Linear and strong convergence of algorithms involving averaged nonexpansive operators

Heinz H. Bauschke[*], Dominikus Noll[†] and Hung M. Phan[‡]

June 14, 2014

Abstract

We introduce regularity notions for averaged nonexpansive operators. Combined with regularity notions of their fixed point sets, we obtain linear and strong convergence results for quasicyclic, cyclic, and random iterations. New convergence results on the Borwein–Tam method (BTM) and on the cylically anchored Douglas–Rachford algorithm (CADRA) are also presented. Finally, we provide a numerical comparison of BTM, CADRA and the classical method of cyclic projections for solving convex feasibility problems.

2010 Mathematics Subject Classification: Primary 65K05; Secondary 47H09, 90C25.

Keywords: Averaged nonexpansive mapping, Borwein–Tam method, bounded linear regularity, convex feasibility problem, convex set, cyclically anchored Douglas–Rachford algorithm, Douglas–Rachford algorithm, nonexpansive operator, projection, random method, transversality.

1 Overview

Throughout this paper, X is a real Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and induced norm $\|\cdot\|$. The convex feasibility problem asks to find a point in the intersection of convex

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

[†]Université Paul Sabatier, Institut de Mathématiques, 118 route de Narbonne, 31062 Toulouse, France. E-mail: noll@mip.ups-tlse.fr.

[‡]Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada. E-mail: hung.phan@ubc.ca.

sets. This is an important problem in mathematics and engineering; see, e.g., [6], [7], [12], [13], [14], [20], [21], [29], and the references therein.

Oftentimes, the convex sets are given as fixed point sets of projections or (more generally) averaged nonexpansive operators. In this case, weak convergence to a solution is guaranteed but the question arises under which circumstances can we guarantee strong or even linear convergence. The situation is quite clear for projection algorithms; see, e.g., [6] and also [23].

The aim of this paper is to provide verifiable sufficient conditions for strong and linear convergence of algorithms based on iterating convex combinations of averaged nonexpansive operators.

Our results can be nontechnically summarized as follows: *If each operator is well behaved and the fixed point sets relate well to each other, then the algorithm converges strongly or linearly.*

Specifically, we obtain the following main results on iterations of averaged nonexpansive mappings:

- If each operator is boundedly linearly regular and the family of corresponding fixed point sets is boundedly linearly regular, then quasicyclic averaged algorithms converge linearly (Theorem 6.1).
- If each operator is boundedly regular and the family of corresponding fixed point sets is boundedly regular, then cyclic algorithms converge strongly (Theorem 7.11).
- If each operator is boundedly regular and the family of corresponding fixed point sets is innately boundedly regular, then random sequential algorithms converge strongly (Theorem 7.14).

We also focus in particular on algorithms featuring the Douglas–Rachford splitting operator and obtain new convergence results on the Borwein–Tam method and the cyclically anchored Douglas–Rachford algorithm.

The remainder of the paper is organized as follows. In Sections 2 and 3, we discuss (boundedly) linearly regular and averaged nonexpansive operators. The bounded linear regularity of the Douglas–Rachford operator in the transversal case is obtained in Section 4. In Section 5, we recall the key notions of Fejér monotonticity and regularity of collections of sets. Our main convergence result on quasicyclic algorithms is presented in Section 6. In Section 7, we turn to strong convergence results for cyclic and random algorithms. Applications and numerical results are provided in Section 8. Notation in this paper is quite standard and follows mostly [7]. The closed ball of radius r centred at x is denoted by $\operatorname{ball}(x;r)$.

2 Operators that are (boundedly) linearly regular

Our linear convergence results depend crucially on the concepts of (bounded) linear regularity which we introduce now.

Definition 2.1 ((bounded) linear regularity) *Let $T\colon X \to X$ be such that $\operatorname{Fix} T \neq \varnothing$. We say that:*

(i) *T is **linearly regular** with constant $\kappa \geq 0$ if*

(1) $$(\forall x \in X) \quad d_{\operatorname{Fix} T}(x) \leq \kappa \|x - Tx\|.$$

(ii) *T is **boundedly linearly regular** if*

(2) $$(\forall \rho > 0)(\exists\, \kappa \geq 0)(\forall x \in \operatorname{ball}(0;\rho)) \quad d_{\operatorname{Fix} T}(x) \leq \kappa \|x - Tx\|;$$

note that in general κ depends on ρ, which we sometimes indicate by writing $\kappa = \kappa(\rho)$.

We clearly have the implication

(3) $$\text{linearly regular} \Rightarrow \text{boundedly linearly regular.}$$

Example 2.2 (relaxed projectors) Let C be a nonempty closed convex subset of X and let $\lambda \in \,]0,2]$. Then $T = (1-\lambda)\operatorname{Id} + \lambda P_C$ is linearly regular with constant λ^{-1}.

Proof. Indeed, $\operatorname{Fix} T = C$ and $(\forall x \in X)\ d_C(x) = \|x - P_C x\| = \lambda^{-1}\|x - Tx\|$. ■

The following example shows that an operator may be boundedly linearly regular yet not linearly regular. This illustrates that the converse of the implication (3) fails.

Example 2.3 (thresholder) Suppose that $X = \mathbb{R}$ and set

(4) $$Tx = \begin{cases} 0, & \text{if } |x| \leq 1; \\ x - 1, & \text{if } x > 1; \\ x + 1, & \text{if } x < -1. \end{cases}$$

Then T is boundedly linearly regular with $\kappa(\rho) = \max\{\rho, 1\}$; however, T is not linearly regular.

Proof. Let $x \in X$. Since $\operatorname{Fix} T = \{0\}$, we deduce

(5) $$d_{\operatorname{Fix} T}(x) = |x| = \max\big\{|x|, 1\big\} \min\big\{|x|, 1\big\}$$

and

$$|x - Tx| = \begin{cases} |x|, & \text{if } |x| \le 1; \\ 1, & \text{if } |x| > 1 \end{cases} = \min\{|x|, 1\}. \tag{6}$$

If $x \notin \operatorname{Fix} T$, then $d_{\operatorname{Fix} T}(x)/|x - Tx| = \max\{|x|, 1\}$ and the result follows. ■

Theorem 2.4 *Let $T\colon X \to X$ be linear and nonexpansive with* $\operatorname{ran}(\operatorname{Id} - T)$ *closed. Then T is linearly regular.*

Proof. Set $A = \operatorname{Id} - T$. Then A is maximally monotone by [7, Example 20.26], and $(\operatorname{Fix} T)^\perp = (\ker A)^\perp = \overline{\operatorname{ran}}\, A^* = \overline{\operatorname{ran}}\, A = \overline{\operatorname{ran}}\,(\operatorname{Id} - T) = \operatorname{ran}(\operatorname{Id} - T)$ using [7, Proposition 20.17]. By the Closed Graph Theorem (see, e.g., [16, Theorem 8.18]), there exists $\beta > 0$ such that

$$\big(\forall z \in \ker(A)^\perp)\big) \quad \|Az\| \ge \beta \|z\|. \tag{7}$$

Now let $x \in X$ and split x into $x = y + z$, where $y = P_{\ker A} x = P_{\operatorname{Fix} T} x$ and $z = P_{(\ker A)^\perp} x = P_{\operatorname{ran} A} x = P_{\operatorname{ran}(\operatorname{Id} - T)} x$. Then

$$\|x - Tx\| = \|Ax\| = \|A(y + z)\| = \|Az\| \ge \beta\|z\| = \beta\|x - P_{\operatorname{Fix} T} x\| = \beta d_{\operatorname{Fix} T}(x) \tag{8}$$

and the result follows. ■

Example 2.5 (Douglas–Rachford operator for two subspaces) Let U and V be closed subspaces of X such that $U + V$ is closed, and set $T = P_V P_U + P_{V^\perp} P_{U^\perp}$. Then $\operatorname{Fix} T = (U \cap V) + (U^\perp \cap V^\perp)$, and $\operatorname{ran}(\operatorname{Id} - T) = (U + V) \cap (U^\perp + V^\perp)$ is closed; consequently, T is linearly regular.

Proof. The formula for $\operatorname{Fix} T$ is in, e.g., [4]. On the one hand, it is well known (see, e.g., [7, Corollary 15.35]) that $U^\perp + V^\perp$ is closed as well. On the other hand, [10, Corollary 2.14] implies that $\operatorname{ran}(\operatorname{Id} - T) = (U + V) \cap (U^\perp + V^\perp)$. Altogether, $\operatorname{ran}(\operatorname{Id} - T)$ is closed. Finally, apply Theorem 2.4. ■

Example 2.6 Suppose that $X = \mathbb{R}^2$, let $\theta \in \left]0, \pi/2\right]$, set $U = \mathbb{R} \cdot (1, 0)$, $V = \mathbb{R} \cdot (\cos\theta, \sin\theta)$, and $T = P_V P_U + P_{V^\perp} P_{U^\perp}$. Then T is linearly regular with rate $1/\sin(\theta)$.

Proof. Let $x \in X$. A direct computation (or [4, Section 5]) yields

$$T = \cos(\theta) \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}, \tag{9}$$

i.e., T shrinks the vector by $\cos(\theta) \in [0,1[$ and rotates it by θ. Hence $\operatorname{Fix} T = \{0\}$ and

$$d_{\operatorname{Fix} T}(x) = \|x\|. \tag{10}$$

On the other hand, using $1 - \cos^2(\theta) = \sin^2(\theta)$, we obtain

$$\operatorname{Id} - T = \sin(\theta) \begin{pmatrix} \sin(\theta) & \cos(\theta) \\ -\cos(\theta) & \sin(\theta) \end{pmatrix} \tag{11}$$

and hence

$$\|x - Tx\| = \sin(\theta)\|x\|. \tag{12}$$

Altogether, $d_{\operatorname{Fix} T}(x) = \|x\| = (1/\sin(\theta))\sin(\theta)\|x\| = (1/\sin(\theta))\|x - Tx\|$. ■

We conclude this section by comparing our notion of bounded linear regularity to metric regularity of set-valued operators.

Remark 2.7 Suppose that T is firmly nonexpansive and thus the resolvent of a maximally monotone operator A. Suppose that $\bar{x} \in X$ is such that $0 \in A\bar{x}$, i.e., $\bar{x} \in \operatorname{Fix} T$. Then *metric subregularity* of A at $\bar{x}$ means that there exists $\delta > 0$ and $\gamma > 0$ such that $x \in \operatorname{ball}(\bar{x};\delta) \Rightarrow d_{A^{-1}0}(x) \leq \gamma d_{Ax}(0)$. In terms of T, this is expressed as $x \in \operatorname{ball}(\bar{x};\delta) \Rightarrow d_{\operatorname{Fix} T}(x) \leq \gamma \inf \|x - T^{-1}x\|$. If $x = Ty \in \operatorname{ball}(\bar{x};\delta)$, then

$$d_{\operatorname{Fix} T}(Ty) \leq \gamma\|y - Ty\|; \tag{13}$$

moreover, $d_{\operatorname{Fix} T}(y) \leq (1+\gamma)\|y - Ty\|$. This is related to bounded linear regularity of T. The interested reader is referred to [18] for further information on metric subregularity; see also [1] and [24].

3 Averaged nonexpansive operators

We work mostly within the class of averaged nonexpansive mappings which have proven to be a good compromise between generality and usability.

Definition 3.1 *The mapping $T\colon X \to X$ is* averaged nonexpansive *if there exists $\lambda \in [0,1[$ and $N\colon X \to X$ nonexpansive such that $T = (1-\lambda)\operatorname{Id} + \lambda N$.*

The class of averaged nonexpansive operators is closed under compositions and convex combinations, and it includes all firmly nonexpansive mappings; see, e.g., [15] for further information.

Example 3.2 Let $T\colon X \to X$ be β-Lipschitz with $\beta \in \,]0,1[$. Then T is averaged.

Proof. Let $\varepsilon \in \,]0,(1-\beta)/2[\, \subset \,]0,1[$. Then $(\beta+\varepsilon)/(1-\varepsilon) \in \,]0,1[$. Now $(1-\varepsilon)^{-1}T$ is $(1-\varepsilon)^{-1}\beta$-Lipschitz and $-\varepsilon(1-\varepsilon)^{-1}\operatorname{Id}$ is $\varepsilon(1-\varepsilon)^{-1}$-Lipschitz, hence

$$N = (1-\varepsilon)^{-1}T - \varepsilon(1-\varepsilon)^{-1}\operatorname{Id} \tag{14}$$

is nonexpansive. Set $\lambda = 1-\varepsilon \in \,]0,1[$. Then $(1-\lambda)\operatorname{Id} + \lambda N = \varepsilon \operatorname{Id} + (1-\varepsilon)N = T$ and T is therefore averaged. ■

Fact 3.3 (See, e.g., [7, Proposition 4.25(iii)].) *Let $T\colon X \to X$ be averaged nonexpansive. Then there exists $\sigma > 0$ such that*

$$(\forall x \in X)(\forall z \in \operatorname{Fix} T) \quad \sigma\|x - Tx\|^2 \le \|x - z\|^2 - \|Tx - z\|^2. \tag{15}$$

The following two properties are crucial to our subsequent analysis.

Corollary 3.4 ($\sigma(T)$ notation) *Let $T\colon X \to X$ be averaged nonexpansive. Then there exists $\sigma = \sigma(T) > 0$ such that for every nonempty subset C of $\operatorname{Fix} T$, we have*

$$(\forall x \in X) \quad \sigma\|x - Tx\|^2 \le d_C^2(x) - d_C^2(Tx). \tag{16}$$

Corollary 3.5 *Let I be a finite ordered index set, let $(T_i)_{i\in I}$ be family of averaged nonexpansive operators with $\sigma_i = \sigma(T_i)$, and let $(\omega_i)_{i\in I}$ be in $[0,1]$ such that $\sum_{i\in I}\omega_i = 1$. Set $I_+ = \big\{i \in I \;\big|\; \omega_i > 0\big\}$, and set $\sigma_+ = \min_{i\in I_+}\sigma_i$. Let $x \in X$, and set $y = \sum_{i\in I}\omega_i T_i x$. Then*

$$\Big(\forall z \in \bigcap_{i\in I_+}\operatorname{Fix} T_i\Big) \quad \|x - z\|^2 \ge \|y - z\|^2 + \sum_{i\in I}\omega_i\sigma_i\|x - T_i x\|^2 \tag{17a}$$

$$\ge \|y - z\|^2 + \sigma_+\|x - y\|^2. \tag{17b}$$

Proof. Indeed, we have

$$\|y - z\|^2 \le \sum_{i\in I}\omega_i\|T_i x - z\|^2 \le \sum_{i\in I}\omega_i\big(\|x - z\|^2 - \sigma_i\|x - T_i x\|^2\big) \tag{18a}$$

$$= \|x - z\|^2 - \sum_{i\in I}\omega_i\sigma_i\|x - T_i x\|^2 \le \|x - z\|^2 - \sigma_+\|x - y\|^2, \tag{18b}$$

as required. ■

Lemma 3.6 *Let $T\colon X \to X$ be averaged nonexpansive such that*

$$(\forall \rho > 0)(\exists\, \theta < 1)(\forall x \in \operatorname{ball}(0;\rho))(\exists\, y \in \operatorname{Fix} T) \quad \langle x - y, Tx - y\rangle \le \theta\|x - y\|\,\|Tx - y\|. \tag{19}$$

Then T is boundedly linearly regular; moreover, T is linearly regular if θ does not depend on ρ.

Proof. We abbreviate $\sigma(T)$ by σ. Let $\rho > 0$ and let $x \in \operatorname{ball}(0;\rho)$. Obtain θ and $y \in \operatorname{Fix} T$ as in (19). Then

$$\|x - Tx\|^2 = \|x-y\|^2 + \|y - Tx\|^2 + 2\langle x-y, y-Tx\rangle \tag{20a}$$

$$\geq \|x-y\|^2 + \|y-Tx\|^2 - 2\theta\|x-y\|\|Tx-y\| \tag{20b}$$

$$= (1-\theta)\big(\|x-y\|^2 + \|y-Tx\|^2\big) + \theta\big(\|x-y\| - \|y-Tx\|\big)^2 \tag{20c}$$

$$\geq (1-\theta)\|x-y\|^2. \tag{20d}$$

Hence $(1-\theta)^{-1}\|x-Tx\|^2 \geq d^2_{\operatorname{Fix} T}(x)$. ∎

The following example can be viewed as a generalization of Example 2.6.

Example 3.7 Suppose that $S\colon X \to X$ is linear such that $S^* = -S$ and $(\forall x \in X)$ $\|Sx\| = \|x\|$. Let $\alpha \in \,]0,\pi/2]$, let $\beta \in \,]-1,1[$, and set $T = \beta(\cos(\alpha)\operatorname{Id} + \sin(\alpha)S)$. Then T is linearly regular.

Proof. Set $R = \cos(\alpha)\operatorname{Id} + \sin(\alpha)S$. Then $T = \beta R$ and $(\forall x \in X)$ $\|Rx\| = \|Sx\| = \|x\|$; hence $\|T\| = |\beta| < 1$. By Example 3.2, T is averaged. Furthermore, $(\forall x \in X)$ $\langle x, Tx\rangle = \beta\cos(\alpha)\|x\|^2 = \cos(\alpha)\|x\|\|\beta Rx\| = \cos(\alpha)\|x\|\|Tx\|$. The linear regularity of T thus follows from Lemma 3.6. ∎

We conclude this section with some key inequalities.

Lemma 3.8 (key inequalities) *Let $T\colon X \to X$ be averaged nonexpansive and boundedly linearly regular, and let $\rho > 0$. Suppose that C is a nonempty subset of* $\operatorname{Fix} T$. *Then there exist $\alpha \in [0,1[$, $\beta \in \,]0,1]$, and $\gamma > 0$ such that for every $x \in \operatorname{ball}(0;\rho)$, we have*

$$d_{\operatorname{Fix} T}(Tx) \leq \alpha d_{\operatorname{Fix} T}(x); \tag{21}$$

$$\beta d^2_{\operatorname{Fix} T}(x) \leq \big(d_{\operatorname{Fix} T}(x) - d_{\operatorname{Fix} T}(Tx)\big)^2 \leq \|x - Tx\|^2; \tag{22}$$

$$d^2_C(Tx) \leq d^2_C(x) - \gamma d^2_{\operatorname{Fix} T}(x). \tag{23}$$

If T is linearly regular, then these constants do not depend on ρ.

Proof. Let us obtain the constants $\kappa = \kappa(\rho) \geq 0$ from bounded linear regularity and $\sigma = \sigma(T)$ from the averaged nonexpansiveness. Abbreviate $Z = \operatorname{Fix} T$, and let $x \in \operatorname{ball}(0;\rho)$. Then $d^2_Z(Tx) \leq d^2_Z(x) \leq \kappa^2\|x - Tx\|^2 \leq \sigma^{-1}\kappa^2(d^2_Z(x) - d^2_Z(Tx))$ by Corollary 3.4. Hence (21) holds with

$$\alpha = \sqrt{\frac{\sigma^{-1}\kappa^2}{1+\sigma^{-1}\kappa^2}} \in [0,1[\,. \tag{24}$$

Note that α depends only on T when T is in addition linearly regular. Next, we set

$$\beta = (1-\alpha)^2 \in \left]0,1\right], \quad \text{and} \quad \gamma = \sigma\kappa^{-2}, \tag{25}$$

which again depend only on T in the presence of linear regularity. Then, by (21), $d_Z(x) - d_Z(Tx) \geq (1-\alpha)d_Z(x)$. Since d_Z is nonexpansive, we deduce

$$\beta d_Z^2(x) \leq \big(d_Z(x) - d_Z(Tx)\big)^2 \leq \|x - Tx\|^2, \tag{26}$$

i.e., (22). Finally, using Corollary 3.4, we conclude that

$$d_C^2(Tx) \leq d_C^2(x) - \sigma\|x - Tx\|^2 \leq d_C^2(x) - \sigma\kappa^{-2}d_Z^2(x), \tag{27}$$

i.e., (23) holds. ■

4 The Douglas–Rachford Operator for Tranversal Sets

In this section, X is finite-dimensional, A and B are nonempty closed convex subsets of X with $A \cap B \neq \varnothing$. Moreover, $L = \operatorname{aff}(A \cup B)$, $Y = L - L = \operatorname{span}(B - A)$, denote the affine span of $A \cup B$ and the corresponding parallel space, respectively. We also set

$$T = P_B R_A + \operatorname{Id} - P_A, \tag{28}$$

i.e., T is the Douglas–Rachford operator for (A, B). Note that $T(L) \subseteq L$. Our next two results are essentially contained in [26], where even nonconvex settings were considered. In our present convex setting, the proofs become much less technical.

Proposition 4.1 *The following hold:*

(i) $\operatorname{Fix} T = (A \cap B) + N_{A-B}(0) = (A \cap B) + \big(Y \cap N_{A-B}(0)\big) + Y^{\perp}$.

(ii) $L \cap \operatorname{Fix} T = (A \cap B) + (Y \cap N_{A-B}(0))$.

(iii) *If* $\operatorname{ri} A \cap \operatorname{ri} B \neq \varnothing$, *then* $\operatorname{Fix} T = (A \cap B) + Y^{\perp}$ *and* $L \cap \operatorname{Fix} T = A \cap B$.

(iv) *If* $\operatorname{ri} A \cap \operatorname{ri} B \neq \varnothing$, *then* $P_{\operatorname{Fix} T} = \operatorname{Id} - P_L + P_{A\cap B}P_L$.

(v) $(\forall n \in \mathbb{N})\ T^n = \operatorname{Id} - P_L + T^n P_L$.

(vi) $\operatorname{Id} - T = P_L - TP_L$.

(vii) *If* $\operatorname{ri} A \cap \operatorname{ri} B \neq \varnothing$, *then* $d_{\operatorname{Fix} T} = d_{A\cap B} \circ P_L$.

Proof. (i): This follows from [8, Corollary 3.9] and [9, Theorem 3.5]. (ii): Clear from (i). (iii): (See also [26, Lemma 6.5 and Theorem 6.12].) Use (i), (ii), and [9, Theorem 3.5 and Theorem 3.13]. (iv): Write $L = \ell + Y$, where $\ell \in Y^{\perp}$. Then $P_L(A \cap B) = A \cap B = \ell + P_Y(A \cap B)$ and hence $\operatorname{Fix} T = P_Y(A \cap B) \oplus (\ell + Y^{\perp})$. Now use [7, Proposition 28.1(i) and Proposition 28.6]. (v): (See also [26, Theorem 3.16].) By [9, Lemma 3.3], $P_A = P_A P_L$ and $P_B = P_B P_L$. Moreover, P_L is affine. This implies $R_A = R_A P_L + P_L - \mathrm{Id}$, $P_L R_A = R_A P_L$, and $P_B R_A = P_B P_L R_A = P_B R_A P_L$. It follows that $T = \mathrm{Id} - P_L + T P_L = \mathrm{Id} - P_L + P_L T P_L$. The result follows then by induction. (vi): (See also [26, Theorem 3.16].) Clear from (v). (vii): Clear from (iv). ■

Lemma 4.2 *Suppose* $\operatorname{ri} A \cap \operatorname{ri} B \neq \varnothing$, *and let* $c \in A \cap B$. *Then there exists* $\delta > 0$ *and* $\theta < 1$ *such that*

$$(29) \qquad \big(\forall x \in L \cap \operatorname{ball}(c;\delta)\big) \quad \langle P_A x - R_A x, P_B R_A x - R_A x \rangle \leq \theta d_A(x) d_B(R_A x);$$

consequently,

$$(30) \qquad \big(\forall x \in L \cap \operatorname{ball}(c;\delta)\big) \quad \|x - Tx\|^2 \geq \frac{1-\theta}{5} \max\big\{d_A^2(x), d_B^2(x)\big\}.$$

Proof. Since $\operatorname{ri} A \cap \operatorname{ri} B \neq \varnothing$, we deduce from [9, Lemma 3.1 and Theorem 3.13] that

$$(31) \qquad N_A(c) \cap \big(-N_B(c)\big) \cap Y = \{0\}.$$

Now suppose that (29) fails. Noting that $P_A - R_A = \mathrm{Id} - P_A$, we obtain a sequence $(x_n)_{n\in\mathbb{N}}$ in L converging to c and a sequence $\theta_n \to 1^-$ such that for every $n \in \mathbb{N}$,

$$(32) \qquad \langle P_A x_n - R_A x_n, P_B R_A x_n - R_A x_n \rangle > \theta_n \|P_A x_n - R_A x_n\| \|P_B R_A x_n - R_A x_n\|.$$

Hence

$$(33) \qquad \left\langle \frac{x_n - P_A x_n}{\|x_n - P_A x_n\|}, \frac{P_B R_A x_n - R_A x_n}{\|P_B R_A x_n - R_A x_n\|} \right\rangle \to 1^-.$$

Set $u_n = (x_n - P_A x_n)/\|x_n - P_A x_n\| \in Y \cap N_A(P_A x_n)$ and $v_n = (P_B R_A x_n - R_A x_n)/\|P_B R_A x_n - R_A x_n\| \in Y \cap -N_B(P_B R_A x_n)$. After passing to subsequences if necessary we assume that $u_n \to u$ and $v_n \to v$. Then $\langle u, v\rangle = 1$ and thus $v = u$. Since $x_n \to c$, we deduce that $P_A x_n \to P_A c = c$, $R_A x_n \to c$, and $P_B R_A x_n \to c$. Thus, $u \in N_A(c)$ and $-u \in N_B(c)$. Altogether, $u \in N_A(c) \cap (-N_B(c)) \cap Y \smallsetminus \{0\}$, which contradicts (31). We thus have proved (29).

Now let $x \in \operatorname{ball}(c;\delta) \cap L$. Because d_B is nonexpansive and $R_A - \mathrm{Id} = 2(P_A - \mathrm{Id})$, we deduce with the Cauchy–Schwarz inequality that

$$(34a) \qquad d_B^2(x) \leq \big(\|x - R_A x\| + d_B(R_A x)\big)^2 = \big(2 d_A(x) + d_B(R_A x)\big)^2$$

$$\leq 5\big(d_A^2(x) + d_B^2(R_A x)\big). \tag{34b}$$

Using (29), we have

$$
\begin{aligned}
\|x - Tx\|^2 &= \|P_A x - P_B R_A x\|^2 && (35a)\\
&= \|(P_A x - R_A x) + (R_A x - P_B R_A x)\|^2 && (35b)\\
&= \|P_A x - R_A x\|^2 + \|R_A x - P_B R_A x\|^2 + 2\,\langle P_A x - R_A x, R_A x - P_B R_A x\rangle && (35c)\\
&\geq d_A^2(x) + d_B^2(R_A x) - 2\theta d_A(x) d_B(R_A x) && (35d)\\
&= (1-\theta)\big(d_A^2(x) + d_B^2(R_A x)\big) + \theta\big(d_A(x) - d_B(R_A x)\big)^2 && (35e)\\
&\geq (1-\theta)\big(d_A^2(x) + d_B^2(R_A x)\big) && (35f)\\
&\geq \frac{1-\theta}{5}\max\big\{d_A^2(x), d_B^2(x)\big\}, && (35g)
\end{aligned}
$$

as claimed. ■

Lemma 4.3 *Suppose that* $\operatorname{ri} A \cap \operatorname{ri} B \neq \varnothing$. *Then*

$$(\forall \rho > 0)(\exists\, \kappa > 0)(\forall x \in L \cap \operatorname{ball}(0;\rho)) \quad \|x - Tx\| \geq \kappa d_{A\cap B}(x). \tag{36}$$

Proof. We argue by contradiction and assume the conclusion fails. Then there exists a bounded sequence $(x_n)_{n\in\mathbb{N}}$ in L and a sequence $\varepsilon_n \to 0^+$ such that

$$(\forall n \in \mathbb{N}) \quad \|x_n - Tx_n\| < \varepsilon_n d_{A\cap B}(x_n) \to 0. \tag{37}$$

In particular, $d_{A\cap B}(x_n) > 0$ and $x_n - Tx_n \to 0$. After passing to subsequences if necessary, we assume that $x_n \to \bar{x}$. Then $\bar{x} \in L \cap \operatorname{Fix} T$. By Proposition 4.1(iii), $\bar{x} \in A \cap B$. Using Lemma 4.2 and after passing to another subsequence if necessary, we obtain $\theta < 1$ such that

$$(\forall n \in \mathbb{N}) \quad \|x_n - Tx_n\|^2 \geq \frac{1-\theta}{5}\max\big\{d_A^2(x_n), d_B^2(x_n)\big\}. \tag{38}$$

Next, bounded linear regularity of (A, B) (see Fact 5.8(viii) below) yields $\mu > 0$ such that $(\forall n \in \mathbb{N})\ d_{A\cap B}(x_n) \leq \mu \max\{d_A(x_n), d_B(x_n)\}$. Combining this with (37) and (38) yields

$$
\begin{aligned}
(\forall n \in \mathbb{N}) \quad \varepsilon_n^2 d_{A\cap B}^2(x_n) > \|x_n - Tx_n\|^2 &\geq \frac{1-\theta}{5}\max\big\{d_A^2(x_n), d_B^2(x_n)\big\} && (39)\\
&\geq \frac{1-\theta}{5\mu^2} d_{A\cap B}^2(x_n). && (40)
\end{aligned}
$$

This is absurd since $\varepsilon_n \to 0^+$. ■

We are now ready for the main result of this section.

Theorem 4.4 (Douglas–Rachford operator for two transversal sets) *Suppose that the pair* (A, B) *is* transversal, *i.e.,* $\operatorname{ri} A \cap \operatorname{ri} B \neq \varnothing$. *Then* T *is boundedly linearly regular.*

Proof. Write $L = \ell + Y$, where $\ell \in Y^{\perp}$, let $\rho > 0$, and set $\rho_L = \|\ell\| + \rho$. Now obtain κ as in Lemma 4.3, applied to ρ_L. Let $x \in \operatorname{ball}(0; \rho)$. Then $\|P_L x\| = \|\ell + P_Y x\| \leq \|\ell\| + \|P_Y x\| \leq \|\ell\| + \|x\| \leq \rho_L$. Hence $\|P_L x - TP_L x\| \geq \kappa d_{A\cap B}(P_L x)$. On the other hand, $\|P_L x - TP_L x\| = \|x - Tx\|$ and $d_{A\cap B}(P_L x) = d_{\operatorname{Fix} T}(x)$ by Proposition 4.1(vi)&(vii). Altogether, $\|x - Tx\| \geq \kappa d_{\operatorname{Fix} T}(x)$. ■

Remark 4.5 Lemma 4.2, which lies at the heart of this section, is proved in much greater generality in the recent paper [26]. The novelty here is to deduce bounded linear regularity of the Douglas–Rachford operator (see Theorem 4.4) in order to make it a useful building block to obtain other linear and strong convergence results.

5 Fejér Monotonicity and Set Regularities

5.1 Fejér monotone sequences and convergence for one operator

Since all algorithms considered in this paper generate Fejér monotone sequences, we review this key notion next.

Definition 5.1 (Fejér monotone sequence) *Let* C *be a nonempty subset of* X, *and let* $(x_n)_{n\in\mathbb{N}}$ *be a sequence in* X. *Then* $(x_n)_{n\in\mathbb{N}}$ *is* Fejér monotone *with respect to* C *if*

(41)
$$(\forall c \in C)(\forall n \in \mathbb{N}) \quad \|x_{n+1} - c\| \leq \|x_n - c\|.$$

Clearly, every Fejér monotone sequence is bounded. Let us now review some results concerning norm and linear convergence of Fejér monotone sequences.

Fact 5.2 (See, e.g., [6, Proposition 1.6].) *Let* $(x_n)_{n\in\mathbb{N}}$ *be a sequence in* X, *let* $\bar{x} \in X$, *and let* $p \in \{1, 2, \ldots\}$. *Suppose that* $(x_n)_{n\in\mathbb{N}}$ *is Fejér monotone with respect to* $\{\bar{x}\}$, *and that* $(x_{pn})_{n\in\mathbb{N}}$ *converges linearly to* $\bar{x}$. *Then* $(x_n)_{n\in\mathbb{N}}$ *itself converges linearly to* $\bar{x}$.

Fact 5.3 *Let* $(x_n)_{n\in\mathbb{N}}$ *be a sequence in* X *that is Fejér monotone with respect to a nonempty closed convex subset* C *of* X. *Then the following hold:*

(i) *If there exists* $\alpha \in [0, 1[$ *such that* $(\forall n \in \mathbb{N})\ d_C(x_{n+1}) \leq \alpha d_C(x_n)$, *then* $(x_n)_{n\in\mathbb{N}}$ *converges linearly to some point* $\bar{x} \in C$; *in fact,*

(42)
$$(\forall n \in \mathbb{N}) \quad \|x_n - \bar{x}\| \leq 2\alpha^n d_C(x_0).$$

(ii) *If C is an affine subspace and all weak cluster points of $(x_n)_{n\in\mathbb{N}}$ belong to C, then* $x_n \rightharpoonup P_C x_0$.

Proof. (i): See, e.g., [7, Theorem 5.12]. (ii): See, e.g., [7, Proposition 5.9(ii)]. ■

Corollary 5.4 *Let $T\colon X \to X$ be averaged nonexpansive and boundedly linearly regular, with* $\operatorname{Fix} T \neq \varnothing$. *Then for every $x_0 \in X$, the sequence $(T^n x_0)_{n\in\mathbb{N}}$ converges linearly to some point* $\bar{x} \in \operatorname{Fix} T$. *If* $\operatorname{Fix} T$ *is an affine subspace, then* $\bar{x} = P_{\operatorname{Fix} T} x_0$.

Proof. Let $x_0 \in X$. The sequence $(T^n x_0)_{n\in\mathbb{N}}$ is bounded because $\operatorname{Fix} T \neq \varnothing$. By (21) of Lemma 3.8, there exists $\alpha \in [0,1[$ such that $(\forall n \in \mathbb{N})\ d_{\operatorname{Fix} T}(x_{n+1}) \leq \alpha d_{\operatorname{Fix} T}(x_n)$. Hence Fact 5.3(i) implies linear convergence of $(T^n x_0)_{n\in\mathbb{N}}$. The remainder of the theorem follows from Fact 5.3(ii). ■

Corollary 5.4 implies the following example, which was analyzed in much greater detail in [4].

Example 5.5 (Douglas–Rachford operator for two subspaces) Let U and V be closed subspaces such that $U + V$ is closed, let $x_0 \in X$, and set $T = P_V P_U + P_{V^\perp} P_{U^\perp}$. Then $(T^n x_0)_{n\in\mathbb{N}}$ converges linearly to $P_{\operatorname{Fix} T} x_0$.

Proof. T is averaged (even firmly nonexpansive), and linearly regular by Example 2.5. Now apply Corollary 5.4. ■

Example 5.6 (Douglas–Rachford operator for transversal sets) Suppose that X is finite-dimensional, and let U and V be closed convex subsets of X such that $\operatorname{ri} U \cap \operatorname{ri} V \neq \varnothing$. Let $x_0 \in X$, and set $T = P_V R_U + \operatorname{Id} - P_U$. Then $(T^n x_0)_{n\in\mathbb{N}}$ converges linearly to some point $\bar{x} \in \operatorname{Fix} T$ such that $P_U \bar{x} \in U \cap V$.

Proof. Combine Theorem 4.4 with Corollary 5.4. ■

5.2 Regularities for families of sets

We now recall the notion of a collection of regular sets and key criteria. (The literature on regularity is vast and surveying it is outside the scope of this paper. Instead, we refer the interested reader to [28, Section 6] as a starting point for very recent work on regularity and constraint qualifications.) This will be crucial in the formulation of the linear convergence results.

Definition 5.7 ((bounded) (linear) regularity) *Let $(C_i)_{i\in I}$ be a finite family of closed convex subsets of X with $C = \bigcap_{i\in I} C_i \neq \varnothing$. We say that*[1]*:*

(i) $(C_i)_{i\in I}$ *is* linearly regular *if* $(\exists\, \mu > 0)$ $(\forall x \in X)$ $d_C(x) \leq \mu \max_{i\in I} d_{C_i}(x)$.

(ii) $(C_i)_{i\in I}$ *is* boundedly linearly regular *if* $(\forall \rho > 0)$ $(\exists\, \mu > 0)$ $(\forall x \in \mathrm{ball}(0;\rho))$ $d_C(x) \leq \mu \max_{i\in I} d_{C_i}(x)$.

(iii) $(C_i)_{i\in I}$ *is* regular *if for every sequence* $(x_n)_{n\in\mathbb{N}}$ *in* X*, we have* $\max_{i\in I} d_{C_i}(x_n) \to 0 \Rightarrow d_C(x_n) \to 0$.

(iv) $(C_i)_{i\in I}$ *is* boundedly regular *if for every bounded sequence* $(x_n)_{n\in\mathbb{N}}$ *in* X*, we have* $\max_{i\in I} d_{C_i}(x_n) \to 0 \Rightarrow d_C(x_n) \to 0$.

Fact 5.8 *Suppose that $I = \{1,\ldots,m\}$, and let $(C_i)_{i\in I}$ be a finite family of closed convex subsets of X with $C = \bigcap_{i\in I} C_i \neq \varnothing$. Then the following hold:*

(i) *Suppose each C_i is a subspace. Then $(C_i)_{i\in I}$ is regular in any of the four senses if and only if $\sum_{i\in I} C_i^{\perp}$ is closed.*

(ii) *Suppose each C_i is a cone. Then $(C_i \cap C^{\ominus})_{i\in I}$ is regular in any of the four senses if and only if $\sum_{i\in I}(C_i \cap C^{\ominus})^{\ominus}$ is closed.*

(iii) *Suppose each C_i is a cone and $C = \{0\}$. Then $(C_i)_{i\in I}$ is regular in any of the four senses if and only if $\sum_{i\in I} C_i^{\ominus}$ is closed.*

(iv) *If $C_m \cap \mathrm{int}(C_1 \cap \cdots \cap C_{m-1}) \neq \varnothing$, then $(C_i)_{i\in I}$ is boundedly linearly regular.*

(v) *If (C_1, C_2), $(C_1 \cap C_2, C_3)$, $\ldots$, $(C_1 \cap \cdots \cap C_{m-1}, C_m)$ are (boundedly) linearly regular, then so is $(C_i)_{i\in I}$.*

(vi) *If $0 \in \mathrm{sri}(C_1 - C_2)$, then (C_1, C_2) is boundedly linearly regular.*

(vii) *If each C_i is a polyhedron, then $(C_i)_{i\in I}$ is linearly regular.*

(viii) *If X is finite-dimensional, $C_1,\ldots,C_k$ are polyhedra, and $C_1 \cap \cdots C_k \cap \mathrm{ri}(C_{k+1}) \cap \cdots \cap \mathrm{ri}(C_m) \neq \varnothing$, then $(C_i)_{i\in I}$ is boundedly linearly regular.*

(ix) *If X is finite-dimensional, then $(C_i)_{i\in I}$ is boundedly regular.*

Proof. (i): [6, Theorem 5.19]. (ii): [17, Theorem 3.28]. (iii): [17, Corollary 3.30]. (iv): [6, Corollary 5.14]. (v): [6, Theorem 5.11]. (vi): [5, Corollary 4.5]. (vii): [6, Corollary 5.26]. (viii): [3, Theorem 5.6.2]. (ix): [5, Proposition 5.4.(iii)]. ■

[1]For each notion, one may replace the maximum by a sum because all norms on Euclidean spaces are equivalent. As the results in this work are *qualitative*, all conclusions remain unchanged.

Definition 5.9 (innate regularity) *Let $(C_i)_{i\in I}$ be a finite family of closed convex subsets of X with $C = \bigcap_{i\in I} C_i \neq \varnothing$. We say that $(C_i)_{i\in I}$ is* innately boundedly regular *if $(C_j)_{j\in J}$ is boundedly regular for every nonempty subset J of I. Innate regularity and innate (bounded) linear regularity are defined analogously.*

Fact 5.8 allows to formulate a variety of conditions sufficient for innate regularity. Here, we collect only some that are quite useful.

Corollary 5.10 *Let $(C_i)_{i\in I}$ be a finite family of closed convex subsets of X with $C = \bigcap_{i\in I} C_i \neq \varnothing$. Then the following hold:*

(i) *If X is finite-dimensional, then $(C_i)_{i\in I}$ is innately boundedly regular.*

(ii) *If X is finite-dimensional and $\bigcap_{i\in I} \operatorname{ri} C_i \neq \varnothing$, then $(C_i)_{i\in I}$ is innately linearly regular.*

(iii) *If each C_i is a subspace and $\sum_{j\in J} C_j^{\perp}$ is closed for every nonempty subset J of I, then $(C_i)_{i\in I}$ is innately linearly regular.*

Proof. (i): Fact 5.8(ix). (ii): Fact 5.8(viii). (iii): Fact 5.8(i). ■

6 Convergence Results for Quasi-Cyclic Algorithms

Unless otherwise stated, we assume from now on that

$$(T_i)_{i\in I} \tag{43}$$

is a finite family of nonexpansive operators from X to X with common fixed point set

$$Z = \bigcap_{i\in I} Z_i \neq \varnothing, \quad \text{where} \quad (Z_i)_{i\in I} = (\operatorname{Fix} T_i)_{i\in I}. \tag{44}$$

We are now ready for our first main result.

Theorem 6.1 (quasi-cyclic algorithm) *Suppose that each T_i is boundedly linearly regular and averaged nonexpansive. Suppose furthermore that $(Z_i)_{i\in I}$ is boundedly linearly regular. Let $(\omega_{i,n})_{(i,n)\in I\times\mathbb{N}}$ be such that $(\forall n\in\mathbb{N})$ $\sum_{i\in I}\omega_{i,n} = 1$ and $(\forall i\in I)$ $\omega_{i,n}\in[0,1]$. Set $(\forall n\in\mathbb{N})$ $I_n = \{i\in I \mid \omega_{i,n} > 0\}$ and suppose that $\omega_+ = \inf_{n\in\mathbb{N}}\inf_{i\in I_n}\omega_{i,n} > 0$. Suppose that there exists $p\in\{1,2,\ldots\}$ such that $(\forall n\in\mathbb{N})$ $I_n\cup I_{n+1}\cup\cdots\cup I_{n+p-1} = I$. Let $x_0\in X$ and generate a sequence $(x_n)_{n\in\mathbb{N}}$ in X by*

$$(\forall n\in\mathbb{N})\quad x_{n+1} = \sum_{i\in I}\omega_{i,n}T_i x_n. \tag{45}$$

Then $(x_n)_{n\in\mathbb{N}}$ converges linearly to some point in Z.

Proof. Set $\sigma_+ = \min_{i\in I}\sigma_i$, where $\sigma_i = \sigma(T_i)$. Let $i \in I$. By assumption,

$$(\forall k \in \mathbb{N})(\exists\, m_k \in \{kp, \ldots, (k+1)p-1\}) \quad i \in I_{m_k}. \tag{46}$$

Then

$$d_{Z_i}(x_{kp}) \le d_{Z_i}(x_{m_k}) + \|x_{kp} - x_{m_k}\| \le d_{Z_i}(x_{m_k}) + \sum_{n=kp}^{m_k-1}\|x_n - x_{n+1}\|. \tag{47}$$

Hence, by using Cauchy–Schwarz,

$$d^2_{Z_i}(x_{kp}) \le (m_k + 1 - kp)\Big(d^2_{Z_i}(x_{m_k}) + \sum_{n=kp}^{m_k-1}\|x_n - x_{n+1}\|^2\Big). \tag{48}$$

Get β_j as in (22) (with T replaced by T_j) and set $\beta_+ = \min_{j\in I}\beta_j > 0$. Let $z \in Z$. In view of Corollary 3.5, it follows that

$$\begin{aligned}
\|x_{kp} - z\|^2 - \|x_{(k+1)p} - z\|^2 &\ge \|x_{m_k} - z\|^2 - \|x_{m_k+1} - z\|^2 && \text{(49a)}\\
&\ge \omega_+\sigma_+\|x_{m_k} - T_i x_{m_k}\|^2 && \text{(49b)}\\
&\ge \omega_+\sigma_+\beta_+ d^2_{Z_i}(x_{m_k}). && \text{(49c)}
\end{aligned}$$

On the other hand, by Corollary 3.5,

$$(\forall n \in \mathbb{N}) \quad \|x_n - z\|^2 - \|x_{n+1} - z\|^2 \ge \sigma_+\|x_n - x_{n+1}\|^2. \tag{50}$$

In particular, $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to Z. Now we combine all of the above:

$$\begin{aligned}
d^2_{Z_i}(x_{kp}) &\le (m_k + 1 - kp)\big(d^2_{Z_i}(x_{m_k}) + \sum_{n=kp}^{m_k-1}\|x_n - x_{n+1}\|^2\big) && \text{(51a)}\\
&\le \underbrace{p\big(\omega_+^{-1}\sigma_+^{-1}\beta_+^{-1} + \sigma_+^{-1}\big)}_{=\lambda}\Big(\|x_{kp} - z\|^2 - \|x_{(k+1)p} - z\|^2\Big). && \text{(51b)}
\end{aligned}$$

Applying this with $z = P_Z x_{kp}$ (and releasing i) yields

$$\max_{i\in I} d^2_{Z_i}(x_{kp}) \le \lambda\big(d^2_Z(x_{kp}) - d^2_Z(x_{(k+1)p})\big). \tag{52}$$

On the other hand, bounded linear regularity yields $\mu > 0$ such that $(\forall n \in \mathbb{N})\ d_Z(x_n) \le \mu \max_{i\in I} d_{Z_i}(x_n)$. Altogether,

$$d^2_Z(x_{kp}) \le \lambda\mu^2\big(d^2_Z(x_{kp}) - d^2_Z(x_{(k+1)p})\big). \tag{53}$$

By Fact 5.3(i), the sequence $(x_{kp})_{k\in\mathbb{N}}$ converges linearly to some point $\bar{z} \in Z$. It now follows from Fact 5.2 that $(x_n)_{n\in\mathbb{N}}$ converges linearly to $\bar{z}$. ■

Theorem 6.1 is quite flexible in the amount of control a user has in generating sequences. We point out two very popular instances next.

Corollary 6.2 (cyclic algorithm) *Suppose that $I = \{1, \ldots, m\}$, and that each T_i is boundedly linearly regular and averaged nonexpansive. Suppose furthermore that $(Z_i)_{i \in I}$ is boundedly linearly regular. Let $x_0 \in X$ and generate a sequence $(x_n)_{n \in \mathbb{N}}$ in X by*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = T_m \cdots T_2 T_1 x_n. \tag{54}$$

Then $(x_n)_{n \in \mathbb{N}}$ converges linearly to some point in Z.

Corollary 6.3 (parallel algorithm) *Suppose that $I = \{1, \ldots, m\}$, and that each T_i is boundedly linearly regular and averaged nonexpansive. Suppose furthermore that $(Z_i)_{i \in I}$ is boundedly linearly regular. Let $x_0 \in X$ and generate a sequence $(x_n)_{n \in \mathbb{N}}$ in X by*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = \frac{1}{m} \sum_{i \in I} T_i x_n. \tag{55}$$

Then $(x_n)_{n \in \mathbb{N}}$ converges linearly to some point in Z.

Some concrete and new results will be considered in Section 8; there are already several known results that can be deduced from this framework (see, e.g., [6] and [23]).

Remark 6.4 We mention here the related frameworks by Kiwiel and Łopuch [23] who bundled regularity of the fixed point sets together with regularity of the operators to study accelerated generalizations of projection methods. Theirs and our techniques find their roots in [6]; see also [3]. We feel that the approach presented here is more convenient for applications; indeed, one first checks that the operators are well behaved — the algorithms will be likewise if the fixed point sets relate well to each other.

We end this section with the following probabilistic result whose basic form is due to Leventhal [25]. The proof presented here is somewhat simpler and the conclusion is stronger.

Corollary 6.5 (probabilistic algorithm) *Suppose that each T_i is boundedly linearly regular and averaged nonexpansive. Suppose furthermore that $(Z_i)_{i \in I}$ is boundedly linearly regular. Let $x_0 \in X$ and generate a sequence $(x_n)_{n \in \mathbb{N}}$ in X by*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = T_i x_n \tag{56}$$

with probability $\pi_i > 0$. Then $(x_n)_{n \in \mathbb{N}}$ converges linearly almost surely to a solution in the sense that there exists a constant $\theta < 1$, depending only on $\|x_0\|$, such that

$$(\forall n \in \mathbb{N}) \quad \mathbf{E}\, d_Z^2(x_{n+1}) \leq \theta d_Z^2(x_n). \tag{57}$$

Proof. Let $z \in Z$, and let $n \in \mathbb{N}$. Then $\|x_{n+1}\| = \|T_i x_n\| \leq \|T_i x_n - z\| + \|z\| \leq \|x_n - z\| + \|z\| \leq \|x_0 - z\| + \|z\|$, hence every instance of $(x_n)_{n\in\mathbb{N}}$ satisfies $\sup_{n\in\mathbb{N}} \|x_n\| \leq \|x_0 - z\| + \|z\| = \rho$. Hence, by (23) of Lemma 3.8, we obtain γ_i such that

$$\gamma_i d^2_{Z_i}(x_n) \leq d^2_Z(x_n) - d^2_Z(T_i x_n). \tag{58}$$

On the other hand, by bounded linear regularity of $(Z_1, \ldots, Z_m)$, we get $\mu > 0$ such that

$$\mu d^2_Z(x_n) \leq \sum_i \pi_i \gamma_i d^2_{Z_i}(x_n). \tag{59}$$

Combining and taking the expected value, we deduce

$$\mu d^2_Z(x_n) \leq d^2_Z(x_n) - \mathbf{E}\, d^2_Z(x_{n+1}), \tag{60}$$

and the result follows with $\theta = 1 - \mu$. ■

7 Convergence Results for Cyclic and Random Algorithms

In this section, we focus on strong convergence results for algorithms which utilize the operators either cyclically or in a more general, not necessarily quasicyclic, fashion. Simple examples involving projectors show that linear convergence results are not to be expected. Accordingly, the less restrictive notion of (bounded) regularity is introduced — it is sufficient for strong convergence.

We start our analysis with the following notion which can be seen as a qualitative variant of (bounded) linear regularity.

Definition 7.1 ((bounded) regularity) *Let $T\colon X \to X$ be such that* $\operatorname{Fix} T \neq \varnothing$. *We say that:*

(i) *T is **regular** if for every sequence $(x_n)_{n\in\mathbb{N}}$ in X, we have*

$$x_n - Tx_n \to 0 \quad \Rightarrow \quad d_{\operatorname{Fix} T}(x_n) \to 0. \tag{61}$$

(ii) *T is **boundedly regular** if for every sequence $(x_n)_{n\in\mathbb{N}}$ in X, we have*

$$(x_n)_{n\in\mathbb{N}} \text{ bounded and } x_n - Tx_n \to 0 \quad \Rightarrow \quad d_{\operatorname{Fix} T}(x_n) \to 0. \tag{62}$$

Comparing with Definition 2.1, we note that

$$\text{linear regularity} \Rightarrow \text{regularity} \tag{63}$$

and that

$$\text{bounded linear regularity} \Rightarrow \text{bounded regularity}. \tag{64}$$

These notions are much less restrictive than their quantitative linear counterparts:

Proposition 7.2 *Let $T\colon X \to X$ be continuous, suppose that X is finite-dimensional*[2] *and that* $\operatorname{Fix} T \neq \varnothing$. *Then T is boundedly regular.*

We now turn to "property (S)", a notion first considered by Dye *et al.* in [19].

Definition 7.3 (property (S)) *Let $T\colon X \to X$ be nonexpansive such that* $\operatorname{Fix} T \neq \varnothing$. *Then T has property (S) with respect to $z \in \operatorname{Fix} T$ if for every bounded sequence $(x_n)_{n\in\mathbb{N}}$ such that $\|x_n - z\| - \|Tx_n - z\| \to 0$, we have $x_n - Tx_n \to 0$.*

Proposition 7.4 *Let $T\colon X \to X$ be averaged nonexpansive such that* $\operatorname{Fix} T \neq \varnothing$. *Then T has property (S) with respect to* $\operatorname{Fix} T$.

Proof. Let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence in X such that $\|x_n - z\| - \|Tx_n - z\| \to 0$, where $z \in \operatorname{Fix} T$. Clearly, $(\|x_n - z\| + \|Tx_n - z\|)_{n\in\mathbb{N}}$ is bounded since $(x_n)_{n\in\mathbb{N}}$ and $(Tx_n)_{n\in\mathbb{N}}$ are. It follows that $\|x_n - z\|^2 - \|Tx_n - z\|^2 \to 0$. By Fact 3.3, $x_n - Tx_n \to 0$. ■

Definition 7.5 (projective) *Let $T\colon X \to X$ be nonexpansive such that* $\operatorname{Fix} T \neq \varnothing$, *and let $z \in \operatorname{Fix} T$. Then T is* projective *with respect to $z \in \operatorname{Fix} T$ if for every bounded sequence $(x_n)_{n\in\mathbb{N}}$ such that $\|x_n - z\| - \|Tx_n - z\| \to 0$, we have $d_{\operatorname{Fix} T}(x_n) \to 0$. We say that T is* projective *if it is projective with respect to all its fixed points.*

Projectivity implies property (S):

Lemma 7.6 *Let $T\colon X \to X$ be nonexpansive and suppose that T is projective with respect to $z \in \operatorname{Fix} T$. Then T has property (S) with respect to z.*

Proof. Observe that

$$(\forall x \in X) \quad \|x - Tx\| \leq \|x - P_{\operatorname{Fix} T}x\| + \|P_{\operatorname{Fix} T}x - Tx\| \tag{65a}$$

$$\leq 2\|x - P_{\operatorname{Fix} T}x\| = 2d_{\operatorname{Fix} T}(x). \tag{65b}$$

Now let $(x_n)_{n\in\mathbb{N}}$ be a bounded sequence such that $\|x_n - z\| - \|Tx_n - z\| \to 0$. Since T is projective with respect to z, we have $d_{\operatorname{Fix} T}(x_n) \to 0$. By (65), $x_n - Tx_n \to 0$. ■

The importance of projectivity stems from the following observation.

[2] Or, more generally, that $\operatorname{ran} T$ is boundedly compact.

Fact 7.7 *Let $T\colon X \to X$ be nonexpansive such that T is projective with respect to some fixed point of T. Then $(T^n x_0)_{n\in\mathbb{N}}$ converges strongly to a fixed point for every starting point $x_0 \in X$.*

Proof. See [2, Lemma 2.8.(iii)]. ■

Proposition 7.8 *Let $I = \{1, \ldots, m\}$, and let $(T_i)_{i\in I}$ be nonexpansive mappings with fixed point sets $(Z_i)_{i\in I}$. Set $Z = \bigcap_{i\in I} Z_i$ and suppose that there exists $z \in Z$ such that each T_i is projective with respect to z and that $(Z_i)_{i\in I}$ is boundedly regular. Then $T = T_m \cdots T_2 T_1$ is projective with respect to z as well. Consquently, for every $x_0 \in X$, $(T^n x_0)_{n\in\mathbb{N}}$ converges strongly to some point in Z.*

Proof. Suppose that $(x_n)_{n\in\mathbb{N}}$ is a bounded sequence in X such that $\|x_n - z\| - \|Tx_n - z\| \to 0$. Note that

$$0 \le \sum_{i=1}^{m} \|T_{i-1} \cdots T_1 x_n - z\| - \|T_i T_{i-1} \cdots T_1 x_n - z\| = \|x_n - z\| - \|Tx_n - z\| \to 0, \tag{66}$$

that each sequence $(T_{i-1} \cdots T_1 x_n)_{n\in\mathbb{N}}$ is bounded, and that

$$(\forall i \in I) \quad \|T_{i-1} \cdots T_1 x_n - z\| - \|T_i T_{i-1} \cdots T_1 x_n - z\| \to 0. \tag{67}$$

Combining this with the assumption that each T_i is projective with respect to z, we deduce two consequences. First,

$$(\forall i \in I) \quad T_{i-1} \cdots T_1 x_n - T_i T_{i-1} \cdots T_1 x_n \to 0 \tag{68}$$

by Lemma 7.6. Second,

$$(\forall i \in I) \quad d_{Z_i}(T_{i-1} \cdots T_1 x_n) \to 0. \tag{69}$$

Altogether, $(\forall i \in I)$ $d_{Z_i}(x_n) \to 0$. Since $(Z_i)_{i\in I}$ is boundedly regular, it follows that $d_Z(x_n) \to 0$. Now $Z \subseteq \operatorname{Fix} T$ yields $d_{\operatorname{Fix} T} \le d_Z$; consequently, $d_{\operatorname{Fix} T}(x_n) \to 0$. Hence T is projective with respect to z and the result now follows from Fact 7.7. ■

Property (S) in tandem with bounded regularity implies projectivity, which turns out to be crucial for the results on random algorithms.

Proposition 7.9 *Let $T\colon X \to X$ be nonexpansive such that $\operatorname{Fix} T \neq \varnothing$, and let $z \in \operatorname{Fix} T$. Suppose that T satisfies property (S) with respect to z, and that T is boundedly regular. Then T is projective with respect to z.*

Proof. Let $(x_n)_{n\in\mathbb{N}}$ be bounded such that $\|x_n - z\| - \|Tx_n - z\| \to 0$. By property (S), $x_n - Tx_n \to 0$. By bounded regularity, $d_{\operatorname{Fix} T}(x_n) \to 0$, as required. ■

The next result is quite useful.

Corollary 7.10 *Let $T\colon X \to X$ be averaged nonexpansive and boundedly regular such that* $\operatorname{Fix} T \neq \varnothing$. *Then T is projective with respect to* $\operatorname{Fix} T$.

Proof. Combine Proposition 7.4 and Proposition 7.9. ■

We now obtain a powerful strong convergence result for cyclic algorithms.

Theorem 7.11 (cyclic algorithm) *Set $I = \{1, \dots, m\}$, and let $(T_i)_{i\in I}$ be family of averaged nonexpansive mappings from X to X with fixed point sets $(Z_i)_{i\in i}$, respectively. Suppose that each T_i is boundedly regular, that $Z = \bigcap_{i\in I} Z_i \neq \varnothing$, and that $(Z_i)_{i\in I}$ is boundedly regular. Then for every $x_0 \in X$, the sequence $((T_m \cdots T_1)^n x_0)_{n\in\mathbb{N}}$ converges strongly to some point in Z.*

Proof. By Corollary 7.10, each T_i is projective with respect to every point in Z. The result thus follows from Proposition 7.8. ■

Let us now turn to random algorithms.

Definition 7.12 (random map) *The map $r\colon \mathbb{N} \to I$ is a* random map *for I if $(\forall i \in I)$ $r^{-1}(i)$ contains infinitely many elements.*

Fact 7.13 (See [2, Theorem 3.3].) *Suppose that $(T_i)_{i\in I}$ are projective with respect to a common fixed point, and that $(Z_i)_{i\in I}$ is innately boundedly regular. Let $x_0 \in X$, let r be a random map for I, and generate a sequence $(x_n)_{n\in\mathbb{N}}$ in X by*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = T_{r(n)} x_n. \tag{70}$$

Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to some point in Z.

We are ready for our last main result.

Theorem 7.14 (random algorithm) *Suppose that each T_i is averaged nonexpansive and boundedly regular, and that $(Z_i)_{i\in I}$ is innately boundedly regular. Let $x_0 \in X$, let r be a random map for I, and generate a sequence $(x_n)_{n\in\mathbb{N}}$ in X by*

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = T_{r(n)} x_n. \tag{71}$$

Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to some point $\bar{z} \in Z$. If Z is an affine subspace, then $\bar{z} = P_Z x_0$.

Proof. By Corollary 7.10, each T_i is projective with respect to Z_i and hence with respect to Z. Now apply Fact 7.13 and Fact 5.3(ii). ■

8 Applications and Numerical Results

8.1 The Borwein–Tam Method (BTM)

In this section, $I = \{1, \ldots, m\}$ and $(U_i)_{i\in I}$ is a family of closed convex subsets of X with

(72) $$U = \bigcap_{i\in I} U_i \neq \varnothing.$$

Now set $U_{m+1} = U_1$,

(73) $$(\forall i \in I) \quad T_i = T_{U_{i+1},U_i} = P_{U_{i+1}} R_{U_i} + \mathrm{Id} - P_{U_i}, \quad Z_i = \operatorname{Fix} T_i, \quad Z = \bigcap_{i\in I} Z_i$$

and define the *Borwein–Tam operator* by

(74) $$T = T_m T_{m-1} T_{m-2} \cdots T_2 T_1.$$

The following result is due to Borwein and Tam (see [11, Theorem 3.1]):

Fact 8.1 (Borwein–Tam method (BTM)) *Let $x_0 \in X$ and generate the sequence $(x_n)_{n\in\mathbb{N}}$ by*

(75) $$(\forall n \in \mathbb{N}) \quad x_{n+1} = T^n x_0.$$

Then $(x_n)_{n\in\mathbb{N}}$ converges weakly to a point $\bar{x} \in Z$ such that $P_{U_1}\bar{x} = \cdots = P_{U_m}\bar{x} \in U$.

The following new results now follow from our analysis.

Corollary 8.2 (transversal sets) *Suppose that X is finite-dimensional and that $\bigcap_{i\in I} \mathrm{ri}(U_i) \neq \varnothing$. Then the convergence of the Borwein–Tam method is with a linear rate.*

Proof. Theorem 4.4 implies that each T_i is boundedly linearly regular. Now set $(\forall i \in I)$ $Y_i = \operatorname{span}(U_{i+1} - U_i)$. By Proposition 4.1(iii), $(\forall i \in I)$ $Z_i = U_i \cap U_{i+1} + Y_i^{\perp}$. It thus follows from [27, Theorem 6.5 and Corollary 6.6.2] that $(\forall i \in I)$ $\mathrm{ri}(U_i) \cap \mathrm{ri}(U_{i+1}) \subseteq \mathrm{ri}(U_i \cap U_{i+1}) + \mathrm{ri}(Y_i^{\perp}) = \mathrm{ri}(Z_i)$. Hence $\bigcap_{i\in I} \mathrm{ri}(Z_i) \supseteq \bigcap_{i\in I} \mathrm{ri}(U_i) \cap \mathrm{ri}(U_{i+1}) \neq \varnothing$. Therefore, $(Z_i)_{i\in I}$ is boundedly linearly regular by Fact 5.8(viii). The conclusion now follow from Corollary 6.2. ■

Corollary 8.3 (subspaces) *Suppose that each U_i is a subspace*[3] *with $U_i + U_{i+1}$ is closed, and that $(Z_i)_{i\in I}$ is boundedly linearly regular. Then the convergence of the Borwein–Tam method is with a linear rate.*

[3] A simple translation argument yields a version for affine subspaces with a nonempty intersection.

Proof. Combine Example 2.5 with Corollary 6.2. ■

Of course, using Theorem 6.1, we can formulate various variants for a general quasi-cyclic variant. We conclude this section with a random version.

Example 8.4 (subspaces — random version) Suppose the hypothesis of Corollary 8.3 holds. Assume in addition that $(Z_i)_{i\in I}$ is innately boundedly regular. Let r be a random map for I, let $x_0 \in X$, and set $(\forall n \in \mathbb{N})\ x_{n+1} = T_{r(n)}x_n$. Then $(x_n)_{n\in\mathbb{N}}$ converges strongly to $P_Z x_0$.

Proof. Combine Example 2.5 with Theorem 7.14. ■

8.2 The Cyclically Anchored Douglas–Rachford Algorithm (CADRA)

In this section, we assume that $I = \{1, \ldots, m\}$, that A is a closed convex subset of X, also referred to as the *anchor*, and that $(B_i)_{i\in I}$ is a family of closed convex subsets of X such that

$$C = A \cap \bigcap_{i\in I} B_i \neq \varnothing. \tag{76}$$

We set

$$(\forall i \in I) \quad T_i = P_{B_i} R_A + \operatorname{Id} - P_A, \quad Z_i = \operatorname{Fix} T_i; \quad Z = \bigcap_{i\in I} Z_i. \tag{77}$$

The *Cyclically Anchored Douglas–Rachford Algorithm (CADRA)* with starting point $x_0 \in X$ generates a sequence $(x_n)_{n\in\mathbb{N}}$ by iterating

$$(\forall n \in \mathbb{N}) \quad x_{n+1} = Tx_n, \quad \text{where} \ \ T = T_m \cdots T_2 T_1. \tag{78}$$

Note that when $m = 1$, then CADRA coincides with the classical Douglas–Rachford algorithm[4].

Let us record a central convergence result concerning the CADRA.

Theorem 8.5 (CADRA) *The sequence $(x_n)_{n\in\mathbb{N}}$ generated by CADRA converges weakly to a point $\bar{x} \in Z$ such that $P_A\bar{x} \in C$. Furthermore, the convergence is linear provided that one of the following holds:*

(i) *X is finite-dimensional and that* $\operatorname{ri}(A) \cap \bigcap_{i\in I} \operatorname{ri}(B_i) \neq \varnothing$.

[4] This is not the case for the BTM considered in the previous subsection.

(ii) *A and each B_i is a subspace with $A + B_i$ closed and that $(Z_i)_{i\in I}$ is boundedly linearly regular.*

Proof. The weak convergence follows from e.g. [7, Theorem 5.22]. (i): Now combine Theorem 4.4 with Corollary 6.2. (ii): Combine Example 2.5 with Corollary 6.2. ■

One may also obtain a random version of CADRA by using Theorem 7.14.

8.3 Numerical experiments

We now work in $X = \mathbb{R}^{100}$. We set $A = \mathbb{R}^{50}_{+} \times \{0\} \subset X$, and we let each B_i be a hyperplane with normal vector in $\mathbb{R}^{100}_{++}$, where $1 \leq i \leq m$ and $1 \leq m \leq 50$. Using the programming language `julia` [22], we generated these data randomly, where for each $m \in \{1, \ldots, 50\}$, the problem

$$\text{find } x \in A \cap \bigcap_{i\in\{1,\ldots,m\}} B_i \tag{79}$$

has a solution in $\operatorname{ri} A$. We then choose 10 random starting points in $\mathbb{R}^{100}_{+}$, each with Euclidean norm equal to 100. Altogether, we obtain 50 problems and 500 instances for each of the algorithms Cyclic Projections (CycP), BTM, and CADRA applied to the sets $A, B_1, \ldots, B_m$. If $(x_n)_{n\in\mathbb{N}}$ is the main sequence generated by one of these algorithms and $(z_n)_{n\in\mathbb{N}} = (P_A x_n)_{n\in\mathbb{N}}$, then we terminate at stage n when

$$\max\left\{d_{B_1}(z_n), \ldots, d_{B_m}(z_n)\right\} \leq 10^{-3}. \tag{80}$$

We divide the 50 problems into 5 groups, depending on the value of m. In Table 1, we record the median of the number of iterations required for each algorithm to terminate, and we also list the number of wins[5] that each algorithm is the fastest among the three.

Finally, we observe that CADRA performs quite well compared to CycP and BTM, especially when the range of parameters keep the problems moderately underdetermined.

Acknowledgments

We are indebted to the referees for very careful reading and constructive comments. We also would like to thank a referee for bringing [28] to our attention. HHB was partially

[5] If algorithms were tied for first place, they were both declared winner. In our experiment, we had 5 ties for first place between CycP and BTM when $m \in \{1, \ldots, 10\}$.

Range of m	CycP		BTM		CADRA	
	Iterations	Wins	Iterations	Wins	Iterations	Wins
1–10	79.5	10	78.5	52	80.0	43
11–20	391.0	0	384.0	0	179.5	100
21–30	932.0	0	942.5	2	370.5	98
31–40	1,645.0	8	1,690.5	6	959.5	86
41–50	4,749.0	35	4,482.5	35	5,151.5	30

Table 1: Median of number of iterations and number of wins

supported by the Natural Sciences and Engineering Research Council of Canada and by the Canada Research Chair Program. DN acknowledges hospitality of the University of British Columbia in Kelowna and support by the Pacific Institute of the Mathematical Sciences during the preparation of this paper. HMP was partially supported by an NSERC accelerator grant of HHB.

References

[1] F.J. Aragón Artacho, A.L. Dontchev, and M.H. Geoffroy, Convergence of the proximal point method for metrically regular mappings, *ESAIM Proceedings* 17 (2007), 1–8.

[2] H.H. Bauschke, A norm convergence result on random products of relaxed projections in Hilbert space, *Transactions of the AMS* 347(4) (April 1995), 1365–1373.

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

[4] H.H. Bauschke, J.Y. Bello Cruz, T.T.A. Nghia, H.M. Phan, and X. Wang, The rate of linear conergence of the Douglas–Rachford algorithm for subspaces is the cosine of the Friedrichs angle, *Journal of Approximation Theory*, in press. `http://arxiv.org/abs/1309.4709`.

[5] H.H. Bauschke and J.M. Borwein, On the convergence of von Neumann's alternating projection algorithm for two sets, *Set-Valued Analysis* 1 (1993), 185–212.

[6] H.H. Bauschke and J.M. Borwein, On projection algorithms for solving convex feasibility problems, *SIAM Review* 38(3) (1996), 367–426.

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

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

[9] H.H. Bauschke, D.R. Luke, H.M. Phan, and X. Wang, Restricted normal cones and the method of alternating projections: theory, *Set-Valued and Variational Analysis* 21 (2013), 431–473.

[10] H.H. Bauschke, W.L. Hare, and W.M. Moursi, Generalized solutions for the sum of two maximally monotone operators, *SIAM Journal on Control and Optimization*, in press.

[11] J.M Borwein and M.K. Tam, A cyclic Douglas–Rachford iteration scheme, *Journal of Optimization Theory and Applications* 160 (2014), 1–29.

[12] A. Cegielski, *Iterative Methods for Fixed Point Problems in Hilbert Spaces*, Springer, 2012.

[13] Y. Censor and S.A. Zenios, *Parallel Optimization*, Oxford University Press, 1997.

[14] P.L. Combettes, The convex feasibility problem in image recovery, *Advances in Imaging and Electronic Physics* 95 (1996), 155–270.

[15] P.L. Combettes, Solving monotone inclusions via compositions of nonexpansive averaged operators, *Optimization* 53 (2004), 475–504.

[16] F. Deutsch, *Best Approximation in Inner Product Spaces*, Springer, 2001.

[17] F. Deutsch and H. Hundal, The rate of convergence for the cyclic projections algorithm III: regularity of convex sets, *Journal of Approximation Theory* 155 (2008), 155–184.

[18] A.L. Dontchev and R.T. Rockafellar, *Implicit Functions and Solution Mappings*, Springer 2009.

[19] J.M. Dye, T. Kuczumow, P.-K. Lin, and S. Reich, Random products of nonexpansive mappings in spaces with the Opial property, *Contemporary Mathematics* 144 (1993), 87–93.

[20] K. Goebel and W.A. Kirk, *Topics in Metric Fixed Point Theory*, Cambridge University Press, 1990.

[21] K. Goebel and S. Reich, *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings*, Marcel Dekker, 1984.

[22] `http://julialang.org`

[23] K.C. Kiwiel and B. Łopuch, Surrogate projection methods for finding fixed points of firmly nonexpansive mappings, *SIAM Journal on Optimiztion* 7 (1997), 1084–1102.

[24] A.Y. Kruger, About regularity of collections of sets, *Set-Valued Analysis* 14 (2006), 187–206.

[25] D. Leventhal, Metric subregularity and the proximal point method, *Journal of Mathematical Analysis and Applications* 360 (2009), 681–688.

[26] H.M. Phan, Linear convergence of the Douglas–Rachford method for two closed sets, preprint 2014, `http://arxiv.org/abs/1401.6509`

[27] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

[28] Z. Wei, J.-C. Yao, and X.Y. Zheng, Strong Abadie CQ, ACQ, calmness and linear regularity, *Mathematical Programming (Series A)* 145 (2014), 97–131.

[29] I. Yamada, The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings, *Studies in Computational Mathematics* 8 (2001), 473–504.