, 2]$, and let $u, v \in X$ be such that $\langle u, v \rangle \geq -\theta \|u\| \|v\|$. Then*

$$\|u + \mu v\|^2 \geq \max \left\{ \mu^2 (1 - \theta^2) \|v\|^2, \frac{4\mu^2 (1 - \theta^2)}{(|1 - \mu| + \sqrt{(1 - \mu)^2 + 4\mu(1 - \theta^2)})^2} \|u + v\|^2 \right\}. \quad (53)$$

Proof. First, it follows from $\langle u, v \rangle \geq -\theta \|u\| \|v\|$ that

$$\|u + \mu v\|^2 \geq \|u\|^2 + \mu^2 \|v\|^2 - 2\mu\theta \|u\| \|v\| \quad (54a)$$

$$= (\|u\| - \mu\theta \|v\|)^2 + \mu^2 (1 - \theta^2) \|v\|^2 \geq \mu^2 (1 - \theta^2) \|v\|^2. \quad (54b)$$

Next, we show that

$$\|u + \mu v\|^2 \geq \xi \|u + v\|^2, \quad \text{where } \xi := \frac{4\mu^2 (1 - \theta^2)}{(|1 - \mu| + \sqrt{(1 - \mu)^2 + 4\mu(1 - \theta^2)})^2}. \quad (55)$$

Observe that $|1 - \mu| + \sqrt{(1 - \mu)^2 + 4\mu(1 - \theta^2)} \geq \sqrt{4\mu(1 - \theta^2)} > 0$, so

$$\xi = \frac{4\mu^2 (1 - \theta^2)}{(|1 - \mu| + \sqrt{(1 - \mu)^2 + 4\mu(1 - \theta^2)})^2} \leq \frac{4\mu^2 (1 - \theta^2)}{4\mu(1 - \theta^2)} = \mu. \quad (56)$$

Now (55) is equivalent to

$$(1 - \xi) \|u\|^2 + (\mu^2 - \xi) \|v\|^2 + 2(\mu - \xi) \langle u, v \rangle \geq 0. \quad (57)$$

Since $\mu - \xi \geq 0$ and $\langle u, v \rangle \geq -\theta \|u\| \|v\|$, it suffices to prove that

$$\forall u \in X, \forall v \in X, \quad (1 - \xi) \|u\|^2 + (\mu^2 - \xi) \|v\|^2 - 2\theta(\mu - \xi) \|u\| \|v\| \geq 0. \quad (58)$$

The latter one can be written as

$$\forall u \in X, \forall v \in X, \quad \begin{bmatrix} \|u\| & \|v\| \end{bmatrix} M \begin{bmatrix} \|u\| \\ \|v\| \end{bmatrix} \geq 0 \quad \text{with} \quad M := \begin{bmatrix} 1 - \xi & \theta(\xi - \mu) \\ \theta(\xi - \mu) & \mu^2 - \xi \end{bmatrix}, \quad (59)$$

which is equivalent to positive semidefiniteness of M . Because M is a symmetric 2×2 matrix whose trace $(1 - \xi) + (\mu^2 - \xi) = (1 - \mu)^2 + 2(\mu - \xi) \geq 0$, it is positive semidefinite if and only if its determinant

$$(1 - \xi)(\mu^2 - \xi) - \theta^2(\mu - \xi)^2 = (1 - \theta^2)\xi^2 - (1 + \mu^2 - 2\mu\theta^2)\xi + \mu^2(1 - \theta^2) \geq 0. \quad (60)$$

Finally, one can directly check that ξ in (55) is a solution of (60). The proof is complete. \blacksquare

Lemma 4.4. *Let A and B be closed subsets of X with $A \cap B \neq \emptyset$, U a nonempty subset of X , $\lambda, \mu \in]0, 2]$, $\alpha \in \mathbb{R}_{++}$, and $\nu \in \mathbb{R}_{++}$. Suppose that $T_{\lambda, \mu}^\alpha$ is $(A \cap B, \nu)$ -quasi coercive on $U \cap L$ with $L := \text{aff}(A \cup B)$. Then $T_{\lambda, \mu}^\alpha$ is $((A \cap B) + (L - L)^\perp, \nu)$ -quasi coercive on $P_L^{-1}(U \cap L)$. Moreover, if $\min\{\lambda, \mu\} < 2$, then $T_{\lambda, \mu}^\alpha$ is also $(A \cap B, \nu')$ -quasi coercive on $P_L^{-1}(U \cap L)$ with $\nu' := \min\{\nu, \alpha(\lambda + \mu - \lambda\mu)\}$.*

Proof. Let $x \in P_L^{-1}(U \cap L)$ and $x_+ \in T_{\lambda, \mu}^\alpha x$. Define $y = P_L x$ and $y_+ = P_L x_+$, then $y \in U \cap L$ and, by Lemma 2.9(i), $y_+ \in T_{\lambda, \mu}^\alpha y$. By assumption,

$$\|y - y_+\| \geq \nu d_{A \cap B}(y). \quad (61)$$

Setting $\eta := (1 - \alpha) + \alpha(1 - \lambda)(1 - \mu)$, we obtain from Lemma 2.9(iii) that

$$\|x - x_+\|^2 = \|y - y_+\|^2 + (1 - \eta)^2 \|x - y\|^2 \geq \|y - y_+\|^2. \quad (62)$$

By Lemma 2.1(i), $x - y = x - P_L x \in (L - L)^\perp$ and then, by Lemma 2.2(v),

$$d_{A \cap B}(y) = d_{(A \cap B) + (L - L)^\perp}(y) = d_{(A \cap B) + (L - L)^\perp}(y + (x - y)) = d_{(A \cap B) + (L - L)^\perp}(x). \quad (63)$$

This together with (61) and (62) yields

$$\|x - x_+\| \geq \|y - y_+\| \geq \nu d_{A \cap B}(y) = \nu d_{(A \cap B) + (L - L)^\perp}(x), \quad (64)$$

i.e., $T_{\lambda, \mu}^\alpha$ is $((A \cap B) + (L - L)^\perp, \nu)$ -quasi coercive on $P_L^{-1}(U \cap L)$.

Now assume that $\min\{\lambda, \mu\} < 2$. Then $\eta < 1$. Combining (61), (62), and Lemma 2.2(iv), we deduce that

$$\|x - x_+\|^2 \geq \nu^2 d_{A \cap B}^2(y) + (1 - \eta)^2 d_L^2(x) \quad (65a)$$

$$\geq \min\{\nu^2, (1 - \eta)^2\} (d_{A \cap B}^2(y) + d_L^2(x)) = (\nu')^2 d_{A \cap B}^2(x), \quad (65b)$$

which means that $T_{\lambda, \mu}^\alpha$ is $(A \cap B, \nu')$ -quasi coercive on $P_L^{-1}(U \cap L)$. ■

Proposition 4.5 (quasi coercivity of gDR operators under affine-hull regularity). *Let A and B be closed subsets of X such that $A \cap B \neq \emptyset$. Let also $w \in A \cap B$, $\varepsilon \in [0, 1/3]$, $\delta \in \mathbb{R}_{++}$, $\kappa \in \mathbb{R}_{++}$, $\lambda, \mu \in]0, 2]$, and $\alpha \in \mathbb{R}_{++}$. Suppose that A is (ε, δ) -regular at w , that $\{A, B\}$ is κ -linearly regular on $\mathbf{B}(w; \delta/2)$, and that the CQ-number $\theta := \theta_{A, B, L}(w, \sqrt{2}\delta) < 1$ with $L := \text{aff}(A \cup B)$. Define*

$$\nu := \frac{\alpha \sqrt{1 - \theta^2}}{\kappa} \min \left\{ \lambda, \frac{2\mu}{|1 - \mu| + \sqrt{(1 - \mu)^2 + 4\mu(1 - \theta^2)}} \right\}. \quad (66)$$

Then the following hold:

- (i) $T_{\lambda, \mu}^\alpha$ is $(A \cap B, \nu)$ -quasi coercive on $\mathbf{B}(w; \delta/2) \cap L$.
- (ii) $T_{\lambda, \mu}^\alpha$ is $((A \cap B) + (L - L)^\perp, \nu)$ -quasi coercive on $\mathbf{B}(w; \delta/2)$.
- (iii) If $\min\{\lambda, \mu\} < 2$, then $T_{\lambda, \mu}^\alpha$ is $(A \cap B, \nu')$ -quasi coercive on $\mathbf{B}(w; \delta/2)$ with $\nu' := \min\{\nu, \alpha(\lambda + \mu - \lambda\mu)\}$.

Additionally, if $\lambda \neq 1$, A is an affine subspace of X , and $\{A, B\}$ is κ -linearly regular on $\mathbf{B}(w; \sqrt{2}\delta/2)$, then we can choose

$$\nu = \frac{\alpha \sqrt{1 - \theta^2}}{\kappa} \min\{\lambda, \mu|1 - \lambda|\}. \quad (67)$$

Proof. First, it follows from the definition of the CQ-number θ that

$$\left. \begin{aligned} a \in A \cap \mathbf{B}(w; \sqrt{2}\delta), \quad b \in B \cap \mathbf{B}(w; \sqrt{2}\delta), \\ u \in N_A^{\text{prox}}(a) \cap (L - a), \quad v \in N_B^{\text{prox}}(b) \cap (L - b) \end{aligned} \right\} \implies \langle u, v \rangle \geq -\theta \|u\| \|v\|. \quad (68)$$

To prove (i), in view of Lemma 4.1, it suffices to prove that $P_B^\mu P_A^\lambda$ is $(A \cap B, \nu/\alpha)$ -quasi coercive on $\mathbf{B}(w; \delta/2) \cap L$. Let $x \in \mathbf{B}(w; \delta/2) \cap L$ and $s \in P_B^\mu P_A^\lambda x$. Then there exist $a \in P_A x$, $r \in P_A^\lambda x$, and $b \in P_B r$ such that

$$x - r = \lambda(x - a) \quad \text{and} \quad r - s = \mu(r - b). \quad (69)$$

We note that $r \in P_A^\lambda(L) \subseteq L$ due to Lemma 2.2(i). By Fact 2.3(i), $a \in P_A x \subseteq A \cap \mathbf{B}(w; \delta)$. Since $\varepsilon \in [0, 1/3]$, Fact 2.3(ii) yields $r \in \mathbf{B}(w; \sqrt{2}\delta/2)$. Again by Fact 2.3(i), $b \in P_B r \subseteq B \cap \mathbf{B}(w; \sqrt{2}\delta)$. Now as $x - r = \lambda(x - a) \in N_A^{\text{prox}}(a) \cap (L - a)$ and $r - s = \mu(r - b) \in N_B^{\text{prox}}(b) \cap (L - b)$, it follows from (68) that

$$\langle x - a, r - s \rangle \geq -\theta \|x - a\| \|r - s\| \quad \text{and} \quad \langle x - r, r - b \rangle \geq -\theta \|x - r\| \|r - b\|. \quad (70)$$

Applying Lemma 4.3 with $(u, v) = (r - s, x - a)$, we obtain

$$\|x - s\|^2 = \|(x - r) + (r - s)\|^2 = \|\lambda(x - a) + (r - s)\|^2 \quad (71a)$$

$$\geq \lambda^2(1 - \theta^2) \|x - a\|^2 = \lambda^2(1 - \theta^2) d_A^2(x); \quad (71b)$$

and with $(u, v) = (x - r, r - b)$, we obtain

$$\|x - s\|^2 = \|(x - r) + \mu(r - b)\|^2 \geq \max\{\mu^2(1 - \theta^2) \|r - b\|^2, \xi \|x - b\|^2\} \quad (72a)$$

$$\geq \max\{\mu^2(1 - \theta^2) d_B^2(r), \xi d_B^2(x)\}, \quad (72b)$$

where $\xi := \frac{4\mu^2(1-\theta^2)}{(|1-\mu| + \sqrt{(1-\mu)^2 + 4\mu(1-\theta^2)})^2}$. Combining with the κ -linear regularity of $\{A, B\}$ yields

$$\|x - s\| \geq \min\{\lambda\sqrt{1-\theta^2}, \sqrt{\xi}\} \max\{d_A(x), d_B(x)\} \quad (73a)$$

$$\geq \frac{\sqrt{1-\theta^2}}{\kappa} \min\left\{\lambda, \frac{2\mu}{|1-\mu| + \sqrt{(1-\mu)^2 + 4\mu(1-\theta^2)}}\right\} d_{A \cap B}(x) = \frac{\nu}{\alpha} d_{A \cap B}(x), \quad (73b)$$

and (i) is proved. Now applying Lemma 4.4 and noting from Lemma 2.1(iii) that $\mathbf{B}(w; \delta/2) \subseteq P_L^{-1}(\mathbf{B}(w; \delta/2) \cap L)$, we get (ii) and (iii).

In addition, suppose that $\lambda \neq 1$, that A is an affine subspace of X , and that $\{A, B\}$ is κ -linearly regular on $\mathbf{B}(w; \sqrt{2}\delta/2)$. Then $x - a = \frac{1}{1-\lambda}(r - a)$ and, by (71),

$$\|x - s\|^2 \geq \lambda^2(1 - \theta^2) \|x - a\|^2 = \frac{\lambda^2(1 - \theta^2)}{(1 - \lambda)^2} \|r - a\|^2 \geq \frac{\lambda^2(1 - \theta^2)}{(1 - \lambda)^2} d_A^2(r). \quad (74)$$

Together with (72) and the κ -linear regularity of $\{A, B\}$ on $\mathbf{B}(w; \sqrt{2}\delta/2)$, we get

$$\|x - s\| \geq \sqrt{1 - \theta^2} \min\left\{\frac{\lambda}{|1 - \lambda|}, \mu\right\} \max\{d_A(r), d_B(r)\} \quad (75a)$$

$$\geq \frac{\sqrt{1 - \theta^2}}{\kappa} \min\left\{\frac{\lambda}{|1 - \lambda|}, \mu\right\} d_{A \cap B}(r). \quad (75b)$$

By applying Lemma 2.2(iv) (with $C = A \cap B$ and $L = A$),

$$d_{A \cap B}^2(r) = d_{A \cap B}^2(P_A x) + (1 - \lambda)^2 d_A^2(x) \quad (76a)$$

$$\geq (1 - \lambda)^2 (d_{A \cap B}^2(P_A x) + d_A^2(x)) = (1 - \lambda)^2 d_{A \cap B}^2(x), \quad (76b)$$

and the conclusion follows. \blacksquare

Corollary 4.6. *Let A and B be closed subsets of X such that $A \cap B \neq \emptyset$. Let also $w \in A \cap B$, $\lambda, \mu \in]0, 2]$, and $\alpha \in \mathbb{R}_{++}$. Suppose that A is superregular at w and $\{A, B\}$ is affine-hull regular at w . Define $L := \text{aff}(A \cup B)$. Then the following hold:*

- (i) $T_{\lambda, \mu}^\alpha$ is $(A \cap B)$ -quasi coercive on $\mathbf{B}(w; \delta) \cap L$ for some $\delta \in \mathbb{R}_{++}$.
- (ii) $T_{\lambda, \mu}^\alpha$ is $((A \cap B) + (L - L)^\perp)$ -quasi coercive around w .
- (iii) If $\min\{\lambda, \mu\} < 2$, then $T_{\lambda, \mu}^\alpha$ is $(A \cap B)$ -quasi coercive around w .

Proof. By superregularity, affine-hull regularity, Proposition 2.5(i), and Proposition 2.8, we can find $\varepsilon \in [0, 1/3]$, $\delta \in \mathbb{R}_{++}$, and $\kappa \in \mathbb{R}_{++}$ such that A is (ε, δ) -regular at w , that $\{A, B\}$ is κ -linearly regular on $\mathbf{B}(w; \delta/2)$, and that the CQ-number $\theta_{A, B, L}(w, \sqrt{2}\delta) < 1$. The conclusion then follows from Proposition 4.5. \blacksquare

Remark 4.7. In Proposition 4.5 and Corollary 4.6, the term $(L - L)^\perp$ is indeed indispensable when $\lambda = \mu = 2$. For instance, suppose A and B are two arbitrary closed sets in $X = \mathbb{R}^3$ with $w \in A \cap B$ and $L = \text{aff}(A \cup B) = \mathbb{R}^2 \times \{0\}$. Then $(L - L)^\perp = \{(0, 0, \xi) \mid \xi \in \mathbb{R}\}$. For any $\nu, \delta \in \mathbb{R}_{++}$, taking $x = w + (0, 0, \delta)$, we have $x_+ = x = T_{2, 2}^\alpha x$ due to Lemma 2.10(ii), and $d_{A \cap B}(x) = d_L(x) = \delta$. Hence, $0 = \|x_+ - x\| < \nu d_{A \cap B}(x) = \nu \delta$, i.e., $T_{2, 2}^\alpha$ fails to be $(A \cap B, \nu)$ -quasi coercive on $\mathbf{B}(w; \delta)$.

Remark 4.8 (improved quasi coercivity constant for DR operator). As we will see in the next part (see Remark 5.5), larger quasi coercivity constant will improve the linear convergence rate. In this aspect, Proposition 4.5 indeed subsumes [32, Lemma 4.2] and [24, Lemma 3.14]; and moreover, provides *larger (local) quasi coercivity constant* for the classical DR operator ($\lambda = \mu = 2$ and $\alpha = 1/2$). To see this, we consider the DR operator $T := T_{2, 2}^{1/2}$ for the pair (A, B) of closed subsets of X and $w \in A \cap B$. Assume that A is superregular at w and that $\{A, B\}$ is strongly regular at w , i.e., the limiting CQ-number $\bar{\theta}_{A, B}(w) < 1$ (due to Proposition 2.8). Then for any $\theta \in]\bar{\theta}, 1[$, there exist $\varepsilon \in [0, 1/4[$, $\delta \in \mathbb{R}_{++}$, and $\kappa \in \mathbb{R}_+$ such that

- (i) A is $(\varepsilon, 2\delta)$ -regular at w ;
- (ii) θ is the CQ-number at w associated with (A, B) and 3δ , which implies that

$$\begin{cases} a \in A \cap \mathbf{B}(w; 3\delta), b \in B \cap \mathbf{B}(w; 3\delta), \\ u \in N_A^{\text{prox}}(a), v \in N_B^{\text{prox}}(b) \end{cases} \implies \langle u, v \rangle \geq -\theta \|u\| \|v\|; \quad (77)$$

- (iii) $\{A, B\}$ is κ -linearly regular on $\mathbf{B}(w; 2\delta)$.

Now, on the one hand, [32, Lemma 4.2] derives that T is $(A \cap B, \hat{\nu})$ -quasi coercive on $\mathbf{B}(w; \delta)$ with $\hat{\nu} = \frac{\sqrt{1-\theta}}{\kappa\sqrt{5}}$. On the other hand, Proposition 4.5 derives that T is $(A \cap B, \nu)$ -quasi coercive on $\mathbf{B}(w; \delta)$ with $\nu = \frac{\sqrt{1-\theta^2}}{\kappa} \frac{2}{1+\sqrt{1+8(1-\theta^2)}}$. It is clear that $\nu \geq \frac{\sqrt{1-\theta^2}}{\kappa} \frac{2}{1+\sqrt{9}} > \frac{\sqrt{1-\theta}}{\kappa} \frac{1}{\sqrt{5}} = \hat{\nu}$ regardless of $\theta \in]0, 1[$.

In addition, if A is an affine subspace, then Proposition 4.5 also improves the estimate in [24, Lemma 3.14]. Indeed, under the assumptions made, on the one hand, [24, Lemma 3.14] implies that T is $(A \cap B, \hat{\nu})$ -quasi coercive on $\mathbf{B}(w; \delta)$ with $\hat{\nu} = \frac{\sqrt{1-\theta}}{\kappa}$. On the other hand, Proposition 4.5 yields that T is $(A \cap B, \nu)$ -quasi coercive on $\mathbf{B}(w; \delta)$ with $\nu = \frac{\sqrt{1-\theta^2}}{\kappa} > \hat{\nu}$.

Similarly, we can show that Proposition 4.5 also improves quasi coercivity constant derived in [19, Proposition 3.8] for gDR operators.

The following global version of Proposition 4.5 is an extension of [14, Lemma 4.3].

Proposition 4.9 (global quasi coercivity of convex gDR operators). *Let $\lambda, \mu \in]0, 2]$ and $\alpha \in \mathbb{R}_{++}$. Then the following hold:*

- (i) If A and B are closed convex subsets of X such that $\text{ri } A \cap \text{ri } B \neq \emptyset$, then $T_{\lambda,\mu}^\alpha$ is $((A \cap B) + (L - L)^\perp)$ -quasi coercive on every bounded set S of X with $L := \text{aff}(A \cup B)$, and if additionally $\min\{\lambda, \mu\} < 2$, then $T_{\lambda,\mu}^\alpha$ is $(A \cap B)$ -quasi coercive on every bounded set S of X .
- (ii) If A and B are polyhedral subsets of X such that $A \cap B \neq \emptyset$, then $T_{\lambda,\mu}^\alpha$ is $\text{Fix } T_{\lambda,\mu}^\alpha$ -quasi coercive on every bounded set S of X .

Proof. First, by Lemma 2.12(i), $T_{\lambda,\mu}^\alpha$ is continuous and single-valued.

(i): Set $L := \text{aff}(A \cup B)$. Since A and B are convex with $\text{ri } A \cap \text{ri } B \neq \emptyset$, Lemma 2.13(iii) yields $\text{Fix } T_{\lambda,\mu}^\alpha \cap L = A \cap B$. Now for every $w \in A \cap B$, we derive from [11, Remark 8.2(v)] that A is superregular at w , from [11, Example 7.2(i)–(ii) and Proposition 7.5] that $\{A, B\}$ is affine-hull regular at w , and then from Corollary 4.6 that $T_{\lambda,\mu}^\alpha$ is $(A \cap B)$ -quasi coercive on $\mathbf{B}(w; \delta) \cap L$ for some $\delta \in \mathbb{R}_{++}$. In turn, Lemma 4.2 implies that $T_{\lambda,\mu}^\alpha$ is $(A \cap B)$ -quasi coercive on $S \cap L$ for every bounded set S of X . Finally, apply Lemma 4.4 and note from Lemma 2.1(ii) that $\mathbf{B}(0; \delta) \subseteq P_L^{-1}(\mathbf{B}(0; \delta) \cap L)$.

(ii): By [35, Example 12.31(a)&(d)], P_A and P_B are piecewise linear in the sense of [35, Definition 2.47(a)]. Notice that compositions of piecewise linear operators are also piecewise linear (see [36, Corollary 2.3]), and that linear combinations of piecewise linear operators are also piecewise linear. Thus, $Q := \text{Id} - T_{\lambda,\mu}^\alpha = \alpha(\text{Id} - P_B^\mu P_A^\lambda)$ is piecewise linear. By [35, Example 2.48], Q is also polyhedral (i.e., the graph of Q is a union of finitely many polyhedral sets). Noting that $Q^{-1}(0) = \text{Fix } T_{\lambda,\mu}^\alpha$ and $d_{Qx}(0) = \|Qx\| = \|x - T_{\lambda,\mu}^\alpha x\|$, and using [33, Corollary of Proposition 1], there exists $\nu \in \mathbb{R}_{++}$ and $\varepsilon \in \mathbb{R}_{++}$ such that

$$\|x - T_{\lambda,\mu}^\alpha x\| \geq \nu d_{\text{Fix } T_{\lambda,\mu}^\alpha}(x) \quad \text{whenever } \|x - T_{\lambda,\mu}^\alpha x\| < \varepsilon. \quad (78)$$

Now for every $w \in \text{Fix } T_{\lambda,\mu}^\alpha$, since $T_{\lambda,\mu}^\alpha$ is continuous, there exists $\delta \in \mathbb{R}_{++}$ such that $\|x - T_{\lambda,\mu}^\alpha x\| < \varepsilon$ for all $x \in \mathbf{B}(w; \delta)$. It follows that $T_{\lambda,\mu}^\alpha$ is $(\text{Fix } T_{\lambda,\mu}^\alpha, \nu)$ -quasi coercive on $\mathbf{B}(w; \delta)$. Applying Lemma 4.2 with $C = \text{Fix } T_{\lambda,\mu}^\alpha$ and $L = X$ completes the proof. \blacksquare

4.2. In the presence of linear regularity

Corollary 4.6 shows that superregularity and affine-hull regularity assumption is sufficient for quasi coercivity of gDR operators. We will show in Proposition 4.12 that if $\min\{\lambda, \mu\} < 2$, then affine-hull regularity can be replaced by linear regularity, a *milder* assumption (see Remark 2.6). This is a *new* result that obtains quasi coercivity for gDR operators via linear regularity and operator parameters.

For $x, y, z \in X$, we denote \widehat{xyz} the angle between two vectors $x - y$ and $z - y$, i.e.,

$$\widehat{xyz} \in [0, \pi] \quad \text{such that} \quad \cos \widehat{xyz} = \frac{\langle x - y, z - y \rangle}{\|x - y\| \cdot \|z - y\|}, \quad (79)$$

with the convention that $\widehat{xyz} = 0$ if $x = y$ or $z = y$. The following two lemmas are crucial for our analysis.

Lemma 4.10. *Let $(\varepsilon_n)_{n \in \mathbb{N}}$ be a sequence in \mathbb{R}_+ convergent to 0. Let $(x_n)_{n \in \mathbb{N}}$, $(r_n)_{n \in \mathbb{N}}$, $(s_n)_{n \in \mathbb{N}}$ and $(p_n)_{n \in \mathbb{N}}$ be sequences in X such that, for all $n \in \mathbb{N}$, $r_n \notin \{x_n, s_n, p_n\}$, and that*

$$\|x_n - s_n\| \leq \varepsilon_n (\|x_n - r_n\| + \|s_n - r_n\|). \quad (80)$$

Then $\widehat{x_n r_n s_n} \rightarrow 0$ and $\cos \widehat{x_n r_n p_n} - \cos \widehat{s_n r_n p_n} \rightarrow 0$ as $n \rightarrow +\infty$.

Proof. By assumption and Cauchy–Schwarz inequality,

$$2 \langle x_n - r_n, s_n - r_n \rangle = \|x_n - r_n\|^2 + \|s_n - r_n\|^2 - \|x_n - s_n\|^2 \quad (81a)$$

$$\geq \|x_n - r_n\|^2 + \|s_n - r_n\|^2 - \varepsilon_n^2 (\|x_n - r_n\| + \|s_n - r_n\|)^2 \quad (81b)$$

$$\geq (2 - 4\varepsilon_n^2) \|x_n - r_n\| \|s_n - r_n\|. \quad (81c)$$

It follows that

$$\cos \widehat{x_n r_n s_n} = \frac{\langle x_n - r_n, s_n - r_n \rangle}{\|x_n - r_n\| \|s_n - r_n\|} \geq 1 - 2\varepsilon_n^2 \rightarrow 1 \quad \text{as } n \rightarrow +\infty, \quad (82)$$

hence $\cos \widehat{x_n r_n s_n} \rightarrow 1$ and $\widehat{x_n r_n s_n} \rightarrow 0$ as $n \rightarrow +\infty$. Now we compute

$$\left\| \frac{x_n - r_n}{\|x_n - r_n\|} - \frac{s_n - r_n}{\|s_n - r_n\|} \right\|^2 = 2 - 2 \frac{\langle x_n - r_n, s_n - r_n \rangle}{\|x_n - r_n\| \|s_n - r_n\|} = 2(1 - \cos \widehat{x_n r_n s_n}) \rightarrow 0, \quad (83)$$

and again by Cauchy–Schwarz inequality,

$$|\cos \widehat{x_n r_n p_n} - \cos \widehat{s_n r_n p_n}| = \left| \left\langle \frac{x_n - r_n}{\|x_n - r_n\|} - \frac{s_n - r_n}{\|s_n - r_n\|}, \frac{p_n - r_n}{\|p_n - r_n\|} \right\rangle \right| \quad (84a)$$

$$\leq \left\| \frac{x_n - r_n}{\|x_n - r_n\|} - \frac{s_n - r_n}{\|s_n - r_n\|} \right\| \cdot 1 \rightarrow 0, \quad (84b)$$

which completes the proof. \blacksquare

Lemma 4.11. *Let x, r, s and p be in X and set $u := P_{L_1} p$ and $v := P_{L_2} p$ with $L_1 := \text{aff}\{x, r\}$ and $L_2 := \text{aff}\{s, r\}$. Suppose that $r \notin \{x, s, u, v\}$. Then $\|u - v\| \leq \|p - r\| \sin \widehat{urv} = \|p - r\| \sin \widehat{xrs}$.*

Proof. First, since $u \in \text{aff}\{x, r\}$ and $v \in \text{aff}\{s, r\}$, we have

$$|\cos \widehat{urv}| = \left| \left\langle \frac{u - r}{\|u - r\|}, \frac{v - r}{\|v - r\|} \right\rangle \right| = \left| \left\langle \frac{x - r}{\|x - r\|}, \frac{s - r}{\|s - r\|} \right\rangle \right| = |\cos \widehat{xrs}|, \quad (85)$$

which yields $\sin \widehat{urv} = \sin \widehat{xrs}$.

Set $q := \frac{1}{2}(p + r)$. Then $\|q - r\| = \frac{1}{2}\|p - r\|$. Using Lemma 2.1(i), $\langle p - u, r - u \rangle = 0$ and $\langle p - v, r - v \rangle = 0$. We compute

$$\|q - u\|^2 = \frac{1}{4}\|(p - u) + (r - u)\|^2 = \frac{1}{4}\|(p - u) - (r - u)\|^2 = \frac{1}{4}\|p - r\|^2 \quad (86)$$

and get $\|q - u\| = \frac{1}{2}\|p - r\|$. By the same argument for $\|q - v\|^2$, we obtain

$$\|q - r\| = \|q - u\| = \|q - v\| = \frac{1}{2}\|p - r\|. \quad (87)$$

Now define $\hat{q} := P_L q$ with $L := \text{aff}\{r, u, v\}$. By Lemma 2.1(iii),

$$\|q - r\| \geq \|\hat{q} - r\| = \|\hat{q} - u\| = \|\hat{q} - v\|, \quad (88)$$

i.e., \hat{q} is the center of the circumcircle passing r, u , and v . Applying the law of sines, we get

$$\|u - v\| = 2\|\hat{q} - r\| \sin \widehat{urv} \leq 2\|q - r\| \sin \widehat{urv} = \|p - r\| \sin \widehat{urv}. \quad (89)$$

The lemma is proved. \blacksquare

We now ready to prove our *new* result on quasi coercivity of gDR operators under linear regularity assumption.

Proposition 4.12 (quasi coercivity of gDR operators under linear regularity). *Let A and B be closed subsets of X such that $A \cap B \neq \emptyset$. Let also $w \in A \cap B$, $\lambda, \mu \in]0, 2]$, and $\alpha \in \mathbb{R}_{++}$. Suppose that $\{A, B\}$ is superregular at w and linearly regular around w and that $\min\{\lambda, \mu\} < 2$. Then $T_{\lambda, \mu}^\alpha$ is $(A \cap B)$ -quasi coercive around w .*

Proof. By assumption, there exist $\kappa \in \mathbb{R}_{++}$ and $\delta \in \mathbb{R}_{++}$ such that

$$\forall x \in \mathbf{B}(w; \delta), \quad d_{A \cap B}(x) \leq \kappa \max\{d_A(x), d_B(x)\}. \quad (90)$$

Now suppose on the contrary that $T_{\lambda, \mu}^\alpha$ is not $(A \cap B)$ -quasi coercive around w . By Lemma 4.1, $P_B^\mu P_A^\lambda$ is not $(A \cap B)$ -quasi coercive around w , which means that there exist sequences $\zeta_n \downarrow 0$, $x_n \rightarrow w$, $s_n \in P_B^\mu P_A^\lambda x_n$ such that

$$0 \leq \|x_n - s_n\| < \zeta_n d_{A \cap B}(x_n). \quad (91)$$

We find $a_n \in P_A x_n$, $r_n \in P_A^\lambda x_n$, and $b_n \in P_B r_n$ such that

$$x_n - r_n = \lambda(x_n - a_n) \quad \text{and} \quad s_n - r_n = \mu(b_n - r_n). \quad (92)$$

Without loss of generality, we assume that $\zeta_n < \min\{\frac{1}{\kappa}, \frac{\lambda}{\kappa}, \frac{\mu}{\kappa}\}$ and $x_n \in \mathbf{B}(w; \delta/(1 + \lambda)) \subset \mathbf{B}(w; \delta)$ for all $n \in \mathbb{N}$. It then follows from (90) and (91) that

$$\forall n \in \mathbb{N}, \quad \|x_n - s_n\| < \zeta_n d_{A \cap B}(x_n) \leq \zeta_n \kappa \max\{d_A(x_n), d_B(x_n)\}. \quad (93)$$

Let $n \in \mathbb{N}$. Since $a_n \in A$ and $b_n \in B$, we have $d_A(x_n) \leq \|x_n - a_n\| = \frac{1}{\lambda} \|x_n - r_n\|$ and

$$d_B(x_n) \leq \|x_n - b_n\| \leq \|x_n - r_n\| + \|b_n - r_n\| = \|x_n - r_n\| + \frac{1}{\mu} \|s_n - r_n\|. \quad (94)$$

Therefore,

$$\forall n \in \mathbb{N}, \quad \|x_n - s_n\| < \varepsilon_n (\|x_n - r_n\| + \|s_n - r_n\|), \quad \text{where } \varepsilon_n := \zeta_n \kappa \max\{1, \frac{1}{\lambda}, \frac{1}{\mu}\} \rightarrow 0. \quad (95)$$

Noting that $\|x_n - s_n\| \geq \|x_n - r_n\| - \|s_n - r_n\|$ and that $\varepsilon_n < 1$, we obtain

$$\|x_n - r_n\| \leq \frac{1 + \varepsilon_n}{1 - \varepsilon_n} \|s_n - r_n\|, \quad (96)$$

which combined with (92) yields

$$\forall n \in \mathbb{N}, \quad \|a_n - r_n\| = \frac{|\lambda - 1|}{\lambda} \|x_n - r_n\| \leq \sigma_n \|b_n - r_n\|, \quad \text{where } \sigma_n := \frac{\mu |\lambda - 1|}{\lambda} \cdot \frac{1 + \varepsilon_n}{1 - \varepsilon_n}. \quad (97)$$

Next, let $p_n \in P_{A \cap B} r_n$. Since $x_n \rightarrow w$, Fact 2.3(i) implies that $a_n, r_n, b_n, s_n, p_n \rightarrow w$. In turn, $x_n - a_n \in N_A^{\text{prox}}(a_n)$ and $r_n - b_n \in N_B^{\text{prox}}(b_n)$. By the superregularity of A and B at w , taking subsequences if necessary and without relabeling, we can assume that

$$\langle x_n - a_n, p_n - a_n \rangle \leq \varepsilon_n \|x_n - a_n\| \|p_n - a_n\|, \quad \text{and} \quad (98a)$$

$$\langle r_n - b_n, p_n - b_n \rangle \leq \varepsilon_n \|r_n - b_n\| \|p_n - b_n\|. \quad (98b)$$

Now we divide the proof into several parts.

Part 1: We claim that

$$\forall n \in \mathbb{N}, \quad x_n \notin \{r_n, a_n\} \quad \text{and} \quad r_n \notin \{s_n, b_n, p_n\} \quad (99)$$

and that

$$\widehat{s_n r_n x_n} \rightarrow 0 \quad \text{and} \quad \cos \widehat{x_n r_n p_n} - \cos \widehat{s_n r_n p_n} \rightarrow 0 \quad \text{as } n \rightarrow +\infty. \quad (100)$$

Indeed, if $x_n = r_n$ or $x_n = a_n$ for some n , then $x_n = a_n = r_n \in A$ and, by (91),

$$\zeta_n d_{A \cap B}(x_n) > \|x_n - s_n\| = \|r_n - s_n\| = \mu \|r_n - b_n\| \quad (101a)$$

$$= \mu d_B(r_n) = \mu d_B(x_n) = \mu \max\{d_A(x_n), d_B(x_n)\} \geq \frac{\mu}{\kappa} d_{A \cap B}(x_n), \quad (101b)$$

which contradicts the fact that $\zeta_n < \frac{\mu}{\kappa}$. Similarly, if $r_n = s_n$ or $r_n = b_n$ for some n , then $r_n = b_n = s_n \in B$ and, by (91),

$$\zeta_n d_{A \cap B}(x_n) > \|x_n - s_n\| = \|x_n - r_n\| = \lambda \|x_n - a_n\| \quad (102a)$$

$$= \lambda d_A(x_n) = \lambda \max\{d_A(x_n), d_B(x_n)\} \geq \frac{\lambda}{\kappa} d_{A \cap B}(x_n), \quad (102b)$$

which contradicts the fact that $\zeta_n < \frac{\lambda}{\kappa}$. So we complete (99) due to $\|p_n - r_n\| \geq d_B(r_n) = \|b_n - r_n\| = \frac{1}{\mu} \|s_n - r_n\|$. Now combining (95) and Lemma 4.10, we arrive at (100).

Part 2: Let $\bar{\sigma}_n := \max\{1, \sigma_n\}$. Since $\varepsilon_n \rightarrow 0$, the sequence $(\sigma_n)_{n \in \mathbb{N}}$ is bounded, and so is $(\bar{\sigma}_n)_{n \in \mathbb{N}}$. We claim that

$$\forall n \in \mathbb{N}, \quad \|p_n - r_n\| \leq \bar{\sigma}_n \kappa \|b_n - r_n\|, \quad (103a)$$

$$\max\{\|p_n - a_n\|, \|p_n - b_n\|\} \leq \bar{\sigma}_n (\kappa + 1) \|b_n - r_n\|. \quad (103b)$$

Let $n \in \mathbb{N}$. Since $x_n \in \mathbf{B}(w; \delta/(1 + \lambda))$ and $r_n \in P_A^\lambda x_n$, Fact 2.3(i) yields $r_n \in \mathbf{B}(w; \delta)$. Using linear regularity and (97), we estimate

$$\|p_n - r_n\| = d_{A \cap B}(r_n) \leq \kappa \max\{d_A(r_n), d_B(r_n)\} \quad (104a)$$

$$\leq \kappa \max\{\|a_n - r_n\|, \|b_n - r_n\|\} \quad (104b)$$

$$\leq \kappa \max\{\sigma_n \|b_n - r_n\|, \|b_n - r_n\|\} = \kappa \bar{\sigma}_n \|b_n - r_n\|. \quad (104c)$$

So (103a) holds. Combining with (97) gives

$$\|p_n - a_n\| \leq \|p_n - r_n\| + \|a_n - r_n\| \leq (\kappa \bar{\sigma}_n + \sigma_n) \|b_n - r_n\|, \quad (105a)$$

$$\|p_n - b_n\| \leq \|p_n - r_n\| + \|b_n - r_n\| \leq (\kappa \bar{\sigma}_n + 1) \|b_n - r_n\|. \quad (105b)$$

Thus, (103b) holds.

Part 3: We claim that $\lambda > 1$ and that there exists $\sigma \in]0, 1[$ satisfying

$$\sigma_n \leq \sigma < 1 \quad \text{and} \quad \bar{\sigma}_n = 1 \quad \text{for all } n \text{ sufficiently large.} \quad (106)$$

Let $\bar{\sigma}$ be an upper bound of the bounded sequence $(\bar{\sigma}_n)_{n \in \mathbb{N}}$. Using (98b) and (103b), we have

$$\cos \widehat{s_n r_n p_n} = \frac{\langle s_n - r_n, p_n - r_n \rangle}{\|s_n - r_n\| \|p_n - r_n\|} = \frac{\langle b_n - r_n, p_n - r_n \rangle}{\|b_n - r_n\| \|p_n - r_n\|} \quad (107a)$$

$$= \frac{\langle b_n - r_n, p_n - b_n \rangle + \|b_n - r_n\|^2}{\|b_n - r_n\| \|p_n - r_n\|} \quad (107b)$$

$$\geq \frac{-\varepsilon_n \|b_n - r_n\| \|p_n - b_n\| + \|b_n - r_n\|^2}{\|b_n - r_n\| \|p_n - r_n\|} \quad (107c)$$

$$\geq \frac{1 - \varepsilon_n \bar{\sigma} (\kappa + 1)}{\kappa \bar{\sigma}} \rightarrow \frac{1}{\kappa \bar{\sigma}} > 0. \quad (107d)$$

It follows that

$$\cos \widehat{s_n r_n p_n} > \frac{1}{2\kappa\bar{\sigma}} > 0 \quad \text{for all } n \text{ sufficiently large.} \quad (108)$$

Combining with (100) yields

$$\cos \widehat{x_n r_n p_n} > \frac{1}{4\kappa\bar{\sigma}} > 0 \quad \text{for all } n \text{ sufficiently large.} \quad (109)$$

Suppose that $\lambda \leq 1$. Using (98a), (103b) and noting that $\|p_n - r_n\| = d_{A \cap B}(r_n) \geq d_B(r_n) = \|r_n - b_n\|$, we obtain that

$$\cos \widehat{x_n r_n p_n} = \frac{\langle x_n - r_n, p_n - r_n \rangle}{\|x_n - r_n\| \|p_n - r_n\|} = \frac{\langle x_n - a_n, p_n - r_n \rangle}{\|x_n - a_n\| \|p_n - r_n\|} \quad (110a)$$

$$= \frac{\langle x_n - a_n, p_n - a_n \rangle + (\lambda - 1)\|x_n - a_n\|^2}{\|x_n - a_n\| \|p_n - r_n\|} \quad (110b)$$

$$\leq \varepsilon_n \frac{\|p_n - a_n\|}{\|p_n - r_n\|} + (\lambda - 1) \frac{\|x_n - a_n\|}{\|p_n - r_n\|} \leq \varepsilon_n \bar{\sigma}(\kappa + 1) \rightarrow 0, \quad (110c)$$

which contradicts (109). We must therefore have $\lambda > 1$. Now notice that $\frac{|\lambda-1|}{\lambda} = \frac{\lambda-1}{\lambda} \leq \frac{1}{2}$ and $\mu \leq 2$, where only one equality can happen. Thus,

$$\sigma_n = \frac{\mu(\lambda-1)}{\lambda} \cdot \frac{1+\varepsilon_n}{1-\varepsilon_n} \rightarrow \frac{\mu(\lambda-1)}{\lambda} < 1. \quad (111)$$

The claim then follows.

Part 4: Define lines $L_{1,n} := \text{aff}\{x_n, r_n\}$, $L_{2,n} := \text{aff}\{r_n, s_n\}$ and projections $u_n := P_{L_{1,n}}p_n$, $v_n := P_{L_{2,n}}p_n$. We have that

$$u_n := a_n + \eta_1 \frac{x_n - a_n}{\|x_n - a_n\|} \quad \text{with} \quad \eta_1 := \frac{\langle x_n - a_n, p_n - a_n \rangle}{\|x_n - a_n\|} \leq \varepsilon_n \|p_n - a_n\| \quad (112)$$

and that

$$v_n := b_n + \eta_2 \frac{r_n - b_n}{\|r_n - b_n\|}, \quad \text{with} \quad \eta_2 := \frac{\langle r_n - b_n, p_n - b_n \rangle}{\|r_n - b_n\|} \leq \varepsilon_n \|p_n - b_n\|, \quad (113)$$

where the upper bound for η_1 and η_2 follows from (98). By using (97), (103b), and (106), for all n sufficiently large,

$$\|u_n - r_n\| \leq \|a_n - r_n\| + \eta_1 \leq \|a_n - r_n\| + \varepsilon_n \|p_n - a_n\| \leq (\sigma + \varepsilon_n(\kappa + 1)) \|b_n - r_n\|, \quad (114a)$$

$$\|v_n - r_n\| = \| \|b_n - r_n\| - \eta_2 \|r_n - b_n\| \| \geq \|b_n - r_n\| - \varepsilon_n \|p_n - b_n\| \geq (1 - \varepsilon_n(\kappa + 1)) \|b_n - r_n\|, \quad (114b)$$

and so

$$\|u_n - v_n\| \geq \|v_n - r_n\| - \|u_n - r_n\| \geq \frac{1-\sigma}{2} \|b_n - r_n\|, \quad (115)$$

which together with (99), (103a), and (106) yields

$$\|u_n - v_n\| \geq \frac{1-\sigma}{2\kappa} \|p_n - r_n\| > 0. \quad (116)$$

On the other hand, for all n sufficiently large, $r_n \notin \{u_n, v_n\}$ due to (108) and (109). Noting also from (99) that $r_n \notin \{x_n, s_n\}$, we then apply Lemma 4.11 to get $\|u_n - v_n\| \leq \|p_n - r_n\| \sin \widehat{x_n r_n s_n}$, hence

$$0 < \frac{1-\sigma}{2\kappa} \|p_n - r_n\| \leq \|p_n - r_n\| \sin \widehat{x_n r_n s_n}. \quad (117)$$

Using (100), we derive that

$$0 < \frac{1-\sigma}{2\kappa} \leq \sin \widehat{x_n r_n s_n} \rightarrow 0, \quad (118)$$

which is a contradiction. ■

Remark 4.13. In Proposition 4.12, the parameter condition $\min\{\lambda, \mu\} < 2$ cannot be removed. For example, we consider two convex (hence, superregular) sets $A := \text{epi}|\cdot| = \{(s, t) \in \mathbb{R}^2 \mid s \leq |t|\}$ and $B := \mathbb{R} \times \{0\}$ in \mathbb{R}^2 . Clearly, $\{A, B\}$ is linearly regular at $(0, 0) \in A \cap B$. Consider the DR operator $T := T_{2,2}^\alpha$ for some $\alpha \in]0, 1]$ and $x = (0, t)$ for $t < 0$. In this case $\lambda = \mu = 2$, and we check that $x = Tx$, while $d_{A \cap B}(x) = |t|$. Therefore, T is not $(A \cap B)$ -quasi coercive at $(0, 0)$.

We also obtain a global version of Proposition 4.12.

Proposition 4.14 (global quasi coercivity of convex gDR operators). *Let A and B be closed convex subsets of X such that $A \cap B \neq \emptyset$ and that $\{A, B\}$ is boundedly linearly regular. Let $\lambda, \mu \in]0, 2]$ be such that $\min\{\lambda, \mu\} < 2$, and let $\alpha \in \mathbb{R}_{++}$. Then $T_{\lambda, \mu}^\alpha$ is $(A \cap B)$ -quasi coercive on every bounded set S of X .*

Proof. On the one hand, by assumption, $T_{\lambda, \mu}^\alpha$ is a continuous single-valued operator and $\text{Fix } T_{\lambda, \mu}^\alpha = A \cap B$ due to Lemma 2.12(i) and Lemma 2.13(ii). On the other hand, for every $w \in A \cap B$, $\{A, B\}$ is superregular at w (see [11, Remark 8.2(v)]) and linearly regular around w , hence, by Proposition 4.12, $T_{\lambda, \mu}^\alpha$ is $(A \cap B)$ -quasi coercive around w . Applying Lemma 4.2 with $L = X$, we obtain that $T_{\lambda, \mu}^\alpha$ is $(A \cap B)$ -quasi coercive on every bounded set S of X . \blacksquare

As a supplement for the above result, we refer to [4, Corollary 5] for the most common *sufficient condition* that guarantees bounded linear regularity for finite systems of convex sets.

5. Convergence rate analysis

In this section, let m be a positive integer, set $I := \{1, \dots, m\}$, and let $\{C_i\}_{i \in I}$ be a system of closed (possibly nonconvex) subsets of X with $C := \bigcap_{i \in I} C_i \neq \emptyset$. Given an ordered tuple $(T_i)_{i \in I}$ of set-valued operators from X to X , the *cyclic algorithm* associated with $(T_i)_{i \in I}$ generates *cyclic sequences* $(x_n)_{n \in \mathbb{N}}$ by

$$\forall n \in \mathbb{N}, \quad x_{n+1} \in T_{n+1}x_n, \quad \text{where } x_0 \in X. \quad (119)$$

Here we adopt the convention that

$$\forall n \in \mathbb{N}, \forall i \in I, \quad T_{mn+i} := T_i. \quad (120)$$

Recall that a sequence $(x_n)_{n \in \mathbb{N}}$ is said to converge to a point \bar{x} with *R-linear rate* $\rho \in [0, 1[$ if there exists a constant $\sigma \in \mathbb{R}_+$ such that

$$\forall n \in \mathbb{N}, \quad \|x_n - \bar{x}\| \leq \sigma \rho^n. \quad (121)$$

In what follows, we denote $[\rho]_+ := \max\{0, \rho\}$ for $\rho \in \mathbb{R}$.

Theorem 5.1 (sufficient condition for linear convergence). *Let $w \in C$, $\delta \in \mathbb{R}_{++}$, and $\nu \in]0, 1]$. For every $i \in I$, let $\gamma_i \in [1, +\infty[$ and $\beta_i \in \mathbb{R}_{++}$. Let $(x_n)_{n \in \mathbb{N}}$ be a cyclic sequence generated by $(T_i)_{i \in I}$. Suppose that*

- (a) $\{C_i\}_{i \in I}$ is κ -linearly regular on $\mathbf{B}(w; \delta/2)$ for some $\kappa \in \mathbb{R}_{++}$.
- (b) For every $i \in I$, T_i is $(C_i \cap \mathbf{B}(w; \delta), \gamma_i, \beta_i)$ -quasi firmly Fejér monotone and (C_i, ν) -quasi coercive on $\mathbf{B}(w; \delta/2)$.

Set $\Gamma := (\gamma_1 \cdots \gamma_m)^{1/2}$, $\delta_0 := \frac{\delta}{2\Gamma} \gamma_m^{1/2}$, and

$$\rho := \left[\Gamma^2 - \frac{\nu^2}{\kappa^2} \left(\sum_{i \in I} \frac{1}{\beta_i} \right)^{-1} \right]_+^{1/2}. \quad (122)$$

Then the following hold:

- (i) $\forall x_0 \in \mathbf{B}(w; \delta_0), \quad d_C(x_m) \leq \rho d_C(x_0).$
- (ii) If $\rho < 1$, then whenever $(x_{mn})_{n \in \mathbb{N}} \subset \mathbf{B}(w; \delta_0)$ or $x_0 \in \mathbf{B}(w; \frac{\delta_0(1-\rho)}{2+\Gamma-\rho})$, the sequence $(x_n)_{n \in \mathbb{N}}$ converges R -linearly to a point $\bar{x} \in C$ with rate $\rho^{1/m}$.
- (iii) If $\gamma_i = 1$ for every $i \in I$, then whenever $x_0 \in \mathbf{B}(w; \delta/2)$, the sequence $(x_n)_{n \in \mathbb{N}}$ converges R -linearly to a point $\bar{x} \in C$ with rate

$$\left[1 - \frac{\nu^2}{\kappa^2} \left(\sum_{i \in I} \frac{1}{\beta_i} \right)^{-1} \right]_+^{1/2m}. \quad (123)$$

Proof. (i)&(ii): This follows from [19, Theorem 4.5].

(iii): Assume that $\gamma_i = 1$ for every $i \in I$. Then $\Gamma = 1$, $\delta_0 = \delta/2$, and

$$\rho = \left[1 - \frac{\nu^2}{\kappa^2} \left(\sum_{i \in I} \frac{1}{\beta_i} \right)^{-1} \right]_+^{1/2m} < 1. \quad (124)$$

Let $x_0 \in \mathbf{B}(w; \delta/2)$. Since for every $i \in I$, T_i is $(C_i \cap \mathbf{B}(w; \delta), 1, \beta_i)$ -quasi firmly Fejér monotone and hence $(C_i \cap \mathbf{B}(w; \delta), 1)$ -quasi Fejér monotone on $\mathbf{B}(w; \delta/2)$, we obtain from [19, Lemma 3.4(ii)] that $(x_n)_{n \in \mathbb{N}} \subset \mathbf{B}(w; \delta/2)$. Now apply (ii). \blacksquare

From now on, let ℓ be a positive integer and set $J := \{1, \dots, \ell\}$. For every $j \in J$, let

$$\lambda_j, \mu_j \in]0, 2], \quad \widehat{\beta}_j := \left(\frac{\lambda_j}{2 - \lambda_j} + \frac{\mu_j}{2 - \mu_j} \right)^{-1}, \quad \alpha_j \in]0, 1 + \widehat{\beta}_j[, \quad \text{and} \quad (125a)$$

$$s_j, t_j \in I \quad \text{such that } s_j \neq t_j \text{ and } \{s_j\}_{j \in J} \cup \{t_j\}_{j \in J} = I. \quad (125b)$$

Setting

$$\forall j \in J, \quad T_j := (1 - \alpha_j) \text{Id} + \alpha_j P_{C_{t_j}}^{\mu_j} P_{C_{s_j}}^{\lambda_j}, \quad (126)$$

we study the *cyclic generalized Douglas–Rachford algorithm* defined by $(T_j)_{j \in J}$, which includes several algorithms in the literature, for example, the *cyclically anchored DR algorithm* [14] and the *cyclic DR algorithm* [17]; see [19, Section 5.3] for more details. We also say that the cyclic gDR algorithm is *connected* if for every $i, k \in I$, there exists a path

$$\{(i_1, i_2), (i_2, i_3), \dots, (i_{q-1}, i_q)\} \subseteq \{(s_j, t_j), (t_j, s_j)\}_{j \in J} \quad \text{such that } i_1 = i, i_q = k; \quad (127)$$

in other words, $I = \{1, \dots, m\}$ and $\{(s_j, t_j)\}_{j \in J}$ respectively represent the vertices and edges of a *connected undirected graph*. Here, a graph is undirected if every edge is bidirectional, and is connected if every two vertices can be linked by some path composed by the edges. It is clear that the cyclically anchored DR algorithm and the cyclic DR algorithm are connected.

Next, for every $j \in J$, we define

$$L_j := \text{aff}(C_{s_j} \cup C_{t_j}) \quad \text{and} \quad Z_j := \begin{cases} (C_{s_j} \cap C_{t_j}) + (L_j - L_j)^\perp & \text{if } \lambda_j = \mu_j = 2, \\ C_{s_j} \cap C_{t_j} & \text{otherwise} \end{cases} \quad (128)$$

and note from Lemma 2.10(iii) that $Z_j \subseteq \text{Fix } T_j$. A relationship between $\{Z_j\}_{j \in J}$ and $\{C_i\}_{i \in I}$ is given as follows.

Lemma 5.2 ($\{Z_j\}_{j \in J}$ vs. $\{C_i\}_{i \in I}$). *Let $w \in \bigcap_{i \in I} C_i$. Suppose that $\{C_{s_j}, C_{t_j}\}$ is strongly regular at w whenever $\lambda_j = \mu_j = 2$. Then*

$$\bigcap_{j \in J} Z_j = \bigcap_{i \in I} C_i. \quad (129)$$

If, in addition, $\{C_i\}_{i \in I}$ is κ -linearly regular around w , then so is $\{Z_j\}_{j \in J}$.

Proof. By assumption and Proposition 2.5, $L_j = X$ whenever $\lambda_j = \mu_j = 2$. Thus, (128) implies that $Z_j = C_{s_j} \cap C_{t_j}$ for every $i \in J$. So it follows from (125b) that

$$\bigcap_{j \in J} Z_j = \bigcap_{j \in J} (C_{s_j} \cap C_{t_j}) = \bigcap_{i \in I} C_i. \quad (130)$$

Next, we note that

$$\forall j \in J, \forall x \in X, \quad \max\{d_{C_{s_j}}(x), d_{C_{t_j}}(x)\} \leq d_{Z_j}(x). \quad (131)$$

Taking the maximum over all $j \in J$ and using (125b) yield

$$\forall x \in X, \quad \max_{i \in I} d_{C_i}(x) \leq \max_{j \in J} d_{Z_j}(x). \quad (132)$$

Now suppose in addition that $\{C_i\}_{i \in I}$ is κ -linearly regular around w . Then combining with (130) and (132), we deduce that $\{Z_j\}_{j \in J}$ is also κ -linearly regular around w . \blacksquare

Lemma 5.3 (shadows of common fixed points). *Suppose that the cyclic gDR algorithm is connected. Then*

$$\forall \bar{x} \in \bigcap_{j \in J} Z_j, \quad P_{C_1} \bar{x} = \cdots = P_{C_m} \bar{x} = P_C \bar{x} \in C. \quad (133)$$

Proof. Let $\bar{x} \in \bigcap_{j \in J} Z_j$. By Lemma 2.10(ii), $P_{C_{s_j}} \bar{x} = P_{C_{t_j}} \bar{x} \in C_{s_j} \cap C_{t_j}$ for every $j \in J$. Since the algorithm is connected, in view of (125b) and (127), we conclude that $P_{C_1} \bar{x} = \cdots = P_{C_m} \bar{x} \in C$, which also implies that $P_{C_i} \bar{x} = P_C \bar{x}$ for every $i \in I$. \blacksquare

We note from Propositions 2.5 and 2.8 that $\{C_{s_j}, C_{t_j}\}$ is affine-hull regular at w if and only if the CQ-number $\theta_{C_{s_j}, C_{t_j}, L_j}(w, \delta) < 1$ for some $\delta \in \mathbb{R}_{++}$, in which case $\{C_{s_j}, C_{t_j}\}$ is linearly regular around w . This perspective supports the use of our assumptions in the following.

Theorem 5.4 (linear convergence under affine-hull regularity). *Let $w \in C := \bigcap_{i \in I} C_i$, $\varepsilon \in [0, 1/3]$, $\delta \in \mathbb{R}_{++}$, $\kappa \in \mathbb{R}_{++}$, and $\kappa_j \in \mathbb{R}_{++}$ for every $j \in J$. Suppose that*

- (a) $\{Z_j\}_{j \in J}$ is κ -linearly regular on $B(w; \delta/2)$ and for every $j \in J$, $\{C_{s_j}, C_{t_j}\}$ is κ_j -linearly regular on $B(w; \delta/2)$.
- (b) For every $i \in I$, C_i is $(\varepsilon, \sqrt{2}\delta)$ -regular at w .
- (c) For every $j \in J$, the CQ-number $\theta_j := \theta_{C_{s_j}, C_{t_j}, L_j}(w, \sqrt{2}\delta) < 1$.
- (d) Setting for every $j \in J$,

$$\gamma_j := 1 - \alpha_j + \alpha_j \left(1 + \frac{\lambda_j \varepsilon}{1 - \varepsilon}\right) \left(1 + \frac{\mu_j \varepsilon}{1 - \varepsilon}\right), \quad (134a)$$

$$\nu_j := \frac{\alpha_j \sqrt{1 - \theta_j^2}}{\kappa} \min \left\{ \lambda_j, \frac{2\mu_j}{|1 - \mu_j| + \sqrt{(1 - \mu_j)^2 + 4\mu_j(1 - \theta_j^2)}} \right\}, \quad \text{and} \quad (134b)$$

$$\nu'_j := \begin{cases} \nu_j & \text{if } \lambda_j = \mu_j = 2 \text{ or } L_j = X, \\ \min\{\nu_j, \alpha_j(\lambda_j + \mu_j - \lambda_j \mu_j)\} & \text{otherwise,} \end{cases} \quad (134c)$$

it holds that

$$\rho := \left[\Gamma^2 - \frac{\nu^2}{\kappa^2} \left(\sum_{j \in J} \frac{\alpha_j}{1 - \alpha_j + \hat{\beta}_j} \right)^{-1} \right]_+^{\frac{1}{2\varepsilon}} < 1, \quad (135)$$

where $\Gamma := (\gamma_1 \cdots \gamma_\ell)^{1/2}$ and $\nu := \min_{j \in J} \{\nu'_j, 1\}$.

Then if either $(x_{\ell n})_{n \in \mathbb{N}} \subset \mathbf{B}(w; \delta_0)$ or $x_0 \in \mathbf{B}(w; \frac{\delta_0(1-\rho)}{2+\Gamma-\rho})$, where $\delta_0 := \frac{\delta}{2\Gamma} \gamma_\ell^{1/2}$, the cyclic sequence $(x_n)_{n \in \mathbb{N}}$ generated by $(T_j)_{j \in J}$ converges R -linearly with rate ρ to a point

$$\bar{x} \in \bigcap_{j \in J} Z_j \subseteq \bigcap_{j \in J} \text{Fix } T_j. \quad (136)$$

Additionally, if the cyclic gDR algorithm is connected, then $P_{C_1} \bar{x} = \dots = P_{C_m} \bar{x} \in C$.

Proof. Let $j \in J$. We have that $w \in C_{s_j} \cap C_{t_j}$ and, by Proposition 4.5, that T_j is (Z_j, ν'_j) - and therefore (Z_j, ν) -quasi coercive on $\mathbf{B}(w; \delta/2)$. Noting that

$$(C_{s_j} \cap C_{t_j} + (L_j - L_j)^\perp) \cap \mathbf{B}(w; \delta) \subseteq (C_{s_j} \cap C_{t_j} \cap \mathbf{B}(w; \delta)) + (L_j - L_j)^\perp \quad (137)$$

and using Proposition 3.4(ii)–(iii), we derive that T_j is $(Z_j \cap \mathbf{B}(w; \delta), \gamma_j, \frac{1-\alpha_j+\hat{\beta}_j}{\alpha_j})$ -quasi firmly Fejér monotone on $\mathbf{B}(w; \delta/2)$. Thus, applying Theorem 5.1(ii) to $(T_j)_{j \in J}$ and the corresponding sets $(Z_j)_{j \in J}$, we obtain R -linear convergence of the cyclic sequence $(x_n)_{n \in \mathbb{N}}$ to a point \bar{x} satisfying (136). Now Lemma 5.3 completes the proof. \blacksquare

Remark 5.5 (sharper convergence rate). Theorem 5.4 indeed generalizes [19, Theorem 5.21], and moreover, provides sharper convergence rate under the *same* assumptions. To see this, let $w \in \bigcap_{i \in I} C_i$ and suppose all assumptions [19, Theorem 5.21] are fulfilled. By Proposition 2.5, Proposition 2.8, and Lemma 5.2, all assumptions in Theorem 5.4 are also satisfied on some neighborhood of w . It can be seen that the linear convergence rate ρ in Theorem 5.4 is smaller than the one obtained in the proof of [19, Theorem 5.21] because its corresponding quasi coercivity constant ν is better (see Remark 4.8).

Remark 5.6 (shadows of common fixed points). As shown in Theorem 5.4, the cyclic gDR algorithm converges (locally) to the set of common fixed points. However, without *additional* conditions, one should not expect the limit points or their projections (or “shadows”) onto C_i ’s to lie in the intersection $C := \bigcap_{i \in I} C_i$, which means that those points might not solve the feasibility problem! We will illustrate this phenomenon in the following example.

Suppose that $C_1 = \mathbb{R}_+ \times \mathbb{R} \times \{0\}$, $C_2 = \mathbb{R} \times \mathbb{R}_+ \times \{0\}$, $C_3 = \{(\xi, \zeta, \zeta) \mid \xi, \zeta \in \mathbb{R}\}$, and $C_4 = \mathbb{R}_+ \times \mathbb{R}^2$ in $X = \mathbb{R}^3$. Consider $J = \{1, 2\}$, $\lambda_1 = \mu_1 = 2$, $\min\{\lambda_2, \mu_2\} < 2$, $(s_1, t_1) = (1, 2)$, and $(s_2, t_2) = (3, 4)$. So $T_1 := T_{2,2}^{\alpha_1}$ and $T_2 := T_{\lambda_2, \mu_2}^{\alpha_2}$ are the gDR operators for (C_1, C_2) and (C_3, C_4) , respectively. Then

$$L_1 = \text{aff}(C_1 \cup C_2) = \mathbb{R}^2 \times \{0\}, \quad Z_1 = (C_1 \cap C_2) + (L_1 - L_1)^\perp = \mathbb{R}_+^2 \times \mathbb{R}, \quad (138a)$$

$$L_2 = \text{aff}(C_3 \cup C_4) = \mathbb{R}^3, \quad Z_2 = (C_3 \cap C_4) = \{(\xi, \zeta, \zeta) \mid \xi, \zeta \in \mathbb{R}_+\}. \quad (138b)$$

Now take $\bar{x} = (1, 1, 1) \in Z_1 \cap Z_2 \subseteq \text{Fix } T_1 \cap \text{Fix } T_2$. Then $P_{C_1} \bar{x} = P_{C_2} \bar{x} = (1, 1, 0) \neq (1, 1, 1) = P_{C_3} \bar{x} = P_{C_4} \bar{x}$. So these projections are not identical and neither of them lies in the intersection $C_1 \cap C_2 \cap C_3 \cap C_4 = \mathbb{R}_+ \times \{0\} \times \{0\}$.

Remark 5.7 (on linear regularity moduli). To the best of our knowledge, there is no known results on the relationship between the linear regularity modulus κ of the entire system $\{C_i\}_{i \in I}$ and those of its subsystems. So we present here two simple examples showing that one modulus can be arbitrarily large while others remain bounded.

We will need the following formula whose proof is elementary: For two intersecting hyperplanes A and B of X with two nonparallel unit normal vectors n_A and n_B , the system $\{A, B\}$ is linearly regular on X with modulus

$$\sqrt{\frac{2}{1 - |\langle n_A, n_B \rangle|}}. \quad (139)$$

- (i) **Arbitrarily large linear regularity modulus for subsystem:** Let $\varepsilon \in \mathbb{R}_{++}$ and suppose that $C_1 = \mathbb{R} \times \{0\}$, $C_2 = \{\xi(1, \varepsilon) \mid \xi \in \mathbb{R}\}$, $C_3 = \{0\} \times \mathbb{R}$ are lines in $X = \mathbb{R}^2$. One can check that $\bigcap_{i=1}^3 C_i = \{(0, 0)\}$ and that $\{C_i\}_{i \in \{1, 2, 3\}}$ is κ -linearly regular on X with $\kappa = \sqrt{2}$. As noticed, $\{C_1, C_2\}$ is linearly regular on X . Let κ' be a linear regularity modulus of $\{C_1, C_2\}$ around $w = (0, 0)$ and take $x = (\varepsilon, \varepsilon^2) \in C_2$. Then, as ε is sufficiently small, $\sqrt{\varepsilon^2 + \varepsilon^4} = d_{C_1 \cap C_2}(x) \leq \kappa' \max\{d_{C_1}(x), d_{C_2}(x)\} = \kappa' \varepsilon^2$. We deduce that $\kappa' \geq \sqrt{1 + 1/\varepsilon^2}$ and so κ' can be arbitrarily large while κ remains constant.
- (ii) **Arbitrarily large linear regularity modulus for entire system:** Let $\varepsilon \in \mathbb{R}_{++}$ and suppose that $X = \mathbb{R}^3$. Consider the planes $C_1 = \mathbb{R}^2 \times \{0\}$, $C_2 = \{0\} \times \mathbb{R}^2$, C_3 being the plane defined by $\{(0, 0, 0), (\varepsilon, 1, 0), (0, 1, \varepsilon)\}$, and let $w = (0, 0, 0) \in C := C_1 \cap C_2 \cap C_3 = \{(0, 0, 0)\}$. We see that C_1 , C_2 , and C_3 respectively have unit normal vectors

$$n_1 = (0, 0, 1), \quad n_2 = (1, 0, 0), \quad \text{and} \quad n_3 = \frac{1}{\sqrt{2+\varepsilon^2}}(1, -\varepsilon, 1). \quad (140)$$

So $\{C_i, C_j\}$ is $\kappa_{i,j}$ -linearly regular on X , where $\kappa_{i,j}$ is computed by (139) as $\kappa_{1,2} = \sqrt{2}$ and $\kappa_{1,3} = \kappa_{2,3} = \sqrt{2}/\sqrt{1 - 1/\sqrt{2 + \varepsilon^2}} \in]\sqrt{2}, 2[$. Now assume that $\{C_1, C_2, C_3\}$ is κ -linearly regular around w for some $\kappa \in \mathbb{R}_+$ and let $x = (\varepsilon^2, \varepsilon, 0) \in C_1 \cap C_3$. Then, as ε is sufficiently small,

$$\sqrt{\varepsilon^2 + \varepsilon^4} = d_C(x) \leq \kappa \max_{i \in \{1, 2, 3\}} d_{C_i}(x) = \kappa \varepsilon^2, \quad (141)$$

which yields $\kappa \geq \sqrt{1 + 1/\varepsilon^2}$. Hence, κ can be arbitrarily large while $\kappa_{i,j}$ remains bounded.

We note from Proposition 2.8(i)–(ii) that assumption (c) in Theorem 5.4 means that for every $j \in J$, $\{C_{s_j}, C_{t_j}\}$ is affine-hull regular ar w . Nevertheless, in the following linear convergence result, we only require linear regularity for pairs $\{C_{s_j}, C_{t_j}\}$ corresponding to $\min\{\lambda_j, \mu_j\} < 2$.

Theorem 5.8 (linear convergence under linear regularity). *Let $w \in C := \bigcap_{i \in I} C_i$. Suppose that $\{Z_j\}_{j \in J}$ is linearly regular around w , that $\{C_i\}_{i \in I}$ is superregular at w , and that for every $j \in J$, $\{C_{s_j}, C_{t_j}\}$ is linearly regular around w if $\min\{\lambda_j, \mu_j\} < 2$ and affine-hull regular at w otherwise. Then the cyclic gDR algorithm converges R -linearly locally to a point*

$$\bar{x} \in \bigcap_{j \in J} Z_j \subseteq \bigcap_{j \in J} \text{Fix } T_j. \quad (142)$$

Moreover, $P_{C_1} \bar{x} = \dots = P_{C_m} \bar{x} \in C$ provided that the cyclic gDR algorithm is connected.

Proof. Combining Corollary 4.6(ii) and Proposition 4.12, there exist $\nu \in]0, 1]$ and $\delta \in \mathbb{R}_{++}$ such that, for every $j \in J$, T_j is (Z_j, ν) -quasi coercive on $B(w; \delta/2)$. Let $\varepsilon \in]0, 1/3]$. Since $\{C_i\}_{i \in I}$ is superregular at w , we shrink δ if necessary so that C_i is $(\varepsilon, \sqrt{2}\delta)$ -regular at w for every $i \in I$. Now let $j \in J$. Then C_{s_j} and C_{t_j} are respectively (ε, δ) - and $(\varepsilon, \sqrt{2}\delta)$ -regular at w . Using Proposition 3.4(ii)–(iii) and noting that

$$(C_{s_j} \cap C_{t_j} + (L_j - L_j)^\perp) \cap B(w; \delta) \subseteq (C_{s_j} \cap C_{t_j} \cap B(w; \delta)) + (L_j - L_j)^\perp, \quad (143)$$

we have that T_j is $(Z_j \cap B(w; \delta), \gamma_j, \frac{1 - \alpha_j + \hat{\beta}_j}{\alpha_j})$ -quasi firmly Fejér monotone on $B(w; \delta/2)$, where $\gamma_j := 1 - \alpha_j + \alpha_j(1 + \frac{\lambda_j \varepsilon}{1 - \varepsilon})(1 + \frac{\mu_j \varepsilon}{1 - \varepsilon})$. Since $\gamma_j \rightarrow 1^+$ as $\varepsilon \rightarrow 0^+$, we can choose ε sufficiently small so that

$$\rho := \left[\gamma_1 \cdots \gamma_m - \frac{\nu^2}{\kappa^2} \left(\sum_{j \in J} \frac{\alpha_j}{1 - \alpha_j + \hat{\beta}_j} \right)^{-1} \right]_+^{1/2} < 1. \quad (144)$$

Now R -linear convergence of the cyclic sequence $(x_n)_{n \in \mathbb{N}}$ is obtained by applying Theorem 5.1(ii) to $(T_j)_{j \in J}$ and the corresponding sets $(Z_j)_{j \in J}$. The rest then follows from Lemma 5.3. \blacksquare

In Theorem 5.8, if $\{C_{s_j}, C_{t_j}\}$ is strongly regular instead of affine-hull regular at w whenever $\lambda_j = \mu_j = 2$, then the limit point $\bar{x} \in C$. In this case, by Lemma 5.2, the linear regularity of $\{Z_j\}_{j \in J}$ is a consequence of that of $\{C_i\}_{i \in I}$. We summarize this observation in the following corollary, which indeed extends [19, Theorem 5.21].

Corollary 5.9 (linear convergence to a common point). *Let $w \in C := \bigcap_{i \in I} C_i$. Suppose that $\{C_i\}_{i \in I}$ is superregular at w and linearly regular around w , and that for every $j \in J$, $\{C_{s_j}, C_{t_j}\}$ is linearly regular around w if $\min\{\lambda_j, \mu_j\} < 2$ and strongly regular at w otherwise. Then the cyclic gDR algorithm converges R -linearly locally to a point $\bar{x} \in C$.*

Proof. Since strong regularity implies affine-hull regularity, the conclusion follows from Theorem 5.8 and Lemma 5.2. \blacksquare

We say that the cyclic gDR algorithm is *fully connected* if there exist positive integers r, q such that $1 \leq r \leq q \leq \ell$ and (not necessarily distinct) indices $i_1, \dots, i_q \in I$ such that $\{i_1, \dots, i_q\} = I$ and that

$$\{(i_1, i_2), (i_2, i_3), \dots, (i_{r-1}, i_r), (i_r, i_1)\} \cup \{(i_1, i_{r+1}), \dots, (i_1, i_q)\} \subseteq \{(s_j, t_j)\}_{j \in J}. \quad (145)$$

Here we adopt the following convention. If $r = 1$, then (145) reads as

$$\{(i_1, i_2), \dots, (i_1, i_q)\} \subseteq \{(s_j, t_j)\}_{j \in J}, \text{ i.e., } \{(i_1, k)\}_{k \in I \setminus \{i_1\}} \subseteq \{(s_j, t_j)\}_{j \in J}, \quad (146)$$

which is a generalization of the cyclically anchored DR algorithm. If $r = q$, then (145) reads as

$$\{(i_1, i_2), (i_2, i_3), \dots, (i_{r-1}, i_r), (i_r, i_1)\} \subseteq \{(s_j, t_j)\}_{j \in J}, \quad (147)$$

which is a generalization of the cyclic DR algorithm.

Lemma 5.10 (shadows of common fixed points under convexity). *Suppose that C_i is convex for every $i \in I$. Let T_{C_i, C_j} denote a gDR operator for the pair (C_i, C_j) . Let $i_1, \dots, i_r \in I$ (not necessarily distinct). Then the following hold:*

- (i) *If $\bar{x} \in \bigcap_{k=2}^r \text{Fix } T_{C_{i_1}, C_{i_k}}$, then $P_{C_{i_1}} \bar{x} \in \bigcap_{k=1}^r C_{i_k}$.*
- (ii) *If $\bar{x} \in \bigcap_{k=1}^r \text{Fix } T_{C_{i_k}, C_{i_{k+1}}}$, where $i_{r+1} := i_1$, then $P_{C_{i_1}} \bar{x} = \dots = P_{C_{i_r}} \bar{x} \in \bigcap_{k=1}^r C_{i_k}$.*
- (iii) *If the cyclic gDR algorithm is fully connected and $\bar{x} \in \bigcap_{j \in J} \text{Fix } T_j$, then there exists $k \in I$ such that $P_{C_k} \bar{x} \in \bigcap_{i \in I} C_i$.*

Proof. (i): For every $k \in \{2, \dots, r\}$, since $C_{i_1} \cap C_{i_k} \neq \emptyset$, Lemma 2.13(i) implies that $P_{C_{i_1}} \bar{x} \in C_{i_1} \cap C_{i_k}$. Hence, $P_{C_{i_1}} \bar{x} \in \bigcap_{k=1}^r C_{i_k}$.

(ii): This is basically similar to the argument of [17, Theorem 3.1]. For every $k \in \{1, \dots, r\}$, by Lemma 2.13(i), $P_{C_{i_k}} \bar{x} \in C_{i_k} \cap C_{i_{k+1}} \subseteq C_{i_{k+1}}$, and then, by [5, Theorem 3.16], $\langle x - P_{C_{i_{k+1}}}, P_{C_{i_k}} - P_{C_{i_{k+1}}} \rangle \leq 0$. It follows that

$$0 \leq \sum_{k=1}^r \|P_{C_{i_k}} - P_{C_{i_{k+1}}}\|^2 = 2 \sum_{k=1}^r \langle -P_{C_{i_{k+1}}}, P_{C_{i_k}} - P_{C_{i_{k+1}}} \rangle \quad (148a)$$

$$= 2 \sum_{k=1}^r \langle x - P_{C_{i_{k+1}}}, P_{C_{i_k}} - P_{C_{i_{k+1}}} \rangle \leq 0, \quad (148b)$$

which yields $P_{C_{i_k}} = P_{C_{i_{k+1}}}$ for all $k \in \{1, \dots, r\}$.

(iii): Combine (i) and (ii). \blacksquare

As one would hope for, the linear convergence of the cyclic gDR algorithm is global in the convex case. The next result encompasses [14, Corollary 8.2 and Theorem 8.5].

Theorem 5.11 (global linear convergence under convexity). *Suppose that for every $i \in I$, C_i is convex and that $\bigcap_{i \in I_p} C_i \cap \bigcap_{i \in I \setminus I_p} \text{ri} C_i \neq \emptyset$, where I_p is the set of all $i \in I$ such that C_i and C_k are polyhedral whenever $(i, k) \in \{(s_j, t_j), (t_j, s_j)\}_{j \in J}$. Set*

$$\forall j \in J, \quad Y_j := \begin{cases} C_{s_j} \cap C_{t_j} & \text{if } \min\{\lambda_j, \mu_j\} < 2, \\ (C_{s_j} \cap C_{t_j}) + (L_j - L_j)^\perp & \text{if } \lambda_j = \mu_j = 2 \text{ and } \text{ri} C_{s_j} \cap \text{ri} C_{t_j} \neq \emptyset, \\ \text{Fix } T_j & \text{otherwise.} \end{cases} \quad (149)$$

Then the following hold:

- (i) *Regardless of the starting point x_0 , the cyclic gDR sequence $(x_n)_{n \in \mathbb{N}}$ generated by $(T_j)_{j \in J}$ converges R -linearly to a point*

$$\bar{x} \in \bigcap_{j \in J} Y_j \subseteq \bigcap_{j \in J} \text{Fix } T_j, \quad (150)$$

while the “shadow sequence” $(P_{C_i} x_n)_{n \in \mathbb{N}}$ also converges R -linearly to $P_{C_i} \bar{x}$ for every $i \in I$.

- (ii) *If $0 \in \text{int}(C_{t_j} - C_{s_j})$ whenever $\lambda_j = \mu_j = 2$, then the limit point \bar{x} in (i) even satisfies $\bar{x} \in C$.*
(iii) *If the cyclic gDR algorithm is connected and $\text{ri} C_{s_j} \cap \text{ri} C_{t_j} \neq \emptyset$ whenever $\lambda_j = \mu_j = 2$, then the limit point \bar{x} in (i) satisfies $P_{C_1} \bar{x} = \dots = P_{C_m} \bar{x} \in C$.*
(iv) *If the cyclic gDR algorithm is fully connected, then the limit point \bar{x} in (i) satisfies $P_{C_k} \bar{x} \in C$ for some $k \in I$.*

Proof. First, it follows from Lemma 2.10(iii) that

$$\forall j \in J, \quad C_{s_j} \cap C_{t_j} \subseteq Y_j \subseteq \text{Fix } T_j. \quad (151)$$

(i): For every $j \in J$, by the convexity of C_{s_j} and C_{t_j} and by noting from Lemma 2.13(i) that $\text{Fix } T_j = C_{s_j} \cap C_{t_j} + N_{C_{s_j} - C_{t_j}}(0)$ whenever $\lambda_j = \mu_j = 2$, we have that Y_j is convex. Set $J_p := \{j \in J \mid C_{s_j} \text{ and } C_{t_j} \text{ are polyhedral}\}$. Then Y_j is polyhedral for every $j \in J_p$.

Let $j \in J \setminus J_p$. Since $j \notin J_p$, we must have $s_j, t_j \notin I_p$ and, by assumption, $\text{ri} C_{s_j} \cap \text{ri} C_{t_j} \neq \emptyset$, so $Y_j = C_{s_j} \cap C_{t_j}$ if $\min\{\lambda_j, \mu_j\} < 2$, and $Y_j = (C_{s_j} \cap C_{t_j}) + (L_j - L_j)^\perp$ otherwise. Using [34, Corollary 6.6.2],

$$\text{ri} Y_j \supseteq \begin{cases} \text{ri}(C_{s_j} \cap C_{t_j}) & \text{if } \min\{\lambda_j, \mu_j\} < 2, \\ \text{ri}(C_{s_j} \cap C_{t_j}) + \text{ri}(L_j - L_j)^\perp & \text{otherwise,} \end{cases} \quad (152)$$

thus $\text{ri} Y_j \supseteq \text{ri}(C_{s_j} \cap C_{t_j}) = \text{ri} C_{s_j} \cap \text{ri} C_{t_j}$ due to [34, Theorem 6.5]. By combining with (125b) and noting that if $i \in I_p$, then $i \in \{s_j, t_j\}$ for some $j \in J_p$,

$$\bigcap_{j \in J_p} Y_j \cap \bigcap_{j \in J \setminus J_p} \text{ri} Y_j \supseteq \bigcap_{j \in J_p} (C_{s_j} \cap C_{t_j}) \cap \bigcap_{j \in J \setminus J_p} (\text{ri} C_{s_j} \cap \text{ri} C_{t_j}) \supseteq \bigcap_{i \in I_p} C_i \cap \bigcap_{i \in I \setminus I_p} \text{ri} C_i \neq \emptyset. \quad (153)$$

Let $x_0 \in X$ and let $(x_n)_{n \in \mathbb{N}}$ be the cyclic sequence generated by $(T_j)_{j \in J}$ with starting point x_0 . Choose $\delta \in \mathbb{R}_{++}$ and $w \in C$ such that $\delta \geq 2\|x_0 - w\| \geq 2d_C(x_0)$. Then $x_0 \in \mathbf{B}(w; \delta/2)$. From (153) and [4, Corollary 5], there is $\kappa \in \mathbb{R}_{++}$ such that $\{Y_j\}_{j \in J}$ is κ -linearly regular on $\mathbf{B}(w; \delta/2)$. Let $j \in J$. By assumption, either $\text{ri} C_{s_j} \cap \text{ri} C_{t_j} \neq \emptyset$ or C_{s_j} and C_{t_j} are polyhedral with $C_{s_j} \cap C_{t_j} \neq \emptyset$, so we derive from Proposition 4.9 that T_j is (Y_j, ν_j) -quasi coercive on $\mathbf{B}(w; \delta/2)$ for some $\nu_j \in \mathbb{R}_{++}$. By convexity and Lemma 2.12(ii), T_j is $\alpha_j/(1 + \hat{\beta}_j)$ -averaged, which implies that T_j is $(\text{Fix } T_j, 1, \frac{1 - \alpha_j + \hat{\beta}_j}{\alpha_j})$ -

and hence $(Y_j, 1, \frac{1-\alpha_j+\widehat{\beta}_j}{\alpha_j})$ -quasi firmly Fejér monotone on X . Now Theorem 5.1(iii) yields the R -linear convergence of the sequence $(x_n)_{n \in \mathbb{N}}$ to a point $\bar{x} \in \bigcap_{j \in J} Y_j$.

Next, for every $i \in I$, by Fact 2.11(ii), P_{C_i} is nonexpansive, so $\|P_{C_i}x_n - P_{C_i}\bar{x}\| \leq \|x_n - \bar{x}\|$ for all n , and we deduce that $P_{C_i}x_n$ converges R -linearly to $P_{C_i}\bar{x}$.

(ii): In the case where $\lambda_j = \mu_j = 2$, since $0 \in \text{int}(C_{t_j} - C_{s_j})$, Lemma 2.13(ii) yields $\text{Fix } T_j = C_{s_j} \cap C_{t_j}$, and then, by (151), $Y_j = C_{s_j} \cap C_{t_j}$. We deduce that $Y_j = C_{s_j} \cap C_{t_j}$ for every $j \in J$, which together with (125b) yields $\bigcap_{j \in J} Y_j = \bigcap_{i \in I} C_i = C$.

(iii): Combine (i) and Lemma 5.3.

(iv): This follows from (i) and Lemma 5.10(iii). ■

When specialized to the case of two sets, our results cover Theorems 4.3, 4.7, and 4.14 in [32] where R -linear convergence is proved for the classical DR algorithm.

Corollary 5.12 (linear convergence of gDR algorithm). *Let A and B be closed subsets of X , $L := \text{aff}(A \cup B)$, and $w \in A \cap B \neq \emptyset$. Let $\lambda, \mu \in]0, 2]$, $\widehat{\beta} := (\frac{\lambda}{2-\lambda} + \frac{\mu}{2-\mu})^{-1}$, $\alpha \in]0, 1 + \widehat{\beta}[$, and T the gDR operator for (A, B) with parameters (λ, μ, α) . Then the gDR algorithm generated by T*

(i) *locally converges with R -linear rate to a point \bar{x} in each of the following cases:*

- (a) *$\{A, B\}$ is superregular and affine-hull regular at w , in which case $\bar{x} \in A \cap B + (L - L)^\perp$ with $P_A\bar{x} = P_B\bar{x} \in A \cap B$, and if additionally $\{A, B\}$ is strongly regular at w , then $\bar{x} \in A \cap B$.*
- (b) *$\min\{\lambda, \mu\} < 2$ and $\{A, B\}$ is superregular at w and linearly regular around w , in which case $\bar{x} \in A \cap B$.*

(ii) *globally converges with R -linear rate to a point \bar{x} in each of the following cases:*

- (a) *A and B are convex and $\text{ri } A \cap \text{ri } B \neq \emptyset$, in which case $\bar{x} \in A \cap B + (L - L)^\perp$ with $P_A\bar{x} = P_B\bar{x} \in A \cap B$, and if additionally $0 \in \text{int}(B - A)$, then $\bar{x} \in A \cap B$.*
- (b) *A and B are polyhedral, in which case $\bar{x} \in \text{Fix } T \subseteq A \cap B + N_{A-B}(0)$ with $P_A\bar{x} \in A \cap B$.*
- (c) *$\min\{\lambda, \mu\} < 2$, A and B are convex and $\{A, B\}$ is boundedly linearly regular, in which case $\bar{x} \in A \cap B$.*

Proof. Applying Theorem 5.8 (noting that affine-hull regularity implies linear regularity due to Proposition 2.5) and Theorem 5.11 with $m = 2$, $\ell = 1$, and $(s_1, t_1) = (1, 2)$, we get (i) and (ii)(a)–(ii)(b).

We now prove (ii)(c). Let $x_0 \in X$ and let $(x_n)_{n \in \mathbb{N}}$ be the cyclic sequence generated by $(T_j)_{j \in J}$ with starting point x_0 . Take $\delta \in \mathbb{R}_{++}$ and $w \in A \cap B$ such that $\delta \geq 2\|x_0 - w\| \geq 2d_{A \cap B}(x_0)$. Then $x_0 \in B(w; \delta/2)$. Since A and B are convex, Lemma 2.12(ii) implies that T is $\alpha/(1 + \widehat{\beta})$ -averaged, hence it is $(\text{Fix } T, 1, \frac{1-\alpha+\widehat{\beta}}{\alpha})$ - and also $(A \cap B, 1, \frac{1-\alpha+\widehat{\beta}}{\alpha})$ -quasi firmly Fejér monotone on X . According to Proposition 4.14, T is $(A \cap B, \nu)$ -quasi coercive on $B(w; \delta/2)$ for some $\nu \in \mathbb{R}_{++}$. Now apply Theorem 5.1(iii) to T and the corresponding set $A \cap B$. ■

6. Conclusion

In this paper, we have presented a diverse collection of improvements on the linear convergence of the (cyclic) generalized Douglas–Rachford algorithm for solving feasibility problems. Our results indicate that one has great flexibility in choosing suitable parameters and still achieve convergence with linear rate. For instance, we have proved that the generalized DR algorithm involving at most one reflection converges R -linearly locally assuming only that the system of superregular sets $\{A, B\}$ is linearly regular around the reference point. Interestingly, it remains open even in the convex case whether the classical DR algorithm (i.e., $\lambda = \mu = 2$ and $\alpha = 1/2$) converges with R -linear rate under the same assumption.

Acknowledgments

The authors are thankful to the editors and the referees for their constructive comments. MND was partially supported by the Australian Research Council (ARC) under Discovery Project 160101537 and by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant 101.02-2016.11. This work was essentially completed during MND's visit to UMass Lowell in September 2017 to whom he acknowledges the hospitality.

References

- [1] H. H. Bauschke, J. Y. Bello Cruz, T. T. A. Nghia, H. M. Phan, and X. Wang. The rate of linear convergence of the Douglas–Rachford algorithm for subspaces is the cosine of the Friedrichs angle. *J. Approx. Theory*, 185:63–79, 2014.
- [2] H. H. Bauschke and J. M. Borwein. Dykstra's alternating projection algorithm for two sets. *J. Approx. Theory*, 79(3):418–443, 1994.
- [3] H. H. Bauschke and J. M. Borwein. On projections algorithms for solving convex feasibility problems. *SIAM Rev.*, 38(3):367–426, 1996.
- [4] H. H. Bauschke, J. M. Borwein, and W. Li. Strong conical hull intersection property, bounded linear regularity, Jameson's property (G), and error bounds in convex optimization. *Math. Program.*, 86(1):135–160, 1999.
- [5] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York, 2nd edition, 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*, 127:178–192, 2004.
- [7] H. H. Bauschke and M. N. Dao. On the finite convergence of the Douglas–Rachford algorithm for solving (not necessarily convex) feasibility problems in Euclidean spaces. *SIAM J. Optim.*, 27(1):507–537, 2017.
- [8] H. H. Bauschke, M. N. Dao, and W. M. Moursi. The Douglas–Rachford algorithm in the affine-convex case. *Oper. Res. Lett.*, 44(3):379–382, 2016.
- [9] H. H. Bauschke, M. N. Dao, D. Noll, and H. M. Phan. On Slater's condition and finite convergence of the Douglas–Rachford algorithm for solving convex feasibility problems in Euclidean spaces. *J. Global Optim.*, 65(2):329–349, 2016.
- [10] H. H. Bauschke, M. N. Dao, D. Noll, and H. M. Phan. Proximal point algorithm, Douglas–Rachford algorithm and alternating projections: a case study. *J. Convex Anal.*, 23(1):237–261, 2016.
- [11] H. H. Bauschke, D. R. Luke, H. M. Phan, and X. Wang. Restricted normal cones and the method of alternating projections: theory. *Set-Valued Var. Anal.*, 21(3):431–473, 2013.
- [12] H. H. Bauschke, D. R. Luke, H. M. Phan, and X. Wang. Restricted normal cones and the method of alternating projections: applications. *Set-Valued Var. Anal.*, 21(3):475–501, 2013.
- [13] H. H. Bauschke and W. M. Moursi. On the Douglas–Rachford algorithm. *Math. Program.*, 164(1–2):263–284, 2017.
- [14] H. H. Bauschke, D. Noll, and H. M. Phan. Linear and strong convergence of algorithms involving averaged nonexpansive operators. *J. Math. Anal. Appl.*, 421(1):1–20, 2015.
- [15] H. H. Bauschke, H. M. Phan, and X. Wang. The method of alternating relaxed projections for two nonconvex sets. *Vietnam J. Math.*, 42(4):421–450, 2014.

- [16] J. M. Borwein, B. Sims, and M. K. Tam. Norm convergence of realistic projection and reflection methods. *Optim.*, 64(1):161–178, 2015.
- [17] J. M. Borwein and M. K. Tam. A cyclic Douglas–Rachford iteration scheme. *J. Optim. Theory Appl.*, 160(1):1–29, 2014.
- [18] L. M. Bregman. The method of successive projection for finding a common point of convex sets. *Soviet Math. Dokl.*, 6:688–692, 1965.
- [19] M. N. Dao and H. M. Phan. Linear convergence of projection algorithms, 2016, [arXiv:1609.00341](https://arxiv.org/abs/1609.00341).
- [20] M. N. Dao and M. K. Tam. A Lyapunov-type approach to convergence of the Douglas–Rachford algorithm, 2017, [arXiv:1706.04846](https://arxiv.org/abs/1706.04846).
- [21] J. Douglas and H. H. Rachford. On the numerical solution of heat conduction problems in two and three space variables. *Trans. Amer. Math. Soc.*, 82:421–439, 1956.
- [22] D. Drusvyatskiy, A. D. Ioffe, and A. S. Lewis. Transversality and alternating projections for nonconvex sets. *Found. Comput. Math.*, 15(6):1637–1651, 2015.
- [23] V. Elser, I. Rankenburg, and P. Thibault. Searching with iterated maps. *Proc. Natl. Acad. Sci. USA*, 104(2):418–423, 2007.
- [24] R. Hesse and D. R. Luke. Nonconvex notions of regularity and convergence of fundamental algorithms for feasibility problems. *SIAM J. Optim.*, 23(4):2397–2419, 2013.
- [25] A.Y. Kruger. About regularity of collections of sets. *Set-Valued Anal.*, 14(2):187–206, 2006.
- [26] A. S. Lewis, D. R. Luke, and J. Malick. Local linear convergence for alternating and averaged nonconvex projections. *Found. Comput. Math.*, 9(4):485–513, 2009.
- [27] P.-L. Lions and B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.*, 16(6):964–979, 1979.
- [28] D. R. Luke. Finding best approximation pairs relative to a convex and prox-regular set in a Hilbert space. *SIAM J. Optim.*, 19(2):714–739, 2008.
- [29] D. R. Luke, N. H. Thao, and M. K. Tam. Quantitative convergence analysis of iterated expansive, set-valued mappings. *Math. Oper. Res.*, to appear. [arXiv:1605.05725](https://arxiv.org/abs/1605.05725).
- [30] B. S. Mordukhovich. *Variational Analysis and Generalized Differentiation I*. Springer, Berlin, 2006.
- [31] D. Noll and A. Rondepierre. On local convergence of the method of alternating projections. *Found. Comput. Math.*, 16(2):425–455, 2016.
- [32] H. M. Phan. Linear convergence of the Douglas–Rachford method for two closed sets. *Optim.*, 65(2):369–385, 2016.
- [33] S.M. Robinson. Some continuity properties of polyhedral multifunctions. *Math. Programming Stud.*, 14:206–214, 1981.
- [34] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, 1970.
- [35] R. T. Rockafellar and R. J-B Wets. *Variational Analysis*. Springer, Berlin, 1998.
- [36] E. D. Sontag. Remarks on piecewise-linear algebra. *Pacific J. Math.*, 98(1):183–201, 1982.
- [37] B. F. Svaiter. On weak convergence of the Douglas–Rachford method. *SIAM J. Control Optim.*, 49(1):280–287, 2011.