

Convergence Analysis of the Relaxed Douglas-Rachford Algorithm

D. Russell Luke^{*} Anna-Lena Martins[†]

May 24, 2019

Motivated by nonconvex, inconsistent feasibility problems in imaging, the relaxed alternating averaged reflections algorithm, or relaxed Douglas-Rachford algorithm (DR λ), was first proposed over a decade ago. Convergence results for this algorithm are limited either to 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 DR λ for feasibility problems that are both nonconvex and inconsistent. We introduce a new type of regularity of sets, called super-regular 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. super-regular, inconsistent feasibility problem, projection, relaxed averaged alternating reflections, fixed point, linear convergence, metric subregularity, nonconvex, subtransversality

Mathematics Subject Classification: 65K10, 49K40, 49M05, 65K05, 90C26, 49M20, 49J53

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 [25], and its many set version, the cyclic projection algorithm, or averaged projections and, in the case of two-set feasibility, the Douglas-Rachford algorithm [11] as formulated by Lions and Mercier [17].

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 last 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 [4, 12, 17]. 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* [4, 5]. For consistent nonconvex feasibility, Hesse and Luke [13] were the first to prove meaningful local convergence results for Douglas-Rachford.

^{*}Institut für Numerische und Angewandte Mathematik, Universität Göttingen, Lotzestr. 16-18, 37083 Göttingen, Germany, Email: r.luke@math.uni-goettingen.de. The research of DRL was supported in part by the German Research Foundation grant SFB755-C02 and in part by the German Research Foundation grant GRK2088-B5.

[†]Institut für Numerische und Angewandte Mathematik, Universität Göttingen, Lotzestr. 16-18, 37083 Göttingen, Germany, Email: a.martins@math.uni-goettingen.de. The research of A.-L.M. was supported by the German Research Foundation grant SFB755-C02.

This was quickly followed by several generalizations [6, 9, 16, 22, 24]. For inconsistent feasibility, since Douglas-Rachford cannot converge, weak convergence follows generically if the iterates are bounded, 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 [18] with a proof of convergence for convex feasibility – inconsistent and consistent. Given $x^0 \in \mathcal{E}$ and $\lambda \in (0, 1)$, for $k = 0, 1, 2, \dots$, the iteration takes the form

$$x^{k+1} \in T_{\text{DR}\lambda}(x^k) := \left\{ \frac{\lambda}{2} (R_A(2b - x^k) + x^k) + (1 - \lambda)b \mid b \in P_B(x^k) \right\}. \quad (1)$$

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, $T_{\text{DR}} := \frac{1}{2}(R_A R_B + \text{Id})$, where Id denotes the identity. From here on, we will refer to the algorithm as $\text{DR}\lambda$. A characterization of the fixed points in the nonconvex inconsistent case and a first attempt at a local convergence result was given in [19]. The analysis required one of the sets to be convex and the other set to be *prox-regular*. More recently, Li and Pong [16] 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 semi-algebraic, they showed global convergence, [16, Corollary 1]. Under still stronger assumptions (that one of the sets is linear, the other semi-algebraic, and that the intersection is *strongly regular* [15]), local linear convergence can be shown, [16, Proposition 2]. Minh and Phan [9, Theorem 5.8] show local linear convergence of a generalized Douglas-Rachford algorithm (which includes $T_{\text{DR}\lambda}$ as a special case) when the two sets are super-regular, and have sufficiently regular nonempty intersection. Noteworthy here is that none of these approaches can explain convergence in the case of two affine halfspaces 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 *super-regularity at a distance* that will be our only assumption on the sets themselves. Super-regularity at a distance falls into the spectrum of other regularity notions like ϵ -subregularity (cf. [8, 14]) and, as the name suggests, super-regularity [15]. 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 [19], 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 [22], 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 [14], the sets in the present analysis need not have points in common. The analysis of [22] uses two properties of fixed point mappings. The first property, *pointwise almost averagedness*, follows from the regularity of the sets and, as shown in [21, Proposition 4] is an important ingredient in guaranteeing convergence of the iterates to fixed points. In Theorem 3.7 we establish that the $T_{\text{DR}\lambda}$ mapping is almost averaged at its fixed points when the sets A and B are super-regular at a distance. The second property, *metric subregularity* of the fixed point mapping at its fixed points, was subsequently shown in [21, 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 [22, 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 $\text{DR}\lambda$. 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 $\text{dist}(x, C) := \inf_{y \in C} \|x - y\|$ and the *projector* onto C is the *set-valued mapping* $P_C(x) := \{z \mid \|z - x\| = \text{dist}(x, C)\}$. 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*

$$N_C^{\text{P}}(\bar{x}) = \text{cone}(P_C^{-1}\bar{x} - \bar{x}).$$

Equivalently, $\bar{x}^ \in N_C^{\text{P}}(\bar{x})$ 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*

$$N_C(\bar{x}) = \text{Lim sup}_{x \rightarrow \bar{x}} N_C^{\text{P}}(x),$$

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 Super-regularity at a Distance

We limit our attention in this study to *super-regular* sets and their extension to sets with the corresponding properties relative to points *not* belonging to the sets.

Definition 2.1 (super-regularity [15, Definition 4.3]). *Let $\Omega \subseteq \mathbb{R}^n$ and $\bar{x} \in \Omega$. The set Ω is said to be super-regular at \bar{x} if it is locally closed at \bar{x} and for every $\epsilon > 0$ there is a $\delta > 0$ such that for all $(x, 0) \in \text{gph } N_\Omega \cap \{(\mathbb{B}_\delta(\bar{x}), 0)\}$*

$$\langle y' - y, x - y \rangle \leq \epsilon \|y' - y\| \|x - y\|, \quad (\forall y' \in \mathbb{B}_\delta(\bar{x})) (\forall y \in P_\Omega(y')). \quad (2)$$

Rewriting the above leads to the following equivalent characterization of super-regularity, which might be more useful for our purposes.

Proposition 2.2. [15, Proposition 4.4] *The set $\Omega \subseteq \mathcal{E}$ is super-regular at $\bar{x} \in \Omega$ if and only if it is locally closed at \bar{x} and for every $\epsilon > 0$ there exists $\delta > 0$ such that*

$$\langle v, x - y \rangle \leq \epsilon \|v\| \|x - y\|, \quad (\forall (x, v) \in \text{gph } N_\Omega \cap (\mathbb{B}_\delta(\bar{x}) \times \mathcal{E})) (\forall y \in \Omega \cap \mathbb{B}_\delta(\bar{x})). \quad (3)$$

To extend super-regularity to super-regularity at a distance, we employ the more general framework of ϵ -subregular sets first introduced in [14]. 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 \bar{x} for $(x, v) \in \text{gph } N_\Omega$ if it is locally closed at \bar{x} and there exists an $\epsilon > 0$ together with a neighborhood U_ϵ of \bar{x} such that*

$$\langle v - (y' - y), y - x \rangle \leq \epsilon \|v - (y' - y)\| \|y - x\| \quad (\forall y' \in \Lambda \cap U_\epsilon) (\forall y \in P_\Omega(y')). \quad (4)$$

Ω is subregular relative to Λ at \bar{x} for $(x, v) \in \text{gph } N_\Omega$ if it is locally closed and for all $\epsilon > 0$ there exists U_ϵ such that (4) holds.

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

$$V_\epsilon := \{(x, v) \in \text{gph } N_\Omega^P \mid x + v \in U_\epsilon, x \in P_\Omega(x + v)\}, \quad (5)$$

and U_ϵ is a neighborhood of \bar{x} with respect to an $\epsilon > 0$. The set Ω is called super-regular at a distance relative to Λ at \bar{x} if it is ϵ -super-regular at a distance relative to Λ at \bar{x} for all $\epsilon > 0$.

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

Remark 2.5 (super-regularity at a distance relative to \mathcal{E} implies super-regularity). *Being super-regular at a distance relative to $\Lambda = \mathcal{E}$ at some point $\bar{x} \in \Omega$ implies that the set is super-regular at \bar{x} . To see this, let Ω be super-regular at a distance relative to $\Lambda = \mathcal{E}$ at $\bar{x} \in \Omega$. For fixed $\epsilon > 0$ note that $(x, 0) \in V_\epsilon$ for all $x \in \Omega \cap U_\epsilon$. With these, (4) becomes*

$$\langle y - y', y - x \rangle \leq \epsilon \|y - y'\| \|y - x\| \quad (6)$$

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 $\Lambda = \mathcal{E}$ (6) 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 super-regularity of Ω at \bar{x} .

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

Proof. Fix $\bar{x} \in \Omega$. 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 $U_\epsilon \subset \mathcal{E}$ of \bar{x} , $y' \in U$, $y \in P_\Omega(y')$, which implies that $y' - y \in N_\Omega(y)$, we deduce that $\langle y' - y, x - y \rangle \leq 0$ and thus

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

This shows super-regularity at a distance of Ω relative to \mathcal{E} at all $\bar{x} \in \Omega$ as claimed. \square

Example 2.7 (circle). *Consider the set*

$$\Omega := \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\}.$$

This set is ϵ -subregular relative to $\Lambda = P_\Omega^{-1}(\bar{x})$ at any $\bar{x} \in \Omega$ for all $(\bar{x}, v) \in \text{gph } N_\Omega$ with $\epsilon = 0$ (which implies that Ω is in fact subregular relative to Λ for all $(\bar{x}, v) \in \text{gph } N_\Omega$). Indeed, for any $\delta \in (0, 1)$ we have, for any $y' \in \Lambda \cap \mathbb{B}_\delta(\bar{x})$, that $y \in P_\Omega(y')$ is given by $y = \bar{x}$ and (4) specializes to

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

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

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

where

$$V_\delta := \{(x, v) \in \text{gph } N_\Omega^P \mid x + v \in \mathbb{B}_\delta(\bar{x}), x \in P_\Omega(x + v)\},$$

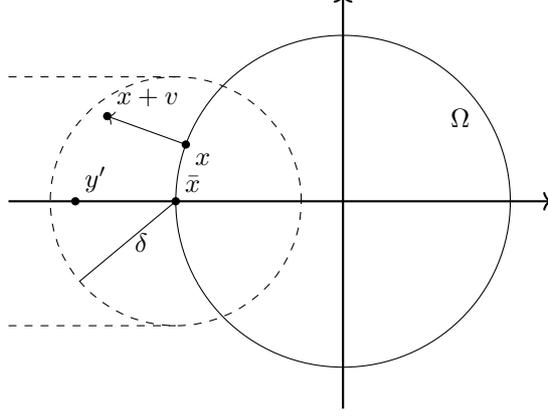


Figure 1: Example 2.7

which shows that Ω is ϵ -super-regular at a distance relative to $P_{\Omega}^{-1}(\bar{x})$ at \bar{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 \bar{x} , where $\delta < 1$. Thus, Ω is super-regular relative to $P_{\Omega}^{-1}(\bar{x})$ at \bar{x} .

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

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

is a neighborhood for \bar{x} such that ϵ -super-regularity relative to $\Lambda = P_{\Omega}^{-1}(\bar{x})$ is satisfied for Ω . Fortunately, our violation ϵ will not be worse compared to the neighborhood being a ball with radius δ around \bar{x} . This allows us to include more points in $\Lambda \cap U$ without violating (4).

Proposition 2.8 (characterization of super-regularity at a distance).

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

$$\|x - y\|^2 \leq \epsilon \|(y' - y) - (x' - x)\| \|x - y\| + \langle x' - y', x - y \rangle \quad (\forall y' \in U_{\epsilon} \cap \Lambda)(\forall y \in P_{\Omega}(y')) \quad (7)$$

holds with $x' = x + v \in U_{\epsilon}$ for all $(x, v) \in V_{\epsilon}$ for V_{ϵ} defined by (5).

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

$$\|x - y\| \leq \epsilon \|(y' - y) - (x' - x)\| + \|x' - y'\| \quad (\forall y' \in U_{\epsilon} \cap \Lambda)(\forall y \in P_{\Omega}(y')) \quad (8)$$

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

Proof. (i). Let $\Omega \subset \mathcal{E}$ be ϵ -super-regular at a distance relative to $\Lambda \subset \mathcal{E}$ at \bar{x} . Then, for fixed $\epsilon > 0$, there exists a neighborhood U_{ϵ} of \bar{x} such that for all $(x, v) \in V_{\epsilon}$ for V_{ϵ} defined by (5) and $x' = x + v \in U_{\epsilon}$ the following hold:

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

This proves the first part of the equivalence. For the remaining one let U_{ϵ} be a neighborhood of \bar{x} such that

$$\|x - y\|^2 \leq \epsilon \|(y' - y) - (x' - x)\| \|x - y\| + \langle x' - y', x - y \rangle \quad (\forall y' \in U_{\epsilon} \cap \Lambda)(\forall y \in P_{\Omega}(y')) \quad (9)$$

holds with $x' = x + v \in U_\epsilon$ for all $(x, v) \in V_\epsilon$ for V_ϵ defined by (5). (9) is equivalent to

$$\langle y' - y - (x' - x), x - y \rangle + \langle x' - y', x - y \rangle \leq \epsilon \|(y' - y) - (x' - x)\| \|x - y\| + \langle x' - y', x - y \rangle,$$

($\forall y' \in U_\epsilon \cap \Lambda$) ($\forall y \in P_\Omega(y')$) by the calculations made before. Subtracting $\langle x' - y', x - y \rangle$ from both sides and inserting $v = x' - x$ yields

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

Reordering the left-hand side we deduce the definition of ϵ -super-regular for Ω at \bar{x}

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

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

3 Properties of $T_{\text{DR}\lambda}$ and Characterization of its Fixed Points

Our convergence analysis is based on the framework established in [22] 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 fixed 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* 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 Definition 4.3.

3.1 Almost Averaged Mappings

Definition 3.1 (almost nonexpansive/averaged mappings, [22], 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*

$$\|x^+ - y^+\| \leq \sqrt{1 + \epsilon} \|x - y\| \quad \forall y^+ \in Ty \text{ and } \forall x^+ \in Tx \text{ whenever } x \in D. \quad (10)$$

If (10) 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*

$$\tilde{T} = \frac{1}{\alpha} T + \frac{(\alpha-1)}{\alpha} \text{Id}$$

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) *$(1 - \frac{1}{\alpha}) \text{Id} + \frac{1}{\alpha} T$ 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*

$$\|x^+ - y^+\|^2 \leq (1 + \epsilon) \|x - y\|^2 - \frac{1 - \alpha}{\alpha} \|(x - x^+) - (y - y^+)\|. \quad (11)$$

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 [22, Proposition 2.1]. \square

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 $j = 1, 2, \dots, m$ 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 $j = 1, 2, \dots, m$.*

(i) *If $U := U_1 = U_2 = \dots = U_m$ and $S := S_1 = S_2 = \dots = S_m$ 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 $\alpha = \max_{j=1,2,\dots,m} \{\alpha_j\}$ on U .*

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

$$\epsilon = \prod_{j=1}^m (1 + \epsilon_j) - 1. \quad (12)$$

(iii) *If $T_j U_j \subseteq U_{j-1}$ and $T_j S_j \subseteq S_{j-1}$ for $j = 2, 3, \dots, m$, then the composite mapping $T := T_1 \circ T_2 \circ \dots \circ T_m$ is pointwise almost averaged at all $y \in S_m$ on U_m with violation at most ϵ given by (12) and averaging constant at least*

$$\alpha = \frac{m}{m - 1 + \frac{1}{\max_{j=1,2,\dots,m} \{\alpha_j\}}}. \quad (13)$$

Proof. The proof of this statement can be found in [22, Proposition 2.4]. \square

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 $T_{\text{DR}\lambda}$ 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 $\bar{x} \in \Omega$. Let $\Lambda := P_\Omega^{-1}(\bar{x}) \cap U$. If Ω is ϵ -super-regular at a distance ϵ 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 $\tilde{\epsilon}$ on U for $\tilde{\epsilon} := 4\epsilon / (1 - \epsilon)^2$. That is, at each $y' \in \Lambda$*

$$\|x - y\| \leq \sqrt{1 + \tilde{\epsilon}} \|x' - y'\| = \frac{1 + \epsilon}{1 - \epsilon} \|x' - y'\| \quad (\forall x' \in U) (\forall x \in P_\Omega(x')) (\forall y \in P_\Omega(y')).$$

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

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

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

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

Proof. Our proof follows that of [22, 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 $(x, x' - x) \in \text{gph } N_\Omega^P \subset \text{gph } N_\Omega$. Moreover, by construction $(x, x' - x) \in V_\epsilon$ where V_ϵ is defined by (5).

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

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

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

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

and thus conclude the claimed result.

(ii). By super-regularity at a distance relative to Λ of Ω and Proposition 2.8(i) we have

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

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

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

Using part (i) we deduce

$$\begin{aligned} & \|x - y\|^2 + \|(x' - x) - (y' - y)\|^2 \\ &\leq \left(1 + 4\epsilon \frac{1 + \epsilon}{(1 - \epsilon)^2}\right) \|x' - y'\|^2 \end{aligned} \tag{14}$$

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 $x \in P_\Omega(x')$, (14) 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, $R_\Omega := 2P_\Omega - \text{Id}$, is pointwise almost nonexpansive at each $y' \in \Lambda$ with violation $8\epsilon(1 + \epsilon)/(1 - \epsilon)^2$ on U . \square

3.2 $T_{\text{DR}\lambda}$ is Almost Averaged at $\text{Fix } T_{\text{DR}\lambda}$

For general multivalued mappings $T : \mathcal{E} \rightrightarrows \mathcal{E}$ the set of fixed points is defined as

$$\text{Fix } T := \{x \in \mathcal{E} \mid x \in T(x)\}. \tag{15}$$

Note that, by this definition, the set $T(x)$ need not consist entirely of fixed points (see [22, Example 2.1]). If T is pointwise almost averaged, however, the mapping T is single-valued on its fixed point set.

Proposition 3.5. [22, Proposition 2.2] *If $T : \mathcal{E} \rightrightarrows \mathcal{E}$ is pointwise almost nonexpansive on $D \subseteq \mathcal{E}$ at $\bar{x} \in D$ with violation $\varepsilon \geq 0$, then T is single-valued at \bar{x} . In particular, if $\bar{x} \in \text{Fix } T$ then $T(\bar{x}) = \{\bar{x}\}$.*

The mapping $T_{\text{DR}\lambda}$ 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 $\bar{x} \in \mathcal{E}$ and let $\Omega \subset \mathcal{E}$ be super-regular at a distance relative to $\Lambda \subset P_\Omega^{-1}(\bar{\omega})$ at $\bar{\omega}$ where $\bar{\omega} \in P_\Omega(\bar{x})$ and $\bar{x} \in \Lambda$. In addition, for each $\epsilon > 0$, let $\bar{x} \in U_\epsilon(\bar{\omega})$ where $U_\epsilon(\bar{\omega})$ is a neighborhood of $\bar{\omega}$ on which (4) holds. Then $P_\Omega(\bar{x}) = \{\bar{\omega}\}$.*

Proof. For some fixed $\epsilon > 0$, we get by the assumptions on super-regularity 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 $\bar{x} \in \Lambda \cap U_\epsilon(\bar{\omega})$ on $U_\epsilon(\bar{\omega})$ with violation $\tilde{\epsilon} = 4\epsilon/(1-\epsilon)^2$. This implies single-valuedness of P_Ω at \bar{x} by Proposition 3.5, i.e. that $\{\bar{\omega}\} = P_\Omega(\bar{x})$, as claimed. \square

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

- (i) *For each $\epsilon > 0$, $\bar{x} \in U_\epsilon(\bar{b})$ where $U_\epsilon(\bar{b})$ is a neighborhood of \bar{b} on which (4) holds for ϵ .*
- (ii) *For each $\epsilon > 0$, $2\bar{b} - \bar{x} \in U_\epsilon(\bar{a})$ where $U_\epsilon(\bar{a})$ is a neighborhood of \bar{a} on which (4) holds for ϵ .*
- (iii) $R_B(\Lambda_{\bar{b}}) \subset \Lambda_{\bar{a}}$.
- (iv) $R_B(U_\epsilon(\bar{b})) \subset U_\epsilon(\bar{a})$ for all $\epsilon > 0$.

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

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 $T_{\text{DR}\lambda}$ at \bar{x} holds on open subsets containing both \bar{x} and $\bar{b} = P_B(\bar{x})$. This follows from assumption (i). The conclusion of the theorem could have been equivalently stated: *for all $\epsilon > 0$ there exists a neighborhood $U(B, \epsilon, \bar{x})$ of \bar{x} such that $T_{\text{DR}\lambda}$ is pointwise almost averaged at \bar{x} with violation at most ϵ on U .* We have presented the statement with neighborhood $U(B, \epsilon, \bar{x})$ containing \bar{b} to emphasize the fact that the open sets on which the regularity of $T_{\text{DR}\lambda}$ holds is constructed relative to points \bar{b} at a distance from the point of interest $\bar{x} \in \text{Fix } T_{\text{DR}\lambda}$. 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 $\{\bar{b}\} = P_B(\bar{x})$ and $\{\bar{a}\} = P_A(R_B(\bar{x}))$, as claimed. From this one can immediately conclude that $T_{\text{DR}\lambda}$ is single-valued at \bar{x} .

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

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 $U(A, \epsilon_{R_A R_B}, \bar{x})$ such that $R_A R_B$ is pointwise almost nonexpansive at \bar{x} with violation at most $\epsilon_{R_A R_B} = \epsilon_{R_A} + \epsilon_{R_B} + \epsilon_{R_A} \epsilon_{R_B}$ on $U(B, \epsilon, \bar{x})$.

By Proposition 3.2(ii) we get that $1/2(R_A R_B + \text{Id})$ is almost firmly nonexpansive at \bar{x} with violation $\epsilon_{R_A R_B}/2$ on $U_{\epsilon_B}(\bar{b})$. Applying Proposition 3.3(i) yields pointwise almost firm nonexpansivity of $T_{\text{DR}\lambda}$ at \bar{x} on $U_{\epsilon_B}(\bar{b})$ 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 $U(B, \epsilon, \bar{x})$ such that $T_{\text{DR}\lambda}$ is pointwise almost nonexpansive at \bar{x} on $U(B, \epsilon, \bar{x})$ with violation at most ϵ (the corresponding neighborhood $U(A, \epsilon_{R_A R_B}, \bar{x})$ of \bar{a} will be denoted by $U(A, \epsilon, \bar{x})$). This conclusion is consistent with the fact established above that $T_{\text{DR}\lambda}$ is single-valued, which completes the proof. \square

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

Corollary 3.8 implies that $T_{\text{DR}\lambda}$ is pointwise almost firmly nonexpansive at \bar{x} with violation ϵ on $U(B, \epsilon, \bar{x})$. The strength of Corollary 3.8, however, is hidden in the proof given below and the explicit construction of the neighborhoods $U(B, \epsilon, \bar{x})$ and $U(A, \epsilon, \bar{x})$. Thus, under the assumptions of Theorem 3.7, and given the neighborhoods for some fixed violation $\bar{\epsilon}$, we are always able to restrict these neighborhoods to smaller sets where (4) holds with some violation smaller than $\bar{\epsilon}$.

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

Let $\epsilon < \bar{\epsilon}$. Theorem 3.7(i) implies that there exists a neighborhood $U(B, \epsilon, \bar{x})$ of \bar{b} where (4) holds such that $U(B, \epsilon, \bar{x}) \subset U(B, \bar{\epsilon}, \bar{x})$. To see this, note that by (i) the existence of $U(B, \epsilon, \bar{x})$ is guaranteed and thus only $U(B, \epsilon, \bar{x}) \subset U(B, \bar{\epsilon}, \bar{x})$ has to be proven. Let $\tilde{U}(B, \epsilon, \bar{x})$ be a neighborhood for ϵ where (4) holds. Then (4) is satisfied for $\bar{\epsilon}$ as well. Thus, $U(B, \epsilon, \bar{x}) := \tilde{U}(B, \epsilon, \bar{x}) \cap U(B, \bar{\epsilon}, \bar{x})$ is a neighborhood of \bar{b} where Theorem 3.7(iv) holds and (4) is satisfied for both ϵ and $\bar{\epsilon}$, which shows that $U(B, \epsilon, \bar{x}) \subset U(B, \bar{\epsilon}, \bar{x})$ as required. Next, applying the reflection on B on both of neighborhoods $U(B, \epsilon, \bar{x})$ and $U(B, \bar{\epsilon}, \bar{x})$ yields

$$R_B(U(B, \epsilon, \bar{x})) \subset R_B(U(B, \bar{\epsilon}, \bar{x})). \quad (16)$$

Let $\tilde{U}(A, \epsilon, \bar{x})$ be a neighborhood of \bar{a} where (4) holds for ϵ such that (iv) of Theorem 3.7 is satisfied. That is

$$R_B(U(B, \epsilon, \bar{x})) \subset \tilde{U}(A, \epsilon, \bar{x}). \quad (17)$$

Combining Theorem 3.7(iv) for the neighborhoods $U(B, \bar{\epsilon}, \bar{x})$ and $U(A, \bar{\epsilon}, \bar{x})$ and (16) we deduce

$$R_B(U(B, \epsilon, \bar{x})) \subset R_B(U(B, \bar{\epsilon}, \bar{x})) \subset U(A, \bar{\epsilon}, \bar{x}). \quad (18)$$

This and (17) imply that

$$R_B(U(B, \epsilon, \bar{x})) \subset \tilde{U}(A, \epsilon, \bar{x}) \cap U(A, \bar{\epsilon}, \bar{x}). \quad (19)$$

Set

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

Then, $U(A, \epsilon, \bar{x})$ is neighborhood of \bar{a} where (4) holds with ϵ , since it is a subset of $\tilde{U}(A, \epsilon, \bar{x})$. Moreover, $U(B, \epsilon, \bar{x})$ and $U(A, \epsilon, \bar{x})$ satisfy (iv) of Theorem 3.7 by (19). By the construction of $U(A, \bar{\epsilon}, \bar{x})$ and the choice of $U(B, \epsilon, \bar{x})$ both sets satisfy the inclusions $U(A, \epsilon, \bar{x}) \subset U(A, \bar{\epsilon}, \bar{x})$ and $U(B, \epsilon, \bar{x}) \subset U(B, \bar{\epsilon}, \bar{x})$. This completes the proof. \square

Example 3.9. *The following examples illustrate the assumptions of the above theorem. For these examples it is easy to determine the sets of fixed points of the mapping $T_{\text{DR}\lambda}$. In Theorem 3.13 we give a precise characterization of $\text{Fix } T_{\text{DR}\lambda}$. 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 super-regular relative to \mathcal{E} at any of their points, i.e. ϵ -super-regular for all $\epsilon > 0$. In fact, the violation can be set to 0. Thus, as long as $\text{Fix } T_{\text{DR}\lambda} \neq \emptyset$ the mappings P_B , R_B and R_A are nonexpansive (i.e. no violation) at \bar{x} on the whole space \mathcal{E} by Proposition 3.4 and we can apply Theorem 3.7 to conclude that $T_{\text{DR}\lambda}$ is firmly nonexpansive at \bar{x} on the neighborhood \mathcal{E} . For instance, consider the two sets

$$A := \{x = (x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\} \quad \text{and} \quad B := \{x = (x_1, x_2) \in \mathbb{R}^2 \mid (x_1 - 3)^2 + x_2^2 \leq 1\}.$$

The set of fixed points is given by the unique point

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

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

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

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

The sets A and B are both non-convex, but still super-regular. The set of fixed points is again given by the unique point

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

for fixed $\lambda \in (0, 1)$. As seen in Example 2.7 both sets are super regular relative to radial directions. Thus, applying Theorem 3.7 we deduce that $T_{\text{DR}\lambda}$ for some fixed $\lambda \in (0, 1)$ is only almost firmly nonexpansive at \bar{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 Fig. 2.

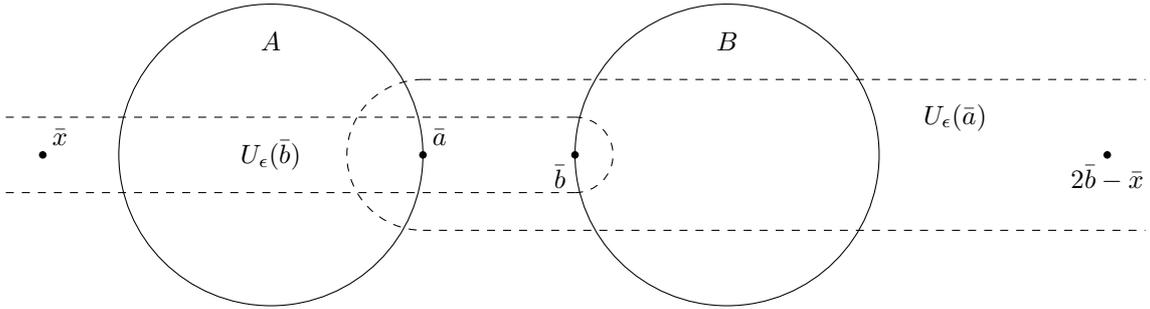


Figure 2: Example 3.9(ii)

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

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

The fixed point set then reduces to $\text{Fix } T_{\text{DR}\lambda} = \{(1, 0)\} = A \cap B$. 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 $\text{Fix } T_{\text{DR}\lambda} \neq \emptyset$. If A and B are closed and convex, $T_{\text{DR}\lambda}$ is firmly nonexpansive on \mathcal{E} .*

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

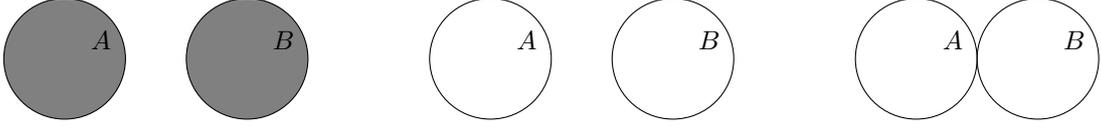


Figure 3: Example 3.9(i)-(iii)

3.3 Characterization of $\text{Fix } T_{\text{DR}\lambda}$

We collect some facts and identities that will be useful throughout.

Lemma 3.11. *Let A and B be closed and $T_{\text{DR}\lambda}$ given by (1) with $\lambda \in (0, 1)$. Let $x \in \text{Fix } T_{\text{DR}\lambda} \neq \emptyset$ such that $T_{\text{DR}\lambda}$ 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 $\text{Fix } T_{\text{DR}\lambda}$;
- (ii) $P_A(R_B(x))$ is single-valued, hence so is $R_A(R_B(x))$;
- (iii) $P_A(2f - x) = P_A(R_B(x))$;
- (iv) $T_{\text{DR}}(x) - x = P_A(R_B(x)) - P_B(x)$;
- (v) $f + \frac{1-\lambda}{\lambda}y = P_A(2f - x)$;
- (vi) if A is convex, then, for $e = P_A(f)$

$$P_A\left(e + \frac{1}{1-\lambda}(f - e)\right) = e. \quad (20)$$

Proof. (i)-(ii). Since

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

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

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

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

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

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

(v). To see this, note that

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

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

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

by the definition of y .

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

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 $T_{\text{DR}\lambda}$ at x .

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

$$\text{Fix } T_{\text{DR}\lambda} \cap U \subset \mathcal{M} := \left\{ f - \frac{\lambda}{1-\lambda}(f - e) \mid f \in P_B \left(f - \frac{\lambda}{1-\lambda}(f - e) \right), \text{ and } e \in P_A(f) \right\} \cap U. \quad (21)$$

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.

Proof. Let $x \in \text{Fix } T_{\text{DR}\lambda} \cap U$. By the assumptions $T_{\text{DR}\lambda}$ 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 .

$$\begin{aligned} x \in \text{Fix } T_{\text{DR}\lambda} &\implies f + \frac{1-\lambda}{\lambda}y = P_A(2f - x) \\ &\iff f + \frac{1-\lambda}{\lambda}(x - f) = P_A(2f - x) \\ &\iff x = \frac{\lambda}{1-\lambda}P_A(2f - x) - \frac{2\lambda-1}{1-\lambda}f \\ &\iff x = f - \frac{\lambda}{1-\lambda}(f - P_A(2f - x)). \end{aligned} \quad (22)$$

Comparing with (21) we have to show that $P_A(f) = P_A(2f - x)$.

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

$$\begin{aligned} x \in \text{Fix } T_{\text{DR}\lambda} &\iff x = \frac{\lambda}{2}(R_A(R_B(x)) + x) + (1-\lambda)P_B(x) \\ &\iff f = \lambda P_A(2f - x) + (1-\lambda)(2f - x). \end{aligned} \quad (23)$$

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 $P_A(f) = P_A(2f - x)$.

Using this fact, then (22) becomes

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

Finally, (21) 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 $\tilde{x} \in \mathcal{M} \cap U$ in (21). Then $\tilde{x} := f - \frac{\lambda}{1-\lambda}(f - e)$ for some $e \in P_A(f)$ and $f \in P_B(\tilde{x})$, and

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

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

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 $P_A\left(e + \frac{1}{1-\lambda}(f-e)\right) = e$ for all $\lambda \in (0,1)$ and, by Theorem 3.13 the inclusion (21) is in fact equality for all λ for which $f \in P_B\left(f - \frac{\lambda}{1-\lambda}(f-e)\right)$. Compare this to the statement in [19, 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 [19] that f is a local best approximation point is actually incorrect. Where our description includes $f \in P_B\left(f - \frac{\lambda}{1-\lambda}(f-e)\right)$, and $e \in P_A(f)$, the version in [19, Lemma 3.8] states that f is a local best approximation point [19, 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 $T_{\text{DR}\lambda}$ with A the line and B the circle, the point $\left(0, 1 - \frac{\lambda}{4(1-\lambda)}\right)$ 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 $e \in P_A\left(f + \frac{\lambda}{1-\lambda}(f-e)\right)$ 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 $e \notin P_A\left(f + \frac{\lambda}{1-\lambda}(f-e)\right)$. 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 (21), namely $f \in P_B\left(f - \frac{\lambda}{1-\lambda}(f-e)\right)$.

(iii) By Theorem 3.7 we know that $T_{\text{DR}\lambda}$ is single-valued at its fixed points if both A and B are super-regular at a distance and assumptions (i)-(iv) of Theorem 3.7 hold. The local gap $f - P_A(f)$ is therefore unique. In [19] 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 $\text{DR}\lambda$ and the corresponding gap). *In the setting of Theorem 3.13 let $x \in \text{Fix } T_{\text{DR}\lambda} \cap U$. Then*

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

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

In our statements we require that $\text{Fix } T_{\text{DR}\lambda} \neq \emptyset$. 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 $\text{Fix } T_{\text{DR}\lambda} \neq \emptyset$. Moreover, $\text{Fix } T_{\text{DR}\lambda} = \mathcal{M}$ where \mathcal{M} is given by (21).*

Proof. The proof follows the pattern of proof in [19, Lemma 2.1] which establishes existence of fixed points for $T_{\text{DR}\lambda}$ 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.

To show the tight characterization of $\text{Fix } T$, let $e \in \text{Fix } T$. Then $P_B(e) = f$ and $P_A(f) = e$ and $T_{\text{DR}\lambda}$ 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

$P_A\left(e + \frac{1}{1-\lambda}(f - e)\right) = e$ for all $\lambda \in (0, 1)$. Moreover, for all such λ we have $f = P_B\left(f - \frac{\lambda}{1-\lambda}(f - e)\right)$. Together, for $\bar{x} = f - \frac{\lambda}{1-\lambda}(f - e)$, this yields

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

Now, applying Theorem 3.13 immediately yields $\text{Fix } T_{\text{DR}\lambda} = \mathcal{M}$ where \mathcal{M} is given by (21). This completes the proof. \square

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

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

$$A := \{x = (x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\},$$

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

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

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

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

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

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

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

If x is a fixed point of $T_{\text{DR}\lambda}$, that is $x = T_{\text{DR}\lambda}(x)$, 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 $x \notin \text{Fix } T_{\text{DR}\lambda}$, and therefore $\text{Fix } T_{\text{DR}\lambda} = \emptyset$.

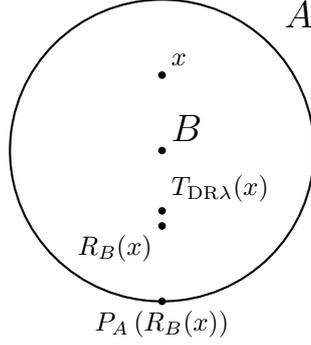


Figure 4: Example 3.17 for a point $x \in \mathbb{R}^2$ and $\lambda = 0.8$.

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

Proposition 3.18 (monotonicity of $\text{Fix } T_{DR\lambda}$ 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 $\text{Fix } T_{DR\lambda_2} \neq \emptyset$. Moreover, let $T_{DR\lambda_2}$ be single-valued at its fixed points. Then*

$$P_B(\text{Fix } T_{DR\lambda_2}) \subseteq P_B(\text{Fix } T_{DR\lambda_1}). \quad (24)$$

If (21) holds for λ_2 with equality instead of just set inclusion, then (24) holds with equality.

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

$$x = P_B(x) - \frac{\lambda_2}{1 - \lambda_2} (P_B(x) - P_A(P_B(x))). \quad (25)$$

Consider $\tilde{x} := P_B(x) - \frac{\lambda_1}{1 - \lambda_1} (P_B(x) - P_A(P_B(x)))$ 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 $T_{DR\lambda_2} \neq \emptyset$. Set $f := P_B(x)$. Then $f \in B$ and $P_B(\tilde{x}) = f$. To see this, note that $\frac{\lambda_2}{1 - \lambda_2} (P_A(f) - f) \in N_B^P(f)$. 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 $P_B(\tilde{x}) = f$. Moreover, since $P_B(x) = f = P_B(\tilde{x})$, we can conclude that $\tilde{x} \in \text{Fix } T_{DR\lambda_1}$. To see this, evaluate $T_{DR\lambda_1}(\tilde{x})$

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

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

$$\begin{aligned} 2f - \tilde{x} &= f + \frac{\lambda_1}{1 - \lambda_1} (f - P_A(f)) \\ &= P_A(f) + \frac{1}{1 - \lambda_1} (f - P_A(f)). \end{aligned}$$

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

$$\begin{aligned} \lambda_1 (P_A(f) + \tilde{x}) + (1 - 2\lambda_1) f &\in T_{DR\lambda_1}(\tilde{x}) \\ \iff \\ \lambda_1 \left(P_A(f) + f - \frac{\lambda_1}{1 - \lambda_1} (f - P_A(f)) \right) + (1 - 2\lambda_1) f &\in T_{DR\lambda_1}(\tilde{x}) \\ \iff \\ f - \frac{\lambda_1}{1 - \lambda_1} (f - P_A(f)) &\in T_{DR\lambda_1}(\tilde{x}) \\ \iff \\ \tilde{x} &\in T_{DR\lambda_1}(\tilde{x}), \end{aligned}$$

and therefore $\tilde{x} \in \text{Fix } T_{\text{DR}\lambda_1}$. In conclusion,

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

which proves the claim. \square

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, $T_{\text{DR}\lambda}$, have been established in Section 3. The main idea for convergence goes back to Opial, [23]. 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 [21], 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 $U \subset \mathcal{E}$, $\bar{y} \in \mathcal{Y}$. The mapping Φ is called metrically subregular for \bar{y} on U with constant κ relative to $\Lambda \subset \mathcal{E}$ if*

$$\text{dist}(x, \Phi^{-1}(\bar{y}) \cap \Lambda) \leq \kappa \text{dist}(\bar{y}, \Phi(x)) \quad (26)$$

holds for all $x \in U \cap \Lambda$. When $\bar{y} \in \Phi(\bar{x})$, then Φ is said to be metrically regular at \bar{x} for \bar{y} relative to Λ when there exists a neighborhood U of \bar{x} and a constant κ such that (26) holds for all $x \in U \cap \Lambda$.

When $\Lambda = \mathcal{E}$, the quantifier “relative to” is dropped. The smallest constant κ for which (26) holds is called modulus of metric subregularity.

The abstract result that allows us to quantify the convergence of $T_{\text{DR}\lambda}$ follows. It is a simplified version of [22, Corollary 2.3] which was later refined to show convergence to a specific point in [21, Corollary 1]. The convergence result Theorem 4.2 is later specialized to $T_{\text{DR}\lambda}$ in Theorem 4.11 presented in Section 4.2.

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

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

(ii) for

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

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 $\text{Fix } T \cap \Lambda$, the iterates $x^{j+1} \in Tx^j$ satisfy

$$\text{dist}(x^{j+1}, \text{Fix } T \cap \Lambda) \leq c \text{dist}(x^j, \text{Fix } T \cap \Lambda) \quad \forall x^j \in \bar{S} \quad (27)$$

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

$$\kappa < \sqrt{\frac{1-\alpha}{\epsilon\alpha}}. \quad (28)$$

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

We have already shown in Theorem 3.7 that $T_{\text{DR}\lambda}$ is almost averaged, *with any desired violation constant* $\epsilon > 0$, at its fixed points on certain neighborhoods when A and B are super-regular at a distance. To achieve local linear convergence, inequality (28) must hold, and this is where uniformity of almost averagedness with respect to ϵ is crucial: as long as the mapping $T_{\text{DR}\lambda} - \text{Id}$, or a related mapping (see the discussion below), can be shown to be relatively metrically subregular on a neighborhood of $\text{Fix } T_{\text{DR}\lambda}$ - *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 (28) 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 $T_{\text{DR}\lambda}$. 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 $T_{\text{DR}\lambda}$ mapping itself, that guarantees metric subregularity from properties of the regularity of the sets A and B both individually (e.g. relative super-regularity at a distance) or 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 $T_{\text{DR}\lambda}$ 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 $T_{\text{DR}\lambda}$. 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 $T_{\text{DR}\lambda}$ at $\text{Fix } T_{\text{DR}\lambda}$: Metric Subregularity via Subtransversality

Direct verification of metric subregularity is notoriously difficult and verifying this for $T_{\text{DR}\lambda}$ is no different. In principle, one must show that the coderivative (the generalized Jacobian) of the (multi-valued) $T_{\text{DR}\lambda}$ mapping is injective on neighborhoods of $\text{Fix } T_{\text{DR}\lambda}$ [10, Theorems 4B.1 and 4C.2]. We were unable to compute the coderivative of the $T_{\text{DR}\lambda}$ mapping, let alone determine whether this is injective.

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* [14].¹ This was expanded in [22, Definition 3.2] to account for *inconsistent set feasibility*. Based on this more general notion of subtransversality of non-overlapping sets Luke et al. 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 $\{\Omega_1, \dots, \Omega_m\}$ is subtransversal, and an additional technical assumption is satisfied. We follow this approach here, but for the mapping $T_{\text{DR}\lambda}$. 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 $\{\Omega_1, \Omega_2, \dots, \Omega_m\}$ 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 $\Pi : x = (x_1, x_2, \dots, x_m) \mapsto (x_2, x_3, \dots, x_m, x_1)$ is the permutation mapping on the product space \mathcal{E}^m for $x_j \in \mathcal{E}$ ($j = 1, 2, \dots, m$). Let $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_m) \in \mathcal{E}^m$ and $\bar{y} \in \Psi(\bar{x})$. The collection of sets is said to be subtransversal with constant κ relative to $\Lambda \subset \mathcal{E}^m$ at \bar{x} for \bar{y} if Ψ is metrically subregular for \bar{y} on some neighborhood U of \bar{x} with constant κ relative to Λ .*

In contrast to the original model setting, where $\{\Omega_1, \Omega_2, \dots, \Omega_m\}$ 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 $\{\Omega_1, \Omega_2, \dots, \Omega_m\} \subset \mathcal{E}$ be a subtransversal collection of sets at a point $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_m)$ for $\bar{y} \in \Psi(\bar{x})$ relative to $\Lambda \subset \mathcal{E}^m$ with modulus κ . Then the collection*

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

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

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

$$\tilde{x} = (\bar{x}_1 - g, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_m, \bar{x}_1, \bar{x}_2 - g, \bar{x}_3 - g, \dots, \bar{x}_m - g) \in \mathcal{E}^{2m} \quad (29)$$

for

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

relative to

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

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 $\Pi_{\mathcal{E}}^s$ the permutation mapping on \mathcal{E}^s .

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

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

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 $\Psi = P_{\Omega}(\Pi_{\mathcal{E}}^4) - \Pi_{\mathcal{E}}^4$ 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. $\tilde{y} \in P_{\Omega \times \Omega - (0,0,g,g)}(\Pi_{\mathcal{E}}^4(\tilde{x})) - \Pi_{\mathcal{E}}^4(\tilde{x})$. Let \tilde{x} be defined by (29) then

$$\begin{aligned} & P_{\Omega \times \Omega - (0,0,g,g)}(\Pi_{\mathcal{E}}^4(\tilde{x})) - \Pi_{\mathcal{E}}^4(\tilde{x}) \\ &= (P_{\Omega}(\tilde{x}_2, \tilde{x}_3) - (\tilde{x}_2, \tilde{x}_3), P_{\Omega - (g,g)}(\tilde{x}_4, \tilde{x}_1) - (\tilde{x}_4, \tilde{x}_1)) \\ &= (P_{\Omega}(\bar{x}_2, \bar{x}_1) - (\bar{x}_2, \bar{x}_1), P_{\Omega - (g,g)}(\bar{x}_2 - g, \bar{x}_1 - g) - (\bar{x}_2 - g, \bar{x}_1 - g)) \\ &= (P_{\Omega}(\bar{x}_2, \bar{x}_1) - (\bar{x}_2, \bar{x}_1), P_{\Omega}(\bar{x}_2, \bar{x}_1) - (\bar{x}_2, \bar{x}_1)), \end{aligned} \quad (30)$$

where the last equality holds since $P_{C-g}(x-g) = P_C(x) - g$ for any set C . Then (30) yields

$$\begin{aligned} & P_{\Omega \times \Omega - (0,0,g,g)}(\Pi_{\mathcal{E}}^4(\tilde{x})) - \Pi_{\mathcal{E}}^4(\tilde{x}) \\ &= (P_{\Omega}(\bar{x}_2, \bar{x}_1) - (\bar{x}_2, \bar{x}_1), P_{\Omega}(\bar{x}_2, \bar{x}_1) - (\bar{x}_2, \bar{x}_1)) \\ &\ni (\bar{y}, \bar{y}) = \tilde{y}, \end{aligned}$$

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

It remains to prove that inequality (26) 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:

$$\kappa^2 \text{dist}^2 \left(P_{\Omega \times \Omega - (0,0,g,g)} \left(\Pi_{\mathcal{E}}^4 \left(\begin{pmatrix} x_1 - g \\ x_2 \\ x_1 \\ x_2 - g \end{pmatrix} \right) - \Pi_{\mathcal{E}}^4 \left(\begin{pmatrix} x_1 - g \\ x_2 \\ x_1 \\ x_2 - g \end{pmatrix}, \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} \right) \right) \right) \quad (31)$$

$$= \kappa^2 \left(\text{dist}^2 \left(P_{\Omega} \left(\begin{pmatrix} x_2 \\ x_1 \end{pmatrix} \right) - \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}, \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} \right) + \text{dist}^2 \left(P_{\Omega - (g,g)} \left(\begin{pmatrix} x_2 - g \\ x_1 - g \end{pmatrix} \right) - \begin{pmatrix} x_2 - g \\ x_1 - g \end{pmatrix}, \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} \right) \right), \quad (32)$$

by rewriting the distance on \mathcal{E}^4 in terms of the distance on \mathcal{E}^2 . Using again that $P_{C-g}(x-g) = P_C(x) - g$ for an arbitrary set C , (32) ends up as

$$\begin{aligned} & \kappa^2 \left(\text{dist}^2 \left(P_{\Omega} \left(\Pi_{\mathcal{E}}^2 \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) \right) - \Pi_{\mathcal{E}}^2 \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} \right) \right) + \text{dist}^2 \left(P_{\Omega} \left(\begin{pmatrix} x_2 \\ x_1 \end{pmatrix} \right) - \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}, \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} \right) \right) \\ &= 2\kappa^2 \left(\text{dist}^2 \left(P_{\Omega} \left(\begin{pmatrix} x_2 \\ x_1 \end{pmatrix} \right) - \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}, \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} \right) \right) \end{aligned} \quad (33)$$

$$\geq 2 \text{dist}^2 \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \left(P_{\Omega} \left(\Pi_{\mathcal{E}}^2(\cdot) \right) - \Pi_{\mathcal{E}}^2(\cdot) \right)^{-1} \left(\begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} \right) \right), \quad (34)$$

where the last inequality holds by subtransversality of $\{\Omega_1, \Omega_2\}$ at (\bar{x}_1, \bar{x}_2) for (\bar{y}_1, \bar{y}_2) relative to Λ with modulus κ on U . Rewriting (34) 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_{\mathcal{E}}^4 \begin{pmatrix} x_1 - g \\ x_2 \\ x_1 \\ x_2 - g \end{pmatrix} \right) - \Pi_{\mathcal{E}}^4 \begin{pmatrix} x_1 - g \\ x_2 \\ x_1 \\ x_2 - g \end{pmatrix}, \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} \right) \\
& \geq \operatorname{dist}^2 \left(\begin{pmatrix} x_1 \\ x_2 \\ x_1 \\ x_2 \end{pmatrix}, \left\{ \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix} \mid P_{\Omega} \begin{pmatrix} z_2 \\ z_1 \end{pmatrix} - \begin{pmatrix} z_2 \\ z_1 \end{pmatrix} \ni \begin{pmatrix} \bar{y}_1 \\ \bar{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} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} \right\} \right) \\
& = \operatorname{dist}^2 \left(\begin{pmatrix} x_1 \\ x_2 \\ x_1 \\ x_2 \end{pmatrix}, \left\{ \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix} \mid P_{\Omega} \begin{pmatrix} z_2 \\ z_1 \end{pmatrix} - \begin{pmatrix} z_2 \\ z_1 \end{pmatrix} \ni \begin{pmatrix} \bar{y}_1 \\ \bar{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} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} \right\} \right) \\
& = \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} \mid 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} \bar{y}_1 \\ \bar{y}_2 \\ \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} \right\} \right) \\
& = \operatorname{dist}^2 \left(\begin{pmatrix} x_1 - g \\ x_2 \\ x_1 \\ x_2 - g \end{pmatrix}, \left\{ \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix} \mid 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} \bar{y}_1 \\ \bar{y}_2 \\ \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} \right\} \right) \\
& = \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_{\mathcal{E}}^4 (\cdot) \right) - \Pi_{\mathcal{E}}^4 (\cdot) \right)^{-1} \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \bar{y}_1 \\ \bar{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. \square

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 in the following way*

$$(B - g) \times (A - g) \times A \times B \subset \mathcal{E}^4, \quad (35)$$

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

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

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 \tilde{y} emerged from the changed order of B and A in comparison to Lemma 4.4.

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 . Our aim is to show local linear convergence of $T_{\text{DR}\lambda}$ by adapting the approach developed in [22] 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 $T_{\text{DR}\lambda}$ mapping. As established in Theorem 3.13 and Corollary 3.15 fixed points \bar{x} of $T_{\text{DR}\lambda}$ at which $T_{\text{DR}\lambda}$ is single-valued can be described as

$$\{\bar{x}\} = P_B(\bar{x}) - \frac{\lambda}{1-\lambda} (P_B(\bar{x}) - P_A(P_B(\bar{x}))),$$

which is an element in $B - \frac{\lambda}{1-\lambda}g$ when $g = P_B(\bar{x}) - P_A(P_B(\bar{x}))$. Thus, locally $B - \frac{\lambda}{1-\lambda}g$ contains fixed points of $T_{\text{DR}\lambda}$. 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

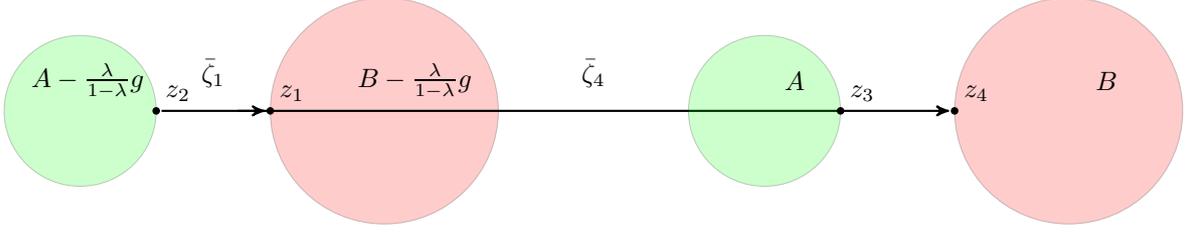


Figure 5: Framework for the convergence analysis illustrated in \mathcal{E} .

$$W_0(g) := \left\{ u = (u_1, u_2, u_3, u_4) \in \mathcal{E}^4 \mid \right. \\ \left. u_1 \in P_{B - \frac{\lambda}{1-\lambda}g}(u_2), u_2 \in P_{A - \frac{\lambda}{1-\lambda}g}(u_3), u_3 \in P_A(u_4), u_4 \in P_B(u_1) \right\}. \quad (36)$$

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 $P_{B - \frac{\lambda}{1-\lambda}g} P_{A - \frac{\lambda}{1-\lambda}g} P_A P_B$. 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

$$\mathcal{Z}(x, g) := \left\{ \zeta := z - \Pi z \mid z \in W_0(g) \subset \mathcal{E}^4, z_1 = x \right\}. \quad (37)$$

The last set to introduce is

$$W(\bar{\zeta}) := \left\{ u \in \mathcal{E}^4 \mid u - \Pi u = \bar{\zeta} \right\}. \quad (38)$$

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 $T_{\text{DR}\lambda}$.

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 $T_{\text{DR}\lambda}$ 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 $T_{\text{DR}\lambda}$ at which $T_{\text{DR}\lambda}$ 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 $T_{\text{DR}\lambda}$ on the individual sets. This enables us to distinguish different fixed points of $T_{\text{DR}\lambda}$ according to their respective difference vectors.

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

- (i) $z_3 - z_4 \in P_A(P_B(z_1)) - P_B(z_1) = \{-g\}$
- (ii) $z_4 - z_1 \in P_B(z_1) - (P_B(z_1) - \frac{\lambda}{1-\lambda}g) = \left\{ \frac{\lambda}{1-\lambda}g \right\},$

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

- (iii) $z_1 - z_2 \in P_B(\bar{u}) - \frac{\lambda}{1-\lambda}g - P_{A - \frac{\lambda}{1-\lambda}g}(z_3)$

$$= P_B(\bar{u}) - \frac{\lambda}{1-\lambda}g - P_A(z_3 + \frac{\lambda}{1-\lambda}g) + \frac{\lambda}{1-\lambda}g$$

$$= P_B(\bar{u}) - z_3$$

$$= P_B(\bar{u}) - P_A(P_B(\bar{u}))$$

$$= \{g\};$$

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

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

$$T_{\bar{\zeta}} : \mathcal{E}^4 \rightrightarrows \mathcal{E}^4 : u \mapsto \{(u_1^+, u_1^+ - \bar{\zeta}_1, u_1^+ - \bar{\zeta}_1 - \bar{\zeta}_2, u_1^+ - \bar{\zeta}_1 - \bar{\zeta}_2 - \bar{\zeta}_3) \mid u_1^+ \in T_{\text{DR}\lambda}u_1\},$$

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

$$T_{\bar{\zeta}} : \mathcal{E}^4 \rightrightarrows \mathcal{E}^4 : u \mapsto \{(u_1^+, u_1^+ - \bar{\zeta}_1, u_1^+ - \bar{\zeta}_1 - \bar{\zeta}_2, u_1^+ + \bar{\zeta}_4) \mid u_1^+ \in T_{\text{DR}\lambda}u_1\}. \quad (39)$$

The mapping $T_{\bar{\zeta}}$ represents the iterates of $T_{\text{DR}\lambda}$ 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 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 $T_{\text{DR}\lambda}$.

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

(i) $T_{\bar{\zeta}}$ maps $W(\bar{\zeta})$ to itself. Moreover $u \in \text{Fix } T_{\bar{\zeta}}$ if and only if $u \in W(\bar{\zeta})$ with $u_1 \in \text{Fix } T_{\text{DR}\lambda}$.

(ii)

$$\Psi_g^{-1}(\bar{\zeta}) \cap W(\bar{\zeta}) \cap \mathcal{N} \subseteq \Phi_{\bar{\zeta}}^{-1}(0) \cap W(\bar{\zeta}),$$

$$\text{where } \mathcal{N} := \left\{ z \in \mathcal{E}^4 \mid P_A(z_4 + \frac{\lambda}{1-\lambda}g) = z_3 \right\}.$$

(iii) If the distance is with respect to the Euclidean norm, then $\text{dist}(0, \Phi_{\bar{\zeta}}(u)) = 2 \text{dist}(u_1, T_{\text{DR}\lambda}u_1)$ 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 assure 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 $u \in \text{Fix } T_{\bar{\zeta}}$,

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

which proves the rest of (i).

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

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

which is equivalent to

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

This implies

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

The mapping $\Phi_{\bar{\zeta}}(z) = T_{\bar{\zeta}}z - z$ has the image $(0, 0)$ if $z_1 \in \text{Fix } T_{\text{DR}\lambda} z_1$. 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 $P_A(R_B(z_1)) \ni P_A(2z_4 - z_1) = P_A\left(z_4 + \frac{\lambda}{1-\lambda}g\right) = z_3$. Inserting this in $T_{\text{DR}\lambda}(z_1)$ yields

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

since $\bar{\zeta}$ is generated by a fixed point of $T_{\text{DR}\lambda}$. Thus $z_1 \in \text{Fix } T_{\text{DR}\lambda}$, 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:

$$\begin{aligned} &\text{dist}(0, \Phi_{\bar{\zeta}}(u)) \\ &= \text{dist}(0, T_{\bar{\zeta}}u - u) \\ &= \sqrt{\text{dist}^2(0, T_{\text{DR}\lambda}u_1 - u_1) + \text{dist}^2(0, T_{\text{DR}\lambda}u_1 - \bar{\zeta}_1 - u_2) + \cdots + \text{dist}^2\left(0, T_{\text{DR}\lambda}u_1 - \sum_{i=1}^3 \bar{\zeta}_i - u_4\right)} \\ &= \sqrt{4 \text{dist}^2(0, T_{\text{DR}\lambda}u_1 - u_1)} \\ &= 2 \text{dist}(0, T_{\text{DR}\lambda}u_1 - u_1). \end{aligned}$$

□

We are now ready for the main result of this subsection. We show that the mapping $T_{\bar{\zeta}} - \text{Id}$ 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 $T_{\text{DR}\lambda}$.

Proposition 4.8 (metric subregularity of $T_{\bar{\zeta}}$ by subtransversality). *Let $\lambda \in (0, 1)$, $\bar{x} \in \text{Fix } T_{\text{DR}\lambda}$ with $T_{\text{DR}\lambda}$ being single-valued at \bar{x} and set $g := P_B(\bar{x}) - P_A(P_B(\bar{x}))$. Furthermore, let $\bar{\zeta} \in \mathcal{Z}(\bar{x}, g)$ and $\bar{u} = (\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4) \in W_0(g)$ satisfy $\bar{\zeta} = \bar{u} - \Pi\bar{u}$ with $\bar{u}_1 = \bar{x}$. Let $T_{\bar{\zeta}}$ be defined by (39) and define $\Phi_{\bar{\zeta}} := T_{\bar{\zeta}} - \text{Id}$. Suppose the following hold:*

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

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

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

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

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

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

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 [22, 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, super-regularity is enough to show convergence of the Douglas-Rachford algorithm, [24, Theorem 6.8] [13, Theorem 3.18]. For alternating projections one only needs transversality at points of intersection and super-regularity of *one* of the sets [15, Theorem 5.16]. In any case, the additional assumption (ii), is not needed when the 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 $\bar{x} \in A \cap B \subset \text{Fix } T_{\text{DR}\lambda}$ we have $g := P_B(\bar{x}) - P_A(P_B(\bar{x})) = 0$. Furthermore, let $\bar{\zeta} \in \mathcal{Z}(\bar{x}, g)$. Then (ii) in Proposition 4.8 is always satisfied on $\Lambda \subset W(\bar{\zeta})$ with $\sigma = \frac{1}{\sqrt{2\lambda}}$.*

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

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

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

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

and thus

$$\text{dist}^2(0, \Phi_{\bar{\zeta}}(u)) = \text{dist}^2(0, T_0(u) - u) = 4 \text{dist}^2(0, \lambda(P_A(u_1) - u_1)). \quad (40)$$

On the other hand

$$\begin{aligned} \text{dist}^2(\bar{\zeta}, \Psi_g(u)) &= \text{dist}^2((0, 0, 0, 0), \Psi_0(u)) \\ &= \text{dist}^2((0, 0, 0, 0), P_{\Omega_0}\Pi(u) - \Pi(u)) \\ &= 2 \text{dist}^2(0, P_A(u_1) - u_1), \end{aligned} \quad (41)$$

since $\Omega_0 = B \times A \times A \times B$. Combining (41) and (40) yields (ii) in Proposition 4.8 with $\sigma = \frac{1}{\sqrt{2\lambda}}$. \square

4.2 Local Linear Convergence of $T_{\text{DR}\lambda}$

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

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

Proof. By definition (37) $\mathcal{Z}(x, g)$ is given by

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

for

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

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

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 Section 3 and Section 4.

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

$$S_0 := \{x \in \text{Fix } T_{\text{DR}\lambda} \mid \{g\} = P_B(x) - P_A(P_B(x))\} \quad (42)$$

and let

$$S_j := \left(S_0 - \sum_{i=1}^{j-1} \bar{\zeta}_i \right) \quad (j = 1, 2, 3, 4). \quad (43)$$

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

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

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 $\bar{u} = (\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4) \in S$:

(i) for all $(\bar{u}_3, \bar{u}_4) \in S_3 \times S_4$, the collection of sets $\{A, B\}$ is subtransversal at (\bar{u}_3, \bar{u}_4) for $(\bar{u}_3, \bar{u}_4) - \Pi(\bar{u}_3, \bar{u}_4)$ relative to $\Lambda' := \left\{ u = (u_1, u_2) \in \mathcal{E}^2 \mid \left(u_2 - \frac{\lambda}{1-\lambda} g, u_1 - \frac{\lambda}{1-\lambda} g, u_1, u_2 \right) \in \Lambda \right\}$ with constant κ on the neighborhood $U_A \times U_B$;

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

$$\text{dist}(\bar{\zeta}, \Psi_g(u)) \leq \sigma \text{dist}(0, \Phi_{\bar{\zeta}}(u)) \quad (44)$$

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

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

Then there exists an $\epsilon' \leq \epsilon$ 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 $(u^k)_{k \in \mathbb{N}}$ generated by $u^{k+1} \in T_{\bar{\zeta}} u^k$ seeded by a point $u^0 \in W(\bar{\zeta}) \cap U'$ with $u_1^0 \in U'_1 \cap \left(B - \frac{\lambda}{1-\lambda} g \right)$ satisfies

$$\text{dist}(u^{k+1}, \text{Fix } T_{\bar{\zeta}} \cap S) \leq c \text{dist}(u^k, S) \quad (\forall k \in \mathbb{N})$$

for

$$c := \sqrt{1 + \epsilon' - \frac{1}{\bar{\kappa}^2}} < 1 \quad (45)$$

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

$$\text{dist}(u_1^k, \tilde{u}_1) \rightarrow 0$$

at least R -linearly with rate $c < 1$. If $\text{Fix } T_{\text{DR}\lambda} \cap S_1$ 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 $T_{\text{DR}\lambda}$.

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 $T_{\text{DR}\lambda}$ 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 $\left\{B - \frac{\lambda}{1-\lambda}g, A - \frac{\lambda}{1-\lambda}g, A, B\right\}$ 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 $U(A, \epsilon, \bar{x})$ and $U(B, \epsilon, \bar{x})$ and violations ϵ_A and ϵ_B of (4) 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 $T_{\text{DR}\lambda}$ is almost firmly nonexpansive at each $\tilde{y} \in S_1$ on U_B by Theorem 3.7 since the 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 $T_{\text{DR}\lambda}$ 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 $\bar{u} = (\bar{u}_1, \bar{u}_2, \bar{u}_3, \bar{u}_4) \in S$, the collection of sets

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

is subtransversal at \bar{u} for $\bar{\zeta} := \bar{u} - \Pi\bar{u}$ 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 $U \cap \tilde{\Lambda} \subset \mathcal{N} := \left\{z \in \mathcal{E}^4 \mid P_A(z_4 + \frac{\lambda}{1-\lambda}g) = z_3\right\}$ by Theorem 3.7(iv) 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 (27) holds.

What remains is to show that (28) holds, which would imply at least R-linear convergence. To achieve this choose some $\epsilon' > 0$ with $\epsilon' < \epsilon$ such that (28) 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

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

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 (28) holds which implies at least R-linear convergence to \tilde{u} . Since $\tilde{u}_1 \in \text{Fix } T_{\text{DR}\lambda} \cap S_1$, this completes the proof. \square

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. $\lambda \rightarrow 1$. Then $\frac{\lambda}{1-\lambda} \rightarrow +\infty$. That is, the phantom sets recede to the horizon in the direction $-g$.
3. $\lambda \rightarrow 0$. 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 is nothing more than alternating projections between the sets A and B . At $\lambda = 0$, however, $\text{Fix } T_{\text{DR}\lambda} = B$, 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, [22, 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). [22, Proposition 3.3] Let \mathcal{E}^m be endowed with the 2-norm, that is, $\|(x_1, \dots, x_m)\|_2 = \left(\sum_{j=1}^m \|x_j\|_{\mathcal{E}}^2\right)^{1/2}$. A collection $\{\Omega_1, \Omega_2, \dots, \Omega_m\}$ of nonempty and closed subsets of \mathcal{E} is subtransversal relative to

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

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

$$\text{dist}(u, \cap_{j=1}^m \Omega_j) \leq \kappa' \max_{j=1, \dots, m} \text{dist}(u, \Omega_j), \quad \forall u \in U'. \quad (46)$$

Conversely, if $\{\Omega_1, \Omega_2, \dots, \Omega_m\}$ is subtransversal relative to Λ at \bar{x} for $\bar{y} = 0$ with constant κ , then (46) is satisfied with any constant κ' for which $\kappa \leq \kappa'$.

The property in (46) 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 $\bar{x} \in A \cap B$ on $\mathbb{B}_\delta(\bar{x})$.
- (ii) The collection of sets is linearly regular at $\bar{x} \in A \cap B$.

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 \bar{x} be a point in $\text{Fix } T_{\text{DR}\lambda}$. Moreover, let both A and B be closed and convex with nonempty intersection, i.e. $A \cap B \neq \emptyset$ and therefore $\text{Fix } T_{\text{DR}\lambda} = A \cap B$. Then $\{g\} = P_B(\bar{x}) - P_A(P_B(\bar{x})) = 0$ and $\{\bar{\zeta}\} = \mathcal{Z}(\bar{x}, g) = \{0\}$ ($\bar{\zeta} = (\bar{\zeta}_1, \bar{\zeta}_2, \bar{\zeta}_3, \bar{\zeta}_4) \in \mathcal{E}^4$). Define the set

$$S_0 := \text{Fix } T_{\text{DR}\lambda} = A \cap B.$$

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

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

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

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

for

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

with κ by (i). Consequently, $\text{dist}(x^k, \tilde{x}) \rightarrow 0$ for some $\tilde{x} \in \text{Fix } T_{\text{DR}\lambda}$ at least R -linearly with rate $c < 1$. If $\text{Fix } T_{\text{DR}\lambda}$ 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 [23] 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 $\text{Fix } T_{\text{DR}\lambda} = A \cap B$.*

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 also the gap vector g for any fixed point in $\text{Fix } T_{\text{DR}\lambda}$. Thus, $S_0 = \text{Fix } T_{\text{DR}\lambda}$. $\text{Fix } T_{\text{DR}\lambda} = A \cap B$ by Theorem 3.13. With these observations we get immediately that the sets involved in Theorem 4.11 simplify to the following

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

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 [3, Theorem 3.14]). Additionally the projection is firmly nonexpansive, [3, Proposition 4.8], and thus the reflector nonexpansive, [3, Proposition 4.2], which implies that $T_{\text{DR}\lambda}$ 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 super-regular 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 $\mathbb{B}_\delta(\bar{x})$. 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. \square

Remark 4.18 (linking our results to already existing literature). *As noted in the introduction, the works [9, 13, 16, 22, 24] all analyze the Douglas-Rachford algorithm for consistent nonconvex feasibility. In [22] 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 [16], which addresses global convergence guarantees for consistent feasibility. The assumptions of that paper, namely compactness and semi-algebraicity (not to mention nonempty intersection) are different than the notions that we work with. Certainly compactness is a regularity assumption, as is semi-algebraicity or its more general Kurdyka-Lojasiewicz-type regularity, but these notions serve a different purpose. Indeed, even convex sets need not be semi-algebraic or compact. This suggests that Kurdyka-Lojasiewicz-type 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 super-regular 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 not only able to present existence results on neighborhoods where we get local convergence, but we are able 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 [20] to be the fundamental geometry for phase retrieval and source location problems. Affine subspaces are included as spheres centered at infinity.

There are 5 distinct cases to consider: 1. intersecting circles, 2. nonintersecting separable circles, 3. nonintersecting, nonseparable, nonconcentric circles, 4. nonintersecting 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,

$$\begin{aligned} \kappa &> \frac{\text{dist}(u, \Psi_g^{-1}(\bar{\zeta}) \cap W(\bar{\zeta}))}{\text{dist}(\bar{\zeta}, \Psi_g(u))}, \\ \sigma &> \frac{\text{dist}(\bar{\zeta}, \Psi_g(u))}{\text{dist}(0, \Phi_{\bar{\zeta}}(u))}, \end{aligned}$$

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. Without loss of generality we can restrict the analysis to the following setting*

$$\begin{aligned} A &:= \{x \in \mathbb{R}^2 \mid \|x\| = 1\} \\ B &:= \{x \in \mathbb{R}^2 \mid \|x - (0, a)\| = R\}, \end{aligned}$$

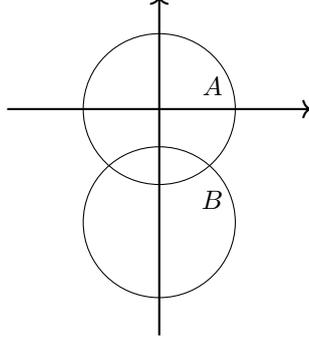


Figure 6: Example 5.1 for $a = -1.5$ and $R = 1$

where $a \in \mathbb{R} \setminus \{0\}$ and $R \in (\min_{y \in A} \text{dist}((0, a), y), \max_{y \in A} \text{dist}((0, a), y))$. 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) $S_0 := \left\{ \left(+\sqrt{1 - \left(\frac{1 - R^2 + a^2}{2a} \right)^2}, \frac{1 - R^2 + a^2}{2a} \right) \right\} \in \text{Fix } T_{\text{DR}\lambda}$

(ii) In $\mathbb{R}_+ \times \mathbb{R}$ there is a unique fixed point. $\bar{x} = (\bar{u}, \bar{u}, \bar{u}, \bar{u})$, where $\bar{u} = \left(+\sqrt{1 - \left(\frac{1 - R^2 + a^2}{2a} \right)^2}, \frac{1 - R^2 + a^2}{2a} \right)$.

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

(iv) The sets A and B satisfy the assumptions of Theorem 3.7 at \bar{u} with neighborhoods U_1 and U_2 being open balls around \bar{u} , that is $\mathbb{B}_\delta(\bar{u})$, 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 in such instances. Our aim is to use this statement to prove that Example 5.1 satisfies (46).

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

$$\text{dist}(u, A \cap B) \leq \kappa' \max \{ \text{dist}(u, A), \text{dist}(u, B) \},$$

simplifies to

$$\text{dist}(u, A \cap B) \leq \kappa' \text{dist}(u, B),$$

which we reformulate in the following to

$$\|u - P_{A \cap B}(u)\| \leq \kappa' \|u - P_B(u)\|. \quad (47)$$

Note that (47) implies (46) since $u \in A$ and thus implies linear regularity.

Next, we show (47).

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

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

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

Inserting this in the above expression yields

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

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

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

Equivalently

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

Thus,

$$\begin{aligned} 1 - \frac{R}{\|u - (0, a)\|} &\geq 1 - \frac{R}{\max_{y \in A} \text{dist}((0, a), y)}, \quad \text{and} \\ \|u - (0, a)\| \left(1 - \frac{R}{\|u - (0, a)\|}\right) &\geq \min_{y \in A} \text{dist}((0, a), y) \left(1 - \frac{R}{\max_{y \in A} \text{dist}((0, a), y)}\right) \\ \Rightarrow \frac{1}{\|u - (0, a)\| \left(1 - \frac{R}{\|u - (0, a)\|}\right)} &\leq \frac{1}{\min_{y \in A} \text{dist}((0, a), y) \left(1 - \frac{R}{\max_{y \in A} \text{dist}((0, a), y)}\right)} =: \kappa'. \end{aligned}$$

Since $R \in (\min_{y \in A} \text{dist}((0, a), y), \max_{y \in A} \text{dist}((0, a), y))$, 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.

Thus, the assumptions of Theorem 4.11 are satisfied and the relaxed Douglas-Rachford algorithm converges locally linearly to \bar{u} with rate $1 > c > \sqrt{1 - \frac{\lambda^2}{\kappa^2}}$ as long as the starting point is close enough to \bar{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 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 $T_{\text{DR}\lambda}$ does not contain any point of the form $(0, y)$ for $y \in \mathbb{R}$.

Example 5.2 (nonintersecting separable circles). This example 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

$$A := \{x \in \mathbb{R}^2 \mid \|x\| = 1\}$$

$$\text{and } B := \{x \in \mathbb{R}^2 \mid \|x - (2 + R, 0)\| = R\}.$$

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

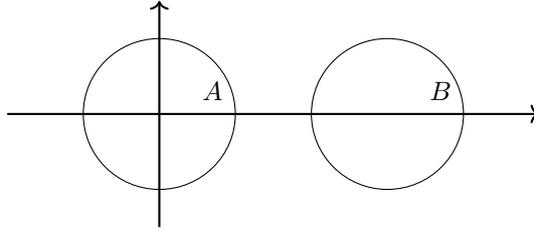


Figure 7: Example 5.2 for $R = 1$

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

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) $S_0 := \text{Fix } T_{\text{DR}\lambda} = \{\bar{u}\}$.
- (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

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

- (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 \bar{u} with rate $1 > c > \sqrt{1 - \frac{1}{(\kappa\sigma)^2}}$ as long as the starting point is close enough to \bar{u} .

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

$$A := \{x \in \mathbb{R}^2 \mid \|x\| = 1\}$$

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

Our analysis considers the fixed point

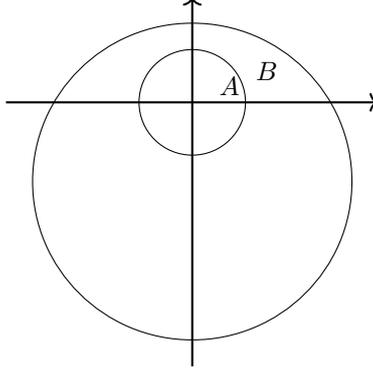


Figure 8: Example 5.3 for $R = 1$

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

of $T_{\text{DR}\lambda}$ 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) $S_0 := \{\bar{u}\} \in \text{Fix } T_{\text{DR}\lambda}$.

(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

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

(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 \bar{u} with rate $1 > c > \sqrt{1 - \frac{1}{(\kappa\sigma)^2}}$ as long as as long as the starting point is close enough to \bar{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*

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

where we restrict R to be strictly greater than 1, i.e. $R > 1^2$. Our analysis focuses on the fixed point

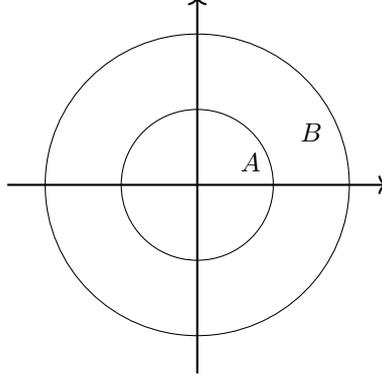


Figure 9: Example 5.4 for $R = 2$

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

of $T_{\text{DR}\lambda}$ on A and B . Note that it is enough to consider \bar{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 2 circles touching at a single point. Let $R > 0$ and define*

$$A := \{x \in \mathbb{R}^2 \mid \|x\| = 1\}$$

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

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

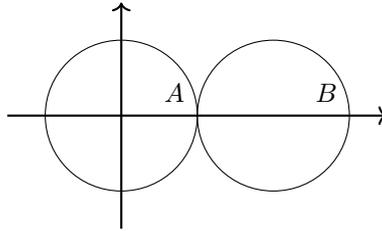


Figure 10: Example 5.5 for $R = 1$

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

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

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

- (i) $S_0 := \{\bar{u}\} \in \text{Fix } T_{\text{DR}\lambda}$.
- (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 \bar{u} with neighborhoods U_1 and U_2 being open balls around \bar{u} , that is $\mathbb{B}_\delta(\bar{u})$, 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), yield a value of

$$\frac{2R}{(R+1)b} \tag{48}$$

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

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

As $b \rightarrow 0$ (in other words, for points close to \bar{u}) the ratio (48) tends to ∞ . This implies that Example 5.5 cannot be linearly regular at the point $\bar{u} = (1, 0)$ 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 [21, Theorem 2], we conclude that $T_{\text{DR}\lambda}$ 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 Example 5.2-Example 5.4 that the presence of subtransversality in the inconsistent case can come as a 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.

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. *Journal of Functional Analysis*, 262(1):400–408, 2012.
- [2] A. Bakan, F. Deutsch, and W. Li. Strong CHIP, normality, and linear regularity of convex sets. *Transactions of the American Mathematical Society*, 357(10):3831–3863, 2005.
- [3] H. H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books Math./Ouvrages Math. SMC. Springer, New York, 2011.
- [4] H. H. Bauschke, P. L. Combettes, and D. R. Luke. Finding best approximation pairs relative to two closed convex sets in Hilbert spaces. *Journal of Approximation Theory*, 127(2):178–192, 2004.
- [5] H. H. Bauschke and W. M. Moursi. The Douglas–Rachford algorithm for two (not necessarily intersecting) affine subspaces. *SIAM Journal on Optimization*, 26:968–985, 2016.
- [6] H.H. Bauschke and D. Noll. On the local convergence of the Douglas–Rachford algorithm. *Archiv der Mathematik (Basel)*, 102:589–600, 2014.
- [7] F. E. Browder. Existence and approximation of solutions of nonlinear variational inequalities. *Proceedings of the National Academy of Sciences*, 56(4):1080–1086, 1966.

- [8] A. Daniilidis, D. R. Luke, and M. K. Tam. *Splitting Algorithms, Modern Operator Theory and Applications*, chapter Characterizations of Super-regularity and its Variants. Springer, 2018.
- [9] M. N. Dao and H. M. Phan. Linear convergence of the generalized Douglas-Rachford algorithm for feasibility problems. *Journal of Global Optimization*, 72(3):443–474, 2018.
- [10] A. L. Dontchev and R. T. Rockafellar. *Implicit Functions and Solution Mappings*. Springer-Verlag, Dordrecht, second edition, 2014.
- [11] J. Douglas and H. H. Rachford. On the numerical solution of heat conduction problems in two and three space variables. *Transactions of the American Mathematical Society*, 82:421–439, 1956.
- [12] J. Eckstein and D. P. Bertsekas. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming*, 55(1-3):293–318, 1992.
- [13] R. Hesse and D. R. Luke. Nonconvex notions of regularity and convergence of fundamental algorithms for feasibility problems. *SIAM Journal on Optimization*, 23(4):2397–2419, 2013.
- [14] A. Y. Kruger, D. R. Luke, and N. H. Thao. Set regularities and feasibility problems. *Mathematical Programming*, 168(1-2):279–311, 2018.
- [15] A. S. Lewis, D. R. Luke, and J. Malick. Local linear convergence for alternating and averaged nonconvex projections. *Foundations of Computational Mathematics*, 9(4):485–513, 2009.
- [16] G. Li and T. K. Pong. Douglas–Rachford splitting for nonconvex optimization with application to nonconvex feasibility problems. *Mathematical Programming*, 159(1-2):371–401, 2016.
- [17] P.-L. Lions and B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis*, 16(6):964–979, 1979.
- [18] D. R. Luke. Relaxed averaged alternating reflections for diffraction imaging. *Inverse problems*, 21(1):37, 2005.
- [19] D. R. Luke. Finding best approximation pairs relative to a convex and prox-regular set in a Hilbert space. *SIAM Journal on Optimization*, 19(2):714–739, 2008.
- [20] D. R. Luke, S. Sabach, and M. Teboulle. Optimization on spheres: Models and proximal algorithms with computational performance comparisons. *arXiv:1810.02893*, 2018.
- [21] D. R. Luke, M. Teboulle, and N. H. Thao. Necessary conditions for linear convergence of iterated expansive, set-valued mappings. *Mathematical Programming*, pages 1–31, 2018.
- [22] D. R. Luke, N. H. Thao, and M. K. Tam. Quantitative convergence analysis of iterated expansive, set-valued mappings. *Mathematics of Operations Research*, 43:1143–1176, 2018.
- [23] Z. Opial. Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bulletin of the American Mathematical Society*, 73(4):591–597, 1967.
- [24] H. Phan. Linear convergence of the Douglas–Rachford method for two closed sets. *Optimization*, 65(2):369–385, 2016.
- [25] J. von Neumann. *Functional Operators, Volume 2: The Geometry of Orthogonal Spaces.*, volume 2. Princeton University Press, 1951.