SIAM J. OPTIM.
Vol. 30, No. 1, pp. 542–584
© 2020 Society for Industrial and Applied Mathematics

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

CONVERGENCE ANALYSIS OF THE RELAXED DOUGLAS–RACHFORD ALGORITHM*

D. RUSSELL LUKE† AND ANNA-LENA MARTINS†

Abstract. Motivated by nonconvex, inconsistent feasibility problems in imaging, the relaxed alternating averaged reflections algorithm, or relaxed Douglas–Rachford algorithm (DRλ), was first proposed over a decade ago. Convergence results for this algorithm are limited to either convex feasibility or consistent nonconvex feasibility with strong assumptions on the regularity of the underlying sets. Using an analytical framework depending only on metric subregularity and pointwise almost averagedness, we analyze the convergence behavior of DRλ for feasibility problems that are both nonconvex and inconsistent. We introduce a new type of regularity of sets, called superregular at a distance, to establish sufficient conditions for local linear convergence of the corresponding sequence. These results subsume and extend existing results for this algorithm.

Key words. superregular, inconsistent feasibility problem, projection, relaxed averaged alternating reflections, fixed point, linear convergence, metric subregularity, nonconvex, subtransversality

AMS subject classifications. 65K10, 49K40, 49M05, 65K05, 90C26, 49M20, 49J53

DOI. 10.1137/18M1229638

1. Introduction.

The *feasibility problem* consists of finding a common point in a collection of closed sets. If no such common point exists, the feasibility problem is called *inconsistent*, and one seeks instead an adequate approximation to the problem. Typically, feasibility problems are solved by projection-based algorithms. Among these are von Neumann's alternating projections [26] and its many-set version, the cyclic projection algorithm, or averaged projections, and, in the case of two-set feasibility, the Douglas–Rachford algorithm [12] as formulated by Lions and Mercier [18].

Alternating and cyclic projections have long been standard iterative procedures. They are stable and reliable in the sense that they always seem to converge to a limit cycle, though the cycle is not always desirable or easy to interpret [1]. Because it has so many different formulations, the Douglas–Rachford algorithm has been rediscovered many times and has become quite popular in the past decade. This algorithm has many curious features, the first of which is that the iterates do not, in general, converge to solutions to the target feasibility problem, when they converge at all. The second unusual feature of the algorithm is that it cannot converge if the feasibility problem is inconsistent. For convex feasibility the iterates *diverge* in the direction of the gap between the sets [18, 13, 5]. In the convex setting this is not too worrisome since the *shadows* of the iterates, defined as the projection of the iterates onto one of the sets (the "inner set"), converge to a *best approximation point* [5, 6]. For consistent nonconvex feasibility, Hesse and Luke [14] were the first to prove meaningful local convergence results for Douglas–Rachford. This was quickly followed by several generalizations [3, 25, 17, 23, 9]. For inconsistent feasibility, since Douglas–Rachford cannot converge, weak convergence follows generically if the iterates are bounded,

*Received by the editors November 29, 2018; accepted for publication (in revised form) November 14, 2019; published electronically February 20, 2020.
https://doi.org/10.1137/18M1229638
Funding: The first author was supported in part by German Research Foundation grants SFB755-C02 and GRK2088-B5. The second author was supported by the German Research Foundation grant SFB755-C02.
†Institut für Numerische und Angewandte Mathematik, Universität Göttingen, Göttingen 37083, Germany (r.luke@math.uni-goettingen.de, a.martins@math.uni-goettingen.de).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

but otherwise meaningful results appear to only be possible for relaxations of the Douglas–Rachford algorithm.

To address failure of convergence of Douglas–Rachford for inconsistent feasibility, Luke introduced the relaxed Douglas–Rachford algorithm in [19] with a proof of convergence for convex feasibility—inconsistent and consistent. Given $x^0 \in \mathcal{E}$ and $\lambda \in (0,1)$ for $k = 0, 1, 2, \ldots$, the iteration takes the form

$$x^{k+1} \in T_{\mathrm{DR}\lambda}(x^k) := \bigcup_{b \in P_B(x^k)} \left\{ \frac{\lambda}{2} \left(R_A(2b - x^k) + x^k \right) + (1-\lambda)\, b \right\}. \tag{1.1}$$

Here R_A is the *reflector* across the set A and P_B is the *projector* onto B (see the next subsection for details). For $\lambda = 1$ this mapping is the Douglas–Rachford fixed point mapping which, in the convex setting, takes the form $T_{\mathrm{DR}} := \frac{1}{2}\left(R_A R_B + \mathrm{Id}\right)$, where Id denotes the identity. From here on, we will refer to the algorithm as DRλ. A characterization of the fixed points in the nonconvex inconsistent case and a first attempt at a local convergence result were given in [20]. The analysis required one of the sets to be convex and the other set to be *prox-regular*. More recently, Li and Pong [17] rediscovered this algorithm and showed convergence results when both sets are closed, one set is convex, at least one of the sets is compact, and the intersection is nonempty (i.e.. consistent feasibility). When, in addition, both sets are semialgebraic, they showed global convergence [17, Corollary 1]. Under still stronger assumptions (that one of the sets is linear and the other semialgebraic and that the intersection is *strongly regular* [16]), local linear convergence can be shown [17, Proposition 2]. Minh and Phan [9, Theorem 5.8] show local linear convergence of a generalized Douglas–Rachford algorithm (which includes $T_{\mathrm{DR}\lambda}$ as a special case) when the two sets are superregular and have sufficiently regular nonempty intersection. Noteworthy here is that none of these approaches can explain convergence in the case of two affine half-spaces with empty intersection, much less for any other inconsistent feasibility problem, convex or otherwise.

In the present work, we extend the results above to inconsistent feasibility for sets with the weakest regularity assumptions to date. We introduce in section 2 a new kind of set regularity, called *superregularity at a distance*, that will be our only assumption on the sets themselves. Superregularity at a distance falls into the spectrum of other regularity notions like ϵ-subregularity (cf. [15, 8]) and, as the name suggests, superregularity [16]. The innovation of this characterization is that it allows one to describe the regularity of a set relative to a point not in that set; that is, it characterizes how the set looks from the outside. This is especially important for the relaxed Douglas–Rachford algorithm, whose fixed points do not usually lie in any of the sets. As in [20], however, the projections of the fixed points are shown to include best approximation points (Theorem 3.13 and Corollary 3.15).

Following the framework established in [23], in section 4 we prove local linear convergence of the algorithm under additional assumptions on the regularity of the *collection of sets* taken together. Unlike previous notions of regularity of collections of sets [15], the sets in the present analysis need not have points in common. The analysis of [23] uses two properties of fixed point mappings. The first property, *pointwise almost averagedness*, follows from the regularity of the sets and, as shown in [22, Proposition 4], is an important ingredient in guaranteeing convergence of the iterates to fixed points. In Theorem 3.7 we establish that the $T_{\mathrm{DR}\lambda}$ mapping is almost averaged at its fixed points when the sets A and B are superregular at a distance. The second property, *metric subregularity* of the fixed point mapping at its fixed points, was

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

subsequently shown in [22, Theorem 2] to be necessary for local linear convergence. In the context of feasibility, this property becomes *subtransversality* of the sets in relation to each other plus an additional technical condition. Under these conditions, [23, Theorem 3.2] establishes local linear convergence of cyclic projections onto sufficiently regular sets that need not have points in common. Following their approach we show in Theorem 4.11 that a similar result is true for DRλ. We conclude our study with a demonstration of this theory in section 5 via several elementary examples that allow explicit evaluation of the relevant constants.

1.1. Notation and definitions. Our notation is standard in variational analysis. Our setting is a finite-dimensional Euclidean space, denoted $\mathcal{E}$, with inner product $\langle\cdot,\cdot\rangle$ and induced norm $\|\cdot\|$. We denote by $\mathbb{B}$ the open unit ball and by $\mathbb{B}_\delta(x)$ the open ball with radius δ around the point x. The model we consider is a feasibility problem, that is, the problem of finding points common to closed subsets of $\mathcal{E}$ or reasonable substitutions thereof when the sets have no points in common. The distance of a point x to a set C is $\operatorname{dist}(x,C) := \inf_{y\in C}\|x-y\|$, and the *projector* onto C is the *set-valued mapping* $P_C(x) := \{z \mid \|z-x\| = \operatorname{dist}(x,C)\}$. A *projection* is a selection from $P_C(x)$. The *reflector* of a point x across C is $R_C(x) := 2P_C(x) - x$, and a *reflection* is a selection from this set-valued mapping. For the purposes of this paper, we define the normal cone to the set C in terms of the projector onto that set.

Definition 1.1 (normal cones). *Let $C \subseteq \mathcal{E}$, and let $\bar{x} \in C$.*

(i) *The* proximal normal cone *of C at $\bar{x}$ is defined by*

$$N_C^{\mathrm{P}}(\bar{x}) = \operatorname{cone}\left(P_C^{-1}\bar{x} - \bar{x}\right).$$

Equivalently, $\bar{x}^* \in N_C^{\mathrm{P}}(\bar{x})$ whenever there exists $\sigma \geq 0$ such that

$$\langle \bar{x}^*, x-\bar{x}\rangle \leq \sigma\|x-\bar{x}\|^2 \quad (\forall x \in C).$$

(ii) *The* limiting (proximal) normal cone *of C at $\bar{x}$ is defined by*

$$N_C(\bar{x}) = \operatorname*{Lim\,sup}_{x\to\bar{x}} N_C^{\mathrm{P}}(x),$$

where the limit superior is taken in the sense of Painlevé–Kuratowski outer limit.

When $\bar{x} \notin C$, all normal cones at $\bar{x}$ are empty (by definition).

2. Superregularity at a distance. We limit our attention in this study to *superregular* sets and their extension to sets with the corresponding properties relative to points *not* belonging to the sets.

Definition 2.1 (superregularity [16, Definition 4.3]). *Let $\Omega \subseteq \mathbb{R}^n$ and $\overline{x} \in \Omega$. The set Ω is said to be* superregular at $\overline{x}$ *if it is locally closed at $\overline{x}$ and for every $\epsilon > 0$ there is a $\delta > 0$ such that for all $(x,0) \in \operatorname{gph} N_\Omega \cap \{(\mathbb{B}_\delta(\overline{x}),0)\}$,*

$$\langle y'-y, x-y\rangle \leq \varepsilon\,||y'-y||\|x-y\|, \quad (\forall y' \in \mathbb{B}_\delta(\overline{x}))\,(\forall y \in P_\Omega(y')). \tag{2.1}$$

Rewriting the above leads to the following equivalent characterization of superregularity, which can be more convenient.

Proposition 2.2 ([16, Proposition 4.4]). *The set $\Omega \subseteq \mathcal{E}$ is* superregular *at $\overline{x} \in \Omega$ if and only if it is locally closed at $\overline{x}$ and for every $\varepsilon > 0$ there exists $\delta > 0$ such that*

$$\begin{aligned}&\langle v, x-y\rangle \leq \varepsilon\,||v||\|x-y\|,\\ &\qquad (\forall (x,v) \in \operatorname{gph} N_\Omega \cap (\mathbb{B}_\delta(\overline{x}) \times \mathcal{E}))\,(\forall y \in \Omega \cap \mathbb{B}_\delta(\overline{x})).\end{aligned} \tag{2.2}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

To extend superregularity to superregularity at a distance, we employ the more general framework of *ϵ-subregular sets* first introduced in [15]. The following terminology follows [8].

DEFINITION 2.3 (ϵ-subregularity [8, Definition 2.2]). *A set $\Omega \subset \mathcal{E}$ is* ϵ-subregular relative to $\Lambda \subset \mathcal{E}$ at $\overline{x}$ for $(x, v) \in \operatorname{gph} N_\Omega$ *if it is locally closed near $\overline{x}$ and there exists an $\epsilon > 0$ together with a neighborhood U_ϵ of $\overline{x}$ such that*

$$\langle v - (y' - y), y - x\rangle \leq \epsilon \, \|v - (y' - y)\| \, \|y - x\| \quad (\forall y' \in \Lambda \cap U_\epsilon)(\forall y \in P_\Omega(y')). \tag{2.3}$$

Ω *is* subregular relative to Λ at $\overline{x}$ for $(x, v) \in \operatorname{gph} N_\Omega$ *if it is locally closed near $\overline{x}$ and for each $\epsilon > 0$ there exists U_ϵ such that* (2.3) *holds.*

Note that the definition above does not require that the reference point $\overline{x}$ be an element of the set Ω. This allows one to characterize the regularity of sets from the perspective of any point, not just those in the set.

DEFINITION 2.4 (superregularity at a distance). *A set $\Omega \subset \mathcal{E}$ is called* ϵ-superregular at a distance relative to $\Lambda \subset \mathcal{E}$ at $\overline{x}$ *if it is ϵ-subregular relative to Λ at $\overline{x}$ for all $(x, v) \in V_\epsilon$, where*

$$V_\epsilon := \left\{(x, v) \in \operatorname{gph} N_\Omega^{\mathrm{P}} \mid x + v \in U_\epsilon, \; x \in P_\Omega(x + v)\right\} \tag{2.4}$$

and U_ϵ is a neighborhood of $\overline{x}$ dependent on a constant $\epsilon > 0$. The set Ω is called superregular at a distance relative to Λ at $\overline{x}$ *if it is ϵ-superregular at a distance relative to Λ at $\overline{x}$ for all $\epsilon > 0$.*

Note that implicitly $U_\epsilon \cap \Lambda \neq \emptyset$ for all $\epsilon > 0$.

Remark 2.5 (superregularity at a distance relative to $\mathcal{E}$ implies superregularity). Being superregular at a distance relative to $\Lambda = \mathcal{E}$ at some point $\overline{x} \in \Omega$ implies that the set is superregular at $\overline{x}$. To see this, let Ω be superregular at a distance relative to $\Lambda = \mathcal{E}$ at $\overline{x} \in \Omega$. For fixed $\epsilon > 0$ note that $(x, 0) \in V_\epsilon$ for all $x \in \Omega \cap U_\epsilon$. With these, (2.3) becomes

$$\langle y - y', y - x\rangle \leq \epsilon \, \|y - y'\| \, \|y - x\| \tag{2.5}$$

for all $y' \in \Lambda \cap U_\epsilon$, $y \in P_\Omega(y')$ and for all $x \in U_\epsilon \cap \Omega$. For sure, there exists an $\delta > 0$ such that $\mathbb{B}_\delta \subset U_\epsilon$. Moreover, since $\Lambda = \mathcal{E}$, (2.5) holds for all $y' \in U_\epsilon$, $y \in P_\Omega(y')$ and for all $x \in U_\epsilon \cap \Omega$, which is by Definition 2.1 superregularity of Ω at $\overline{x}$.

PROPOSITION 2.6 (convex sets are superregular at a distance). *Let $\Omega \subset \mathcal{E}$ be convex and closed. Then Ω is superregular at a distance relative to $\Lambda = \mathcal{E}$ at any $\overline{x} \in \Omega$.*

Proof. Fix $\overline{x} \in \Omega$. For convex sets Ω one has

$$\langle v, x - y\rangle \leq 0 \quad (\forall x, y \in \Omega)\,(\forall v \in N_\Omega(y))\,.$$

Thus, for any neighborhood $U \subset \mathcal{E}$ of $\overline{x}$ with $y' \in U$, $y \in P_\Omega(y')$ and hence $y' - y \in N_\Omega(y)$, $\langle y' - y, x - y\rangle \leq 0$, and thus

$$\langle v - (y' - y), y - x\rangle \leq 0 \quad (\forall y' \in \Lambda \cap U)(\forall y \in P_\Omega(y')).$$

This shows superregularity at a distance of Ω relative to $\mathcal{E}$ at all $\overline{x} \in \Omega$ as claimed. ∎

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

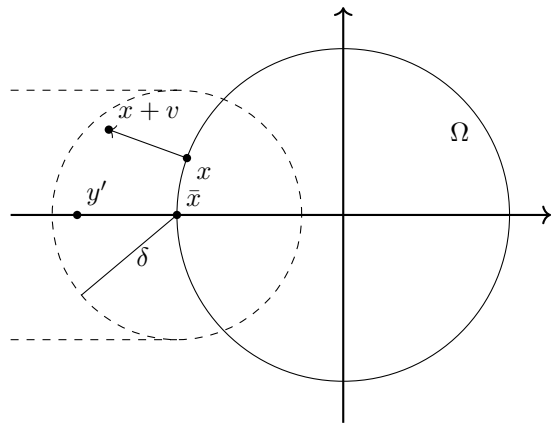

FIG. 1. *Example* 2.7.

Example 2.7 (circle). Consider the set

$$\Omega := \left\{ (x_1, x_2) \in \mathbb{R}^2 \;\middle|\; x_1^2 + x_2^2 = 1 \right\}.$$

This set, depicted in Figure 1, is ϵ-subregular relative to $\Lambda = P_\Omega^{-1}(\overline{x})$ at any $\overline{x} \in \Omega$ for all $(\overline{x}, v) \in \operatorname{gph} N_\Omega$ with $\epsilon = 0$ (which implies that Ω is in fact subregular relative to Λ for all $(\overline{x}, v) \in \operatorname{gph} N_\Omega$). Indeed, for any $\delta \in (0,1)$ we have, for any $y' \in \Lambda \cap \mathbb{B}_\delta(\overline{x})$, that $y \in P_\Omega(y')$ is given by $y = \overline{x}$, and (2.3) specializes to

$$\langle v - (y' - y), y - \overline{x} \rangle = \langle v - (y' - y), \overline{x} - \overline{x} \rangle = 0 \quad (\forall y' \in \Lambda \cap \mathbb{B}_\delta(\overline{x})(\forall v \in N_\Omega(\overline{x})).$$

Moreover, the set Ω is superregular at a distance relative to $\Lambda = P_\Omega^{-1}(\overline{x})$ at any $\overline{x} \in \Omega$. To see this, we will first show that Ω is ϵ-superregular at a distance relative to $P_\Omega^{-1}(\overline{x})$ at $\overline{x}$ for any $\epsilon \in (0, 0.5)$. Fix a $\epsilon \in (0, 0.5)$, and set $\delta = 2\epsilon$. For any $w \in N_\Omega(\overline{x})$ and $x \in \Omega \cap \mathbb{B}_\delta(\overline{x})$ it holds that $\cos\angle(w, x - \overline{x}) \le \cos\angle(-\overline{x}, x - \overline{x})$. By the law of cosine we conclude that $\cos\angle(-\overline{x}, \overline{x} - x) = \|\overline{x} - x\|/2 < \delta/2 = \epsilon$. Since $v - (y' - \overline{x}) \in N_\Omega(\overline{x})$ for all $y' \in \Lambda \cap \mathbb{B}_\delta(\overline{x})$, by the definition of the inner product on $\mathbb{R}^2$ we deduce that

$$\begin{aligned} &\langle v - (y' - \overline{x}), \overline{x} - x \rangle \\ &\qquad = \cos\angle(v - (y' - \overline{x}), \overline{x} - x) \, \|v - (y' - \overline{x})\| \, \|\overline{x} - x\| \\ &\qquad \le \epsilon \, \|v - (y' - \overline{x})\| \, \|\overline{x} - x\| \qquad (\forall y' \in \Lambda \cap \mathbb{B}_\delta(\overline{x}))(\forall (x, v) \in V_\delta), \end{aligned}$$

where

$$V_\delta := \left\{ (x, v) \in \operatorname{gph} N_\Omega^{\mathrm{P}} \mid x + v \in \mathbb{B}_\delta(\overline{x}),\ x \in P_\Omega(x + v) \right\},$$

which shows that Ω is ϵ-superregular at a distance relative to $P_\Omega^{-1}(\overline{x})$ at $\overline{x}$ for any $\epsilon \in (0, 0.5)$. Likewise, the same is true for any $\epsilon > 0.5$ when taking a ball with radius δ around $\overline{x}$, where $\delta < 1$. Thus, Ω is superregular relative to $P_\Omega^{-1}(\overline{x})$ at $\overline{x}$.

In fact, we can even enlarge our neighborhood from a ball to a tube in radial direction. Fix $\overline{x} \in \Omega$, $\epsilon > 0$ and some $\delta \in (0,1)$ such that the above construction is satisfied. Then

$$U := \bigcup_{\substack{z \in P_\Omega^{-1}(\overline{x}) \\ \|z\| \ge 1}} \mathbb{B}_\delta(z)$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

is a neighborhood for $\overline{x}$ such that ϵ-superregularity relative to $\Lambda = P_\Omega^{-1}(\overline{x})$ is satisfied for Ω. Fortunately, our violation ϵ will not be worse compared to the neighborhood being a ball with radius δ around $\overline{x}$. This allows us to include more points in $\Lambda \cap U$ without violating (2.3).

PROPOSITION 2.8 (characterization of superregularity at a distance).

(i) *A nonempty set $\Omega \subset \mathcal{E}$ is ϵ-superregular at a distance relative to $\Lambda \subset \mathcal{E}$ at $\overline{x}$ if and only if there is a neighborhood U_ϵ of $\overline{x}$ such that*

$$\|x-y\|^2 \le \epsilon \|(y'-y)-(x'-x)\| \|x-y\| + \langle x'-y', x-y\rangle \qquad (\forall y' \in U_\epsilon \cap \Lambda)(\forall y \in P_\Omega(y'))$$

holds with $x' = x + v \in U_\epsilon$ for all $(x,v) \in V_\epsilon$ for V_ϵ defined by (2.4).

(ii) *Let $\Omega \subset \mathcal{E}$ be ϵ-superregular at a distance relative to Λ at $\overline{x}$. Then*

$$\|x-y\| \le \epsilon \|(y'-y)-(x'-x)\| + \|x'-y'\| \qquad (\forall y' \in U_\epsilon \cap \Lambda)(\forall y \in P_\Omega(y'))$$

holds with $x' = x + v \in U_\epsilon$ for all $(x,v) \in V_\epsilon$.

Proof. (i) Let $\Omega \subset \mathcal{E}$ be ϵ-superregular at a distance relative to $\Lambda \subset \mathcal{E}$ at $\overline{x}$. Then, for fixed $\epsilon > 0$, there exists a neighborhood U_ϵ of $\overline{x}$ such that for all $(x,v) \in V_\epsilon$ for V_ϵ defined by (2.4) and $x' = x + v \in U_\epsilon$ the following hold:

$$\begin{aligned}\|x-y\|^2 = \langle x-y, x-y\rangle &= \langle y'-y-(x'-x), x-y\rangle + \langle x'-y', x-y\rangle \\ &\le \epsilon \|y'-y-(x'-x)\| \|x-y\| + \langle x'-y', x-y\rangle .\end{aligned}$$

This proves the first part of the equivalence. For the remaining one let U_ϵ be a neighborhood of $\overline{x}$ such that

$$\|x-y\|^2 \le \epsilon \|(y'-y)-(x'-x)\| \|x-y\| + \langle x'-y', x-y\rangle \qquad (\forall y' \in U_\epsilon \cap \Lambda)(\forall y \in P_\Omega(y')) \tag{2.6}$$

holds with $x' = x + v \in U_\epsilon$ for all $(x,v) \in V_\epsilon$ for V_ϵ defined by (2.4). A similar argument shows that (2.6) is equivalent to

$$\begin{aligned}&\langle y'-y-(x'-x), x-y\rangle + \langle x'-y', x-y\rangle \\ &\qquad \le \epsilon \|(y'-y)-(x'-x)\| \|x-y\| + \langle x'-y', x-y\rangle , \\ &\qquad\qquad (\forall y' \in U_\epsilon \cap \Lambda)(\forall y \in P_\Omega(y')).\end{aligned}$$

Subtracting $\langle x'-y', x-y\rangle$ from both sides and inserting $v = x' - x$ yields

$$\langle y'-y-v, x-y\rangle \le \epsilon \|(y'-y)-v\| \|x-y\| , \quad (\forall y' \in U_\epsilon \cap \Lambda)(\forall y \in P_\Omega(y')).$$

Reordering the left-hand side we deduce the definition of ϵ-superregular for Ω at $\overline{x}$:

$$\langle v-(y'-y), y-x\rangle \le \epsilon \|(y'-y)-v\| \|x-y\| , \quad (\forall y' \in U_\epsilon \cap \Lambda)(\forall y \in P_\Omega(y')).$$

(ii) The second part follows from (i) by applying the Cauchy–Schwarz inequality to the vectors $x' - y'$ and $x - y$. ∎

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

3. Properties of $T_{\mathrm{DR}\lambda}$ and characterization of its fixed points. Our convergence analysis is based on the framework established in [23] and relies on two essential properties of fixed point mappings. The first property describes the *expansiveness* of the mapping, or the violation of nonexpansiveness. This is called *almost averaging* in Definition 3.1. This property also implies *single-valuedness* of the fixed point mapping at its (see points Theorem 3.7). The characterization of the fixed points is established in Theorem 3.13. In section 4 we discuss the second property, *metric subregularity* (see Definition 4.1), which describes the (one-sided) Lipschitz continuity of the inverse of the fixed point mapping at its fixed points. When specialized to set feasibility, this takes on the more geometric property of *subtransversality* of the collection of sets (see Definition 4.3).

3.1. Almost averaged mappings.

DEFINITION 3.1 (almost nonexpansive/averaged mappings [23, Definition 2.2]). *Let D be a nonempty subset of $\mathcal{E}$, and let T be a (set-valued) mapping from D to $\mathcal{E}$.*

(i) *T is said to be* pointwise almost nonexpansive *on D at $y \in D$ if there exists a constant $\epsilon \in [0,1)$ such that*

$$\left\|x^+ - y^+\right\| \leq \sqrt{1+\epsilon}\,\|x-y\| \quad \forall\ y^+ \in Ty \text{ and } \forall x^+ \in Tx \text{ whenever } x \in D. \tag{3.1}$$

If (3.1) holds with $\epsilon = 0$, then T is called pointwise nonexpansive *at y on D. If T is pointwise (almost) nonexpansive on D at every point $y \in D$ with the violation constant ϵ, then T is said to be* (almost) nonexpansive on D (with violation ϵ).

(ii) *T is called* pointwise almost averaged on D at y *if there is an averaging constant $\alpha \in (0,1)$ and a violation constant $\epsilon \in [0,1)$ such that the mapping $\tilde{T}$ defined by*

$$\tilde{T} = \tfrac{1}{\alpha}T + \tfrac{(\alpha-1)}{\alpha}\,\mathrm{Id}$$

is almost nonexpansive at y with violation ϵ/α.

PROPOSITION 3.2 (characterization of almost averaged operators). *Let $T : \mathcal{E} \rightrightarrows \mathcal{E}$, $U \subset \mathcal{E}$, and let $\alpha \in (0,1)$. The following are equivalent.*

(i) *T is pointwise almost averaged at y on U with violation ϵ and averaging constant α.*

(ii) *$\left(1 - \frac{1}{\alpha}\right)\mathrm{Id} + \frac{1}{\alpha}T$ is pointwise almost nonexpansive at y on U with violation ϵ/α.*

(iii) *For all $x \in U$, $x^+ \in T(x)$, and $y^+ \in T(y)$ it holds that*

$$\left\|x^+ - y^+\right\|^2 \leq (1+\epsilon)\,\|x-y\|^2 - \frac{1-\alpha}{\alpha}\left\|\left(x - x^+\right) - \left(y - y^+\right)\right\|. \tag{3.2}$$

Therefore, if T is pointwise almost averaged at y on U with violation ϵ and averaging constant α, then T is pointwise almost nonexpansive at y on U with violation at most ϵ.

Proof. The proof of this statement can be found in [23, Proposition 2.1]. ∎

In terms of the above Definition 3.1, *pointwise firmly nonexpansive* mappings are pointwise averaged mappings with averaging constant $\alpha = 1/2$.

PROPOSITION 3.3 (compositions of averages of averaged operators). *Let $T_j : \mathcal{E} \rightrightarrows \mathcal{E}$ for $j = 1, 2, \ldots, m$ be pointwise almost averaged on U_j at all $y_j \in S_j \subset \mathcal{E}$ with violation ϵ_j and averaging constant $\alpha_j \in (0,1)$, where $U_j \supset S_j$ for $j = 1, 2, \ldots, m$.*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

(i) *If* $U := U_1 = U_2 = \cdots = U_m$ *and* $S := S_1 = S_2 = \cdots = S_m$, *then the weighted mapping* $T := \sum_{j=1}^m w_j T_j$ *with weights* $w_j \in [0,1]$, $\sum_{j=1}^m w_j = 1$, *is pointwise almost averaged at all* $y \in S$ *with violation* $\epsilon = \sum_{j=1}^m w_j \epsilon_j$ *and averaging constant* $\alpha = \max_{j=1,2,\ldots,m}\{\alpha_j\}$ *on* U.

(ii) *If* $T_j U_j \subseteq U_{j-1}$ *and* $T_j S_j \subseteq S_{j-1}$ *for* $j = 2,3,\ldots,m$, *then the composite mapping* $T := T_1 \circ T_2 \circ \cdots \circ T_m$ *is pointwise almost nonexpansive at all* $y \in S_m$ *on* U_m *with violation at most*

$$\epsilon = \prod_{j=1}^m (1+\epsilon_j) - 1. \tag{3.3}$$

(iii) *If* $T_j U_j \subseteq U_{j-1}$ *and* $T_j S_j \subseteq S_{j-1}$ *for* $j = 2,3,\ldots,m$, *then the composite mapping* $T := T_1 \circ T_2 \circ \cdots \circ T_m$ *is pointwise almost averaged at all* $y \in S_m$ *on* U_m *with violation at most* ϵ *given by* (3.3) *and averaging constant at least*

$$\alpha = \frac{m}{m - 1 + \frac{1}{\max_{j=1,2,\ldots,m}\{\alpha_j\}}}. \tag{3.4}$$

Proof. The proof of this statement can be found in [23, Proposition 2.4]. ∎

Definition 2.4 allows us to get pointwise almost nonexpansivity of the projector and reflector on a neighborhood of a point in Ω relative to points not in Ω. This is of particular interest for us since the fixed points of $T_{\mathrm{DR}\lambda}$ will be (depending on $\lambda < 1$) in neither of the sets A and B if the problem is inconsistent (see Theorem 3.13, where we do not require that $A \cap B \neq \emptyset$).

PROPOSITION 3.4 (regularity of projectors and reflectors at a distance). *Let* $\Omega \subset \mathcal{E}$ *be nonempty and closed, and let* U *be a neighborhood of* $\overline{x} \in \Omega$. *Let* $\Lambda := P_\Omega^{-1}(\overline{x}) \cap U$. *If* Ω *is* ϵ*-superregular at a distance at* $\overline{x}$ *relative to* Λ *with constant* ϵ *on the neighborhood* U, *then the following hold.*

(i) *If* $\epsilon \in [0,1)$, *then the projector* P_Ω *is pointwise almost nonexpansive at each* $y' \in \Lambda$ *with violation* $\widetilde{\epsilon}$ *on* U *for* $\widetilde{\epsilon} := 4\epsilon/(1-\epsilon)^2$. *That is, at each* $y' \in \Lambda$,

$$\begin{aligned}\|x-y\| &\le \sqrt{1+\widetilde{\epsilon}}\,\|x'-y'\| \\ &= \frac{1+\epsilon}{1-\epsilon}\,\|x'-y'\| \quad (\forall x' \in U)\,(\forall x \in P_\Omega(x'))\,(\forall y \in P_\Omega(y'))\,.\end{aligned}$$

(ii) *If* $\epsilon \in [0,1)$, *then the projector* P_Ω *is pointwise almost firmly nonexpansive at each* $y' \in \Lambda$ *with violation* $\widetilde{\epsilon}_2$ *on* U *for* $\widetilde{\epsilon}_2 := 4\epsilon(1+\epsilon)/(1-\epsilon)^2$. *That is, at each* $y' \in \Lambda$,

$$\|x-y\|^2 + \|(x'-x)-(y'-y)\|^2 \le (1+\widetilde{\epsilon}_2)\,\|x'-y'\| \quad (\forall x' \in U)\,(\forall x \in P_\Omega(x'))\,(\forall y \in P_\Omega(y'))\,.$$

(iii) *The reflector* R_Ω *is pointwise almost nonexpansive at each* $y' \in \Lambda$ *with violation* $\widetilde{\epsilon}_3 := 8\epsilon(1+\epsilon)/(1-\epsilon)^2$ *on* U. *That is, for all* $y' \in \Lambda$,

$$\|x-y\| \le \sqrt{1+\widetilde{\epsilon}_3}\,\|x'-y'\| \quad (\forall x' \in U)\,(\forall x \in R_\Omega(x'))\,(\forall y \in R_\Omega(y'))\,.$$

Proof. Our proof follows that of [23, Theorem 3.1]. Before proving each of the statements individually, note the following. Take any $x' \in U$. Then for each $x \in P_\Omega(x')$ we have $(x, x'-x) \in \operatorname{gph} N_\Omega^{\mathrm{P}} \subset \operatorname{gph} N_\Omega$. Moreover, by construction $(x, x'-x) \in V_\epsilon$, where V_ϵ is defined by (2.4).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

(i) Choosing $x' \in U$ and $x \in P_\Omega(x')$ we get $(x, x'-x) \in \operatorname{gph} N_\Omega^{\mathrm{P}} \subset \operatorname{gph} N_\Omega$. Applying Proposition 2.8(ii) yields

$$\|y-x\| \le \epsilon \|(x'-x)-(y'-y)\| + \|y'-x'\|$$

whenever $y' \in U \cap \Lambda$ and $y \in P_\Omega(y')$. Exploiting the triangle inequality we deduce that

$$\|y-x\| \le \epsilon \left(\|y'-x'\| + \|y-x\|\right) + \|y'-x'\|$$

and thus conclude the claimed result.

(ii) By superregularity at a distance relative to Λ of Ω and Proposition 2.8(i) we have

$$\begin{aligned}
&\|x-y\|^2 + \|(x'-x)-(y'-y)\|^2 \\
&= 2\|x-y\|^2 + \|x'-y'\|^2 - 2\langle x'-y', x-y\rangle \\
&\le \|x'-y'\|^2 + 2\epsilon \|(x'-x)-(y'-y)\| \|x-y\|
\end{aligned}$$

for $(x, x'-x) \in V_\epsilon$ and $y' \in U \cap \Lambda$, $y \in P_\Omega(y')$. Together with the triangle inequality this implies that

$$\begin{aligned}
&\|x-y\|^2 + \|(x'-x)-(y'-y)\|^2 \\
&\le \|x'-y'\|^2 + 2\epsilon \left(\|x'-y'\| + \|x-y\|\right) \|x-y\| .
\end{aligned}$$

Using part (i) we deduce that

$$\begin{aligned}
&\|x-y\|^2 + \|(x'-x)-(y'-y)\|^2 \\
&\le \left(1 + 4\epsilon \frac{1+\epsilon}{(1-\epsilon)^2}\right) \|x'-y'\|^2
\end{aligned} \tag{3.5}$$

for all $(x, x'-x) \in V_\epsilon$ and for all $y \in P_\Omega(y')$ at each $y' \in U \cap \Lambda$. Since, as mentioned in the beginning, for all $x' \in U$ it holds that $(x, x'-x) \in V_\epsilon$ for all $x \in P_\Omega(x')$, (3.5) holds at each $y' \in \Lambda = \Lambda \cap U$ for all $x \in P_\Omega(x')$ whenever $x' \in U$. By Proposition 3.2 with $\alpha = 1/2$ we conclude that P_Ω is pointwise almost firmly nonexpansive at each $y' \in \Lambda$ with violation $4\epsilon(1+\epsilon)/(1-\epsilon)^2$ on U.

(iii) By (ii) P_Ω is pointwise almost firmly nonexpansive at each $y' \in \Lambda$ with violation

$$4\epsilon(1+\epsilon)/(1-\epsilon)^2$$

on U. Thus, by Proposition 3.2 the reflector, $R_\Omega := 2P_\Omega - \mathrm{Id}$, is pointwise almost nonexpansive at each $y' \in \Lambda$ with violation $8\epsilon(1+\epsilon)/(1-\epsilon)^2$ on U. ∎

3.2. $T_{\mathrm{DR}\lambda}$ is almost averaged at $\mathsf{Fix}\, T_{\mathrm{DR}\lambda}$. For general multivalued mappings $T : \mathcal{E} \rightrightarrows \mathcal{E}$ the set of fixed points is defined as

$$\mathsf{Fix}\, T := \{x \in \mathcal{E} \mid x \in T(x)\} . \tag{3.6}$$

Note that, by this definition, the set $T(x)$ need not consist entirely of fixed points (see [23, Example 2.1]). If T is pointwise almost averaged, however, the mapping T is single-valued on its fixed point set.

PROPOSITION 3.5 ([23, Proposition 2.2]). *If $T : \mathcal{E} \rightrightarrows \mathcal{E}$ is pointwise almost nonexpansive on $D \subseteq \mathcal{E}$ at $\overline{x} \in D$ with violation $\varepsilon \ge 0$, then T is single-valued at $\overline{x}$. In particular, if $\overline{x} \in \mathsf{Fix}\, T$, then $T(\overline{x}) = \{\overline{x}\}$.*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

The mapping $T_{\mathrm{DR}\lambda}$ is a composition and convex combination of projectors and reflectors. The almost averaging property is preserved under compositions and convex combinations of pointwise almost averaged mappings, as we have seen in Proposition 3.3.

LEMMA 3.6. *Let $\overline{x} \in \mathcal{E}$, and let $\Omega \subset \mathcal{E}$ be superregular at a distance relative to $\Lambda \subset P_\Omega^{-1}(\bar{\omega})$ at $\bar{\omega}$, where $\overline{\omega} \in P_\Omega(\overline{x})$ and $\overline{x} \in \Lambda$. In addition, for each $\epsilon > 0$, let $\overline{x} \in U_\epsilon(\overline{\omega})$, where $U_\epsilon(\overline{\omega})$ is a neighborhood of $\overline{\omega}$ on which* (2.3) *holds. Then $P_\Omega(\overline{x}) = \{\overline{\omega}\}$.*

Proof. For some fixed $\epsilon > 0$, we get by the assumptions on superregularity at a distance of Ω relative to Λ and Proposition 3.4(i) that there exists some neighborhood $U_\epsilon(\bar{\omega})$ such that P_Ω is pointwise almost nonexpansive at $\overline{x} \in \Lambda \cap U_\epsilon(\bar{\omega})$ on $U_\epsilon(\bar{\omega})$ with violation $\widetilde{\epsilon} = 4\epsilon/(1-\epsilon)^2$. This implies single-valuedness of P_Ω at $\overline{x}$ by Proposition 3.5, i.e., that $\{\bar{\omega}\} = P_\Omega(\overline{x})$, as claimed. ∎

THEOREM 3.7 ($T_{\mathrm{DR}\lambda}$ is pointwise almost firmly nonexpansive at its fixed points). *Let A, B be closed and nonempty, $\lambda \in (0,1)$ and $\overline{x} \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda} \neq \emptyset$. Let $\overline{b} \in P_B(\overline{x})$ and $\overline{a} \in P_A(2\overline{b} - \overline{x})$. Suppose that B is superregular at a distance relative to $\Lambda_{\bar{b}} := P_B^{-1}(\bar{b})$ at $\bar{b}$ and, likewise, that A is superregular at a distance relative to $\Lambda_{\overline{a}} := P_A^{-1}(\bar{a})$ at $\bar{a}$. Suppose, moreover, that the following hold:*

(i) *For each $\epsilon > 0$, $\overline{x} \in U_\epsilon(\overline{b})$, where $U_\epsilon(\overline{b})$ is a neighborhood of $\overline{b}$ on which* (2.3) *holds for ϵ.*
(ii) *For each $\epsilon > 0$, $2\overline{b} - \overline{x} \in U_\epsilon(\overline{a})$, where $U_\epsilon(\overline{a})$ is a neighborhood of $\overline{a}$ on which* (2.3) *holds for ϵ.*
(iii) *$R_B(\Lambda_{\bar{b}}) \subset \Lambda_{\bar{a}}$.*
(iv) *$R_B(U_\epsilon(\overline{b})) \subset U_\epsilon(\overline{a})$ for all $\epsilon > 0$.*

Then, $\{\bar{b}\} = P_B(\overline{x})$, $\{\bar{a}\} = P_A(R_B(\overline{x}))$, $T_{\mathrm{DR}\lambda}$ is single-valued at $\overline{x}$, and for all $\epsilon > 0$ there exists a neighborhood $U(B, \epsilon, \overline{x})$ of $\bar{b}$ such that $T_{\mathrm{DR}\lambda}$ is pointwise almost firmly nonexpansive at $\overline{x}$ with violation at most ϵ on $U(B, \epsilon, \overline{x})$.

Before we begin the proof of this statement, we would like to point out an important feature of our construction. The claimed pointwise almost averagedness of $T_{\mathrm{DR}\lambda}$ at $\overline{x}$ holds on open subsets containing both $\overline{x}$ and $\overline{b} = P_B(\overline{x})$. This follows from assumption (i). The conclusion of the theorem could have been equivalently stated: *For all $\epsilon > 0$ there exists a neighborhood $U(B, \epsilon, \overline{x})$ of $\overline{x}$ such that $T_{\mathrm{DR}\lambda}$ is pointwise almost averaged at $\overline{x}$ with violation at most ϵ on U.* We have presented the statement with neighborhood $U(B, \epsilon, \overline{x})$ containing $\overline{b}$ to emphasize the fact that the open sets on which the regularity of $T_{\mathrm{DR}\lambda}$ holds is constructed relative to points $\overline{b}$ *at a distance* from the point of interest $\overline{x} \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda}$. The usual use of balls for neighborhoods is not the most convenient or appropriate for this setting.

Proof of Theorem 3.7. Under assumptions (i) and (ii), Lemma 3.6 yields $\{\bar{b}\} = P_B(\overline{x})$ and $\{\bar{a}\} = P_A(R_B(\overline{x}))$, as claimed. From this one can immediately conclude that $T_{\mathrm{DR}\lambda}$ is single-valued at $\overline{x}$.

For any fixed $\epsilon_B > 0$, we get by the assumptions on superregularity at a distance of B relative to Λ_b and Proposition 3.4(ii) that there exists some neighborhood $U_{\epsilon_B}(\overline{b})$ such that P_B is pointwise almost firmly nonexpansive at $\overline{x} \in \Lambda_{\bar{b}} \cap U(B, \epsilon_B, \overline{x})$ on $U_{\epsilon_B}(\overline{b})$ with violation $\epsilon_{P_B} = 4\epsilon_B(1+\epsilon_B)/(1-\epsilon_B)^2$. Note that this also shows that P_B is pointwise almost nonexpansive at $\overline{x}$ on $U_{\epsilon_B}(\overline{b})$. Similarly, by Proposition 3.4(iii), R_B is pointwise almost nonexpansive at $\overline{x}$ with violation $\epsilon_{R_B} = 8\epsilon_B(1+\epsilon_B)/(1-\epsilon_B)^2$ on $U_{\epsilon_B}(\overline{b})$. Likewise, for any $\epsilon_A > 0$ there exists a neighborhood $U_{\epsilon_A}(\overline{a})$ of $\overline{a}$ such that R_A is pointwise almost nonexpansive at $\overline{a} = 2\overline{b} - \overline{x}$ with violation $\epsilon_{R_A} = 8\epsilon_A(1+\epsilon_A)/(1-\epsilon_A)^2$ on $U_{\epsilon_A}(\overline{a})$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

By (iii) and (iv), the assumptions of Proposition 3.3(ii) are satisfied; hence, we deduce that for any fixed $\epsilon_A > 0$ there exists a neighborhood $U(A, \epsilon_{R_A R_B}, \overline{x})$ such that $R_A R_B$ is pointwise almost nonexpansive at $\overline{x}$ with violation at most $\epsilon_{R_A R_B} = \epsilon_{R_A} + \epsilon_{R_B} + \epsilon_{R_A}\epsilon_{R_B}$ on $U(B, \epsilon, \overline{x})$.

By Proposition 3.2(ii) we get that $1/2(R_A R_B + \mathrm{Id})$ is almost firmly nonexpansive at $\overline{x}$ with violation $\epsilon_{R_A R_B}/2$ on $U_{\epsilon_B}(\overline{b})$. Applying Proposition 3.3(i) yields pointwise almost firm nonexpansivity of $T_{\mathrm{DR}\lambda}$ at $\overline{x}$ on $U_{\epsilon_B}(\overline{b})$ with violation at most

$$\epsilon' = \lambda \epsilon_{R_A R_B}/2 + (1-\lambda)\epsilon_{P_B}.$$

Since the above properties hold for each $\epsilon_B > 0$ and $\epsilon_A > 0$, then given any $\epsilon > 0$ we can construct the neighborhoods above so that $\epsilon' \leq \epsilon$. We conclude that for any $\epsilon > 0$ there is a neighborhood $U(B, \epsilon, \overline{x})$ such that $T_{\mathrm{DR}\lambda}$ is pointwise almost nonexpansive at $\overline{x}$ on $U(B, \epsilon, \overline{x})$ with violation at most ϵ (the corresponding neighborhood $U(A, \epsilon_{R_A R_B}, \overline{x})$ of $\overline{a}$ will be denoted by $U(A, \epsilon, \overline{x})$). This conclusion is consistent with the fact established above that $T_{\mathrm{DR}\lambda}$ is single-valued, which completes the proof. ∎

COROLLARY 3.8. *In the setting of Theorem* 3.7*, fix* $\overline{\epsilon} > 0$*, and let* $U(B, \overline{\epsilon}, \overline{x})$ *and* $U(A, \overline{\epsilon}, \overline{x})$ *be neighborhoods that satisfy assumptions* (i)*,* (ii)*, and* (iv) *such that* $T_{\mathrm{DR}\lambda}$ *is pointwise almost firmly nonexpansive at* $\overline{x}$ *with violation* $\overline{\epsilon}$ *on* $U(B, \overline{\epsilon}, \overline{x})$*. Then, for all* $\epsilon < \overline{\epsilon}$ *there exists a neighborhood* $U(B, \epsilon, \overline{x})$ *and a neighborhood* $U(A, \epsilon, \overline{x})$ *such that conditions* (i)*,* (ii)*, and* (iv) *hold in addition to the inclusions* $U(A, \epsilon, \overline{x}) \subset U(A, \overline{\epsilon}, \overline{x})$ *and* $U(B, \epsilon, \overline{x}) \subset U(B, \overline{\epsilon}, \overline{x})$*.*

Corollary 3.8 implies that $T_{\mathrm{DR}\lambda}$ is pointwise almost firmly nonexpansive at $\overline{x}$ with violation ϵ on $U(B, \epsilon, \overline{x})$. The strength of Corollary 3.8, however, is hidden in the proof given below and the explicit construction of the neighborhoods $U(B, \epsilon, \overline{x})$ and $U(A, \epsilon, \overline{x})$. Thus, under the assumptions of Theorem 3.7 and given the neighborhoods for some fixed violation $\overline{\epsilon}$, we are always able to restrict these neighborhoods to smaller sets where (2.3) holds with some violation smaller than $\overline{\epsilon}$.

Proof of Corollary 3.8. Our approach to prove this statement is based on an explicit construction of the neighborhood $U(A, \epsilon, \overline{x})$ and $U(B, \epsilon, \overline{x})$.

Let $\epsilon < \overline{\epsilon}$. Theorem 3.7(i) implies that there exists a neighborhood $U(B, \epsilon, \overline{x})$ of $\overline{b}$ where (2.3) holds such that $U(B, \epsilon, \overline{x}) \subset U(B, \overline{\epsilon}, \overline{x})$. To see this, note that by (i) the existence of $U(B, \epsilon, \overline{x})$ is guaranteed, and thus only $U(B, \epsilon, \overline{x}) \subset U(B, \overline{\epsilon}, \overline{x})$ has to be proven. Let $\tilde{U}(B, \epsilon, \overline{x})$ be a neighborhood for ϵ where (2.3) holds. Then (2.3) is satisfied for $\overline{\epsilon}$ as well. Thus, $U(B, \epsilon, \overline{x}) := \tilde{U}(B, \epsilon, \overline{x}) \cap U(B, \overline{\epsilon}, \overline{x})$ is a neighborhood of $\overline{b}$ where Theorem 3.7(iv) holds and (2.3) is satisfied for both ϵ and $\overline{\epsilon}$, which shows that $U(B, \epsilon, \overline{x}) \subset U(B, \overline{\epsilon}, \overline{x})$ as required. Next, applying the reflection on B on both of neighborhoods $U(B, \epsilon, \overline{x})$ and $U(B, \overline{\epsilon}, \overline{x})$ yields

$$R_B\left(U(B, \epsilon, \overline{x})\right) \subset R_B\left(U(B, \overline{\epsilon}, \overline{x})\right). \tag{3.7}$$

Let $\tilde{U}(A, \epsilon, \overline{x})$ be a neighborhood of $\overline{a}$ where (2.3) holds for ϵ such that (iv) of Theorem 3.7 is satisfied. That is,

$$R_B(U(B, \epsilon, \overline{x})) \subset \tilde{U}(A, \epsilon, \overline{x}). \tag{3.8}$$

Combining Theorem 3.7(iv) for the neighborhoods $U(B, \overline{\epsilon}, \overline{x})$ and $U(A, \overline{\epsilon}, \overline{x})$ and (3.7) we deduce that

$$R_B\left(U(B, \epsilon, \overline{x})\right) \subset R_B\left(U(B, \overline{\epsilon}, \overline{x})\right) \subset U(A, \overline{\epsilon}, \overline{x}). \tag{3.9}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

This and (3.8) imply that

$$R_B(U(B,\epsilon,\overline{x})) \subset \tilde{U}(A,\epsilon,\overline{x}) \cap U(A,\overline{\epsilon},\overline{x}). \tag{3.10}$$

Set

$$U(A,\epsilon,\overline{x}) := \tilde{U}(A,\epsilon,\overline{x}) \cap U(A,\overline{\epsilon},\overline{x}).$$

Then, $U(A,\epsilon,\overline{x})$ is a neighborhood of $\overline{a}$ where (2.3) holds with ϵ since it is a subset of $\tilde{U}(A,\epsilon,\overline{x})$. Moreover, $U(B,\epsilon,\overline{x})$ and $U(A,\epsilon,\overline{x})$ satisfy (iv) of Theorem 3.7 by (3.10). By the construction of $U(A,\overline{\epsilon},\overline{x})$ and the choice of $U(B,\epsilon,\overline{x})$ both sets satisfy the inclusions $U(A,\epsilon,\overline{x}) \subset U(A,\overline{\epsilon},\overline{x})$ and $U(B,\epsilon,\overline{x}) \subset U(B,\overline{\epsilon},\overline{x})$. This completes the proof. ∎

Example 3.9. The following examples, illustrated in Figure 2, demonstrate the assumptions of the above theorem. For these examples it is easy to determine the sets of fixed points of the mapping $T_{\mathrm{DR}\lambda}$. In Theorem 3.13 we give a precise characterization of $\mathsf{Fix}\ T_{\mathrm{DR}\lambda}$. More intuitively, the fixed points must lie on lines containing *local best approximation points* between the sets.

(i) (convex sets with empty intersection). Let A and B be closed convex subsets of $\mathcal{E}$. By Proposition 2.6 both sets are superregular relative to $\mathcal{E}$ at any of their points, i.e., ϵ-superregular for all $\epsilon > 0$. In fact, the violation can be set to 0. Thus, as long as $\mathsf{Fix}\ T_{\mathrm{DR}\lambda} \neq \emptyset$, the mappings P_B, R_B, and R_A are nonexpansive (i.e., no violation) at $\overline{x}$ on the whole space $\mathcal{E}$ by Proposition 3.4, and we can apply Theorem 3.7 to conclude that $T_{\mathrm{DR}\lambda}$ is firmly nonexpansive at $\overline{x}$ on the neighborhood $\mathcal{E}$. For instance, consider the two sets

$$A := \left\{x = (x_1,x_2) \in \mathbb{R}^2 \;\middle|\; x_1^2 + x_2^2 \leq 1\right\} \text{ and}$$
$$B := \left\{x = (x_1,x_2) \in \mathbb{R}^2 \;\middle|\; (x_1-3)^2 + x_2^2 \leq 1\right\}.$$

The set of fixed points is given by the unique point

$$\mathsf{Fix}\ T_{\mathrm{DR}\lambda} = \{\overline{x}\} = \left\{(2,0) - \frac{\lambda}{1-\lambda}(1,0)\right\}$$

for fixed $\lambda \in (0,1)$, and by the above discussion, we know that $T_{\mathrm{DR}\lambda}$ built from the projections onto these sets is firmly nonexpansive.

(ii) (superregular sets with empty intersection). Continuing with the concrete example above, suppose that A and B are spheres instead of balls:

$$A := \left\{x = (x_1,x_2) \in \mathbb{R}^2 \;\middle|\; x_1^2 + x_2^2 = 1\right\} \text{ and}$$
$$B := \left\{x = (x_1,x_2) \in \mathbb{R}^2 \;\middle|\; (x_1-3)^2 + x_2^2 = 1\right\}.$$

The sets A and B are both nonconvex but still superregular. The set of fixed points is again given by the unique point

$$\mathsf{Fix}\ T_{\mathrm{DR}\lambda} = \{\overline{x}\} = \left\{(2,0) - \frac{\lambda}{1-\lambda}(1,0)\right\}$$

for fixed $\lambda \in (0,1)$. As seen in Example 2.7 both sets are superregular relative to radial directions. Thus, applying Theorem 3.7 we deduce that $T_{\mathrm{DR}\lambda}$ for some fixed $\lambda \in (0,1)$ is only almost firmly nonexpansive at $\overline{x}$ on some neighborhood U. As noted before in Example 2.7 the neighborhood should be rather chosen as a tube than the more conventional ball. Such a choice of neighborhoods is visualized in Figure 3.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

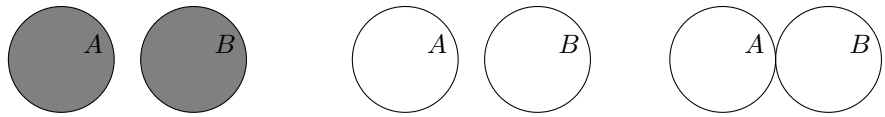

FIG. 2. *Example* 3.9(i)–(iii).

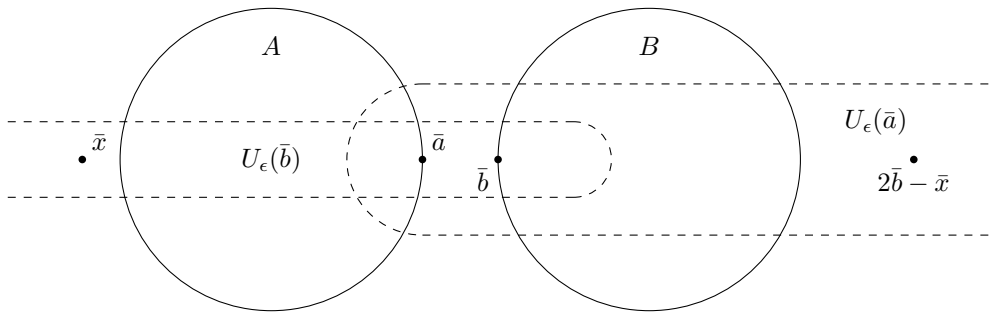

FIG. 3. *Example* 3.9(ii).

(iii) (superregular sets with nonempty intersection). Next we translate the sets in (ii) so that they have exactly one point in common:

$$A := \left\{x = (x_1, x_2) \in \mathbb{R}^2 \;\middle|\; x_1^2 + x_2^2 = 1\right\} \text{ and}$$
$$B := \left\{x = (x_1, x_2) \in \mathbb{R}^2 \;\middle|\; (x_1 - 2)^2 + x_2^2 = 1\right\}.$$

The fixed point set then reduces to $\mathsf{Fix}\, T_{\mathrm{DR}\lambda} = \{(1,0)\} = A \cap B$. By (ii) we know that the assumptions of Theorem 3.7 are satisfied. In contrast to (ii) the fixed point is in the intersection of both sets. Thus, balls as neighborhoods are enough to get pointwise almost firm nonexpansivity. We do not need tubes to include points from a distance.

The examples show that in case of closed balls and circles the assumptions are easily satisfied. Nonetheless, one has to take care of choosing neighborhoods in the right way to get a desired violation.

Example 3.9(i) yields the following specialization of Theorem 3.7.

COROLLARY 3.10. *Let $\lambda \in (0,1)$ and $\mathsf{Fix}\, T_{\mathrm{DR}\lambda} \neq \emptyset$. If A and B are closed and convex, $T_{\mathrm{DR}\lambda}$ is firmly nonexpansive on $\mathcal{E}$.*

Proof. Since A and B are both convex one has by Proposition 2.6 that both sets are superregular at a distance relative to $\mathcal{E}$ at any of their points. Applying Theorem 3.7 we deduce firm nonexpansivity of $T_{\mathrm{DR}\lambda}$ since the violation ϵ can be set to 0, as seen in the proof of Proposition 2.6. ∎

3.3. Characterization of $\mathsf{Fix}\, T_{\mathrm{DR}\lambda}$. We collect some facts and identities that will be useful throughout.

LEMMA 3.11. *Let A and B be closed and $T_{\mathrm{DR}\lambda}$ given by* (1.1) *with $\lambda \in (0,1)$. Let $x \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda} \neq \emptyset$ such that $T_{\mathrm{DR}\lambda}$ is single-valued at x. Take $f \in P_B(x)$ and $y := x - f$. Then the following hold:*

(i) *$P_B(x) = \{f\}$; that is, P_B is single-valued on $\mathsf{Fix}\, T_{\mathrm{DR}\lambda}$;*
(ii) *$P_A(R_B(x))$ is single-valued; hence, so is $R_A(R_B(x))$;*
(iii) *$P_A(2f - x) = P_A(R_B(x))$;*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

(iv) $T_{DR}(x) - x = P_A(R_B(x)) - P_B(x)$;
(v) $f + \frac{1-\lambda}{\lambda} y = P_A(2f - x)$;
(vi) *if A is convex, then for $e = P_A(f)$,*

$$P_A\left(e + \frac{1}{1-\lambda}(f-e)\right) = e. \tag{3.11}$$

Proof. (i)–(ii) Since

$$T_{\mathrm{DR}\lambda}(x) = \bigcup_{b \in P_B(x)} \left\{ \frac{\lambda}{2}\left(R_A(2b-x)+x\right) + (1-\lambda)\, b \right\}$$

is just a single point, we conclude that $P_B(x)$ as well as $P_A\left(R_B(x)\right)$ and $R_A\left(R_B(x)\right)$ have to be single-valued, as claimed.

(iii) This is an easy implication of the single-valuedness of P_B at x:

$$P_A\left(2f - x\right) = P_A\left(2P_B(x) - x\right) = P_A\left(R_B(x)\right).$$

(iv) This also follows from single-valuedness of P_B at x:

$$\begin{aligned} T_{DR}(x) - x &= \frac{1}{2}\left(R_A\left(R_B(x)\right) + x\right) - x \\ &= \frac{1}{2}\left(R_A\left(R_B(x)\right)\right) - \frac{1}{2}x \\ &= P_A\left(R_B(x)\right) - \frac{1}{2}R_B(x) - \frac{1}{2}x \\ &= P_A\left(R_B(x)\right) - P_B(x). \end{aligned}$$

(v) To see this, note that

$$\begin{aligned} & x = T_{\mathrm{DR}\lambda}(x) = \frac{\lambda}{2}\left(R_A\left(R_B(x)\right) + x\right) + (1-\lambda)\, P_B(x) \\ \iff \quad & (1-\lambda)\, x = \lambda\left(T_{DR}(x) - x\right) + (1-\lambda)\, P_B(x) \\ \iff \quad & (1-\lambda)\left(x - P_B(x)\right) = \lambda\left(P_A\left(R_B(x)\right) - P_B(x)\right) \end{aligned}$$

by part (iv). Hence, with $f = P_B(x)$, this yields

$$(1-\lambda)\left(x - f\right) = \lambda\left(P_A\left(2f - x\right) - f\right) \iff f + \frac{1-\lambda}{\lambda} y = P_A(2f - x)$$

by the definition of y.

(vi) This follows from the fact that $f - e \in N_A^{\mathrm{P}}(e)$. Since A is convex, then all points in $e + N_A^{\mathrm{P}}(e)$ project back to e. ∎

Remark 3.12. Note that (i) and (ii) of Lemma 3.11 together at some point $x \in \mathcal{E}$ are equivalent to the single-valuedness of $T_{\mathrm{DR}\lambda}$ at x.

THEOREM 3.13 (fixed points). *Let $A, B \subset \mathcal{E}$ both be closed, and let $\lambda \in (0,1)$. Let $T_{\mathrm{DR}\lambda}$ be single-valued at its fixed points on an open set $U \subset \mathcal{E}$. Then*

(3.12)

$$\begin{aligned} &\mathsf{Fix}\, T_{\mathrm{DR}\lambda} \cap U \subset \mathcal{M} \\ &\quad := \left\{ f - \frac{\lambda}{1-\lambda}\left(f-e\right) \,\middle|\, f \in P_B\left(f - \frac{\lambda}{1-\lambda}\left(f-e\right)\right) \ \textit{and} \ e \in P_A(f) \right\} \cap U. \end{aligned}$$

The inclusion is tight if $e \in P_A\left(f + \frac{\lambda}{1-\lambda}(f-e)\right)$ is true for the right-hand side.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Proof. Let $x \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda} \cap U$. By the assumptions, $T_{\mathrm{DR}\lambda}$ is single-valued at x, and hence the results in Lemma 3.11 can be applied. As before denote by f the projection $P_B(x)$. Reformulating Lemma 3.11(v) yields the desired form of the fixed point x:

$$\begin{aligned} x \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda} \quad &\Longrightarrow \quad f + \frac{1-\lambda}{\lambda} y = P_A(2f - x) \\ &\Longleftrightarrow \quad f + \frac{1-\lambda}{\lambda}(x - f) = P_A(2f - x) \\ &\Longleftrightarrow \quad x = \frac{\lambda}{1-\lambda} P_A(2f - x) - \frac{2\lambda - 1}{1-\lambda} f \\ &\Longleftrightarrow \quad x = f - \frac{\lambda}{1-\lambda}\left(f - P_A(2f - x)\right). \end{aligned} \tag{3.13}$$

Comparing with (3.12) we have to show that $P_A(f) = P_A(2f - x)$.

By Lemma 3.11(i) and (ii) we know that $P_B(x) = \{f\}$ is single-valued as well as $P_A(2f - x)$. Reformulating $x \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda}$ again yields

$$\begin{aligned} x \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda} \quad &\Longleftrightarrow \quad x = \frac{\lambda}{2}\left(R_A\left(R_B(x)\right) + x\right) + (1-\lambda) P_B(x) \\ &\Longleftrightarrow \quad f = \lambda P_A(2f - x) + (1-\lambda)(2f - x). \end{aligned} \tag{3.14}$$

Thus, f is a convex combination of $P_A(2f - x)$ and $2f - x$. The projection of f onto A, therefore, has to be in $P_A(2f - x)$, which was a single point. That is, $P_A(f) \subset \{P_A(2f - x)\}$, which immediately implies that $P_A(f) = P_A(2f - x)$.

Using this fact, (3.13) then becomes

$$x = f - \frac{\lambda}{1-\lambda}\left(f - P_A(f)\right).$$

Finally, (3.12) follows from the fact that $f = P_B(x)$ since x is a fixed point.

It remains to show that the inclusion is in fact an equality when

$$e \in P_A\left(f + \frac{\lambda}{1-\lambda}(f - e)\right)$$

for $e \in P_A(f)$. To see this, let $\widetilde{x} \in \mathcal{M} \cap U$ in (3.12). Then $\widetilde{x} := f - \frac{\lambda}{1-\lambda}(f - e)$ for some $e \in P_A(f)$ and $f \in P_B(\widetilde{x})$ and

$$\begin{aligned} \widetilde{x} - T_{\mathrm{DR}\lambda}\widetilde{x} &= \widetilde{x} - \frac{\lambda}{2}\left(R_A\left(R_B(\widetilde{x})\right) + \widetilde{x}\right) - (1-\lambda) P_B(\widetilde{x}) \\ &= \lambda\widetilde{x} - \frac{\lambda}{2}\left(2P_A\left(R_B(\widetilde{x})\right) - 2P_B(\widetilde{x}) + 2\widetilde{x}\right) + (1-\lambda)\left(\widetilde{x} - P_B(\widetilde{x})\right) \\ &= -\lambda\left(P_A\left(R_B(\widetilde{x})\right) - P_B(\widetilde{x})\right) + (1-\lambda)\left(\widetilde{x} - P_B(\widetilde{x})\right) \\ &\ni -\lambda\left(P_A\left(R_B(\widetilde{x})\right) - f\right) + (1-\lambda)\left(\widetilde{x} - f\right) \\ &= -\lambda\left(P_A\left(R_B(\widetilde{x})\right) - f\right) - \lambda\left(f - e\right) \\ &= -\lambda\left(P_A\left(R_B(\widetilde{x})\right)\right) + \lambda e. \end{aligned}$$

Thus, $0 \in \widetilde{x} - T_{\mathrm{DR}\lambda}\widetilde{x}$ if and only if $e \in P_A\left(R_B(\widetilde{x})\right)$, which is equivalent to $e \in P_A(f + \frac{\lambda}{1-\lambda}(f - e))$. This concludes the proof. ∎

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Remark 3.14. (i) Note that $f + \frac{\lambda}{1-\lambda}(f-e) = e + \frac{1}{1-\lambda}(f-e)$, so that for any $e \in P_A(f)$, $f - e$ is in the normal cone to A at e. It follows immediately that if A is convex, then $P_A(e + \frac{1}{1-\lambda}(f-e)) = e$ for all $\lambda \in (0,1)$, and by Theorem 3.13 the inclusion (3.12) is in fact equality for all λ for which $f \in P_B(f - \frac{\lambda}{1-\lambda}(f-e))$. Compare this to the statement in [20, Lemma 3.8], where the tight fixed point characterization holds for $\lambda \in [0, 1/2]$. This is due to the slightly different characterization. The statement in [20] that f is a local best approximation point is actually incorrect. Where our description includes $f \in P_B(f - \frac{\lambda}{1-\lambda}(f-e))$ and $e \in P_A(f)$, the version in [20, Lemma 3.8] states that f is a local best approximation point [20, Definition 3.3]. Instead, what is needed to correct the statement is $f \in P_B(P_A(f))$, and such a point need not be a local best approximation point. To see this, consider a unit circle in $\mathbb{R}^2$ centered at the origin and a horizontal line passing through the point $(0, 3/4)$. For the fixed point mapping $T_{\mathrm{DR}\lambda}$ with A the line and B the circle, the point $(0, 1 - \frac{\lambda}{4(1-\lambda)})$ is a fixed point for all $\lambda \in (0, 4/5)$. But the corresponding points $f = (0,1)$ and $e = (0, 3/4)$ are not local best approximation points.

(ii) The condition $e \in P_A(f + \frac{\lambda}{1-\lambda}(f-e))$ is easier to interpret with the identity $f + \frac{\lambda}{1-\lambda}(f-e) = e + \frac{1}{1-\lambda}(f-e)$. As $\lambda \nearrow 1$, this vector recedes from A in the direction normal to A at e. The larger the neighborhood on which the projection on A exists and is single-valued, the larger λ can be before $e \notin P_A(f + \frac{\lambda}{1-\lambda}(f-e))$. If A is convex, then λ can be arbitrarily close to 1. Still, λ may need to be bounded away from 1 in order to ensure the other condition in the fixed point characterization (3.12), namely, $f \in P_B(f - \frac{\lambda}{1-\lambda}(f-e))$.

(iii) By Theorem 3.7 we know that $T_{\mathrm{DR}\lambda}$ is single-valued at its fixed points if both A and B are superregular at a distance and assumptions (i)–(iv) of Theorem 3.7 hold. The local gap $f - P_A(f)$ is therefore unique. In [20] uniqueness of such gap vectors was an assumption of the convergence analysis. Our results show that we can remove this assumption.

COROLLARY 3.15 (fixed points of DRλ and the corresponding gap). *In the setting of Theorem* 3.13 *let* $x \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda} \cap U$. *Then*

$$\{x\} = P_B(x) - \frac{\lambda}{1-\lambda}\left(P_B(x) - P_A\left(P_B(x)\right)\right).$$

Proof. The result follows directly from the proof of Theorem 3.13. ∎

In our statements we require that $\mathsf{Fix}\, T_{\mathrm{DR}\lambda} \neq \emptyset$. Although this assumption is very strong, it is not very restrictive and is satisfied under the assumption of compactness of one of the underlying sets and convexity of both sets.

PROPOSITION 3.16 (convexity and compactness imply nonempty fixed point set). *Let* $\lambda \in (0,1)$. *If* A *and* B *are convex and closed and* A *is bounded, then* $\mathsf{Fix}\, T_{\mathrm{DR}\lambda} \neq \emptyset$. *Moreover,* $\mathsf{Fix}\, T_{\mathrm{DR}\lambda} = \mathcal{M}$, *where* $\mathcal{M}$ *is given by* (3.12).

Proof. The proof follows the pattern of proof in [20, Lemma 2.1], which establishes existence of fixed points for $T_{\mathrm{DR}\lambda}$ by first showing the existence of fixed points of the alternating projections mapping $T := P_A P_B$. To see this, note that T is nonexpansive since the projectors P_A and P_B are nonexpansive, and the composition of nonexpansive mappings is nonexpansive by a similar argument as made in Example 3.9(i). Note that $U = \mathcal{E}$. Existence of fixed points of T is then an easy consequence of [7, Theorem 2], which requires that one of the sets, A or B, be compact.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

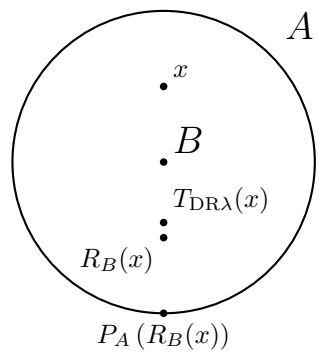

FIG. 4. *Example* 3.17 *for a point* $x \in \mathbb{R}^2$ *and* $\lambda = 0.8$.

To show the tight characterization of $\mathsf{Fix}\, T$, let $e \in \mathsf{Fix}\, T$. Then $P_B(e) = f$ and $P_A(f) = e$ and $T_{\mathrm{DR}\lambda}$ is, by convexity, single-valued. By Remark 3.14(i) we have $f + \frac{\lambda}{1-\lambda}(f-e) = e + \frac{1}{1-\lambda}(f-e)$ and $P_A(e + \frac{1}{1-\lambda}(f-e)) = e$ for all $\lambda \in (0,1)$. Moreover, for all such λ we have $f = P_B(f - \frac{\lambda}{1-\lambda}(f-e))$. Together, for $\overline{x} = f - \frac{\lambda}{1-\lambda}(f-e)$, this yields

$$\begin{aligned}
T_{\mathrm{DR}\lambda}(\overline{x}) &= \frac{\lambda}{2}\left(R_A(R_B(\overline{x})) + \overline{x}\right) + (1-\lambda)P_B(\overline{x}) \\
&= \frac{\lambda}{2}\left(R_A(2f - \overline{x}) + \overline{x}\right) + (1-\lambda)f \\
&= \frac{\lambda}{2}\left(2P_A(e + \frac{1}{1-\lambda}(f-e)) - 2f + \overline{x} + \overline{x}\right) + (1-\lambda)f \\
&= \frac{\lambda}{2}\left(2e - 2f + 2\overline{x}\right) + (1-\lambda)f \\
&= \lambda\left(f - \frac{1}{1-\lambda}(f-e)\right) + (1-\lambda)f \\
&= f - \frac{\lambda}{1-\lambda}(f-e) = \overline{x}.
\end{aligned}$$

Now, applying Theorem 3.13 immediately yields $\mathsf{Fix}\, T_{\mathrm{DR}\lambda} = \mathcal{M}$, where $\mathcal{M}$ is given by (3.12). This completes the proof. ∎

The above result on existence relies heavily on convexity. The next example shows that it is quite easy to construct a scenario where $T_{\mathrm{DR}\lambda}$ has no fixed points.

Example 3.17 (empty fixed point set). Let A be the unit circle in $\mathbb{R}^2$, i.e.,

$$A := \left\{ x = (x_1, x_2) \in \mathbb{R}^2 \;\middle|\; x_1^2 + x_2^2 = 1 \right\},$$

and B its origin, i.e., $B := \{(0,0)\}$. This is depicted in Figure 4. In this setting the fixed point set of $T_{\mathrm{DR}\lambda}$ is empty for all $\lambda \in (0,1)$. To prove this we will show by a case distinction that the fixed point set of $T_{\mathrm{DR}\lambda}$ is empty.

First, note that the projectors and reflectors involved in $T_{\mathrm{DR}\lambda}$ are given by

$$\begin{aligned}
P_B(x) &= (0,0) \quad \forall x \in \mathbb{R}^2 \\
P_A(x) &= \begin{cases} \frac{x}{\|x\|} & \forall x \in \mathbb{R}^2 \setminus (0,0), \\ A & \text{for } x = (0,0). \end{cases}
\end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Now, let $x = (0,0)$. Then

$$T_{\mathrm{DR}\lambda}(x) = \frac{\lambda}{2}\left(R_A R_B(x) + x\right) + (1-\lambda)P_B(x) = \frac{\lambda}{2}\left(R_A(x)\right) = \lambda A.$$

Thus, $x = (0,0)$ cannot be a fixed point of $T_{\mathrm{DR}\lambda}$. For the other case let $x \neq (0,0)$. Then

$$T_{\mathrm{DR}\lambda}(x) = \frac{\lambda}{2}\left(R_A R_B(x) + x\right) + (1-\lambda)P_B(x) = \frac{\lambda}{2}\left(R_A(-x) + x\right) = \lambda\left(x - \frac{x}{\|x\|}\right).$$

If x is a fixed point of $T_{\mathrm{DR}\lambda}$, that is, $x = T_{\mathrm{DR}\lambda}(x)$, then the following has to hold:

$$x = \lambda\left(x - \frac{x}{\|x\|}\right),$$

which is equivalent to

$$\frac{1-\lambda}{\lambda}x = -\frac{x}{\|x\|}.$$

But this is only satisfied when $x = (0,0)$, a contradiction, from which we conclude that $x \notin \mathsf{Fix}\, T_{\mathrm{DR}\lambda}$, and therefore $\mathsf{Fix}\, T_{\mathrm{DR}\lambda} = \emptyset$.

The following proposition provides a comparison of the fixed points for $T_{\mathrm{DR}\lambda}$ for different values of λ.

PROPOSITION 3.18 (monotonicity of $\mathsf{Fix}\, T_{\mathrm{DR}\lambda}$ with respect to λ). *Let A and B be both closed subsets of $\mathcal{E}$ and $\lambda_1, \lambda_2 \in (0,1)$ such that $\lambda_1 \leq \lambda_2$ and $\mathsf{Fix}\, T_{\mathrm{DR}\lambda_2} \neq \emptyset$. Moreover, let $T_{\mathrm{DR}\lambda_2}$ be single-valued at its fixed points. Then*

$$P_B(\mathsf{Fix}\, T_{\mathrm{DR}\lambda_2}) \subseteq P_B(\mathsf{Fix}\, T_{\mathrm{DR}\lambda_1}). \tag{3.15}$$

If (3.12) *holds for λ_2 with equality instead of just set inclusion, then* (3.15) *holds with equality.*

Proof. Let $x \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda_2} \neq \emptyset$. Then, by Corollary 3.15, we have the representation

$$x = P_B(x) - \frac{\lambda_2}{1-\lambda_2}\left(P_B(x) - P_A(P_B(x))\right). \tag{3.16}$$

$\widetilde{x} := P_B(x) - \frac{\lambda_1}{1-\lambda_1}\left(P_B(x) - P_A(P_B(x))\right)$, and note, as in the statements before, that $P_B(x)$ as well as $P_A(P_B(x))$ are single-valued since x is a fixed point of $T_{\mathrm{DR}\lambda_2} \neq \emptyset$. Set $f := P_B(x)$. Then $f \in B$ and $P_B(\widetilde{x}) = f$. To see this, note that $\frac{\lambda_2}{1-\lambda_2}\left(P_A(f) - f\right) \in N_B^{\mathrm{P}}(f)$. Since $0 \leq \frac{\lambda_1}{1-\lambda_1} \leq \frac{\lambda_2}{1-\lambda_2}$, $\tilde{x}$ is a convex combination of f and x, from which we conclude that $P_B(\widetilde{x}) = f$. Moreover, since $P_B(x) = f = P_B(\widetilde{x})$, we can conclude that $\widetilde{x} \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda_1}$. To see this, evaluate $T_{\mathrm{DR}\lambda_1}(\widetilde{x})$

$$\begin{aligned} T_{\mathrm{DR}\lambda_1}(\widetilde{x}) &= \{y \mid y \in \lambda_1\left(P_A(R_B(\widetilde{x})) + \widetilde{x}\right) + (1-2\lambda_1)\, P_B(\widetilde{x})\} \\ &= \{y \mid y \in \lambda_1\left(P_A(2f - \widetilde{x}) + \widetilde{x}\right) + (1-2\lambda_1)\, f\} \end{aligned}$$

since $P_B(\widetilde{x}) = \{f\}$. $2f - \widetilde{x} = 2f - (f - \frac{\lambda_1}{1-\lambda_1}(f - P_A(f)))$, where $P_A(f)$ is single-valued since x is a fixed point of $T_{\mathrm{DR}\lambda_2}$. This yields

$$\begin{aligned} 2f - \widetilde{x} &= f + \tfrac{\lambda_1}{1-\lambda_1}\left(f - P_A(f)\right) \\ &= P_A(f) + \tfrac{1}{1-\lambda_1}\left(f - P_A(f)\right). \end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Analog to what we have seen before, we can argue that $P_A(f) \in P_A(2f - \widetilde{x})$ since $0 \leq \frac{\lambda_1}{1-\lambda_1} \leq \frac{\lambda_2}{1-\lambda_2}$ and $P_A(f) = P_A(2f - x) = P_A(f + \frac{\lambda_2}{1-\lambda_2}(f - P_A(f)))$. This implies that

$$\begin{aligned} \lambda_1 (P_A(f) + \widetilde{x}) + (1 - 2\lambda_1) f \in T_{\mathrm{DR}\lambda_1}(\widetilde{x}) \\ \Longleftrightarrow \\ \lambda_1 \left(P_A(f) + f - \tfrac{\lambda_1}{1-\lambda_1}(f - P_A(f))\right) + (1 - 2\lambda_1) f \in T_{\mathrm{DR}\lambda_1}(\widetilde{x}) \\ \Longleftrightarrow \\ f - \tfrac{\lambda_1}{1-\lambda_1}(f - P_A(f)) \in T_{\mathrm{DR}\lambda_1}(\widetilde{x}) \\ \Longleftrightarrow \\ \widetilde{x} \in T_{\mathrm{DR}\lambda_1}(\widetilde{x}), \end{aligned}$$

and therefore $\widetilde{x} \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda_1}$. In conclusion,

$$P_B(\mathsf{Fix}\, T_{\mathrm{DR}\lambda_2}) \subseteq P_B(\mathsf{Fix}\, T_{\mathrm{DR}\lambda_1}),$$

which proves the claim. ∎

4. Quantitative convergence analysis. We proceed now to the main goal of our study: the convergence analysis of the algorithm. Almost all of the key properties of the relaxed Douglas–Rachford fixed point mapping, $T_{\mathrm{DR}\lambda}$, have been established in section 3. The main idea for convergence goes back to Opial [24]. In our setting nonemptiness of the fixed point set and averagedness of the mapping can be identified as the essential properties yielding convergence of the iterative sequence. It was shown in [22], however, that *gauge metric subregularity* of a fixed point mapping at its fixed points is a necessary condition for quantifiable (by said gauge) rates of convergence of the fixed point iteration. Metric subregularity is still missing from our development, and the main work of this section consists of deriving the conditions on the sets A and B under which (linear) metric subregularity holds.

Definition 4.1 (metric subregularity on a set). *Let $\mathcal{E}$ and $\mathcal{Y}$ be Euclidean spaces, let $\Phi : \mathcal{E} \rightrightarrows \mathcal{Y}$, and let $U \subset \mathcal{E}$, $\overline{y} \in \mathcal{Y}$. The mapping Φ is called* metrically subregular for $\bar{y}$ on U with constant κ relative to $\Lambda \subset \mathcal{E}$ *if*

$$\operatorname{dist}\left(x, \Phi^{-1}(\overline{y}) \cap \Lambda\right) \leq \kappa \operatorname{dist}(\overline{y}, \Phi(x)) \tag{4.1}$$

holds for all $x \in U \cap \Lambda$. When $\overline{y} \in \Phi(\overline{x})$, then Φ is said to be metrically regular at $\overline{x}$ for $\overline{y}$ *relative to Λ when there exists a neighborhood U of $\overline{x}$ and a constant κ such that* (4.1) *holds for all $x \in U \cap \Lambda$.*

When $\Lambda = \mathcal{E}$, the quantifier "relative to" is dropped. The smallest constant κ for which (4.1) *holds is called* modulus *of metric subregularity.*

The abstract result that allows us to quantify the convergence of $T_{\mathrm{DR}\lambda}$ follows. It is a simplified version of [23, Corollary 2.3], which was later refined to show convergence to a specific point in [22, Corollary 1]. The convergence result Theorem 4.2 is later specialized to $T_{\mathrm{DR}\lambda}$ in Theorem 4.11 presented in subsection 4.2.

Theorem 4.2 ((sub)linear convergence with metric subregularity). *Let $T : \Lambda \rightrightarrows \Lambda$ for $\Lambda \subset \mathcal{E}$ with $\mathsf{Fix}\, T$ nonempty and closed, $\Phi := T - \mathrm{Id}$. Denote $(\mathsf{Fix}\, T + \delta\mathbb{B}) \cap \Lambda$ by S_δ for a nonnegative real δ. Suppose that, for all $\bar{\delta} > 0$ small enough, there is a $\gamma \in (0, 1)$, a nonnegative scalar ϵ, and a positive constant α bounded above by 1 such that*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

(i) *T is pointwise almost averaged at all $y \in \mathsf{Fix}\, T \cap \Lambda$ with averaging constant α and violation ϵ on $S_{\gamma\bar{\delta}}$, and*
(ii) *for*

$$\bar{S} := S_{\gamma\bar{\delta}} \setminus \mathsf{Fix}\, T,$$

Φ is metrically subregular for 0 on $\bar{S}$ with constant κ relative to Λ.
Then for any $x^0 \in \Lambda$ close enough to $\mathsf{Fix}\, T \cap \Lambda$, the iterates $x^{j+1} \in Tx^j$ satisfy

$$\operatorname{dist}\left(x^{j+1}, \mathsf{Fix}\, T \cap \Lambda\right) \leq c \operatorname{dist}\left(x^j, \mathsf{Fix}\, T \cap \Lambda\right) \quad \forall x^j \in \bar{S}, \tag{4.2}$$

where $c := \sqrt{1 + \epsilon - \left(\frac{1-\alpha}{\kappa^2 \alpha}\right)}$. If, in addition, κ satisfies

$$\kappa < \sqrt{\frac{1-\alpha}{\epsilon\alpha}}, \tag{4.3}$$

then $\operatorname{dist}\left(x^j, \tilde{x}\right) \to 0$ for some $\tilde{x} \in \mathsf{Fix}\, T \cap \Lambda$ at least R-linearly with rate at most $c < 1$. If $\mathsf{Fix}\, T \cap \Lambda$ is a single point, then convergence is Q-linear.

We have already shown in Theorem 3.7 that $T_{\mathrm{DR}\lambda}$ is almost averaged, *with any desired violation constant* $\epsilon > 0$, at its fixed points on certain neighborhoods when A and B are superregular at a distance. To achieve local linear convergence, inequality (4.3) must hold, and this is where uniformity of almost averagedness with respect to ϵ is crucial: As long as the mapping $T_{\mathrm{DR}\lambda} - \mathrm{Id}$ or a related mapping (see the discussion below) can be shown to be relatively metrically subregular on a neighborhood of $\mathsf{Fix}\, T_{\mathrm{DR}\lambda}$—*regardless of the value of the modulus* κ—then suitable neighborhoods can be found in the context of Theorem 3.7, where the violation, ϵ, is small enough that (4.3) is satisfied, and hence local linear convergence is guaranteed.

The main work before us (section 4.1) is to show metric subregularity of the appropriate mapping at points in the product space corresponding to fixed points of $T_{\mathrm{DR}\lambda}$. There are a number of ways to go about this, but all successful strategies we found are based on a characterization of the iterates on neighborhoods of fixed points lifted to a product space where the tools are applied. We were unable to provide a direct approach, involving the $T_{\mathrm{DR}\lambda}$ mapping itself, that guarantees metric subregularity from properties of the regularity of the sets A and B both individually (e.g., relative superregularity at a distance) and as a collection (e.g., *subtransversality* discussed below). The characterization of the fixed points in Theorem 3.13 allows us to build auxiliary *phantom* sets that are used in the analysis. To adapt the framework above to the present setting we build a product space which represents not only the iterates of $T_{\mathrm{DR}\lambda}$ but also a cyclic projection between the phantom sets. In particular, we will define an operator in the product space $\mathcal{E}^4$ whose first entry is generated by applying $T_{\mathrm{DR}\lambda}$. The remaining three entries are generated by projecting the prior entry onto the sets A and B as well as phantom versions of these sets shifted by a scaling of the local gap vector between A and B at the reference fixed point.

4.1. $T_{\mathrm{DR}\lambda}$ at $\mathsf{Fix}\, T_{\mathrm{DR}\lambda}$: Metric subregularity via subtransversality. Direct verification of metric subregularity is notoriously difficult, and verifying this for $T_{\mathrm{DR}\lambda}$ is no different. In principle, one must show that the coderivative (the generalized Jacobian) of the (multivalued) $T_{\mathrm{DR}\lambda}$ mapping is injective on neighborhoods of $\mathsf{Fix}\, T_{\mathrm{DR}\lambda}$ [11, Theorems 4B.1 and 4C.2]. We were unable to compute the coderivative of the $T_{\mathrm{DR}\lambda}$ mapping let alone determine whether this is injective.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Since our mapping is based on *projectors* to sets, another route is available for showing metric subregularity which uses characterizations of the regularity of sets *in relation to one another*. In the context of consistent set feasibility, metric subregularity of a particular set-valued mapping on the product space has been shown to be equivalent to *subtransversality* of the *collection of sets* [15].[1] This was expanded in [23, Definition 3.2] to account for *inconsistent set feasibility*. Based on this more general notion of subtransversality of nonoverlapping sets Luke et al. [23] were able to show that the cyclic projections mapping, $T_{CP} := P_{\Omega_1} P_{\Omega_2} \cdots P_{\Omega_m}$, is metrically subregular when the collection of sets $\{\Omega_1, \ldots, \Omega_m\}$ is subtransversal, and an additional technical assumption is satisfied. We follow this approach here but for the mapping $T_{\mathrm{DR}\lambda}$. Note that this general definition can simplify in special cases, such as intersecting sets, as is discussed in Proposition 4.14.

DEFINITION 4.3 (subtransversal collection of sets). *Let $\{\Omega_1, \Omega_2, \ldots, \Omega_m\}$ be a collection of nonempty closed subsets of $\mathcal{E}$, and define $\Psi : \mathcal{E}^m \rightrightarrows \mathcal{E}^m$ by $\Psi(x) := P_\Omega(\Pi x) - \Pi x$, where $\Omega := \Omega_1 \times \Omega_2 \times \cdots \times \Omega_m$, the projection P_Ω is with respect to the Euclidean norm on $\mathcal{E}^m$, and $\Pi : x = (x_1, x_2, \ldots, x_m) \mapsto (x_2, x_3, \ldots, x_m, x_1)$ is the permutation mapping on the product space $\mathcal{E}^m$ for $x_j \in \mathcal{E}$ $(j = 1, 2, \ldots, m)$. Let $\overline{x} = (\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_m) \in \mathcal{E}^m$ and $\overline{y} \in \Psi(\overline{x})$. The collection of sets is said to be* subtransversal with constant κ relative to $\Lambda \subset \mathcal{E}^m$ at $\overline{x}$ for $\overline{y}$ *if Ψ is metrically subregular for $\overline{y}$ on some neighborhood U of $\overline{x}$ with constant κ relative to Λ.*

In contrast to the original model setting, where $\{\Omega_1, \Omega_2, \ldots, \Omega_m\}$ is a collection of subsets on $\mathcal{E}$, our definition of subtransversality is formulated on the product space $\mathcal{E}^m$, where $\Omega_1 \times \Omega_2 \times \cdots \times \Omega_m$ in $\mathcal{E}^m$.

LEMMA 4.4 (subtransversality under addition). *Let $\{\Omega_1, \Omega_2 \ldots, \Omega_m\} \subset \mathcal{E}$ be a subtransversal collection of sets at a point $\overline{x} = (\overline{x}_1, \overline{x}_2, \ldots, \overline{x}_m)$ for $\overline{y} \in \Psi(\overline{x})$ relative to $\Lambda \subset \mathcal{E}^m$ with modulus κ. Then the collection*

$$\{\Omega_1, \Omega_2 \ldots, \Omega_m, \Omega_1 - g, \Omega_2 - g, \ldots, \Omega_m - g\} \subset \mathcal{E}$$

for some $g \in \mathcal{E}$ is subtransversal at

$$\tilde{x} = (\overline{x}_1 - g, \overline{x}_2, \overline{x}_3, \ldots, \overline{x}_m, \overline{x}_1, \overline{x}_2 - g, \overline{x}_3 - g, \ldots, \overline{x}_m - g) \in \mathcal{E}^{2m} \tag{4.4}$$

for

$$\tilde{y} = (\overline{y}, \overline{y}) = (\overline{y}_1, \overline{y}_2, \ldots, \overline{y}_m, \overline{y}_1, \overline{y}_2, \ldots, \overline{y}_m) \in \mathcal{E}^{2m}$$

relative to

$$\tilde{\Lambda} = \{z \in \mathcal{E}^{2m} |\, (z_{m+1}, z_2, z_3, \ldots, z_m) \in \Lambda,\ (z_1, z_{m+2}, z_{m+3}, \ldots, z_{2m}) \in \Lambda - (g, g, \ldots, g)\}$$

with modulus κ.

Proof. We will show the result only for $m = 2$ for reasons of simplicity and since one can easily enlarge the number of sets used in the proof by the same pattern shown here. For $s \in \mathbb{N}$ denote by $\Pi^s_{\mathcal{E}}$ the permutation mapping on $\mathcal{E}^s$.

Let $U \subset \mathcal{E}^2$ be a neighborhood of $\overline{x} \in \mathcal{E}^2$ such that subtransversality holds at $\overline{x}$ for $\overline{y}$ relative to Λ. Define $\Omega := \Omega_1 \times \Omega_2$ and therefore $(\Omega_1 - g) \times (\Omega_2 - g) = \Omega - (g, g)$.

[1] The terminology for this property in the literature is in disarray, and there are often several names with snappy prefixes for the same notion.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Likewise set

$$\tilde{U} := \left\{ z \in \mathcal{E}^4 \mid (z_3, z_2) \in U, z_1 = z_3 - g, z_4 = z_2 - g \right\}.$$

Thus, every $z \in \tilde{U} \cap \tilde{\Lambda}$ can be expressed as $(x_1 - g, x_2, x_1, x_2 - g)^T$ for some $(x_1, x_2) \in U \cap \Lambda$.

To show subtransversality of $\{\Omega_1, \Omega_2, \Omega_1 - g, \Omega_2 - g\}$ we have to verify metric subregularity of $\Psi = P_\Omega(\Pi^4_{\mathcal{E}}) - \Pi^4_{\mathcal{E}}$ for $\tilde{y} \in \Psi(\tilde{x})$ relative to $\tilde{\Lambda}$ on $\tilde{U}$, a neighborhood of $\tilde{x}$.

First, we show that $\tilde{y} \in \Psi(\tilde{x})$, i.e., $\tilde{y} \in P_{\Omega\times\Omega-(0,0,g,g)}\left(\Pi^4_{\mathcal{E}}(\tilde{x})\right) - \Pi^4_{\mathcal{E}}(\tilde{x})$. Let $\tilde{x}$ be defined by (4.4); then

$$\begin{aligned} &P_{\Omega\times\Omega-(0,0,g,g)}\left(\Pi^4_{\mathcal{E}}(\tilde{x})\right) - \Pi^4_{\mathcal{E}}(\tilde{x}) \\ &\quad = \left(P_\Omega\left(\tilde{x}_2, \tilde{x}_3\right) - \left(\tilde{x}_2, \tilde{x}_3\right), P_{\Omega-(g,g)}\left(\tilde{x}_4, \tilde{x}_1\right) - \left(\tilde{x}_4, \tilde{x}_1\right)\right) \\ &\quad = \left(P_\Omega\left(\overline{x}_2, \overline{x}_1\right) - \left(\overline{x}_2, \overline{x}_1\right), P_{\Omega-(g,g)}\left(\overline{x}_2 - g, \overline{x}_1 - g\right) - \left(\overline{x}_2 - g, \overline{x}_1 - g\right)\right) \\ &\quad = \left(P_\Omega\left(\overline{x}_2, \overline{x}_1\right) - \left(\overline{x}_2, \overline{x}_1\right), P_\Omega\left(\overline{x}_2, \overline{x}_1\right) - \left(\overline{x}_2, \overline{x}_1\right)\right), \end{aligned} \tag{4.5}$$

where the last equality holds since $P_{C-g}(x - g) = P_C(x) - g$ for any set C. Then (4.5) yields

$$\begin{aligned} &P_{\Omega\times\Omega-(0,0,g,g)}\left(\Pi^4_{\mathcal{E}}(\tilde{x})\right) - \Pi^4_{\mathcal{E}}(\tilde{x}) \\ &\quad = \left(P_\Omega\left(\overline{x}_2, \overline{x}_1\right) - \left(\overline{x}_2, \overline{x}_1\right), P_\Omega\left(\overline{x}_2, \overline{x}_1\right) - \left(\overline{x}_2, \overline{x}_1\right)\right) \\ &\quad \ni \left(\overline{y}, \overline{y}\right) = \tilde{y} \end{aligned}$$

since $\overline{y} \in P_\Omega(\Pi^2_{\mathcal{E}}\overline{x}) - \Pi^2_{\mathcal{E}}\overline{x}$ by the assumptions on subtransversality of $\{\Omega_1, \Omega_2, \ldots, \Omega_m\}$. By $\tilde{x} \in \tilde{\Lambda}$ this shows $\tilde{y} \in \Psi(\tilde{x})$ as claimed.

It remains to prove that inequality (4.1) holds for Ψ and at $\tilde{x}$ for $\tilde{y} \in \Psi(\tilde{x})$ relative to $\tilde{\Lambda}$ on $\tilde{U}$. For this, take a $(x_1 - g, x_2, x_1, x_2 - g)^T \in \tilde{U} \cap \tilde{\Lambda}$; then

$$\operatorname{dist}^2\left(P_{\Omega\times\Omega-(0,0,g,g)}\left(\Pi^4_{\mathcal{E}}\begin{pmatrix} x_1 - g \\ x_2 \\ x_1 \\ x_2 - g \end{pmatrix}\right) - \Pi^4_{\mathcal{E}}\begin{pmatrix} x_1 - g \\ x_2 \\ x_1 \\ x_2 - g \end{pmatrix}, \begin{pmatrix} \overline{y}_1 \\ \overline{y}_2 \\ \overline{y}_1 \\ \overline{y}_2 \end{pmatrix}\right) \tag{4.6}$$

$$\begin{aligned} &= \operatorname{dist}^2\left(P_\Omega\begin{pmatrix} x_2 \\ x_1 \end{pmatrix} - \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}, \begin{pmatrix} \overline{y}_1 \\ \overline{y}_2 \end{pmatrix}\right) \\ &\qquad + \operatorname{dist}^2\left(P_{\Omega-(g,g)}\begin{pmatrix} x_2 - g \\ x_1 - g \end{pmatrix} - \begin{pmatrix} x_2 - g \\ x_1 - g \end{pmatrix}, \begin{pmatrix} \overline{y}_1 \\ \overline{y}_2 \end{pmatrix}\right) \end{aligned} \tag{4.7}$$

by rewriting the distance on $\mathcal{E}^4$ in terms of the distance on $\mathcal{E}^2$. Using again the identity $P_{C-g}(x - g) = P_C(x) - g$ for an arbitrary set C, (4.6) ends up as

$$\begin{aligned} &\left(\operatorname{dist}^2\left(P_\Omega\left(\Pi^2_{\mathcal{E}}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) - \Pi^2_{\mathcal{E}}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} \overline{y}_1 \\ \overline{y}_2 \end{pmatrix}\right) + \operatorname{dist}^2\left(P_\Omega\begin{pmatrix} x_2 \\ x_1 \end{pmatrix} - \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}, \begin{pmatrix} \overline{y}_1 \\ \overline{y}_2 \end{pmatrix}\right)\right) \\ &\quad = 2\left(\operatorname{dist}^2\left(P_\Omega\begin{pmatrix} x_2 \\ x_1 \end{pmatrix} - \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}, \begin{pmatrix} \overline{y}_1 \\ \overline{y}_2 \end{pmatrix}\right)\right) \\ &\quad \geq \frac{2}{\kappa^2}\operatorname{dist}^2\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \left(P_\Omega\left(\Pi^2_{\mathcal{E}}(\cdot)\right) - \Pi^2_{\mathcal{E}}(\cdot)\right)^{-1}\begin{pmatrix} \overline{y}_1 \\ \overline{y}_2 \end{pmatrix}\right), \end{aligned} \tag{4.8}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

where the last inequality holds by subtransversality of $\{\Omega_1, \Omega_2\}$ at $(\overline{x}_1, \overline{x}_2)$ for $(\overline{y}_1, \overline{y}_2)$ relative to Λ with modulus κ on U. Denote

$$\Sigma := \left\{ \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix} \middle| \, P_\Omega \begin{pmatrix} z_2 \\ z_1 \end{pmatrix} - \begin{pmatrix} z_2 \\ z_1 \end{pmatrix} \ni \begin{pmatrix} \overline{y}_1 \\ \overline{y}_2 \end{pmatrix}, \, P_{\Omega-(g,g)} \begin{pmatrix} z_4 \\ z_3 \end{pmatrix} - \begin{pmatrix} z_4 \\ z_3 \end{pmatrix} \ni \begin{pmatrix} \overline{y}_1 \\ \overline{y}_2 \end{pmatrix} \right\}.$$

Rewriting (4.8) in the distance on $\mathcal{E}^4$ yields

$$\begin{aligned}
&\kappa^2 \operatorname{dist}^2 \left(P_{\Omega\times\Omega-(0,0,g,g)} \left(\Pi^4_{\mathcal{E}} \begin{pmatrix} x_1 - g \\ x_2 \\ x_1 \\ x_2 - g \end{pmatrix} \right) - \Pi^4_{\mathcal{E}} \begin{pmatrix} x_1 - g \\ x_2 \\ x_1 \\ x_2 - g \end{pmatrix}, \begin{pmatrix} \overline{y}_1 \\ \overline{y}_2 \\ \overline{y}_1 \\ \overline{y}_2 \end{pmatrix} \right) \\
&\quad \geq \operatorname{dist}^2 \left(\begin{pmatrix} x_1 \\ x_2 \\ x_1 \\ x_2 \end{pmatrix}, \Sigma \right) \\
&\quad = \operatorname{dist}^2 \left(\begin{pmatrix} x_1 \\ x_2 \\ x_1 - g \\ x_2 - g \end{pmatrix}, \left\{ \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix} \middle| \, P_{\Omega\times\Omega-(0,0,g,g)} \begin{pmatrix} z_2 \\ z_1 \\ z_4 \\ z_3 \end{pmatrix} - \begin{pmatrix} z_2 \\ z_1 \\ z_4 \\ z_3 \end{pmatrix} \ni \begin{pmatrix} \overline{y}_1 \\ \overline{y}_2 \\ \overline{y}_1 \\ \overline{y}_2 \end{pmatrix} \right\} \right) \\
&\quad = \operatorname{dist}^2 \left(\begin{pmatrix} x_1 - g \\ x_2 \\ x_1 \\ x_2 - g \end{pmatrix}, \left\{ \begin{pmatrix} z_3 \\ z_2 \\ z_1 \\ z_4 \end{pmatrix} \middle| \, P_{\Omega\times\Omega-(0,0,g,g)} \begin{pmatrix} z_2 \\ z_1 \\ z_4 \\ z_3 \end{pmatrix} - \begin{pmatrix} z_2 \\ z_1 \\ z_4 \\ z_3 \end{pmatrix} \ni \begin{pmatrix} \overline{y}_1 \\ \overline{y}_2 \\ \overline{y}_1 \\ \overline{y}_2 \end{pmatrix} \right\} \right) \\
&\quad = \operatorname{dist}^2 \left(\begin{pmatrix} x_1 - g \\ x_2 \\ x_1 \\ x_2 - g \end{pmatrix}, \left(P_{\Omega\times\Omega-(0,0,g,g)} \left(\Pi^4_{\mathcal{E}} (\cdot) \right) - \Pi^4_{\mathcal{E}} (\cdot) \right)^{-1} \begin{pmatrix} \overline{y}_1 \\ \overline{y}_2 \\ \overline{y}_1 \\ \overline{y}_2 \end{pmatrix} \right),
\end{aligned}$$

where the last three steps were just rearranging the expression to get the claimed result. This completes the proof. ∎

Remark 4.5. The points involved in Lemma 4.4 change if we change the order of the sets involved. Of particular interest for our later analysis is the case of two sets A and B where we change the order on the product space as

$$(B - g) \times (A - g) \times A \times B \subset \mathcal{E}^4 \tag{4.9}$$

in contrast to the order $A \times B \times (A - g) \times (B - g)$ as used in Lemma 4.4. Therefore, the points $\widetilde{x}$ and $\widetilde{y}$ as well as the set $\tilde{\Lambda}$ change to

$$\begin{aligned}
\widetilde{x}' &= (\overline{x}_1, \overline{x}_1 - g, \overline{x}_2 - g, \overline{x}_2) \\
\widetilde{y}' &= (-\overline{y}_1, -\overline{y}_2, \overline{y}_1, \overline{y}_2) \\
\tilde{\Lambda}' &= \left\{ z \in \mathcal{E}^4 \mid (z_3, z_4) \in \Lambda, \ \ (z_2, z_1) \in \Lambda - (g, g) \right\}.
\end{aligned}$$

That is, the collection $\{B - g, A - g, A, B\} \subset \mathcal{E}$ is subtransversal at $\tilde{x}'$ for $\tilde{y}'$ relative to $\tilde{\Lambda}'$. Note that the negative part of $\widetilde{y}$ emerged from the changed order of B and A in comparison to Lemma 4.4.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

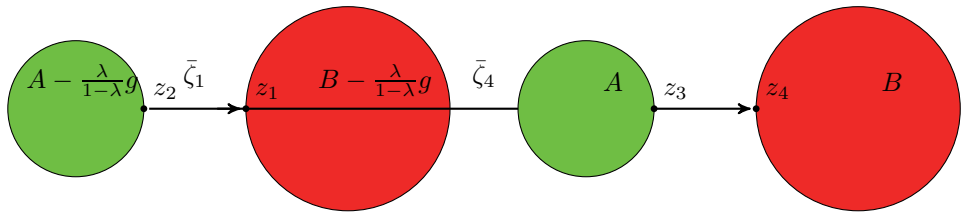

FIG. 5. *Framework for the convergence analysis illustrated in* $\mathcal{E}$.

We are now ready to construct the product space on which we determine metric subregularity via subtransversality. Instead of the two original sets, we consider four sets: the sets A, B and shifted sets $B - \frac{\lambda}{1-\lambda}g$ and $A - \frac{\lambda}{1-\lambda}g$ for some gap vector g. This is illustrated in Figure 5. Our aim is to show local linear convergence of $T_{\mathrm{DR}\lambda}$ by adapting the approach developed in [23] for cyclic projections. There, it was essential that one of the sets involved contains the fixed points of the mapping. The reason for including the set $B - \frac{\lambda}{1-\lambda}g$ in our problem, therefore, lies in the characterization of the fixed point set of the $T_{\mathrm{DR}\lambda}$ mapping. As established in Theorem 3.13 and Corollary 3.15, fixed points $\overline{x}$ of $T_{\mathrm{DR}\lambda}$ at which $T_{\mathrm{DR}\lambda}$ is single-valued can be described as

$$\{\overline{x}\} = P_B(\overline{x}) - \frac{\lambda}{1-\lambda}\left(P_B(\overline{x}) - P_A(P_B(\overline{x}))\right),$$

which is an element in $B - \frac{\lambda}{1-\lambda}g$ when $g = P_B(\overline{x}) - P_A(P_B(\overline{x}))$. Thus, locally, $B - \frac{\lambda}{1-\lambda}g$ contains fixed points of $T_{\mathrm{DR}\lambda}$. To be able to apply our results established in Lemma 4.4 we have to consider the set $A - \frac{\lambda}{1-\lambda}g$ as well.

We denote by Ω_g the product of the collection of sets

$$\left\{B - \frac{\lambda}{1-\lambda}g, A - \frac{\lambda}{1-\lambda}g, A, B\right\}.$$

That is,

$$\Omega_g := \left(B - \frac{\lambda}{1-\lambda}g\right) \times \left(A - \frac{\lambda}{1-\lambda}g\right) \times A \times B.$$

Define

(4.10)
$$W_0(g) := \Big\{u = (u_1, u_2, u_3, u_4) \in \mathcal{E}^4 \ \Big| \ u_1 \in P_{B-\frac{\lambda}{1-\lambda}g}(u_2),\ u_2 \in P_{A-\frac{\lambda}{1-\lambda}g}(u_3),\ u_3 \in P_A(u_4),\ u_4 \in P_B(u_1)\Big\}.$$

This is the set of fixed points of the mapping $P_{\Omega_g} \circ \Pi$ in the product space $\mathcal{E}^4$ corresponding to a cycle of the cyclic projections operator $P_{B-\frac{\lambda}{1-\lambda}g}P_{A-\frac{\lambda}{1-\lambda}g}P_AP_B$. By our construction, the set $W_0(g)$ could be (and for generic g *will be*) empty; this would be the case when g does not correspond to a difference vector. The set of *difference vectors,* ζ, is denoted by $\mathcal{Z}(x, g)$ and defined by

$$\mathcal{Z}(x, g) := \left\{\zeta := z - \Pi z \ \middle| \ z \in W_0(g) \subset \mathcal{E}^4,\ z_1 = x\right\}. \tag{4.11}$$

The last set to introduce is

$$W\left(\bar{\zeta}\right) := \left\{u \in \mathcal{E}^4 \ \middle| \ u - \Pi u = \bar{\zeta}\right\}. \tag{4.12}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

This set is an affine transformation of the diagonal of the product space and serves as a characterization of the local geometry of the sets in relation to each other at fixed points of $T_{\mathrm{DR}\lambda}$.

These sets, of course, only make sense in the context of local nearest points between the components. In particular, we are interested in points $x \in \mathcal{E}$ associated with fixed points of $T_{\mathrm{DR}\lambda}$ and their associated shadow points and gap vectors, respectively $b \in P_B(x)$ and $g \in b - P_A(b)$ (the local gap between A and B). Note that by Theorem 3.7 for fixed points of $T_{\mathrm{DR}\lambda}$ at which $T_{\mathrm{DR}\lambda}$ is single-valued we have $\{b\} = P_B(x)$, and the gap vector g is unique. When x is a fixed point, the set $\mathcal{Z}(x, g)$ characterizes the distance between the cyclically projected iterates of $T_{\mathrm{DR}\lambda}$ on the individual sets. This enables us to distinguish different fixed points of $T_{\mathrm{DR}\lambda}$ according to their respective difference vectors.

Let the assumptions of Theorem 3.7 hold at $\overline{u} \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda}$, and let $z \in W_0(g) \subset \mathcal{E}^4$ for $z_1 = \overline{u}$ and $\{g\} = P_B(\overline{u}) - P_A P_B(\overline{u})$. Then Theorem 3.7 yields

(i) $z_3 - z_4 \in P_A\left(P_B(z_1)\right) - P_B(z_1) = \{-g\}$;

(ii) $z_4 - z_1 \in P_B(z_1) - (P_B(z_1) - \frac{\lambda}{1-\lambda} g) = \{\frac{\lambda}{1-\lambda} g\}$,

where (ii) holds by Theorem 3.13, which is applicable since $T_{\mathrm{DR}\lambda}$ is single-valued at $\overline{u}$ by Theorem 3.7. Moreover, we get by the assumptions of Theorem 3.7 that $P_A(R_B(\overline{u})) = P_A(P_B(\overline{u})) = z_3$. Since $z_3 + \frac{\lambda}{1-\lambda} g$ lies in a straight line between $R_B(\overline{u}) = 2P_B(\overline{u}) - \overline{u} = 2P_B(\overline{u}) - P_B(\overline{u}) + \frac{\lambda}{1-\lambda} g = P_A(P_B(\overline{u})) + (1 + \frac{\lambda}{1-\lambda}) g$ and $P_A(R_B(\overline{u}))$ we also deduce $P_A(z_3 + \frac{\lambda}{1-\lambda} g) = z_3$. Using again Theorem 3.13 yields

(iii)
$$\begin{aligned} z_1 - z_2 &\in P_B(\overline{u}) - \tfrac{\lambda}{1-\lambda} g - P_{A - \frac{\lambda}{1-\lambda} g}(z_3) \\ &= P_B(\overline{u}) - \tfrac{\lambda}{1-\lambda} g - P_A(z_3 + \tfrac{\lambda}{1-\lambda} g) + \tfrac{\lambda}{1-\lambda} g \\ &= P_B(\overline{u}) - z_3 \\ &= P_B(\overline{u}) - P_A(P_B(\overline{u})) \\ &= \{g\}\,; \end{aligned}$$

(iv) $z_2 - z_3 = (z_2 - z_1) + (z_1 - z_4) + (z_4 - z_3) = -\frac{\lambda}{1-\lambda} g$.

Figure 5 illustrates the sets and difference vectors above. The individual entries of z relate to the cyclically projected fixed point x on each of the individual sets.

Along with the definitions above we define the operator

$$\begin{aligned} &T_{\bar{\zeta}} : \mathcal{E}^4 \rightrightarrows \mathcal{E}^4 : \\ &\quad u \mapsto \left\{ \left(u_1^+, u_1^+ - \bar{\zeta}_1, u_1^+ - \bar{\zeta}_1 - \bar{\zeta}_2, u_1^+ - \bar{\zeta}_1 - \bar{\zeta}_2 - \bar{\zeta}_3\right) \;\middle|\; u_1^+ \in T_{\mathrm{DR}\lambda} u_1 \right\} \end{aligned}$$

for $\bar{\zeta} \in \mathcal{Z}(\overline{x}, g)$, where $\overline{x} \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda}$ and $g = P_B(\overline{x}) - P_A P_B(\overline{x})$. Note that $0 = \bar{\zeta}_1 + \bar{\zeta}_2 + \bar{\zeta}_3 + \bar{\zeta}_4$, so the expression above can be simplified to

$$T_{\bar{\zeta}} : \mathcal{E}^4 \rightrightarrows \mathcal{E}^4 : u \mapsto \left\{ \left(u_1^+, u_1^+ - \bar{\zeta}_1, u_1^+ - \bar{\zeta}_1 - \bar{\zeta}_2, u_1^+ + \bar{\zeta}_4\right) \;\middle|\; u_1^+ \in T_{\mathrm{DR}\lambda} u_1 \right\}. \tag{4.13}$$

The mapping $T_{\bar{\zeta}}$ represents the iterates of $T_{\mathrm{DR}\lambda}$ on the space $\mathcal{E}$ by shifting each iterate by some fixed difference vector $\bar{\zeta}$. We presume, in what follows, that $\bar{\zeta}$ is the difference vector corresponding to the fixed point to which our iteration is converging. Of course, when one does not know the location of the fixed points, it is unlikely that the corresponding difference vector will be known, but this situation is no different than other studies which assume that the problem is consistent and that all fixed points correspond to the zero difference vector. Our aim here is not to determine the difference vector or the fixed point but rather to provide a quantification of the

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

convergence based on verifiable regularity of the fixed point mapping *in neighborhoods* of fixed points.

We are now ready to start building our argument. The following lemma establishes a connection between fixed points of $T_{\bar{\zeta}}$ to fixed points of $T_{\mathrm{DR}\lambda}$.

LEMMA 4.6. *Let $\lambda \in (0,1)$ and $A, B \subset \mathcal{E}$ be both nonempty and closed. Fix $\overline{x} \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda} \neq \emptyset$ with $T_{\mathrm{DR}\lambda}$ being single-valued at $\overline{x}$, and set $g := P_B(\overline{x}) - P_A(P_B(\overline{x}))$. Furthermore, let $\bar{\zeta} \in \mathcal{Z}(\overline{x}, g)$, and define $\Psi_g := \left(P_{\Omega_g}\right) \circ \Pi - \Pi$ as well as $\Phi_{\bar{\zeta}} := T_{\bar{\zeta}} - \mathrm{Id}$. Then the following hold:*

(i) *$T_{\bar{\zeta}}$ maps $W\left(\bar{\zeta}\right)$ to itself. Moreover, $u \in \mathsf{Fix}\, T_{\bar{\zeta}}$ if and only if $u \in W\left(\bar{\zeta}\right)$ with $u_1 \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda}$.*

(ii)
$$\Psi_g^{-1}\left(\bar{\zeta}\right) \cap W\left(\bar{\zeta}\right) \cap \mathcal{N} \subseteq \Phi_{\bar{\zeta}}^{-1}(0) \cap W\left(\bar{\zeta}\right),$$
where $\mathcal{N} := \{z \in \mathcal{E}^4 \mid P_A(z_4 + \frac{\lambda}{1-\lambda} g) = z_3\}$.

(iii) *If the distance is with respect to the Euclidean norm, then $\operatorname{dist}\left(0, \Phi_{\bar{\zeta}}(u)\right) = 2 \operatorname{dist}\left(u_1, T_{\mathrm{DR}\lambda} u_1\right)$ for $u \in W(\bar{\zeta})$.*

Remark 4.7. Note that the set $\mathcal{N}$ guarantees equality of the description of the fixed point set in Theorem 3.13. In our main result $\mathcal{N}$ will not appear anymore. This is due to the fact that the assumptions of Theorem 3.7 ensure that the neighborhood we consider will be a subset of $\mathcal{N}$.

Proof of Lemma 4.6. (i) The first part of (i) follows immediately by the definition of $T_{\bar{\zeta}}$ and $W(\bar{\zeta})$. Now let $u \in \mathsf{Fix}\, T_{\bar{\zeta}}$,

$$\begin{aligned}
&\iff \quad u_1 \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda} \text{ and } u_2 = u_1 - \bar{\zeta}_1,\ u_3 = u_1 - \bar{\zeta}_1 - \bar{\zeta}_2,\ u_4 = u_1 + \bar{\zeta}_4 \\
&\iff \quad u_1 \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda} \text{ and } u_2 = u_1 - \bar{\zeta}_1,\ u_3 = u_2 - \bar{\zeta}_2,\ u_4 = u_1 + \bar{\zeta}_4 \\
&\iff \quad u_1 \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda} \text{ and } u \in W(\bar{\zeta}),
\end{aligned}$$

which proves the rest of (i).

(ii) For the second part of the lemma let $z \in \Psi_g^{-1}\left(\bar{\zeta}\right) \cap W(\bar{\zeta}) \cap \mathcal{N}$. This means nothing more than

$$\bar{\zeta} \in \Psi_g\left(z\right) \text{ and } z - \Pi z = \bar{\zeta},$$

which is equivalent to

$$\bar{\zeta} \in P_{\Omega_g} \Pi z - \Pi z \text{ and } z - \Pi z = \bar{\zeta}.$$

This implies that

$$z_1 \in P_{B - \frac{\lambda}{1-\lambda} g} P_{A - \frac{\lambda}{1-\lambda} g} P_A P_B z_1 \text{ and } z - \Pi z = \bar{\zeta}.$$

The mapping $\Phi_{\bar{\zeta}}(z) = T_{\bar{\zeta}} z - z$ has the image $(0,0)$ if $z_1 \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda} z_1$. By $\bar{\zeta}_4 = z_4 - z_1$ and $\bar{\zeta}_4 \in P_B(z_1) - z_1 = \frac{\lambda}{1-\lambda} g$ we know that $z_4 \in P_B(z_1)$. This together with the definition of $\mathcal{N}$ yields $P_A\left(R_B(z_1)\right) \ni P_A\left(2z_4 - z_1\right) = P_A(z_4 + \frac{\lambda}{1-\lambda} g) = z_3$. Inserting this in $T_{\mathrm{DR}\lambda}(z_1)$ yields

$$\begin{aligned}
T_{\mathrm{DR}\lambda}(z_1) &= \lambda\left(P_A\left(R_B(z_1)\right) + z_1\right) + (1 - 2\lambda) P_B(z_1) \\
&\ni \lambda\left(z_3 + z_1\right) + (1 - 2\lambda) z_4 \\
&= z_1 + \lambda(z_3 - z_4) + (1 - \lambda)\left(z_4 - z_1\right) \\
&= z_1 + \bar{\zeta}_3 + (1 - \lambda)\bar{\zeta}_4 \\
&= z_1
\end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

since $\bar{\zeta}$ is generated by a fixed point of $T_{\mathrm{DR}\lambda}$. Thus, $z_1 \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda}$, which proves $z \in \Phi_{\bar{\zeta}}^{-1}(0)$ and completes the proof of (ii).

(iii). This part of the proof is a routine calculation:

$$\begin{aligned}\operatorname{dist}\left(0, \Phi_{\bar{\zeta}}(u)\right) &= \operatorname{dist}\left(0, T_{\bar{\zeta}}u - u\right)\\ &= \Big(\operatorname{dist}^2\left(0, T_{\mathrm{DR}\lambda}u_1 - u_1\right) + \operatorname{dist}^2\left(0, T_{\mathrm{DR}\lambda}u_1 - \bar{\zeta}_1 - u_2\right) + \cdots\\ &\qquad + \operatorname{dist}^2\left(0, T_{\mathrm{DR}\lambda}u_1 - \sum_{i=1}^{3}\bar{\zeta}_i - u_4\right)\Big)^{1/2}\\ &= 2\operatorname{dist}\left(0, T_{\mathrm{DR}\lambda}u_1 - u_1\right). \qquad \square\end{aligned}$$

We are now ready for the main result of this subsection. We show that the mapping $T_{\bar{\zeta}} - \mathrm{Id}$ is metrically subregular on neighborhoods of its zeros; from this we can conclude that the fixed point iteration generated by the mapping $T_{\bar{\zeta}}$ is locally linearly convergent, from which we will be able to deduce local linear convergence of $T_{\mathrm{DR}\lambda}$.

PROPOSITION 4.8 (metric subregularity of $T_{\bar{\zeta}}$ by subtransversality). *Let $\lambda \in (0,1)$, $\overline{x} \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda}$ with $T_{\mathrm{DR}\lambda}$ being single-valued at $\overline{x}$, and set $g := P_B(\overline{x}) - P_A(P_B(\overline{x}))$. Furthermore, let $\bar{\zeta} \in \mathcal{Z}(\overline{x}, g)$ and $\overline{u} = (\overline{u}_1, \overline{u}_2, \overline{u}_3, \overline{u}_4) \in W_0(g)$ satisfy $\bar{\zeta} = \overline{u} - \Pi\overline{u}$ with $\overline{u}_1 = \overline{x}$. Let $T_{\bar{\zeta}}$ be defined by* (4.13)*, and define $\Phi_{\bar{\zeta}} := T_{\bar{\zeta}} - \mathrm{Id}$. Suppose the following hold:*

(i) *the collection of sets $\{B - \frac{\lambda}{1-\lambda}g, A - \frac{\lambda}{1-\lambda}g, A, B\}$ is subtransversal at $\overline{u}$ for $\bar{\zeta}$ relative to $\Lambda \subseteq W\left(\bar{\zeta}\right)$ with constant κ and neighborhood U of $\overline{u}$;*

(ii) *there exists a positive constant σ such that*

$$\operatorname{dist}\left(\bar{\zeta}, \Psi_g(u)\right) \le \sigma \operatorname{dist}\left(0, \Phi_{\bar{\zeta}}(u)\right) \quad \forall u \in \Lambda \cap U \text{ with } u_1 \in B - \tfrac{\lambda}{1-\lambda}g.$$

Then the mapping $\Phi_{\bar{\zeta}} := T_{\bar{\zeta}} - \mathrm{Id}$ is metrically subregular for 0 on U relative to $\Lambda \cap \mathcal{N}$ with constant $\bar{\kappa} = \kappa\sigma$, where $\mathcal{N} := \{z \in \mathcal{E}^4 \mid P_A(z_4 + \frac{\lambda}{1-\lambda}g) = z_3\}$.

Proof. This is an application of the assumptions and Lemma 4.6(ii):

$$\begin{aligned}&\left(\forall u \in U \cap \Lambda \cap \mathcal{N} \text{ with } u_1 \in B - \tfrac{\lambda}{1-\lambda}g\right)\\ &\qquad \operatorname{dist}\left(u, \Phi_{\bar{\zeta}}^{-1}(0) \cap \Lambda \cap \mathcal{N}\right) \le \operatorname{dist}\left(u, \Psi_g^{-1}(0) \cap \Lambda \cap \mathcal{N}\right)\\ &\qquad\qquad \le \kappa \operatorname{dist}\left(\bar{\zeta}, \Psi_g(u)\right)\\ &\qquad\qquad \le \kappa\sigma \operatorname{dist}\left(0, \Phi_{\bar{\zeta}}(u)\right);\end{aligned}$$

i.e., Φ is metrically subregular for 0 on U relative to $\Lambda \cap \mathcal{N}$ with constant $\bar{\kappa}$, as claimed. $\square$

By Theorem 4.2, Proposition 4.8, and Theorem 3.7 the *three* ingredients to get convergence are given by the regularity of the sets A and B, subtransversality of the collection of sets $\{A, B\}$, and the additional assumption (ii) in Proposition 4.8. As seen in [23, Proposition 3.5] this is also true for the alternating projection algorithm. If the intersection $A \cap B$ is nonempty, assuming the stronger property of transversality, superregularity is enough to show convergence of the Douglas–Rachford algorithm, [25, Theorem 6.8] and [14, Theorem 3.18]. For alternating projections one only needs transversality at points of intersection and superregularity of *one* of the sets [16, Theorem 5.16]. In any case, the additional assumption (ii) is not needed when the

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

assumptions on the fixed points are strong enough. This is also the case for consistent feasibility and the relaxed Douglas–Rachford method as seen next.

PROPOSITION 4.9 (intersecting sets). *As before, let $\lambda \in (0,1)$. Moreover, assume that the intersection of A and B is nonempty, i.e., $A \cap B \neq \emptyset$. Thus, for every $\overline{x} \in A \cap B \subset \mathsf{Fix}\, T_{\mathrm{DR}\lambda}$ we have $g := P_B(\overline{x}) - P_A(P_B(\overline{x})) = 0$. Furthermore, let $\bar{\zeta} \in \mathcal{Z}(\overline{x}, g)$. Then* (ii) *in Proposition* 4.8 *is always satisfied on $\Lambda \subset W(\bar{\zeta})$ with $\sigma = \frac{1}{\sqrt{2}\lambda}$.*

Proof. Since $\overline{x} \in A \cap B$ and $g = 0$, we get $\bar{\zeta} = (0,0,0,0)$. Moreover, note that for every $b \in B$ we gather $T_{\mathrm{DR}\lambda}(b) - b = \lambda(P_A(b) - b)$ since

$$\begin{aligned} T_{\mathrm{DR}\lambda}(b) - b &= \frac{\lambda}{2}\left(R_A R_B(b) + b\right) + (1-\lambda)P_B(b) - b \\ &= \frac{\lambda}{2}\left(R_A(b) + b\right) + (1-\lambda)b - b \\ &= \lambda P_A(b) - \lambda b \\ &= \lambda\left(P_A(b) - b\right). \end{aligned}$$

Therefore, we deduce for $u \in \Lambda \subset W(\bar{\zeta}) = \left\{u \in \mathcal{E}^4 \mid u_i = u_j,\ i,j \in \{1,2,3,4\}\right\}$ with $u_1 \in B$

$$\begin{aligned} T_0(u) - u &= (T_{\mathrm{DR}\lambda}(u_1) - u_1, T_{\mathrm{DR}\lambda}(u_1) - u_1, T_{\mathrm{DR}\lambda}(u_1) - u_1, T_{\mathrm{DR}\lambda}(u_1) - u_1) \\ &= \left(\lambda\left(P_A(u_1) - u_1\right), \lambda\left(P_A(u_1) - u_1\right), \lambda\left(P_A(u_1) - u_1\right), \lambda\left(P_A(u_1) - u_1\right)\right), \end{aligned}$$

and thus

$$\operatorname{dist}^2\left(0, \Phi_{\bar{\zeta}}(u)\right) = \operatorname{dist}^2\left(0, T_0(u) - u\right) = 4\operatorname{dist}^2\left(0, \lambda\left(P_A(u_1) - u_1\right)\right). \tag{4.14}$$

On the other hand,

$$\begin{aligned} \operatorname{dist}^2\left(\bar{\zeta}, \Psi_g(u)\right) &= \operatorname{dist}^2\left((0,0,0,0), \Psi_0(u)\right) \\ &= \operatorname{dist}^2\left((0,0,0,0), P_{\Omega_0}\Pi(u) - \Pi(u)\right) \\ &= 2\operatorname{dist}^2\left(0, P_A(u_1) - u_1\right) \end{aligned} \tag{4.15}$$

since $\Omega_0 = B \times A \times A \times B$. Combining (4.15) and (4.14) yields (ii) in Proposition 4.8 with $\sigma = \frac{1}{\sqrt{2}\lambda}$. ∎

4.2. Local linear convergence of $T_{\mathrm{DR}\lambda}$.

LEMMA 4.10 (uniqueness of difference vector for fixed points of $T_{\mathrm{DR}\lambda}$). *Let $\lambda \in (0,1)$, and let $\overline{x}$ be a point in $\mathsf{Fix}\, T_{\mathrm{DR}\lambda}$, where $A, B \subset \mathcal{E}$ satisfy the assumptions of Theorem* 3.7 *with neighborhoods $U(A, \epsilon, \overline{x})$ and $U(B, \epsilon, \overline{x})$. Then $\{\bar{\zeta}\} = \mathcal{Z}(\overline{x}, g) \subset \mathcal{E}^4$ for $\{g\} = P_B(\overline{x}) - P_A(P_B(\overline{x}))$ is unique and given by*

$$\bar{\zeta} = (\bar{\zeta}_1, \ldots, \bar{\zeta}_4) = \left(g, -\tfrac{\lambda}{1-\lambda}g, -g, \tfrac{\lambda}{1-\lambda}g\right).$$

Proof. By definition (4.11), $\mathcal{Z}(x,g)$ is given by

$$\mathcal{Z}(x,g) := \left\{\zeta := z - \Pi z \;\middle|\; z \in W_0(g) \subset \mathcal{E}^4,\ z_1 = x\right\}$$

for

$$\begin{aligned} &W_0(g) \\ &:= \left\{u \in \mathcal{E}^4 \;\middle|\; u_1 \in P_{B - \frac{\lambda}{1-\lambda}g}(u_2),\ u_2 \in P_{A - \frac{\lambda}{1-\lambda}g}(u_3),\ u_3 \in P_A(u_4),\ u_4 \in P_B(u_1)\right\}. \end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Thus, the uniqueness of $\bar{\zeta}$ is a direct implication of the uniqueness of g as seen in Remark 3.14(iii). ∎

Now we are ready to present the main result. The proof is based on the basic convergence result Theorem 4.2 and the facts from sections 3 and 4.

THEOREM 4.11 (local linear convergence of $T_{\mathrm{DR}\lambda}$). *Let $\lambda \in (0,1)$, and let $\overline{x}$ be a point in $\mathsf{Fix}\, T_{\mathrm{DR}\lambda}$, where $A, B \subset \mathcal{E}$ satisfy the assumptions of Theorem 3.7 with neighborhoods $U(A,\epsilon,\overline{x})$ and $U(B,\epsilon,\overline{x})$. Set $\{g\} = P_B(\overline{x}) - P_A(P_B(\overline{x}))$ and $\{\bar{\zeta}\} = \mathcal{Z}(\overline{x}, g)$ ($\bar{\zeta} = (\bar{\zeta}_1, \ldots, \bar{\zeta}_4) \in \mathcal{E}^4$). Suppose that, at all $x \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda}$ with $g \in P_B(x) - P_A P_B(x)$, the sets $A, B \subset \mathcal{E}$ satisfy the assumptions of Theorem 3.7 with corresponding neighborhoods $U(A,\epsilon,x)$ and $U(B,\epsilon,x)$. Define the set*

$$S_0 := \{x \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda} \mid \{g\} = P_B(x) - P_A(P_B(x))\}\,, \tag{4.16}$$

and let

$$S_j := \left(S_0 - \sum_{i=1}^{j-1} \bar{\zeta}_i\right) \quad (j = 1,2,3,4)\,. \tag{4.17}$$

Fix some $\epsilon > 0$, and define the neighborhood $U_A := \cup_{x \in S_0} U(A,\epsilon,x)$ and likewise $U_B := \cup_{x \in S_0} U(B,\epsilon,x)$. Then

$$U := \left(U_B - \tfrac{\lambda}{1-\lambda} g\right) \times \left(U_A - \tfrac{\lambda}{1-\lambda} g\right) \times U_A \times U_B$$

is a neighborhood of $S := S_1 \times S_2 \times S_3 \times S_4$. Suppose that, for $\Lambda \subseteq W(\bar{\zeta})$ satisfying $S \subset \Lambda$ with $T_{\bar{\zeta}} : \Lambda \rightrightarrows \Lambda$, the following hold for all $\overline{u} = (\overline{u}_1, \overline{u}_2, \overline{u}_3, \overline{u}_4) \in S$:

(i) *for all $(\overline{u}_3, \overline{u}_4) \in S_3 \times S_4$, the collection of sets $\{A, B\}$ is subtransversal at $(\overline{u}_3, \overline{u}_4)$ for $(\overline{u}_3, \overline{u}_4) - \Pi(\overline{u}_3, \overline{u}_4)$ relative to*

$$\Lambda' := \left\{u = (u_1, u_2) \in \mathcal{E}^2 \;\middle|\; \left(u_2 - \tfrac{\lambda}{1-\lambda} g, u_1 - \tfrac{\lambda}{1-\lambda} g, u_1, u_2\right) \in \Lambda\right\}$$

with constant κ on the neighborhood $U_A \times U_B$;

(ii) *for $\Phi_{\bar{\zeta}} := T_{\bar{\zeta}} - \mathrm{Id}$ and $\Psi_g := P_{\Omega_g}\Pi - \Pi$ there exists a positive constant σ such that*

$$\operatorname{dist}\left(\bar{\zeta}, \Psi_g(u)\right) \le \sigma \operatorname{dist}\left(0, \Phi_{\bar{\zeta}}(u)\right) \tag{4.18}$$

holds whenever $u \in \tilde{\Lambda} \cap U$ with $u_1 \in B - \frac{\lambda}{1-\lambda} g$ and

$$\tilde{\Lambda} := \left\{u \in \Lambda \;\middle|\; u = \left(x_2 - \tfrac{\lambda}{1-\lambda} g, x_1 - \tfrac{\lambda}{1-\lambda} g, x_1, x_2\right) \text{ for some } x_1, x_2 \in \mathcal{E}\right\}.$$

Then there exists an $\epsilon' \le \epsilon$ and a neighborhood $U' \subset U$ ($U' = U_1' \times U_2' \times U_3' \times U_4' \subset \mathcal{E}^4$) of S on which the sequence $\left(u^k\right)_{k \in \mathbb{N}}$ generated by $u^{k+1} \in T_{\bar{\zeta}} u^k$ seeded by a point $u^0 \in W\left(\bar{\zeta}\right) \cap U'$ with $u_1^0 \in U_1' \cap (B - \frac{\lambda}{1-\lambda} g)$ satisfies

$$\operatorname{dist}\left(u^{k+1}, \mathsf{Fix}\, T_{\bar{\zeta}} \cap S\right) \le c \operatorname{dist}\left(u^k, S\right) \quad (\forall k \in \mathbb{N})$$

for

$$c := \sqrt{1 + \epsilon' - \frac{1}{\bar{\kappa}^2}} < 1, \tag{4.19}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

where $\bar{\kappa} = \kappa\sigma$ *with* κ *and* σ *given by* (i) *and* (ii). *Consequently,* $\operatorname{dist}\left(u^k, \tilde{u}\right) \to 0$ *for some* $\tilde{u} \in \mathsf{Fix}\, T_{\bar{\zeta}} \cap S$, *and hence*

$$\operatorname{dist}\left(u_1^k, \tilde{u}_1\right) \to 0$$

at least R-linearly with rate $c < 1$. *If* $\mathsf{Fix}\, T_{\mathrm{DR}\lambda} \cap S_1$ *is a singleton, then convergence is Q-linear.*

Remark 4.12 (atlas for the assumptions). At first sight the assumptions in Theorem 4.11 might seem overwhelming. To provide some insight into the statement we discuss the most important parts of the setting:

1. The assumptions of Theorem 3.7 are needed to conclude almost averagedness of $T_{\mathrm{DR}\lambda}$.
2. The requirement that the assumptions of Theorem 3.7 hold at all fixed points with the same gap vector is achieved by restricting our analysis to the set S_0. This also implies that we are considering only fixed points that are *isolated relative to* Λ.
3. Although we were not able to prove metric subregularity for a mapping related to $T_{\mathrm{DR}\lambda}$ directly, we can show this property for $T_{\bar{\zeta}}$ on $\mathcal{E}^4$. In particular, assumptions (i) and (ii) are used to guarantee metric subregularity from Proposition 4.8. Assumption (i) guarantees subtransversality of the collection $\{B - \frac{\lambda}{1-\lambda}g, A - \frac{\lambda}{1-\lambda}g, A, B\}$ since we have seen in Lemma 4.4 that subtransversality is preserved under the addition of some constant vector, here $\frac{\lambda}{1-\lambda}g$.
4. The definitions of Λ' and $\tilde{\Lambda}$ relate to the construction of the lifted product space version of the problem.
5. The violation ϵ depends on the violations in Definition 2.4 as seen in Theorem 3.7. Thus, fixing some violation ϵ corresponds to certain choices of neighborhoods $U(A, \epsilon, \overline{x})$ and $U(B, \epsilon, \overline{x})$ and violations ϵ_A and ϵ_B of (2.3) for the sets A and B, respectively.

Proof of Theorem 4.11. First, note that U is a neighborhood of S since $U_A \times U_B$ is a neighborhood of $S_3 \times S_4$ since for every $(u, \tilde{u}) \in S_3 \times S_4$ there exist $x, \tilde{x} \in S_0$ such that $U(A, \epsilon, x) \times U(B, \epsilon, \tilde{x}) \subset U_A \times U_B$ is a neighborhood of $(u, \tilde{u})$.

The neighborhood U can be replaced by an enlargement of S; hence, the result follows from Theorem 4.2 once it can be shown that the assumptions are satisfied for the mapping $T_{\bar{\zeta}}$ on the product space $\mathcal{E}^4$ restricted to $\tilde{\Lambda}$.

To do so, we note that $T_{\mathrm{DR}\lambda}$ is almost firmly nonexpansive at each $\tilde{y} \in S_1$ on U_B by Theorem 3.7 since assumptions (i)–(iv) of Theorem 3.7 are satisfied. Thus, $\alpha = 1/2$. Likewise, the violation is given by ϵ on U_B. Since $T_{\bar{\zeta}}$ is just $T_{\mathrm{DR}\lambda}$ shifted by $\bar{\zeta}$ on the product space, it follows that $T_{\bar{\zeta}}$ is pointwise almost averaged at $y \in S := S_1 \times S_2 \times S_3 \times S_4$ with the same violation ϵ and averaging constant $\alpha = 1/2$ on U.

By Lemma 4.4 and Remark 4.5, therefore, assumption (i) implies that for $\overline{u} = (\overline{u}_1, \overline{u}_2, \overline{u}_3, \overline{u}_4) \in S$, the collection of sets

$$\left\{B - \frac{\lambda}{1-\lambda}g, A - \frac{\lambda}{1-\lambda}g, A, B\right\}$$

is subtransversal at $\overline{u}$ for $\bar{\zeta} := \overline{u} - \Pi\overline{u}$ relative to $\tilde{\Lambda}$ with constant κ on the neighborhood U, hence, Theorem 4.2(i) is satisfied. Moreover, assumption (ii) of Theorem 4.11 and Proposition 4.8 with $U \cap \tilde{\Lambda} \subset \mathcal{N} := \{z \in \mathcal{E}^4 \mid P_A(z_4 + \frac{\lambda}{1-\lambda}g) = z_3\}$ by Theorem 3.7(iv)

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

yield assumption Theorem 4.2(ii). In total, the assumptions of Theorem 4.2 are all satisfied for $T_{\bar{\zeta}}$ on $\mathcal{E}^4$ restricted to $\tilde{\Lambda}$, and thus we conclude that (4.2) holds.

What remains is to show that (4.3) holds, which would imply at least R-linear convergence. To achieve this, choose some $\epsilon' > 0$ with $\epsilon' < \epsilon$ such that (4.3) is satisfied. By Corollary 3.8 we can always find neighborhoods $U(B, \epsilon', x) \subset U(B, \epsilon, x)$ and $U(A, \epsilon', x) \subset U(A, \epsilon, x)$ for all $x \in S_0$ that satisfy the assumptions of Theorem 3.7. Following the constructions above we define $U'_A := \cup_{x \in S_0} U(A, \epsilon', x)$ and $U'_B := \cup_{x \in S_0} U(B, \epsilon', x)$ and get $U'_A \subset U_A$ as well as $U'_B \subset U_B$. Thus, all the properties that we have shown to be true on U also hold on the subset U' defined by

$$U' := \left(U'_B - \tfrac{\lambda}{1-\lambda} g\right) \times \left(U'_A - \tfrac{\lambda}{1-\lambda} g\right) \times U'_A \times U'_B.$$

In particular, the constants κ and σ in (i) and (ii) also suffice for the smaller neighborhoods $U'_A \times U'_B$ and U'. As a consequence, the assumptions of Theorem 4.2 are all satisfied, and (4.3) holds, which implies at least R-linear convergence to $\tilde{u}$. Since $\tilde{u}_1 \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda} \cap S_1$, this completes the proof. ∎

Remark 4.13 (a closer look at the convergence statement). The gap vector g and difference vector $\bar{\zeta}$ in Theorem 4.11 rely on the structure of the intersection of the sets A and B. The consistent case, that is, $A \cap B \neq \emptyset$, leads to a simplification of the problem. Here, the gap is 0. Similarly, the related difference vector is of the form $\bar{\zeta} = \{0, 0, 0, 0\}$. Hence, the assumptions which involve at least one of these vectors can be simplified. When the intersection $A \cap B$ is empty, namely, the inconsistent case, the value of $\bar{\zeta}$ is dependent on the choice of λ. We distinguish three important cases:

1. $\lambda = \frac{1}{2}$. Here $\frac{\lambda}{1-\lambda}$ reduces to 1. As a result, the phantom sets are shifted by the entire gap g such that A and $B - g$ have a common point. The difference vector is of the form $\bar{\zeta} = \{g, -g, -g, g\}$.
2. $\lambda \to 1$. Then $\frac{\lambda}{1-\lambda} \to +\infty$. That is, the phantom sets recede to the horizon in the direction $-g$.
3. $\lambda \to 0$. In this case $\frac{\lambda}{1-\lambda}$ converges to 0, and the phantom sets coincide in the limit with the original ones. So, $\Omega_{u,g} = B \times A \times A \times B$. Cyclic projections on these sets $\{B, A, A, B\}$ in the given order are nothing more than alternating projections between the sets A and B. At $\lambda = 0$, however, $\mathsf{Fix}\, T_{\mathrm{DR}\lambda} = B$, which is clearly larger than the fixed point set for alternating projections.

Although the individual assumptions can be challenging to prove, as we will see in section 5, they can reduce to a simpler form if we consider a convex and consistent setting. The reason for this is twofold. First, subtransversality at points in the intersection is nothing more than *local linear regularity* of the collection of sets [23, Proposition 3.3]. Second, it was shown that local linear regularity is equivalent to the global property of linear regularity in the setting of closed convex sets, as seen in [2, Theorem 6.1]. Thus, assuming (i) locally for closed and convex sets implies that this property holds globally. To prove this statement we will first present the auxiliary statements, which are essential to show the global convergence result.

PROPOSITION 4.14 (subtransversality at common points] [23, Proposition 3.3]). *Let $\mathcal{E}^m$ be endowed with the 2-norm, that is, $\|(x_1, \ldots, x_m)\|_2 = (\sum_{j=1}^m \|x_j\|_{\mathcal{E}}^2)^{1/2}$. A collection $\{\Omega_1, \Omega_2, \ldots, \Omega_m\}$ of nonempty and closed subsets of $\mathcal{E}$ is subtransversal relative to*

$$\Lambda := \{x = (u, u, \ldots, u) \in \mathcal{E}^m \mid |u \in \mathcal{E}\}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

at $\overline{x} = (\overline{u}, \overline{u}, \ldots, \overline{u})$ *with* $\overline{u} \in \cap_{j=1}^m \Omega_j$ *for* $\overline{y} = 0$ *with constant* κ *if there exist a neighborhood* U' *of* $\overline{u}$ *together with a constant* κ' *satisfying* $\sqrt{m}\kappa' \leq \kappa$ *such that*

$$\operatorname{dist}\left(u, \cap_{j=1}^m \Omega_j\right) \leq \kappa' \max_{j=1,\ldots,m} \operatorname{dist}\left(u, \Omega_i\right) \quad \forall\, u \in U'. \tag{4.20}$$

Conversely, if $\{\Omega_1, \Omega_2, \ldots, \Omega_m\}$ *is subtransversal relative to* Λ *at* $\overline{x}$ *for* $\overline{y} = 0$ *with constant* κ*, then* (4.20) *is satisfied with any constant* κ' *for which* $\kappa \leq \kappa'$.

The property in (4.20) is called *local linear regularity*. If the inequality holds for all $u \in \mathcal{E}$, the collection $\{A, B\}$ is said to be *linearly regular*. Bakan, Deutsch, and Li showed in [2] the equivalence of both properties when the sets are closed and convex.

LEMMA 4.15 ([2, Theorem 6.1]). *Let the sets* A *and* B *be nonempty closed convex sets with nonempty intersection, i.e.,* $A \cap B \neq \emptyset$*. Then the following are equivalent:*

(i) *There is a* $\delta > 0$ *such that the collection of sets is locally linearly regular at* $\overline{x} \in A \cap B$ *on* $\mathbb{B}_\delta(\overline{x})$.

(ii) *The collection of sets is linearly regular at* $\overline{x} \in A \cap B$.

Having Proposition 4.14 and Lemma 4.15 we are now ready to state a global convergence result for closed convex sets.

COROLLARY 4.16 (global convergence in the consistent and convex setting). *Let* $\lambda \in (0, 1)$*, and let* $\overline{x}$ *be a point in* $\mathsf{Fix}\, T_{\mathrm{DR}\lambda}$*. Moreover, let both* A *and* B *be closed and convex with nonempty intersection, i.e.,* $A \cap B \neq \emptyset$*, and therefore* $\mathsf{Fix}\, T_{\mathrm{DR}\lambda} = A \cap B$*. Then* $\{g\} = P_B(\overline{x}) - P_A(P_B(\overline{x})) = 0$ *and* $\left\{\bar{\zeta}\right\} = \mathcal{Z}(\overline{x}, g) = \{0\}$ *(*$\bar{\zeta} = \left(\bar{\zeta}_1, \bar{\zeta}_2, \bar{\zeta}_3, \bar{\zeta}_4\right) \in \mathcal{E}^4$*). Define the set*

$$S_0 := \mathsf{Fix}\, T_{\mathrm{DR}\lambda} = A \cap B.$$

Suppose that the following hold for all $\overline{u} = (\overline{u}_1, \overline{u}_2) \in S := S_0 \times S_0$*:*

(i) *the collection of sets* $\{A, B\}$ *is subtransversal at* $(\overline{u}_1, \overline{u}_2)$ *for* $(\overline{u}_1, \overline{u}_2) - \Pi(\overline{u}_1, \overline{u}_2)$ *relative to* $\Lambda' \subset \left\{u \in \mathcal{E}^2 \mid u_1 = u_2\right\}$ *with constant* κ *on some neighborhood* $U' \subset \mathcal{E}^2$ *(*$U' = U_A \times U_B$*).*

Then the sequence $\left(x^k\right)_{k \in \mathbb{N}}$ *generated by* $x^{k+1} \in T_{\mathrm{DR}\lambda}(x^k)$ *seeded by a point* $x^0 \in \Lambda' \cap U_B$ *satisfies*

$$\operatorname{dist}\left(x^{k+1}, \mathsf{Fix}\, T_{\mathrm{DR}\lambda}\right) \leq c \operatorname{dist}\left(x^k, \mathsf{Fix}\, T_{\mathrm{DR}\lambda}\right) \quad (\forall k \in \mathbb{N})$$

for

$$c := \sqrt{1 - \frac{2\lambda^2}{\kappa^2}} < 1.$$

Consequently, $\operatorname{dist}\left(x^k, \tilde{x}\right) \to 0$ *for some* $\tilde{x} \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda}$ *at least R-linearly with rate* $c < 1$*. If* $\mathsf{Fix}\, T_{\mathrm{DR}\lambda}$ *is a singleton, then convergence is Q-linear.*

Remark 4.17 (global convergence for convex sets). There are only two changes from Theorem 4.11 to Corollary 4.16. First, the sets are required to be convex. Thus, convergence in general is guaranteed by Opial [24] as noted in the beginning of this section. Moreover, the local assumption (i) in this case is a global one, i.e., $U' = \mathcal{E}^2$, by Proposition 4.14 and Lemma 4.15. The second difference, assumption (ii) in Theorem 4.11, is always satisfied by Proposition 4.9 since $\mathsf{Fix}\, T_{\mathrm{DR}\lambda} = A \cap B$.

Proof of Corollary 4.16. Since both A and B are convex, not only the difference vector is unique, as seen in Lemma 4.10, but so is the gap vector g for any fixed point

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

in $\mathsf{Fix}\, T_{\mathrm{DR}\lambda}$. Thus, $S_0 = \mathsf{Fix}\, T_{\mathrm{DR}\lambda}$. $\mathsf{Fix}\, T_{\mathrm{DR}\lambda} = A \cap B$ by Theorem 3.13. With these observations we get immediately that the sets involved in Theorem 4.11 simplify to

$$\begin{aligned} S &= S_0 \times S_0 \times S_0 \times S_0, \\ W(\bar{\zeta}) &= \left\{ u \in \mathcal{E}^4 \mid u - \Pi u = 0 \right\} = \left\{ u \in \mathcal{E}^4 \mid u_1 = u_2 = u_3 = u_4 \right\}, \\ U &= U_B \times U_A \times U_A \times U_B, \\ \Lambda' &\subset \left\{ u \in \mathcal{E}^2 \mid u_1 = u_2 \right\} \end{aligned}$$

since $\Lambda \subset W(\bar{\zeta})$. Thus, assuming (i) in Corollary 4.16 is equivalent to assuming Theorem 4.11(i) in the convex and consistent setting. Moreover, since the sets A and B are convex, both the projector and the reflector onto these sets are single-valued (see, for example, [10, Theorem 2.4]). Additionally, the projection is firmly nonexpansive [10, Theorem 5.5], and thus the reflector is nonexpansive (see, for example, [4]), which implies that $T_{\mathrm{DR}\lambda}$ is averaged with constant $\alpha = (1/2)$. The conditions of Theorem 3.7 are therefore satisfied with neighborhoods chosen to be $\mathcal{E}$. Also, since the sets A and B are convex, they are superregular at a distance by Proposition 2.6 with $\epsilon = 0$. Since every fixed point is an element of the intersection $A \cap B$, we deduce by Proposition 4.9 that assumption (ii) of Theorem 4.11 holds. The local convergence result follows then from Theorem 4.11. What is left to show is the global convergence property.

By (i) and Proposition 4.14 the collection of sets $\{A, B\}$ is locally linearly regular on U'. Thus, there exists a $\delta > 0$ such that $\{A, B\}$ is locally linearly regular on $\mathbb{B}_\delta(\overline{x})$. Using Lemma 4.15 we get that $\{A, B\}$ is linearly regular since A and B are convex sets. In total, (i) holds with $U' = \mathcal{E}^2$. That is, the assumption holds globally. Since (ii) of Theorem 4.11 holds globally as well by Proposition 4.9, the assumptions of the underlying convergence framework in Theorem 4.2 hold globally. Therefore, the sequence converges globally, which completes the proof. ∎

Remark 4.18 (linking our results to already existing literature). As noted in the introduction, the works [14, 25, 17, 23, 9] all analyze the Douglas–Rachford algorithm for consistent nonconvex feasibility. In [23] the framework used here was applied to Douglas–Rachford for structured nonconvex optimization. In [9] the authors showed local R-linear convergence for superregular sets intersecting linear regular. Our analysis of relaxed Douglas–Rachford includes or subsumes that of all previous studies in the context of set feasibility, with the exception of [17], which addresses global convergence guarantees for consistent feasibility. The assumptions of that paper, namely, compactness and semialgebraicity (not to mention nonempty intersection), are different than the notions that we work with. Certainly, compactness is a regularity assumption, as is semialgebraicity or its more general Kurdyka–Łojasiewicz type of regularity, but these notions serve a different purpose. Indeed, even convex sets need not be semialgebraic or compact. This suggests that the Kurdyka–Łojasiewicz type of regularity and compactness could be properties *in addition* to the ones we use in order to arrive at global statements. Nevertheless, as shown in Corollary 4.16, in the convex case, the local analysis suffices to infer global convergence properties. A more thorough study of the relationship between the different notions of regularity would be fruitful but is beyond the scope of our paper.

Our results could be extended to sets with even weaker regularity, namely, ϵ-subregular sets instead of superregular sets at a distance under the additional assumption that suitable neighborhoods exist. But the present setting is technical enough–increased generality would have only made the details even more difficult to parse. Moreover, the advantage of this specific type of nonconvexity is that we are able

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

not only to present existence results on neighborhoods where we get local convergence but also to construct the neighborhoods explicitly.

5. Elementary examples. We demonstrate in this section explicit verification of the assumptions of Theorem 4.11 for a typical class of problems. In particular, we consider the configurations that arise with feasibility problems involving intersecting and nonintersecting *spheres*. This is of particular interest for the *source localization problem* and the *phase retrieval problem*, especially the nonintersecting case. The idealized source localization problem amounts to finding the intersection of spheres that are determined by distance measurements to receivers whose locations are known. When the distance measurements are noisy or the given locations of the receivers are inaccurate, the intersection over all spheres will be empty almost always. For phase retrieval, the measurements are pointwise amplitude measurements in the Fourier domain of an unknown object. In other words, the constraint sets are two-dimensional spheres in the image of a linear transformation. Since both the measurements and the object are assumed to have compact support, the phase retrieval problem in diffraction imaging is fundamentally inconsistent. In our development below, we carry out the explicit calculations to verify the assumptions of Theorem 4.11 for circles (spheres in $\mathbb{R}^2$) which was shown in [21] to be the fundamental geometry for phase retrieval and source location problems. Affine subspaces (i.e., lines) are included as spheres with infinite radius and center at infinity.

There are five distinct cases to consider: (1) intersecting circles; (2) nonintersecting separable circles; (3) nonintersecting, nonseparable, nonconcentric circles; (4) nonintersecting, nonseparable, concentric circles; and (5) tangential circles. We show that the verification can be carried out "by hand" in the first example. For the sake of brevity, the verification is carried out in the remaining examples with the help of symbolic computation. We were unable to prove or disprove that the required conditions hold in Example 4. In Example 5 we determine that the assumptions are *not* satisfied, and therefore the algorithm cannot converge linearly. The symbolic worksheets where our calculations were carried out are available at http://vaopt.math.uni-goettingen.de/en/publications.php.

For this entire section let R be a positive real-valued number and $\lambda \in (0,1)$ if not specified. To verify the subtransversality and the technical condition (ii) in Theorem 4.11 we often did not calculate the constants explicitly but bounded them from below. That is,

$$\kappa > \frac{\operatorname{dist}\left(u, \Psi_g^{-1}(\bar{\zeta}) \cap W(\bar{\zeta})\right)}{\operatorname{dist}\left(\bar{\zeta}, \Psi_g(u)\right)},$$

$$\sigma > \frac{\operatorname{dist}\left(\bar{\zeta}, \Psi_g(u)\right)}{\operatorname{dist}\left(0, \Phi_{\bar{\zeta}}(u)\right)},$$

where κ was the constants of subtransversality and σ describes the technical condition. In this subsection we will deal with neighborhoods of fixed points. As a consequence the constants computed bound the rate of linear convergence from below in such cases. Note that we can always find a neighborhood such that the convergence is linear for examples consisting of two circles by Theorem 3.7 and Example 3.9.

Example 5.1 (two intersecting circles). The first example consists of two circles intersecting at exactly two points shown in Figure 6. Without loss of generality we

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

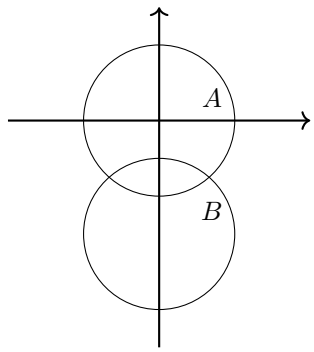

FIG. 6. *Example* 5.1 *for* $a = -1.5$ *and* $R = 1$.

can restrict the analysis to the following setting:

$$A := \left\{ x \in \mathbb{R}^2 \mid \|x\| = 1 \right\},$$
$$B := \left\{ x \in \mathbb{R}^2 \mid \|x - (0,a)\| = R \right\},$$

where $a \in \mathbb{R} \setminus \{0\}$ and $R \in (\min_{y \in A} \operatorname{dist}((0,a), y), \max_{y \in A} \operatorname{dist}((0,a), y))$. Note that the endpoints of the interval for R correspond to the setting of two touching circles; see Example 5.5.

First, we consider the points in the intersection $A \cap B$, namely,

$$\left(\pm\sqrt{1 - \left(\frac{1 - R^2 + a^2}{2a}\right)^2}, \frac{1 - R^2 + a^2}{2a} \right).$$

Due to the symmetry of the problem we restrict the analysis to the point

$$\left(+\sqrt{1 - \left(\frac{1 - R^2 + a^2}{2a}\right)^2}, \frac{1 - R^2 + a^2}{2a} \right).$$

The following statements regarding the assumptions made in Theorem 4.11 are easily verified either by hand or with the help of symbolic computation:

(i) $S_0 := \left\{ \left(+\sqrt{1 - \left(\frac{1-R^2+a^2}{2a}\right)^2}, \frac{1-R^2+a^2}{2a} \right) \right\} \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda}$.

(ii) In $\mathbb{R}_+ \times \mathbb{R}$ there is a unique fixed point

$$\overline{x} = (\overline{u}, \overline{u}, \overline{u}, \overline{u}), \text{ where } \overline{u} = \left(+\sqrt{1 - \left(\frac{1 - R^2 + a^2}{2a}\right)^2}, \frac{1 - R^2 + a^2}{2a} \right).$$

(iii) The difference vector is unique and given by $\bar{\zeta} = ((0,0),(0,0),(0,0),(0,0))$ since $\overline{u} \in A \cap B$.

(iv) The sets A and B satisfy the assumptions of Theorem 3.7 at $\overline{u}$ with neighborhoods U_1 and U_2 being open balls around $\overline{u}$, that is, $\mathbb{B}_\delta(\overline{u})$, for $\delta \in (0,1)$. This can be shown similar to Example 3.9.

(v) This example considers a setting with nonempty intersection. By Proposition 4.14 one can equivalently prove linear regularity to get subtransversality

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

in such instances. Our aim is to use this statement to prove that Example 5.1 satisfies (4.20).

For this we select any $u = (u_1, u_2) \in U_1 \cap A$, where $U_1 = U_2$ and $u_1 > 0$. Such a point exists since the statements in Theorem 4.11 are all with respect to the set Λ, which is a subset of $W(\bar{\zeta})$. Thus, the restriction to one of the sets is no contradiction. The condition $u_1 > 0$ ensures that we always project on the chosen point in the intersection:

$$\left(+\sqrt{1-\left(\frac{1-R^2+a^2}{2a}\right)^2}, \frac{1-R^2+a^2}{2a}\right).$$

Then the condition

$$\operatorname{dist}(u, A\cap B) \le \kappa' \max\{\operatorname{dist}(u,A), \operatorname{dist}(u,B)\}$$

simplifies to

$$\operatorname{dist}(u, A\cap B) \le \kappa' \operatorname{dist}(u,B),$$

which we reformulate in the following to

$$\|u - P_{A\cap B}(u)\| \le \kappa' \|u - P_B(u)\|. \tag{5.1}$$

Note that (5.1) implies (4.20) since $u \in A$ and thus implies linear regularity. Next, we show (5.1):

$$\begin{aligned}\|u - P_{A\cap B}(u)\| &\le \|u - P_B(u)\| + \|P_B(u) - P_{A\cap B}(u)\| \\ &\le \|u - P_B(u)\| \left(1 + \frac{\|P_B(u) - P_{A\cap B}(u)\|}{\|u - P_B(u)\|}\right).\end{aligned}$$

What remains to show is that $1 + \frac{\|P_B(u)-P_{A\cap B}(u)\|}{\|u-P_B(u)\|}$ is bounded above by a nonnegative constant.

By construction we get for the individual projections

$$P_{A\cap B}(u) = \left(+\sqrt{1-\left(\frac{1-R^2+a^2}{2a}\right)^2}, \frac{1-R^2+a^2}{2a}\right),$$

$$P_B(u) = (0,a) + \frac{u-(0,a)}{\|u-(0,a)\|}R.$$

Inserting this into the above expression yields

$$\begin{aligned}1 + \frac{\|P_B(u) - P_{A\cap B}(u)\|}{\|u - P_B(u)\|} &= 1 + \frac{\left\|(0,a) + \frac{u-(0,a)}{\|u-(0,a)\|}R - \left(\sqrt{1-\left(\frac{1-R^2+a^2}{2a}\right)^2}, \frac{1-R^2+a^2}{2a}\right)\right\|}{\left\|u-(0,a)+\frac{u-(0,a)}{\|u-(0,a)\|}R\right\|} \\ &\le 2 + \frac{\left\|u - \left(\sqrt{1-\left(\frac{1-R^2+a^2}{2a}\right)^2}, \frac{1-R^2+a^2}{2a}\right)\right\|}{\left\|u-(0,a)+\frac{u-(0,a)}{\|u-(0,a)\|}R\right\|} \\ &< 2 + \frac{1}{\left\|u-(0,a)+\frac{u-(0,a)}{\|u-(0,a)\|}R\right\|}\end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

since $u \in A$ and thus

$$\left\| u - \left(\sqrt{1 - \left(\frac{1 - R^2 + a^2}{2a} \right)^2}, \frac{1 - R^2 + a^2}{2a} \right) \right\| < 1.$$

The denominator

$$\left\| u - (0,a) + \frac{u - (0,a)}{\|u - (0,a)\|} R \right\|$$

can be bounded as follows. Since $u \in A$ we have

$$0 < \min_{y \in A} \operatorname{dist}\left((0,a), y\right) \leq \|u - (0,a)\| \leq \max_{y \in A} \operatorname{dist}\left((0,a), y\right).$$

Equivalently,

$$\frac{1}{\min_{y \in A} \operatorname{dist}\left((0,a), y\right)} \geq \frac{1}{\|u - (0,a)\|} \geq \frac{1}{\max_{y \in A} \operatorname{dist}\left((0,a), y\right)}.$$

Thus,

$$1 - \frac{R}{\|u - (0,a)\|} \geq 1 - \frac{R}{\max_{y \in A} \operatorname{dist}\left((0,a), y\right)}$$

and

$$\|u - (0,a)\| \left(1 - \frac{R}{\|u - (0,a)\|} \right) \geq \min_{y \in A} \operatorname{dist}\left((0,a), y\right) \left(1 - \frac{R}{\max_{y \in A} \operatorname{dist}\left((0,a), y\right)} \right),$$

which implies that

$$\frac{1}{\|u - (0,a)\| \left(1 - \frac{R}{\|u - (0,a)\|} \right)} \leq \frac{1}{\min_{y \in A} \operatorname{dist}\left((0,a), y\right) \left(1 - \frac{R}{\max_{y \in A} \operatorname{dist}((0,a), y)} \right)} =: \kappa'.$$

Since $R \in (\min_{y \in A} \operatorname{dist}\left((0,a), y\right), \max_{y \in A} \operatorname{dist}\left((0,a), y\right))$, κ' is bounded above. In total, $A \cap B$ is locally linear regular at

$$\left(+ \sqrt{1 - \left(\frac{1 - R^2 + a^2}{2a} \right)^2}, \frac{1 - R^2 + a^2}{2a} \right).$$

By Proposition 4.14 we deduce subtransversality with constant $\kappa := \kappa' \sqrt{2}$.

(vi) The technical condition (ii) in Theorem 4.11 is satisfied with

$$\sigma^2 = \frac{1}{2\lambda^2}$$

by Proposition 4.9.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

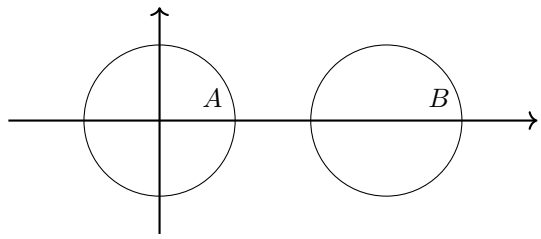

FIG. 7. *Example* 5.2 *for* $R = 1$.

Thus, the assumptions of Theorem 4.11 are satisfied, and the relaxed Douglas–Rachford algorithm converges locally linearly to $\overline{u}$ with rate $1 > c > \sqrt{1 - \frac{\lambda^2}{\kappa'^2}}$ as long as the starting point is close enough to $\overline{u}$.

Similarly, this argument can be repeated for

$$\left(-\sqrt{1 - \left(\frac{1 - R^2 + a^2}{2a}\right)^2}, \frac{1 - R^2 + a^2}{2a}\right),$$

which shows that, in this situation, both subtransversality and the technical condition at the two points in the intersection are satisfied.

Note that the point $(0, -1)$ does not to lead to a fixed point of the relaxed Douglas–Rachford algorithm, whereas for the alternating projection method, defined by the operator $P_A P_B$, $(0, -1)$ is always a fixed point. In particular, for any $\lambda \in (0, 1)$ the fixed point set of $T_{\mathrm{DR}\lambda}$ does not contain any point of the form $(0, y)$ for $y \in \mathbb{R}$.

Example 5.2 (nonintersecting separable circles). This example, illustrated in Figure 7, consists of two circles in $\mathbb{R}^2$ that are shifted by some vector in $\mathbb{R}^2$ such that they do not intersect in any point. Let $R > 0$, and define

$$\begin{aligned} A &:= \left\{x \in \mathbb{R}^2 \mid \|x\| = 1\right\} \\ \text{and} \quad B &:= \left\{x \in \mathbb{R}^2 \mid \|x - (2 + R, 0)\| = R\right\}. \end{aligned}$$

The only fixed point of $T_{\mathrm{DR}\lambda}$ on A and B is given by

$$\overline{u} = (2, 0) - \frac{\lambda}{1 - \lambda}(1, 0)$$

for $\lambda \in (0, 1)$. The following statements regarding the assumptions made in Theorem 4.11 are easily verified either by hand or with the help of symbolic computation:

(i) $S_0 := \mathsf{Fix}\, T_{\mathrm{DR}\lambda} = \{\overline{u}\}$.

(ii) The difference vector is unique as well and given by

$$\bar{\zeta} = \left((1, 0), -\frac{\lambda}{1 - \lambda}(1, 0), (-1, 0), \frac{\lambda}{1 - \lambda}(1, 0)\right).$$

(iii) As noted on Example 3.9(ii) the assumptions of Theorem 3.7 are satisfied for neighborhoods chosen as tubes.

(iv) We bounded the modulus of subtransversality κ from below with the help of symbolic computation as follows:

$$\kappa^2 > \frac{8\left(R^2 + 2R + 1\right)}{R^2 + 2R + 5}.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

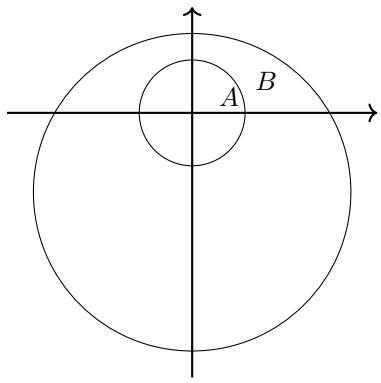

FIG. 8. *Example* 5.3 *for* $R = 1$.

(v) Likewise, we can bound the technical assumption (ii) in Theorem 4.11 from below. Due to the complexity of the constant we omit it here and refer the reader to the Sage worksheet posted at http://vaopt.math.uni-goettingen.de/en/publications.php.

Thus, the assumptions of Theorem 4.11 are satisfied, and the relaxed Douglas–Rachford algorithm converges locally linearly to $\overline{u}$ with rate $1 > c > \sqrt{1 - \frac{1}{(\kappa\sigma)^2}}$ as long as the starting point is close enough to $\overline{u}$.

Example 5.3 (nonintersecting, nonseparable, nonconcentric circles). This example, shown in Figure 8, consists of two sets having not the same center and one of the circles surrounding the other one. Let $R > 0$, and set

$$A := \left\{x \in \mathbb{R}^2 \mid \|x\| = 1\right\}$$
$$B := \left\{x \in \mathbb{R}^2 \;\middle|\; \left\|x - \left(0, -\frac{1}{2} - R\right)\right\| = 2 + R\right\}$$

Our analysis considers the fixed point

$$\overline{u} = \left(0, \frac{3}{2}\right) - \frac{\lambda}{1-\lambda}\left(0, \frac{1}{2}\right)$$

of $T_{\mathrm{DR}\lambda}$ on A and B for $\lambda < 2/3$. The following statements regarding the assumptions made in Theorem 4.11 are easily verified either by hand or with the help of symbolic computation:

(i) $S_0 := \{\overline{u}\} \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda}$.

(ii) The difference vector is unique as well and given by

$$\bar{\zeta} = \left(\left(0, \frac{1}{2}\right), -\frac{\lambda}{1-\lambda}\left(0, \frac{1}{2}\right), -\left(0, \frac{1}{2}\right), -\frac{\lambda}{1-\lambda}\left(0, \frac{1}{2}\right)\right).$$

(iii) Similar to the analysis made in Example 3.9(ii) the assumptions of Theorem 3.7 are satisfied for neighborhoods chosen as tubes.

(iv) We bounded the modulus of subtransversality κ from below with the help of symbolic computation as follows:

$$\kappa^2 > \frac{9\left(4R^2 + 12R + 9\right)}{2R^2 + 6R + 9}.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

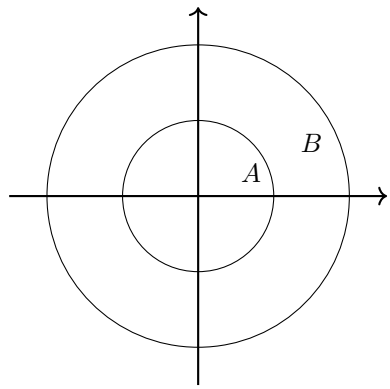

FIG. 9. *Example* 5.4 *for* $R = 2$.

(v) Likewise, we can bound the technical assumption (ii) in Theorem 4.11 from below. Due to the complexity of the constant we omit it here and refer the reader to the Sage worksheet posted at http://vaopt.math.uni-goettingen.de/en/publications.php.

Thus, the assumptions of Theorem 4.11 are satisfied, and the relaxed Douglas–Rachford algorithm converges locally linearly to $\overline{u}$ with rate $1 > c > \sqrt{1 - \frac{1}{(\kappa\sigma)^2}}$ as long as the starting point is close enough to $\overline{u}$.

Example 5.4 (nonintersecting, nonseparable concentric circles). In comparison to Example 5.3, the only thing we change is that we do not allow the circles to have different centers anymore. Let $R > 0$, and define

$$A := \left\{x \in \mathbb{R}^2 \mid \|x\| = 1\right\}, \qquad B := \left\{x \in \mathbb{R}^2 \mid \|x\| = R\right\},$$

where we restrict R to be strictly greater than 1, i.e., $R > 1$.[2] This is illustrated in Figure 9. Our analysis focuses on the fixed point

$$\overline{u} = (0, R) - \frac{\lambda}{1-\lambda}(0, R-1)$$

of $T_{\mathrm{DR}\lambda}$ on A and B. Note that it is enough to consider $\overline{u}$ to get the analysis for any other fixed point due to the symmetry of the problem instance.

Unfortunately, we were unable to verify the technical assumption (ii) in Theorem 4.11.

Nevertheless, this example is subtransversal. The modulus of subtransversality κ is bounded as follows:

$$\kappa^2 > \frac{2R^2}{R^4 - 2R^3 + 2R^2 - 2R + 1}.$$

Example 5.5 (tangential circles). Example 5.5 consists of two circles touching at a single point as shown in Figure 10. Let $R > 0$, and define

$$A := \left\{x \in \mathbb{R}^2 \mid \|x\| = 1\right\},$$
$$B := \left\{x \in \mathbb{R}^2 \mid \|x - (R+1, 0)\| = R\right\}.$$

[2]For $R < 1$ we can change the roles of A and B, which results in the situation presented here.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

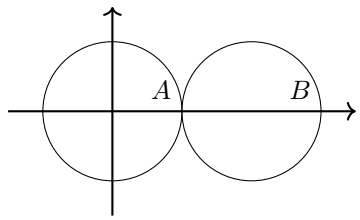

FIG. 10. *Example* 5.5 *for* $R = 1$.

Our convergence analysis focuses on the only point in the intersection of those two sets, namely,

$$\overline{u} = (1, 0)\,.$$

The following statements regarding the assumptions made in Theorem 4.11 are easily verified either by hand or with the help of symbolic computation:

(i) $S_0 := \{\overline{u}\} \in \mathsf{Fix}\, T_{\mathrm{DR}\lambda}$.

(ii) The difference vector is unique as well and given by

$$\bar{\zeta} = ((0,0),(0,0),(0,0),(0,0))\,.$$

(iii) The sets A and B satisfy the assumptions of Theorem 3.7 at $\overline{u}$ with neighborhoods U_1 and U_2 being open balls around $\overline{u}$, that is, $\mathbb{B}_\delta(\overline{u})$, for $\delta \in (0,1)$.

(iv) The technical condition (ii) in Theorem 4.11 is satisfied with

$$\sigma^2 = \frac{1}{2\lambda^2}$$

by Proposition 4.9.

(v) However, this example is not subtransversal when examining it in $\mathbb{R}^2$. Checking equivalently linear regularity, which is fine since we are looking at a point in the intersection of the two sets (see Proposition 4.14), yields a value of

$$\frac{2R}{(R+1)b}, \tag{5.2}$$

where we parametrized a point in the neighborhood of $\overline{u}$ intersected with A as

$$(\sqrt{1-b^2}, b), b \in [1, -1].$$

As $b \to 0$ (in other words, for points close to $\overline{u}$) the ratio (5.2) tends to ∞. This implies that Example 5.5 cannot be linearly regular at the point $\overline{u} = (1, 0)$ and thus is not subtransversal there.

The assumptions of Theorem 4.11 therefore are not satisfied. In light of the necessity of metric subregularity for linear convergence [22, Theorem 2], we conclude that $T_{\mathrm{DR}\lambda}$ *cannot* be linearly convergent in this case (though it might be sublinearly convergent).

Remark 5.6. As shown in the examples above the constants involved for both subtransversality and the technical condition (ii) in Theorem 4.11 can be cumbersome although the actual problem might look relatively easy. We also see in Examples 5.2–5.4 that the presence of subtransversality in the inconsistent case can come as a

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

surprise. Our inability to show the technical condition (ii) in Example 5.4 indicates that this condition characterizes the regularity or nondegeneracy of the underlying model space for the algorithm. Further investigation of this property is needed.

Acknowledgments. The authors would like to thank the referees and the handling editor for their careful reading and constructive comments.

REFERENCES

[1] J.-B. Baillon, P. L. Combettes, and R. Cominetti, *There is no variational characterization of the cycles in the method of periodic projections*, J. Funct. Anal., 262 (2012), pp. 400–408.

[2] A. Bakan, F. Deutsch, and W. Li, *Strong CHIP, normality, and linear regularity of convex sets*, Trans. Amer. Math. Soc., 357 (2005), pp. 3831–3863.

[3] H. Bauschke and D. Noll, *On the local convergence of the Douglas–Rachford algorithm*, Arch. Math. (Basel), 102 (2014), pp. 589–600.

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

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

[6] H. H. Bauschke and W. M. Moursi, *The Douglas–Rachford algorithm for two (not necessarily intersecting) affine subspaces*, SIAM J. Optim., 26 (2016), pp. 968–985.

[7] F. E. Browder, *Existence and approximation of solutions of nonlinear variational inequalities*, Proc. Natl. Acad. Sci., 56 (1966), pp. 1080–1086, https://doi.org/10.1073/pnas.56.4.1080.

[8] A. Daniilidis, D. R. Luke, and M. K. Tam, *Characterizations of super-regularity and its variants*, in Splitting Algorithms, Modern Operator Theory, and Applications, Springer, 2019, Heinz H. Bauschke, Regina S. Burachil, and D. Russell Luke, eds., Springer, Cham, Switzerland, 2019, https://doi.org/10.1007/978-3-030-25939-6.

[9] M. N. Dao and H. M. Phan, *Linear convergence of the generalized Douglas-Rachford algorithm for feasibility problems*, J. Global Optim., 72 (2018), pp. 443–474.

[10] F. Deutsch, *Best Approximation in Inner Product Spaces*, CMS Books Math./Ouvrages Math. SMC, Springer, New York, 2001.

[11] A. L. Dontchev and R. T. Rockafellar, *Implicit Functions and Solution Mappings*, 2nd ed., Springer-Verlag, Dordrecht, 2014.

[12] J. Douglas, Jr., and H. H. Rachford, Jr., *On the numerical solution of heat conduction problems in two or three space variables*, Trans. Amer. Math. Soc., 82 (1956), pp. 421–439.

[13] J. Eckstein and D. P. Bertsekas, *On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators*, Math. Program., 55 (1992), pp. 293–318.

[14] R. Hesse and D. R. Luke, *Nonconvex notions of regularity and convergence of fundamental algorithms for feasibility problems*, SIAM J. Optim., 23 (2013), pp. 2397–2419.

[15] A. Y. Kruger, D. R. Luke, and N. H. Thao, *Set regularities and feasibility problems*, Math. Program., 168 (2018), pp. 279–311, https://doi.org/10.1007/s10107-016-1039-x.

[16] A. S. Lewis, D. R. Luke, and J. Malick, *Local linear convergence of alternating and averaged projections*, Found. Comput. Math., 9 (2009), pp. 485–513.

[17] G. Li and T. K. Pong, *Douglas–Rachford splitting for nonconvex optimization with application to nonconvex feasibility problems*, Math. Program., 159 (2016), pp. 371–401.

[18] P. L. Lions and B. Mercier, *Splitting algorithms for the sum of two nonlinear operators*, SIAM J. Numer. Anal., 16 (1979), pp. 964–979.

[19] D. R. Luke, *Relaxed averaged alternating reflections for diffraction imaging*, Inverse Problems, 21 (2005), pp. 37–50.

[20] D. R. Luke, *Finding best approximation pairs relative to a convex and a prox-regular set in Hilbert space*, SIAM J. Optim., 19 (2008), pp. 714–739.

[21] D. R. Luke, S. Sabach, and M. Teboulle, *Optimization on spheres: Models and proximal algorithms with computational performance comparisons*, SIAM J. Math. Data Sci., 1(3) (2019), pp. 408–445.

[22] D. R. Luke, M. Teboulle, and N. H. Thao, *Necessary conditions for linear convergence of iterated expansive, set-valued mappings*, Math. Program., (2018), pp. 1–31, https://doi.org/10.1007/s10107-018-1343-8.

[23] D. R. Luke, N. H. Thao, and M. K. Tam, *Quantitative convergence analysis of iterated expansive, set-valued mappings*, Math. Oper. Res., 43 (2018), pp. 1143–1176, https://doi.org/10.1287/moor.2017.0898.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[24] Z. Opial, *Weak convergence of the sequence of successive approximations for nonexpansive mappings*, Bull. Amer. Math. Soc. (N.S.), 73 (1967), pp. 591–597.

[25] H. Phan, *Linear convergence of the Douglas–Rachford method for two closed sets*, Optimization, 65 (2016), pp. 369–385.

[26] J. von Neumann, *Functional Operators, Vol* II. *The Geometry of Orthogonal Spaces*, Ann. of Math Stud. 22, Princeton University Press, Princeton, NJ, 1950; reprint of mimeographed lecture notes first distributed in 1933.

Downloaded 02/22/20 to 155.198.30.43. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.