Necessary conditions for linear convergence of iterated expansive, set-valued mappings

Luke, D. Russell; Teboulle, Marc; Nguyen, Thao

DOI
10.1007/s10107-018-1343-8

Publication date
2018

Document Version
Accepted author manuscript

Published in
Mathematical Programming

Citation (APA)
Luke, D. R., Teboulle, M., & Nguyen, T. (2018). Necessary conditions for linear convergence of iterated expansive, set-valued mappings. *Mathematical Programming*, *180 (2020)*, 1-31. https://doi.org/10.1007/s10107-018-1343-8

Noname manuscript No.
(will be inserted by the editor)

Necessary conditions for linear convergence of iterated expansive, set-valued mappings

D. Russell Luke , Marc Teboulle and Nguyen H. Thao

the date of receipt and acceptance should be inserted later

Abstract We present necessary conditions for monotonicity of fixed point iterations of mappings that may violate the usual nonexpansive property. Notions of linear-type monotonicity of fixed point sequences – weaker than the more well-known Fejér monotonicity – are shown to imply *metric subregularity*. This, together with the almost averaging property recently introduced by Luke, Tam and Thao [25], guarantees linear convergence of the sequence of fixed point iterations to a fixed point. We specialize these results to the alternating projections iteration where the metric subregularity property takes on a distinct geometric characterization of sets at points of intersection called *subtransversality*. Subtransversality is shown to be necessary for linear convergence of alternating projections for consistent feasibility.

2010 Mathematics Subject Classification: Primary 49J53, 65K10 Secondary 49K40, 49M05, 49M27, 65K05, 90C26.

Keywords: Almost averaged mappings, averaged operators, calmness, cyclic projections, elemental regularity, feasibility, Fejér monotone, fixed points, fixed point iteration, metric regularity, metric subregularity, nonconvex, nonexpansive, subtransversality, transversality

---

Institut für Numerische und Angewandte Mathematik, Universität Göttingen, 37083 Göttingen, Germany. DRL was supported in part by German Israeli Foundation Grant G-1253-304.6 and Deutsche Forschungsgemeinschaft Research Training Grant 2088 TP-B5. E-mail: r.luke@math.uni-goettingen.de

School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel. MT was supported by German Israeli Foundation Grant G-1253-304.6 and Israel Science Foundation ISF Grant 1844-16. E-mail: teboulle@post.tau.ac.il.

Delft Center for Systems and Control, Delft University of Technology, 2628CD Delft, The Netherlands. Department of Mathematics, Teacher College, Cantho University, Cantho, Vietnam. NHT was supported by German Israeli Foundation Grant G-1253-304.6. E-mail: h.t.nguyen-3@tudelft.nl, h.nguyen@math.uni-goettingen.de, nhthao@ctu.edu.vn

Address(es) of author(s) should be given

1 Introduction

In recent years there has been a lot of progress in determining ever weaker conditions to guarantee local linear convergence of elementary fixed point algorithms, with particular attention given to the method of alternating projections and the Douglas-Rachford iteration [6, 11–13, 23–25, 28, 31]. These works beg the question: what are necessary conditions for linear convergence? We shed some light on this question for iterations generated by not necessarily nonexpansive fixed point mappings and show how our theory specializes for the alternating projections iteration in nonconvex and convex settings.

Our work builds upon the terminology and theoretical framework established in [25]. As much as possible, we have tried to make the present analysis self-contained, but it is not possible to reproduce all the results taken from [25]. After introducing basic notation and definitions in Section 2, we clarify first what we mean by linear convergence, since there are many ways in which a sequence can behave linearly with respect to the set of fixed points. We introduce a generalization of Fejér monotonicity, namely *linear monotonicity* (Definition 2) which is central to our development. We also introduce another generalization, *linearly extendible sequences* in Definition 3, that concerns sequences which can be viewed as the subsequence of some monotone sequence. This is key to the application to alternating projections studied in Sections 4 and 5. In Section 3 we lay the groundwork for the first main result on necessary conditions for linearly monotone fixed point iterations with respect to $\mathsf{Fix}\, T$ (Theorem 2). The result states that *metric subregularity* (Definition 6) is necessary for linearly monotone fixed point iterations. If in addition the fixed point operator T is *almost averaged* at points in $\mathsf{Fix}\, T$ (Definition 5), then metric subregularity is necessary for *linear convergence* of the iterates to a point in $\mathsf{Fix}\, T$ (Corollary 1). Sections 4 and 5 are specializations to the case of alternating projections for *consistent feasibility*. In this setting metric subregularity takes on the more directly geometric interpretation as *subtransversality* of the sets at common points (Definition 8). Theorem 4 establishes the necessity of subtransversality for alternating projections iterations to be linearly monotone with respect to common points. Corollary 2 then shows that for sets with a certain *elemental subregularity* (Definition 7) subtransversality is necessary and sufficient for linear monotonicity of the sequence. For sequences that are R-linearly convergent to a fixed point and satisfy a subsequential linear monotonicity property (condition (28)), Theorem 5 shows that subtransversality is also necessary. Subtransversality is also shown to be necessary for sequences to have linearly extendible subsequences (Theorem 6). These results correspond to our observation in Proposition 13 that subtransversality has appeared in one form or another in all sufficient conditions for linear monotonicity or convergence of alternating projections for consistent feasibility that have appeared previously in the literature. In Section 5 these results are further specialized to the case of convex feasibility. We show in Theorems 8 and 9 that metric subregularity of some form is necessary and sufficient for local and global linear convergence of alternating projections. Moreover, we show in Proposition

14 that R-linear convergence of the sequences in this setting is equivalent to linear monotonicity of the sequence with respect to points of intersection. For Q-linear convergence, we show that linear extendability is necessary (Proposition 15).

Based on the results obtained here we conjecture that, for alternating projections applied to *inconsistent* feasibility, subtransversality as extended in [21, Definition 3.2] is also necessary for R-linear convergence of the iterates to fixed points.

2 Notation and basic definitions

Throughout our discussion $\mathbb{E}$ is a Euclidean space. Given a subset $A \subset \mathbb{E}$, $\operatorname{dist}(x, A)$ stands for the distance from a point $x \in \mathbb{E}$ to A: $\operatorname{dist}(x, A) := \inf_{a\in A} \|x - a\|$. The *projector* onto the set A, $P_A : \mathbb{E} \rightrightarrows A$, is central to algorithms for feasibility and is defined by

$$P_A x := \underset{a\in A}{\operatorname{argmin}} \|a - x\| .$$

A *projection* is a selection from the projector. This exists for any closed set $A \subset \mathbb{E}$, as can be deduced by the continuity and coercivity of the norm. Note that the projector is not, in general, single-valued, and indeed uniqueness of the projector defines a type of regularity of the set A: local uniqueness characterizes *prox-regularity* [32] while in finite dimensional settings global uniqueness characterizes convexity [9].

Given a subset $A \subset \mathbb{E}$ and a point $\overline{x} \in A$, the *Fréchet, proximal and limiting normal cones* to A at $\overline{x}$ are defined, respectively, as follows:

$$\widehat{N}_A(\overline{x}) := \left\{ v \in \mathbb{E} \mid \limsup_{x \overset{A}{\to} \overline{x},\, x\neq\overline{x}} \frac{\langle v, x - \overline{x}\rangle}{\|x - \overline{x}\|} \le 0 \right\},$$
$$N_A^{\text{prox}}(\overline{x}) := \operatorname{cone}\left(P_A^{-1}(\overline{x}) - \overline{x}\right),$$
$$N_A(\overline{x}) := \operatorname*{Lim\,sup}_{x \overset{A}{\to} \overline{x}} N_A^{\text{prox}}(x) := \left\{ v = \lim_{k\to\infty} v_k \mid v_k \in N_A^{\text{prox}}(x_k),\ x_k \overset{A}{\to} \overline{x} \right\}.$$

In the above, $x \overset{A}{\to} \overline{x}$ means that $x \to \overline{x}$ with $x \in A$.

Our other basic notation is standard; cf. [10, 26, 33]. The open unit ball in a Euclidean space is denoted $\mathbb{B}$. $\mathbb{B}_\delta(x)$ stands for the open ball with radius $\delta > 0$ and center x; $\mathbb{B}_\delta$ is the open ball of radius δ centered at the origin.

To quantify convergence of sequences and fixed point iterations, we focus primarily on linear convergence, though sublinear convergence can also be handled in this framework. Linear convergence, however, can come in many forms. We list the more common notions next.

Definition 1 (R- and Q-linear convergence to points, Chapter 9 of [29]) Let $(x_k)_{k\in\mathbb{N}}$ be a sequence in $\mathbb{E}$.

(i) $(x_k)_{k\in\mathbb{N}}$ is said to *converge R-linearly* to $\widetilde{x}$ with rate $c \in [0,1)$ if there is a constant $\gamma > 0$ such that

$$\|x_k - \widetilde{x}\| \le \gamma c^k \quad \forall k \in \mathbb{N}. \tag{1}$$

(ii) $(x_k)_{k\in\mathbb{N}}$ is said to *converge Q-linearly* to $\widetilde{x}$ with rate $c \in [0,1)$ if

$$\|x_{k+1} - \widetilde{x}\| \le c\,\|x_k - \widetilde{x}\| \quad \forall k \in \mathbb{N}.$$

By definition, Q-linear convergence implies R-linear convergence with the same rate. Elementary examples show that the inverse implication does not hold in general.

One of the central concepts in the convergence of sequences is *Fejér monotonicity*: a sequence $(x_k)_{k\in\mathbb{N}}$ is *Fejér monotone* with respect to a nonempty convex set Ω if

$$\|x_{k+1} - x\| \le \|x_k - x\| \quad \forall x \in \Omega, \forall k \in \mathbb{N}.$$

In the context of convergence analysis of fixed point iterations, the following generalization of Fejér monotonicity of sequences is central.

Definition 2 (μ-monotonicity) Let $(x_k)_{k\in\mathbb{N}}$ be a sequence on $\mathbb{E}$, $\Omega \subset \mathbb{E}$ be nonempty and $\mu : \mathbb{R}_+ \to \mathbb{R}_+$ satisfy $\mu(0) = 0$ and

$$\mu(t_1) < \mu(t_2) \le t_2 \text{ whenever } 0 \le t_1 < t_2.$$

(i) $(x_k)_{k\in\mathbb{N}}$ is said to be *μ-monotone with respect to Ω* if

$$\operatorname{dist}(x_{k+1}, \Omega) \le \mu\left(\operatorname{dist}(x_k, \Omega)\right) \quad \forall k \in \mathbb{N}. \tag{2}$$

(ii) $(x_k)_{k\in\mathbb{N}}$ is said to be *linearly monotone with respect to Ω* if (2) is satisfied for $\mu(t) = c \cdot t$ for all $t \in \mathbb{R}_+$ and some constant $c \in [0,1]$.

The focus of our study is linear convergence, so only linear monotonicity will come into play in what follows. A study of other kinds of convergence, particularly *sublinear*, would employ the full generality of μ-monotonicity.

The next result is clear.

Proposition 1 (Fejér monotonicity implies μ-monotonicity) *If the sequence $(x_k)_{k\in\mathbb{N}}$ is Fejér monotone with respect to $\Omega \subset \mathbb{X}$ then it is μ-monotone with respect to Ω with $\mu = \mathrm{Id}$.*

The converse is not true, as the next example shows.

Example 1 (μ-monotonicity is not Fejér monotonicity) *Let $\Omega := \left\{(x,y) \in \mathbb{R}^2 \mid y \le 0\right\}$ and consider the sequence $x_k := \left(1/2^k, 1/2^k\right)$ for all $k \in \mathbb{N}$. This sequence is linearly monotone with respect to Ω with $c = 1/2$, but not Fejér monotone since $\|x_{k+1} - (2,0)\| > \|x_k - (2,0)\|$ for all k.*

The next definition will come into play in Sections 4 and 5. It provides a way to analyze fixed point iterations which, like our main example of alternating projections, are compositions of mappings.

The subset $\Lambda \subset \mathbb{E}$ appearing in Definition 3 and throughout this work is always assumed to be closed and nonempty. We use this set to isolate specific elements of the fixed point set (most often restricted to affine subspaces). This is more than just a formal generalization since in some concrete situations the required assumptions do not hold on $\mathbb{E}$ but they do hold on relevant subsets.

Definition 3 (linearly extendible sequences) A sequence $(x_k)_{k\in\mathbb{N}}$ on $\Lambda \subset \mathbb{E}$ is said to be *linearly extendible* on Λ with frequency $m \geq 1$ ($m \in \mathbb{N}$ is fixed) and rate $c \in [0,1)$ if there is a sequence $(z_k)_{k\in\mathbb{N}}$ on Λ such that $x_k = z_{mk}$ for all $k \in \mathbb{N}$ and the following conditions are satisfied for all $k \in \mathbb{N}$:

$$\|z_{k+2} - z_{k+1}\| \leq \|z_{k+1} - z_k\|, \tag{3}$$

$$\|z_{m(k+1)+1} - z_{m(k+1)}\| \leq c\|z_{mk+1} - z_{mk}\|. \tag{4}$$

When $\Lambda = \mathbb{E}$, the quantifier "on Λ" is dropped.

The requirement on the linear extension sequence $(z_k)_{k\in\mathbb{N}}$ means that the sequence of the distances between its two consecutive iterates is uniformly non-increasing and possesses a subsequence of type $(\|z_{mk+1} - z_{mk}\|)_{k\in\mathbb{N}}$ that converges Q-linearly with a global rate to zero.

The extension of sequences of fixed point iterations $(x_k)_{k\in\mathbb{N}}$ will most often be to the intermediate points generated by the composite mappings. In the case of alternating projections this is $z_{2k} := x_k \in P_A P_B x_{k-1}$, and $z_{2k+1} \in P_B z_{2k}$. This strategy of analyzing alternating projections by keeping track of the intermediate projections has been exploited to great effect in [6, 11, 23–25, 28]. From the Cauchy property of $(z_k)_{k\in\mathbb{N}}$, one can deduce R-linear convergence from linear extendability.

Proposition 2 (linear extendability implies R-linear convergence) *If the sequence $(x_k)_{k\in\mathbb{N}}$ on $\Lambda \subset \mathbb{E}$ is linearly extendible on Λ with some frequency $m \geq 1$ and rate $c \in [0,1)$, then $(x_k)_{k\in\mathbb{N}}$ converges R-linearly to a point $\widetilde{x} \in \Lambda$ with rate c.*

Proof. Let $(z_k)_{k\in\mathbb{N}}$ be a linear extension of $(x_k)_{k\in\mathbb{N}}$ on Λ with frequency m and rate c. Conditions (3) and (4) then imply by induction that

$$\|z_{k+1} - z_k\| \leq \frac{d_0}{c} \sqrt[m]{c}^{\,k} \quad \forall k \in \mathbb{N},$$

where $d_0 := \|z_1 - z_0\|$. This means that $(z_k)_{k\in\mathbb{N}}$ is a Cauchy sequence and hence converges to a limit $\widetilde{x}$, which is in Λ by the closedness of this set. Conditions (3) and (4) also yield that for every $k \in \mathbb{N}$,

$$\|x_k - \widetilde{x}\| = \|z_{mk} - \widetilde{x}\| \leq \sum_{i=mk}^{\infty} \|z_i - z_{i+1}\|$$

$$\leq m \sum_{i=k}^{\infty} \|z_{mi} - z_{mi+1}\| \leq m\|z_0 - z_1\| \sum_{i=k}^{\infty} c^i \leq \gamma c^k,$$

where

$$\gamma := \frac{md_0}{1-c}. \tag{5}$$

This means that $(x_k)_{k\in\mathbb{N}}$ converges R-linearly to $\widetilde{x}$ with rate c. □

3 Linearly monotone fixed point iterations

Quantifying the convergence of fixed point iterations is key to providing error bounds for algorithms. For a multi-valued self-mapping $T : \mathbb{E} \rightrightarrows \mathbb{E}$, the operative inequality leading to linear convergence of the fixed point iteration $x_{k+1} \in Tx_k$ is

$$\operatorname{dist}(x_{k+1}, S) \;\leq\; c \operatorname{dist}(x_k, S) \quad \forall k \in \mathbb{N} \tag{6}$$

for $S \subset \mathsf{Fix}\, T$ and $c \in [0, 1)$. When this holds, the sequence $(x_k)_{k\in\mathbb{N}}$ will be called *linearly monotone* relative to S with constant c.

For multi-valued mappings, however, we need to clarify what is meant in the first place by the fixed point set. We take the least restrictive definition as any point contained in its image via the mapping.

Definition 4 (fixed points of set-valued mappings) The set of fixed points of a possibly set-valued mapping $T : \mathbb{E} \rightrightarrows \mathbb{E}$ is defined by

$$\mathsf{Fix}\, T := \{x \in \mathbb{E} \mid x \in Tx\}.$$

As noted in [25, Example 2.1], for $x \in \mathsf{Fix}\, T$, it is not the case that $Tx \subset \mathsf{Fix}\, T$. This can happen, in particular, when the mapping T is multi-valued on its set of fixed points. Almost averaged mappings detailed next are a generalization of averaged mappings and rule out so-called inhomogeneous fixed point sets.

3.1 Almost averaged mappings

Definition 5 (almost nonexpansive/averaged mappings, Definition 2.2 of [25]) Let Ω be a nonempty subset of $\mathbb{E}$ and let $T : \Omega \rightrightarrows \mathbb{E}$.

(i) T is said to be *pointwise almost nonexpansive on Ω at $y \in \Omega$* if there exists a $\varepsilon \geq 0$ (called the *violation*) such that

$$\begin{aligned} &\left\|x^+ - y^+\right\| \leq \sqrt{1+\varepsilon}\, \|x - y\| \\ &\forall x \in \Omega \quad \forall x^+ \in Tx \quad \forall y^+ \in Ty. \end{aligned} \tag{7}$$

If (7) holds with $\varepsilon = 0$ then T is called *pointwise nonexpansive* at y on Ω.
If T is pointwise (almost) nonexpansive on Ω at every point on a neighborhood of y in Ω (with the same violation ε), then T is said to be *(almost) nonexpansive on Ω at y (with violation ε).*
If T is pointwise (almost) nonexpansive at every point $y \in \Omega$ (with the same violation ε) on Ω, then T is said to be *(almost) nonexpansive on Ω (with violation ε).*

(ii) T is called *pointwise almost averaged on Ω at $y \in \Omega$* with violation $\varepsilon \geq 0$ if there is an averaging constant $\alpha \in (0,1)$ such that the mapping $\widetilde{T}$ defined by

$$\widetilde{T} := \tfrac{1}{\alpha} T + \left(1 - \tfrac{1}{\alpha}\right) \mathrm{Id}$$

is pointwise almost nonexpansive on Ω at y with violation ε/α.
Likewise T is said to be *(pointwise) (almost) averaged on Ω (at y) (with violation $\alpha\varepsilon$)* if $\widetilde{T}$ is (pointwise) (almost) nonexpansive on Ω (at y) (with violation ε).

Remark 1 The following remarks help to place this property in context.

(a) A mapping T is averaged with violation $\varepsilon = 0$ and averaging constant $\alpha = 1/2$ at all points on Ω if and only if it is firmly nonexpansive on Ω. The almost version of this property will be referred to as almost firmly nonexpansive.
(b) As noted in [25], pointwise almost nonexpansiveness of T at $\overline{x}$ with violation ε is related to, but stronger than *calmness* [33, Chapter 8.F] with constant $\lambda = \sqrt{1+\varepsilon}$: for pointwise almost nonexpansiveness the inequality (7) must hold for all pairs $x^+ \in Tx$ and $y^+ \in Ty$, while for calmness the same inequality would hold only for points $x^+ \in Tx$ and their *projections* onto Ty.
(c) See [25, Example 2.2] for concrete examples.

Proposition 3 (characterizations of pointwise averaged mappings) *[25, Proposition 2.1] Let $T : \mathbb{E} \rightrightarrows \mathbb{E}$, $\Omega \subset \mathbb{E}$ and let $\alpha \in (0,1)$. The following are equivalent.*

(i) T is pointwise almost averaged on Ω at $y \in \Omega$ with averaging constant α and violation ε.
(ii)

$$\begin{aligned} \left\|x^+ - y^+\right\|^2 \leq (1+\varepsilon)\left\|x-y\right\|^2 - \tfrac{1-\alpha}{\alpha}\left\|\left(x - x^+\right) - \left(y - y^+\right)\right\|^2 \\ \forall x \in \Omega \quad \forall x^+ \in Tx \quad \forall y^+ \in Ty. \end{aligned} \tag{8}$$

As a consequence, if T is pointwise almost averaged at y with averaging constant $\alpha \in (0,1)$ and violation ε on Ω, then T is pointwise almost nonexpansive at y with violation at most ε on Ω.

Remark 2 Pointwise almost averaged mappings are single-valued on the set of fixed points [25, Proposition 2.2]. If the mapping is actually nonexpansive (that is, almost nonexpansive with violation zero) on Ω, then it must be single-valued on Ω. When this happens, we simply write $x^+ = Tx$ instead of $x^+ \in Tx$.

It was proved in [5, Theorem 5.12] that if $(x_k)_{k\in\mathbb{N}}$ is Fejér monotone with respect to a nonempty closed convex subset Ω and inequality (6) holds true with Ω in place of S, then $(x_k)_{k\in\mathbb{N}}$ converges R-linearly to a point in Ω with rate at most c. The following statement aligns with this fact.

Proposition 4 (linear monotonicity and almost averagedness imply R-linear convergence) *Let* $T : \mathbb{E} \rightrightarrows \mathbb{E}$ *and* $x_{k+1} \in Tx_k \subset \Lambda \subset \mathbb{E}$ *for all* $k \in \mathbb{N}$*. Suppose that* $\mathsf{Fix}\, T \cap \Lambda$ *is closed and nonempty and that* T *is pointwise almost averaged at all points on* $(\mathsf{Fix}\, T + d_0\mathbb{B}) \cap \Lambda$*, where* $d_0 := \operatorname{dist}(x_0, \mathsf{Fix}\, T \cap \Lambda)$*, that is,* T *is almost averaged on* $(\mathsf{Fix}\, T + d_0\mathbb{B}) \cap \Lambda$*. If the sequence* $(x_k)_{k\in\mathbb{N}}$ *is linearly monotone relative to* $\mathsf{Fix}\, T \cap \Lambda$ *with constant* $c \in [0, 1)$*, then* $(x_k)_{k\in\mathbb{N}}$ *converges R-linearly to some point* $\widetilde{x} \in \mathsf{Fix}\, T \cap \Lambda$ *with rate* c*.*

Proof. We use the characterization formulated in Proposition 3(ii) of the almost averagedness of T with averaging constant α and violation ε. Combining (6) with $\mathsf{Fix}\, T \cap \Lambda$ in place of S and (8) (with averaging constant α and violation ε) implies by induction that for every $k \in \mathbb{N}$,

$$\sqrt{\frac{1-\alpha}{\alpha(1+\varepsilon)}}\|x_{k+1} - x_k\| \leq \|x_k - \overline{x}_k\| = \operatorname{dist}(x_k, \mathsf{Fix}\, T \cap \Lambda) \leq d_0 c^k,$$

where $\overline{x}_k$ is any point in $P_{(\mathsf{Fix}\, T\cap\Lambda)}x_k$. Hence, for any natural numbers k and p with $k < p$, we have

$$\begin{aligned} \|x_p - x_k\| \;&\leq\; \sum_{i=k}^{p-1}\|x_{i+1} - x_i\| \;\leq\; \sqrt{\frac{\alpha(1+\varepsilon)}{1-\alpha}}\sum_{i=k}^{p-1} d_0 c^i \\ &\leq\; d_0\sqrt{\frac{\alpha(1+\varepsilon)}{1-\alpha}}(1 + c + \ldots + c^{p-k-1})c^k \;\leq\; \sqrt{\frac{\alpha(1+\varepsilon)}{1-\alpha}}\frac{d_0}{1-c}c^k. \end{aligned} \tag{9}$$

This implies that $(x_k)_{k\in\mathbb{N}}$ is a Cauchy sequence and therefore convergent to some point $\widetilde{x}$.
We claim that $\widetilde{x} \in \mathsf{Fix}\, T \cap \Lambda$. Indeed, let us define

$$\delta := \frac{d_0}{1-c}\max\left\{\sqrt{\frac{\alpha(1+\varepsilon)}{1-\alpha}}, 1\right\}.$$

The sequence $(\overline{x}_k)_{k\in\mathbb{N}}$ is contained in the bounded set $\mathsf{Fix}\, T \cap \Lambda \cap \mathbb{B}_\delta(x_0)$ since

$$\begin{aligned} \|\overline{x}_k - x_0\| &\leq \|\overline{x}_k - x_k\| + \|x_k - x_0\| \\ &\leq d_0 c^k + \sqrt{\frac{\alpha(1+\varepsilon)}{1-\alpha}}\sum_{i=0}^{k-1} d_0 c^i \\ &\leq \max\left\{\sqrt{\frac{\alpha(1+\varepsilon)}{1-\alpha}}, 1\right\} d_0 \sum_{i=0}^{k} c^i < \delta. \end{aligned}$$

Hence it has a subsequence $(\overline{x}_{k_n})_{n\in\mathbb{N}}$ converging to some $\widetilde{x}^* \in \mathsf{Fix}\, T \cap \Lambda$ as $n \to \infty$. Since the corresponding subsequence $(x_{k_n})_{n\in\mathbb{N}}$ converges to $\widetilde{x}$ and

$$\|\overline{x}_{k_n} - x_{k_n}\| = \operatorname{dist}(x_{k_n}, \mathsf{Fix}\, T \cap \Lambda) \leq d_0 c^{k_n} \to 0$$

as $n \to \infty$, we deduce that $\widetilde{x} = \widetilde{x}^* \in \mathsf{Fix}\, T \cap \Lambda$.
Letting $p \to \infty$ in (9) yields (1) with $\gamma = \sqrt{\frac{\alpha(1+\varepsilon)}{1-\alpha}\frac{d_0}{1-c}}$, which completes the proof. □

The converse implication of Proposition 4 is not true in general because condition (1) in principle does not require the distance $\|x_k - \widetilde{x}\|$ to strictly decrease after every iterate while condition (6) does.

Almost nonexpansivity of T and linear extendability of the iteration are sufficient to guarantee that the sequence converges R-linearly to a point in $\mathsf{Fix}\, T$. Compare this to Proposition 2 which, without the additional assumption of almost nonexpansivity of T, only guarantees convergence to a point in Λ.

Proposition 5 (linear extendability and almost nonexpansivity imply R-linear convergence) *Let $T : \mathbb{E} \rightrightarrows \mathbb{E}$ and $(x_k)_{k\in\mathbb{N}}$ be a sequence generated by $x_{k+1} \in Tx_k \subset \Lambda \subset \mathbb{E}$ for all $k \in \mathbb{N}$. Suppose that $(x_k)_{k\in\mathbb{N}}$ is linearly extendible on Λ with some frequency $m \geq 1$ and rate $c \in [0, 1)$ and that T is almost nonexpansive on $\Lambda \cap \mathbb{B}_\gamma(x_0)$, where γ is given by (5). Then $(x_k)_{k\in\mathbb{N}}$ converges R-linearly to a point $\widetilde{x} \in \mathsf{Fix}\, T \cap \Lambda$ with rate c.*

Proof. By Proposition 2 the sequence $(x_k)_{k\in\mathbb{N}}$ converges R-linearly to a point $\widetilde{x} \in \Lambda$ with rate c. It remains to check that $\widetilde{x} \in \mathsf{Fix}\, T$. Suppose to the contrary that there is $\widetilde{x}^+ \in T\widetilde{x}$ with $\rho := \|\widetilde{x}^+ - \widetilde{x}\| > 0$. Since T is almost nonexpansive on $\Lambda \cap \mathbb{B}_\gamma(x_0)$, there is a violation $\varepsilon > 0$ such that

$$\|\widetilde{x}^+ - x_{k+1}\| \leq \sqrt{1+\varepsilon}\|\widetilde{x} - x_k\| \quad \forall k \in \mathbb{N}.$$

This leads to a contradiction since $\|\widetilde{x}^+ - x_{k+1}\| \to \rho > 0$ while $\|\widetilde{x} - x_k\| \to 0$ as $k \to \infty$. As a result, $\widetilde{x}^+ = \widetilde{x} \in \mathsf{Fix}\, T$ and the proof is complete. □

When the fixed point set restricted to Λ is an isolated point, then linear monotonicity of the sequence is equivalent to Q-linear convergence.

Proposition 6 *Let $\widetilde{x}$ be an isolated point of $\mathsf{Fix}\, T \cap \Lambda$, that is $\mathbb{B}_\delta(\widetilde{x}) \cap \mathsf{Fix}\, T \cap \Lambda = \{\widetilde{x}\}$ for $\delta > 0$ small enough. Let $T : \mathbb{E} \rightrightarrows \mathbb{E}$ be almost nonexpansive on a neighborhood of $\widetilde{x}$ relative to $\Lambda \subset \mathbb{E}$. Let $(x_k)_{k\in\mathbb{N}}$ be a sequence generated by $x_{k+1} \in Tx_k \subset \Lambda$ for all $k \in \mathbb{N}$ with $x_0 \in \Lambda$ sufficiently close to $\widetilde{x}$. Then $(x_k)_{k\in\mathbb{N}}$ is linearly monotone with respect to $\mathsf{Fix}\, T \cap \Lambda$ with rate smaller than 1 if and only if it is Q-linearly convergent to $\widetilde{x}$.*

Proof. Since $\widetilde{x}$ is an isolated piont of $\mathsf{Fix}\, T \cap \Lambda$ and T is almost nonexpansive on a neighborhood of $\widetilde{x}$ relative to Λ, there is a $\delta > 0$ small enough that $\mathsf{Fix}\, T \cap \Lambda \cap \mathbb{B}_\delta(\widetilde{x}) = \{\widetilde{x}\}$ and T is almost nonexpansive on $\mathbb{B}_{\delta'}(\widetilde{x}) \cap \Lambda$ with violation ε, where $\delta' = \frac{\delta}{2\sqrt{1+\varepsilon}}$. Let $\rho \in (0, \delta')$. Then by almost nonexpansivity of T on $\mathbb{B}_\rho(\widetilde{x}) \cap \Lambda$ we have that

$$\begin{aligned}\|x^+ - \widetilde{x}\| &\leq \sqrt{1+\varepsilon}\|x - \widetilde{x}\| \leq \rho\sqrt{1+\varepsilon} \\ &< \frac{\delta}{2} \leq \frac{1}{2}\operatorname{dist}(\widetilde{x}, (\mathsf{Fix}\, T \cap \Lambda) \setminus \{\widetilde{x}\}) \quad \forall x \in \mathbb{B}_\rho(\widetilde{x}) \cap \Lambda, \forall x^+ \in Tx.\end{aligned}$$

This implies that

$$\operatorname{dist}(x^+, \mathsf{Fix}\, T \cap \Lambda) = \|x^+ - \widetilde{x}\| \quad \forall x \in \mathbb{B}_\rho(\widetilde{x}) \cap \Lambda, \forall x^+ \in Tx. \tag{10}$$

Hence for any sequence $(x_k)_{k\in\mathbb{N}}$ as described in Proposition 6 with $x_0 \in \Lambda \cap \mathbb{B}_\rho(\widetilde{x})$, the equivalence of linear monotonicity of $(x_k)_{k\in\mathbb{N}}$ relative to $\mathsf{Fix}\, T \cap \Lambda$ with rate smaller than 1 and Q-linear convergence to $\widetilde{x}$ follows from equality (10) and the definitions because each of these properties of $(x_k)_{k\in\mathbb{N}}$ alternatively combined with (10) ensures inductively that the whole sequence $(x_k)_{k\in\mathbb{N}}$ lies in $\mathbb{B}_\rho(\widetilde{x})$. □

It is mainly due to the above proposition that we include the extra technical overhead of making the above statements always relative to some subset Λ. It is not uncommon to have $\mathsf{Fix}\, T$ a singleton relative to Λ, but not on the whole space $\mathbb{E}$. For an example of this, see the analysis of the Douglas-Rachford fixed point iteration in [2].

3.2 Metric subregularity and linear convergence

The following concept of *metric regularity on a set* characterizes the stability of mappings at points in their image and has played a central role, implicitly and explicitly, in our convergence analysis of fixed point iterations [2, 13, 25]. We show in this section that, indeed, metric subregularity is *necessary* to achieve linear convergence. The following definition is a specialized (linear) variant of [25, Definition 2.5] which is a combination with slight modification of those formulated in [16, Definition 2.1 (b)] and [17, Definition 1 (b)] so that the property is relative to relevant sets for iterative methods. Our terminology also follows [10].

Definition 6 (metric regularity on a set) Let $\Phi : \mathbb{E} \rightrightarrows \mathbb{Y}$ be a set-valued mapping between Euclidean spaces and let $U \subset \mathbb{E}$ and $V \subset \mathbb{Y}$. The mapping Φ is called *metrically regular relative to* $\Lambda \subset \mathbb{E}$ on U for V with constant κ if

$$\operatorname{dist}\left(x, \Phi^{-1}(y) \cap \Lambda\right) \le \kappa \operatorname{dist}\left(y, \Phi(x)\right) \tag{11}$$

holds for all $x \in U \cap \Lambda$ and $y \in V$.

When $V = \{y\}$ consists of a single point one says that Φ is *metrically subregular* with constant κ on U for y relative to $\Lambda \subset \mathbb{E}$.

When $\Lambda = \mathbb{E}$, the quantifier "relative to" is dropped.

Remark 3 The conventional concept of *metric regularity* at a point $\overline{x} \in \mathbb{E}$ for $\overline{y} \in \Phi(\overline{x})$ corresponds to setting $\Lambda = \mathbb{E}$, and taking U and V to be neighborhoods of $\overline{x}$ and $\overline{y}$ (as opposed simply to subsets including these points) respectively. Similarly, the conventional *metric subregularity* [10] and *metric hemi/semiregularity* [1, 19, 20] at $\overline{x}$ for $\overline{y}$ correspond to setting $\Lambda = \mathbb{E}$, and respectively either taking U to be a neighborhood of $\overline{x}$ and $V = \{\overline{y}\}$, or taking $U = \{\overline{x}\}$ and taking V to be a neighborhood of $\overline{y}$. This notion can and has been generalized even more. The more general notion of metric subregularity studied by Ioffe [16, 17] for instance, together with μ-monotonicity, would be needed for the study of nonlinear convergence. These more general notions of metric subregularity are the most suitable vehicles to parallel properties like the Kurdyka-Łojasiewicz (KL) property for functions. In fact, for differentiable functions, *metric regularity* of the gradient is equivalent to the KL property [8, Corollary 4 and Remark 5], though from our point of view, metric subregularity is the more general property.

The following convergence criterion is fundamental.

Theorem 1 (linear convergence with metric subregularity) *Let $T : \mathbb{E} \rightrightarrows \mathbb{E}$, let $\Lambda \subset \mathbb{E}$ with $\mathsf{Fix}\, T \cap \Lambda$ closed and nonempty. Suppose that, for some fixed $\delta > 0$, T is pointwise almost averaged at all points $\overline{x} \in \mathsf{Fix}\, T \cap \Lambda$ with averaging constant α and violation ε on $(\mathsf{Fix}\, T + \mathbb{B}_\delta) \cap \Lambda$, and that the mapping $\Phi := T - \mathrm{Id}$ is metrically subregular on $(\mathsf{Fix}\, T + \mathbb{B}_\delta) \setminus \mathsf{Fix}\, T$ for 0 relative to Λ with constant $\kappa > 0$. Then it holds*

$$\begin{aligned} \operatorname{dist}\left(x^+, \mathsf{Fix}\, T \cap \Lambda\right) \le c \operatorname{dist}\left(x, \mathsf{Fix}\, T \cap \Lambda\right) \\ \forall x \in (\mathsf{Fix}\, T + \mathbb{B}_\delta) \cap \Lambda \quad \forall x^+ \in Tx, \end{aligned} \tag{12}$$

where

$$c := \sqrt{1 + \varepsilon - \frac{1-\alpha}{\kappa^2 \alpha}}. \tag{13}$$

In particular, when $c < 1$, every sequence $(x_k)_{k\in\mathbb{N}}$ generated by $x_{k+1} \in Tx_k \subset \Lambda$ with initial point in $(\mathsf{Fix}\, T + \mathbb{B}_\delta) \cap \Lambda$ converges R-linearly to some point in $\mathsf{Fix}\, T \cap \Lambda$ with rate c. If $\mathsf{Fix}\, T \cap \Lambda$ is a singleton, then the convergence is Q-linear.

Proof. The inequality (12) is the content of [25, Corollary 2.3]. Since this is easy to obtain, we reproduce the proof here for convenience. Choose any $x \in (\mathsf{Fix}\, T + \mathbb{B}_\delta) \cap \Lambda$ and select any $x^+ \in Tx$. Metric subregularity of Φ on $(\mathsf{Fix}\, T + \mathbb{B}_\delta) \setminus \mathsf{Fix}\, T$ for 0 relative to Λ with constant $\kappa > 0$ means that

$$\operatorname{dist}\left(x, \Phi^{-1}(0) \cap \Lambda\right) \le \kappa \operatorname{dist}(0, \Phi(x)).$$

Since $\Phi^{-1}(0) \cap \Lambda = \mathsf{Fix}\, T \cap \Lambda$, this then implies that

$$\tfrac{1}{\kappa^2} \operatorname{dist}^2\left(x, \mathsf{Fix}\, T \cap \Lambda\right) \le \|x^+ - x\|^2.$$

Note that T is a single-valued mapping on $\mathsf{Fix}\, T \cap \Lambda$ since T is almost averaged – hence almost nonexpansive – on $(\mathsf{Fix}\, T + \mathbb{B}_\delta) \cap \Lambda$ [25, Proposition 2.2], so we can unambiguously write $\overline{x} = T\overline{x}$ for $\overline{x} \in P_{(\mathsf{Fix}\, T \cap \Lambda)} x$ and rewrite the inequality as

$$\tfrac{1}{\kappa^2} \|x - \overline{x}\|^2 \le \|x^+ - x\|^2 \quad \forall x \in (\mathsf{Fix}\, T + \mathbb{B}_\delta) \setminus \mathsf{Fix}\, T.$$

This inequality, together with the almost averaging property and its characterization Proposition 3(ii) yield

$$\left\|x^+ - \overline{x}\right\|^2 \le \left(1 + \varepsilon - \frac{1-\alpha}{\alpha\kappa^2}\right) \|x - \overline{x}\|^2. \tag{14}$$

Note in particular that $0 \leq 1+\varepsilon-\frac{1-\alpha}{\alpha\kappa^2}$. Since x is any point in $(\mathsf{Fix}\, T+\mathbb{B}_\delta)\cap\Lambda$ this proves (12) with c given by (13).

For $c<1$, it follows by definition that such a sequence $(x_k)_{k\in\mathbb{N}}$ is linearly monotone with respect to $\mathsf{Fix}\, T\cap\Lambda$ with rate c. A combination of Propositions 4 and 6 then shows that the sequence is linearly convergent to a point in $\mathsf{Fix}\, T\cap\Lambda$, R-linearly in general, and Q-linearly if $\mathsf{Fix}\, T\cap\Lambda$ is a singleton. □

When $\delta=\infty$, Theorem 1 provides a criterion for global linear convergence of abstract fixed point iterations. The next result shows that metric subregularity is in fact *necessary* for linearly monotone iterations, without any assumptions about the averaging properties of T, almost or otherwise.

Theorem 2 (necessity of metric subregularity) *Let $T:\mathbb{E}\rightrightarrows\mathbb{E}$ with $\mathsf{Fix}\, T$ nonempty, fix $\Lambda\subset\mathbb{E}$ so that $\mathsf{Fix}\, T\cap\Lambda$ is closed and nonempty, and let $\Omega\subset\Lambda$. If for each $x_0\in\Omega$, every sequence $(x_k)_{k\in\mathbb{N}}$ generated by $x_{k+1}\in Tx_k\subset\Lambda$ is linearly monotone relative to $\mathsf{Fix}\, T\cap\Lambda$ with constant $c\in[0,1)$, then the mapping $\Phi:=T-\mathrm{Id}$ is metrically subregular on Ω for 0 relative to Λ with constant $\kappa\leq\frac{1}{1-c}$.*

Proof. Since every sequence $(x_k)_{k\in\mathbb{N}}$ generated by $x_{k+1}\in Tx_k\subset\Lambda$ starting in Ω is linearly monotone with respect to $\mathsf{Fix}\, T\cap\Lambda$ with rate c, the inequality (6) with $\mathsf{Fix}\, T\cap\Lambda$ in place of S holds true. This together with the triangle inequality implies that for every $k\in\mathbb{N}$,

$$\begin{aligned}\left\|x^+-x_k\right\| &\geq \mathrm{dist}\left(x_k,\mathsf{Fix}\, T\cap\Lambda\right)-\mathrm{dist}\left(x^+,\mathsf{Fix}\, T\cap\Lambda\right)\\ &\geq (1-c)\,\mathrm{dist}\left(x_k,\mathsf{Fix}\, T\cap\Lambda\right)\quad \forall x^+\in Tx_k.\end{aligned}\tag{15}$$

On the other hand, we have from definition of Φ that

$$\begin{aligned}\Phi^{-1}(0)&=\mathsf{Fix}\, T,\\ \mathrm{dist}\left(0,\Phi(x_k)\right)&=\inf_{x^+\in Tx_k}\left\|x^+-x_k\right\|\quad\forall k\in\mathbb{N}.\end{aligned}\tag{16}$$

Combining (15) and (16) yields

$$\mathrm{dist}\left(x_k,\Phi^{-1}(0)\cap\Lambda\right)\leq\frac{1}{1-c}\,\mathrm{dist}\left(0,\Phi(x_k)\right)\quad\forall k\in\mathbb{N}.$$

Consequently, $\frac{1}{1-c}$ is *a* constant of metric subregularity of Φ on Ω for 0 (not necessarily the smallest such constant) as claimed. □

Corollary 1 (necessary conditions for linear convergence) *Let $T:\mathbb{E}\rightrightarrows\mathbb{E}$ with $\mathsf{Fix}\, T$ nonempty. For some $\delta>0$, let T be almost averaged on $(\mathsf{Fix}\, T+\mathbb{B}_\delta)\cap\Lambda$. If, for each $x_0\in((\mathsf{Fix}\, T+\mathbb{B}_\delta)\cap\Lambda)\setminus\mathsf{Fix}\, T$, every sequence $(x_k)_{k\in\mathbb{N}}$ generated by $x_{k+1}\in Tx_k\subset\Lambda$ is linearly monotone relative to $\mathsf{Fix}\, T\cap\Lambda$ with constant $c\in[0,1)$, then all such sequences converge R-linearly with rate c to some point in $\mathsf{Fix}\, T\cap\Lambda$ and $\Phi:=T-\mathrm{Id}$ is metrically subregular on $(\mathsf{Fix}\, T+\mathbb{B}_\delta)\setminus\mathsf{Fix}\, T$ for 0 relative to Λ with constant $\kappa\leq\frac{1}{1-c}$.*

Proof. This is an immediate consequence of Proposition 4 and Theorem 2. □

4 Nonconvex alternating projections

For $x_0 \in \mathbb{E}$ given, the *alternating projections (AP) iteration* for two closed subsets $A, B \subset \mathbb{E}$ is given by

$$x_{k+1} \in T_{AP}x_k := P_A P_B x_k \quad \forall k \in \mathbb{N}. \tag{17}$$

For convenience, we associate $(x_k)_{k\in\mathbb{N}}$ with the sequence $(b_k)_{k\in\mathbb{N}}$ on B such that $b_k \in P_B x_k$ and $x_{k+1} \in P_A b_k$ for all $k \in \mathbb{N}$. In the sequel, we frequently use the joining sequence $(z_k)_{k\in\mathbb{N}}$ given by

$$z_{2k} = x_k \text{ and } z_{2k+1} = b_k \quad \forall k \in \mathbb{N}. \tag{18}$$

We will always assume, without loss of generality, that $x_0 \in A$.

It is well known that every alternating projections iteration for two convex intersecting sets globally converges R-linearly to a feasibility solution if the collection of sets is what we call *subtransversal* [4]. The latter property and its at-point version is a specialization of metric subregularity to the context of set feasibility.

4.1 Elemental regularity and subtransversality

Convexity of the underlying sets has long been the standing assumption in analysis of projection algorithms. The next definition characterizing regularity of nonconvex sets first appeared in [22, Definition 5] and encapsulates many of the regularity notions appearing elsewhere in the literature including Hölder regularity [28, Definition 2], relative (ε, δ)-subregularity [13, Definition 2.9], restricted (ε, δ)-regularity [6, Definition 8.1], Clarke regularity [33, Definition 6.4], super-regularity [24, Definition 4.3], prox-regularity [32, Definition 1.1], and of course convexity. The connection of elemental regularity of sets to the pointwise almost averaging property of their projectors is discussed later.

Definition 7 (elemental regularity of sets) Let $A \subset \mathbb{E}$ be nonempty and let $(a, \overline{v}) \in \operatorname{gph}(N_A)$.

(i) The set A is said to be *elementally subregular relative to* $S \subset A$ *at* $\overline{x} \in A$ *for* $(a, \overline{v})$ *with constant* ε if there exists a neighborhood U of $\overline{x}$ such that

$$\langle \overline{v},\ x - a\rangle \le \varepsilon \|\overline{v}\| \|x - a\|, \quad \forall x \in S \cap U. \tag{19}$$

(ii) The set A is said to be *uniformly elementally subregular relative to* $S \subset A$ *at* $\overline{x}$ for $(a, \overline{v})$ if, for any $\varepsilon > 0$, there exists a neighborhood U (depending on ε) of $\overline{x}$ such that (19) holds.

(iii) The set A is said to be *elementally regular at* $\overline{x}$ *for* $(a, \overline{v})$ *with constant* ε if there exists a neighborhood V of $\overline{v}$ such that, for each $v \in N_A(a) \cap V$, the set A is elementally subregular relative to $S = A$ at $\overline{x}$ for (a, v) with constant ε.

(iv) The set A is said to be *uniformly elementally regular at* $\overline{x}$ *for* $(a, \overline{v})$ if, for any $\varepsilon > 0$, the set A is elementally regular at $\overline{x}$ for $(a, \overline{v})$ with constant ε.

The reference points $\overline{x}$ and a in Definition 7, need not be in S and U, respectively, although these are the main cases of interest for us. The properties are trivial for any constant $\varepsilon \geq 1$, so the only case of interest is elemental (sub)regularity with constant $\varepsilon < 1$.

Proposition 7 (Proposition 4(vii) of [22]) *Let A be closed and nonempty. If A is convex, then it is elementally regular at all points $x \in A$ for all $(a, v) \in \operatorname{gph} N_A$ with constant $\varepsilon = 0$ and any neighborhood in $\mathbb{E}$ for both x and a.*

The next result shows the implications of elemental regularity of sets for regularity of the corresponding projectors.

Theorem 3 (projectors onto elementally subregular sets, Theorem 3.1 of [25]) *Let $A \subset \mathbb{E}$ be nonempty closed, and let U be a neighborhood of $\overline{x} \in A$. Let $S \subset A \cap U$ and $S' := P_A^{-1}(S) \cap U$. If A is elementally subregular at $\overline{x}$ relative to S' for each*

$$(a, v) \in V := \{(z, w) \in \operatorname{gph} N_A^{\text{prox}} \mid z + w \in U \quad \textit{and} \quad z \in P_A(z + w)\}$$

with constant ε on the neighborhood U, then the following hold.

(i) The projector P_A is pointwise almost nonexpansive at each $y \in S$ on U with violation $\varepsilon' := 2\varepsilon + \varepsilon^2$. That is

$$\|x - y\| \leq \sqrt{1 + \varepsilon'}\, \|x' - y\| \quad \forall y \in S \quad \forall x' \in U \quad \forall x \in P_A x'.$$

(ii) The projector P_A is pointwise almost firmly nonexpansive at each $y \in S$ on U with violation $\varepsilon'_2 := 2\varepsilon + 2\varepsilon^2$. That is

$$\|x - y\|^2 + \|x' - x\|^2 \leq (1 + \varepsilon'_2)\, \|x' - y\|^2$$
$$\forall y \in S \quad \forall x' \in U \quad \forall x \in P_A x'.$$

In addition to the pointwise almost averaging property, metric subregularity plays a central role in the general theory of Section 3. In the context of set feasibility, this is translated to what we call subtransversality below. What we present as the definition of subtransversality is the simplified version of [25, Definition 3.2(i)].

Definition 8 (subtransversality) Let A and B be closed subsets of $\mathbb{E}$, let $\mathbb{E}^2$ be endowed with some norm and let $\Gamma \subset \mathbb{E}^2$. The collection of sets $\{A, B\}$ is said to be subtransversal at $\overline{u} = (\overline{x}_1, \overline{x}_2) \in A \times B$ for $\overline{w} = (\overline{y}_1, \overline{y}_2) \in (P_A - \operatorname{Id})(\overline{x}_2) \times (P_B - \operatorname{Id})(\overline{x}_1)$ relative to Γ if there exist numbers $\delta > 0$ and $\kappa \geq 0$ such that

$$\operatorname{dist}\Big(u, \Big(\left(P_B - \operatorname{Id}\right)^{-1}(\overline{y}_2) \times \left(P_A - \operatorname{Id}\right)^{-1}(\overline{y}_1)\Big) \cap \Gamma\Big) \leq$$
$$\kappa \operatorname{dist}\left(\overline{w}, (P_A - \operatorname{Id})(x_2) \times (P_B - \operatorname{Id})(x_1)\right)$$

holds true for all $u = (x_1, x_2) \in \mathbb{B}_\delta(\overline{u}) \cap \Gamma$.

When $\Gamma = \mathbb{E}^2$, the quantifier "relative to" is dropped.

The reference point $\overline{u}$ in Definition 8 need not be in Γ, although this is the only case of interest for us. The following characterization of subtransversality at common points will play a fundamental role in our subsequent analysis.

Proposition 8 (subtransversality at common points, Proposition 3.3 of [25]) *Let A and B be closed subsets of $\mathbb{E}$. Let $\mathbb{E}^2$ be endowed with 2-norm*

$$\|(x_1, x_2)\|_2 = \left(\|x_1\|_{\mathbb{E}}^2 + \|x_2\|_{\mathbb{E}}^2\right)^{1/2} \quad \forall\, (x_1, x_2) \in \mathbb{E}^2.$$

The collection of sets $\{A, B\}$ is subtransversal relative to

$$\Gamma := \left\{u = (x, x) \in \mathbb{E}^2 \mid x \in \Lambda \subset \mathbb{E}\right\} \tag{20}$$

at $\overline{u} = (\overline{x}, \overline{x})$ with $\overline{x} \in A \cap B$ for $\overline{y} = 0_{\mathbb{E}^2}$ if and only if there exist numbers $\delta > 0$ and $\kappa \geq 0$ such that

$$\mathrm{dist}\,(x, A \cap B \cap \Lambda) \leq \kappa \max\left(\mathrm{dist}\,(x, A)\,, \mathrm{dist}\,(x, B)\right) \quad \forall x \in \mathbb{B}_\delta(\overline{x}) \cap \Lambda. \tag{21}$$

The relative set $\Gamma \subset \mathbb{E}^2$ given by (20) which makes the notion of subtransversality consistent in the product space can clearly be identified with the set $\Lambda \subset \mathbb{E}$. We will therefore more often than not use the terms "relative to Λ" instead of "relative to Γ" and "at $\overline{x}$" instead of "at $(\overline{x}, \overline{x})$ for $0_{\mathbb{E}^2}$" when discussing subtransversality at common points where the product-space structure is no longer needed.

Remark 4 It follows from Proposition 8 that the exact lower bound of all numbers κ such that condition (21) is satisfied, denoted $\mathrm{sr}[A, B](\overline{x})$, characterizes the subtransversality of the collection of sets at common points. More specifically, the collection of sets $\{A, B\}$ is subtransversal at $\overline{x}$ if and only if $\mathrm{sr}[A, B](\overline{x}) < +\infty$.

The property (21) with $\Lambda = \mathbb{E}$ has been around for decades under the names of (local) *linear regularity*, *metric regularity*, *linear coherence*, *metric inequality*, and *subtransversality*; cf. [3,4,11,13–15,18,27,30,34]. We refer the reader to the recent articles [21,22] in which a number of necessary and/or sufficient characterizations of subtransversality are discussed. The next characterization of subtransversality, which is the relativized version of [22, Theorem 1(iii)], will play a key role in proving the necessary condition results in Sections 4.2 and 5. This characterization is actually implied in the proof of [11, Theorem 6.2] where the property called *intrinsic transversality* [11, Definition 3.1] was shown to imply subtransversality.

Proposition 9 (subtransversality at common points) *The collection of sets $\{A, B\}$ is subtransversal at $\overline{x} \in A \cap B$ relative to Λ if and only if there exist numbers $\delta > 0$ and $\kappa \geq 0$ such that*

$$\mathrm{dist}(x, A \cap B \cap \Lambda) \leq \kappa\, \mathrm{dist}(x, B) \quad \forall x \in A \cap \mathbb{B}_\delta(\overline{x}) \cap \Lambda. \tag{22}$$

Moreover,

$$\mathrm{sr}'[A, B](\overline{x}) \leq \mathrm{sr}[A, B](\overline{x}) \leq 1 + 2\mathrm{sr}'[A, B](\overline{x}), \tag{23}$$

where $\mathrm{sr}'[A, B](\overline{x})$ is the exact lower bound of all numbers κ such that condition (22) is satisfied.

Remark 5 In light of Remark 4 and the two inequalities in (23), a collection of sets $\{A, B\}$ is subtransversal at $\overline{x} \in A \cap B$ if and only if $\mathrm{sr}'[A, B](\overline{x}) < +\infty$. For the simplicity in terms of presentation, in the sequel, we will frequently use this fact without repeating the argument.

Both inequalities in (23) can be strict as shown in the following example.

Example 2 *Let A and B be two lines in $\mathbb{R}^2$ forming an angle $\pi/3$ at the intersection point $\overline{x}$. One can easily check that*

$$\mathrm{sr}'[A, B](\overline{x}) = 2/\sqrt{3} < 2 = \mathrm{sr}[A, B](\overline{x}) < 1 + 2\mathrm{sr}'[A, B](\overline{x}) = 1 + 4/\sqrt{3}.$$

The connection of subtransversality to metric subregularity was presented for more general cyclic projections in [25, Proposition 3.4]. We present here the simplified version for two sets with possibly empty intersection.

Proposition 10 (metric subregularity for alternating projections) *Let A and B be closed nonempty sets. Let $\overline{x}_1 \in \mathsf{Fix}\, T_{AP}$ and $\overline{x}_2 \in P_B\overline{x}_1$ such that $\overline{x}_1 \in P_A\overline{x}_2$ and let Γ be the affine subspace*

$$\Gamma := \{(x, x + \overline{x}_2 - \overline{x}_1) : x \in \mathbb{E}\} \subset \mathbb{E}^2.$$

Define $\Phi := T_{AP} - \mathrm{Id}$. Suppose the following hold:

(a) the collection of sets $\{A, B\}$ is subtransversal at $\overline{u} = (\overline{x}_1, \overline{x}_2)$ for $\overline{y} = (\overline{x}_1 - \overline{x}_2, \overline{x}_2 - \overline{x}_1)$ relative to Γ with constant κ and neighborhood U of $\overline{u}$;
(b) there exists a positive constant σ such that

$$\begin{aligned}&\forall x = (x_1, x_2) \in U \cap \Gamma,\\ &\mathrm{dist}\left(\overline{y}, (P_A - \mathrm{Id})(x_2) \times (P_B - \mathrm{Id})(x_1)\right) \leq \sigma\, \mathrm{dist}\left(0, \Phi(x)\right).\end{aligned} \tag{24}$$

Then the mapping Φ is metrically subregular on U for 0 relative to Γ with constant $\kappa\sigma$.

4.2 Necessary and sufficient conditions for local linear convergence

It was established in [13, Corollary 13(a)] that *local linear regularity* of the collection of sets (with a reasonably good quantitative constant as always for convergence analysis of nonconvex alternating projections) is sufficient for linear monotonicity of the method for (ε, δ)-subregular sets. This result is updated here in light of more recent results.

Proposition 11 (convergence of consistent alternating projections) *Let S be a nonempty subset of $A \cap B$. Let U be a neighborhood of S such that*

$$P_A(U) \subseteq U \quad \text{and} \quad P_B(U) \subseteq U. \tag{25}$$

Let Λ be an affine subspace containing S such that $T_{AP} : \Lambda \rightrightarrows \Lambda$. Define $\Phi := T_{AP} - \mathrm{Id}$. Let the sets A and B be elementally subregular at all $\widehat{x} \in S$ relative to Λ respectively for each

$$(a, v_A) \in V_A := \{(z,w) \in \operatorname{gph} N_A^{\mathrm{prox}} \mid z + w \in U \quad \textit{and} \quad z \in P_A(z+w)\}$$
$$(b, v_B) \in V_B := \{(z,w) \in \operatorname{gph} N_B^{\mathrm{prox}} \mid z + w \in U \quad \textit{and} \quad z \in P_B(z+w)\}$$

with respective constants $\varepsilon_A, \varepsilon_B \in [0,1)$ on the neighborhood U. Suppose that the following hold:

(a) for each $\widehat{x} \in S$, the collection of sets $\{A, B\}$ is subtransversal at $\widehat{x}$ relative to Λ with constant κ on the neighborhood U;
(b) there exists a positive constant σ such that condition (24) holds true;
(c) $\operatorname{dist}(x, S) \le \operatorname{dist}(x, A \cap B \cap \Lambda)$ for all $x \in U \cap \Lambda$;
(d) $\widetilde{\varepsilon}_A + \widetilde{\varepsilon}_B + \widetilde{\varepsilon}_A\widetilde{\varepsilon}_B < \frac{1}{2(\kappa\sigma)^2}$, where $\widetilde{\varepsilon}_A := 4\varepsilon_A\frac{1+\varepsilon_A}{(1-\varepsilon_A)^2}$ and $\widetilde{\varepsilon}_B := 4\varepsilon_B\frac{1+\varepsilon_B}{(1-\varepsilon_B)^2}$.

Then every sequence $(x_k)_{k\in\mathbb{N}}$ generated by $x_{k+1} \in T_{AP}x_k$ seeded by any point $x_0 \in A \cap U \cap \Lambda$ is linearly monotone relative to S with constant

$$c := \sqrt{1 + \widetilde{\varepsilon}_A + \widetilde{\varepsilon}_B + \widetilde{\varepsilon}_A\widetilde{\varepsilon}_B - \frac{1}{2(\kappa\sigma)^2}} \;<\; 1.$$

Consequently, $\operatorname{dist}(x_k, S) \to 0$ at least Q-linearly with rate c.

Proof. In light of Proposition 10 and the definition of linear monotonicity, Proposition 11 is a specialization of [25, Theorem 3.2] to the case of two sets with nonempty intersection. □

If $S = A \cap B \cap \Lambda$ in Proposition 11, then assumption (c) can obviously be omitted.

The next theorem shows that the converse to Proposition 11 holds more generally without any assumption on the elemental regularity of the individual sets. The proof of the next theorem uses the idea in the proof of [11, Theorem 6.2].

Theorem 4 (subtransversality is necessary for linear monotonicity of subsequences) *Let Λ, A, and B be closed subsets of $\mathbb{E}$, let $\overline{x} \in S \subset A \cap B \cap \Lambda$, and let $1 \le n \in \mathbb{N}$ and $c \in [0,1)$ be fixed. Suppose that for any sequence of alternating projections $(x_k)_{k\in\mathbb{N}}$ starting in Λ and sufficiently close to $\overline{x}$, there exists a subsequence of the form $(x_{j+nk})_{k\in\mathbb{N}}$ for some $j \in \{0, 1, \ldots, n-1\}$ that remains in Λ and is linearly monotone relative to S with constant c. Then the collection of sets $\{A, B\}$ is subtransversal at $\overline{x}$ relative to Λ with constant $\mathrm{sr}'[A,B](\overline{x}) \le \frac{2(2n^2-1-c(n-1))}{1-c}$.*

Proof. Let $\rho > 0$ be so small that any alternating projections sequence $(x_k)_{k\in\mathbb{N}}$ starting in $\mathbb{B}_\rho(\overline{x}) \cap \Lambda$ has a subsequence $(x_{j+nk})_{k\in\mathbb{N}}$ which is linearly monotone relative to S with constant c. Take any $x_0 \in A \cap \mathbb{B}_\rho(\overline{x}) \cap \Lambda$. Let us consider any alternating projections sequence $(x_k)_{k\in\mathbb{N}}$ starting at x_0 and such a subsequence $(x_{j+nk})_{k\in\mathbb{N}}$. On one hand,

$$\begin{aligned}2n\,\mathrm{dist}(x_j,B) &\geq \|x_j - x_{j+n}\| \geq \mathrm{dist}(x_j,S) - \mathrm{dist}(x_{j+n},S)\\ &\geq (1-c)\,\mathrm{dist}(x_j,S) \geq (1-c)\,\mathrm{dist}(x_j, A\cap B\cap \Lambda)\\ &\geq (1-c)\left(\mathrm{dist}(x_0, A\cap B\cap\Lambda) - \|x_0 - x_j\|\right).\end{aligned} \tag{26}$$

On the other hand,

$$\begin{aligned}\mathrm{dist}(x_0,B) &\geq \mathrm{dist}(x_j,B) - \|x_0 - x_j\|\\ &\geq \mathrm{dist}(x_j,B) - 2j\,\mathrm{dist}(x_0,B)\\ &\geq \mathrm{dist}(x_j,B) - 2(n-1)\,\mathrm{dist}(x_0,B).\end{aligned} \tag{27}$$

A combination of (26) and (27) yields

$$\begin{aligned}(2n-1)\,\mathrm{dist}(x_0,B) &\geq \frac{1-c}{2n}\left(\mathrm{dist}(x_0, A\cap B\cap\Lambda) - \|x_0 - x_j\|\right)\\ &\geq \frac{1-c}{2n}\,\mathrm{dist}(x_0, A\cap B\cap\Lambda) - \frac{(1-c)(n-1)}{n}\,\mathrm{dist}(x_0,B).\end{aligned}$$

Hence

$$\mathrm{dist}(x_0, A\cap B\cap\Lambda) \leq \frac{2(2n^2-1-c(n-1))}{1-c}\,\mathrm{dist}(x_0,B) \quad \forall x_0 \in A\cap\mathbb{B}_\rho(\overline{x})\cap\Lambda.$$

This yields subtransversality of $\{A,B\}$ at $\overline{x}$ relative to Λ and $\mathrm{sr}'[A,B](\overline{x}) \leq \frac{2(2n^2-1-c(n-1))}{1-c}$ as claimed. □

The next statement is an immediate consequence of Proposition 11 and Theorem 4.

Corollary 2 (subtransversality is necessary and sufficient for linear monotonicity) *Let $\Lambda \subset \mathbb{E}$ be an affine subspace and let A and B be closed subsets of $\mathbb{E}$ that are elementally subregular relative to $S \subset A\cap B\cap\Lambda$ at $\overline{x}\in S$ with constant ε and neighborhood $\mathbb{B}_\delta(\overline{x})\cap\Lambda$ for all $(a,v)\in \mathrm{gph}\,N_A^{\mathrm{prox}}$ with $a\in\mathbb{B}_\delta(\overline{x})\cap\Lambda$.*

Suppose that every sequence of alternating projections with the starting point sufficiently close to $\overline{x}$ is contained in Λ. All such sequences of alternating projections are linearly monotone relative to S with constant $c\in[0,1)$ if and only if the collection of sets is subtransversal at $\overline{x}$ relative to Λ (with an adequate balance of quantitative constants).

The next technical lemma allows us formally avoid the restriction "monotone" in Theorem 4.

Lemma 1 *Let $(x_k)_{k\in\mathbb{N}}$ be a sequence generated by T_{AP} that converges R-linearly to $\overline{x}\in A\cap B$ with rate $c\in[0,1)$. Then there exists a subsequence $(x_{k_n})_{n\in\mathbb{N}}$ that is linearly monotone with respect to any set $S\subset A\cap B$ with $\overline{x}\in S$.*

Proof. By definition of R-linear convergence, there is $\gamma < +\infty$ such that $\|x_k - \overline{x}\| \le \gamma c^k$ for all $k \in \mathbb{N}$. Let S be any set such that $\overline{x} \in S \subset A \cap B$. If $x_{k_0} := x_0 \notin S$, i.e., $\operatorname{dist}(x_{k_0}, S) > 0$, then there exists an iterate of $(x_k)_{k\in\mathbb{N}}$ (we choose the first one) relabeled x_{k_1} such that

$$\operatorname{dist}(x_{k_1}, S) \le \|x_{k_1} - \overline{x}\| \le \gamma c^{k_1} \le c \operatorname{dist}(x_{k_0}, S). \tag{28}$$

Repeating this argument for x_{k_1} in place of x_{k_0} and so on, we extract a subsequence $(x_{k_n})_{n\in\mathbb{N}}$ satisfying

$$\operatorname{dist}(x_{k_{n+1}}, S) \le c \operatorname{dist}(x_{k_n}, S) \quad \forall n \in \mathbb{N}.$$

The proof is complete. □

The above observation allows us to obtain the statement about necessary conditions for linear convergence of the alternating projections algorithm which extends Theorem 4. Here, the index number k_1 depending on the sequence $(x_k)_{k\in\mathbb{N}}$ will come into play in determining the constant of linear regularity.

Theorem 5 (subtransversality is necessary for linear convergence) *Let $m \in \mathbb{N}$ be fixed and $c \in [0, 1)$. Let Λ, A and B be closed subsets of $\mathbb{E}$ and let $\overline{x} \in S \subset A \cap B \cap \Lambda$. Suppose that any alternating projections sequence $(x_k)_{k\in\mathbb{N}}$ starting in $A\cap\Lambda$ and sufficiently close to $\overline{x}$ is contained in Λ, converges R-linearly to a point in S with rate c, and the index $k_1 \le m$ where k_1 satisfies* (28)*. Then the collection of sets $\{A, B\}$ is subtransversal at $\overline{x}$ relative to Λ with constant $\operatorname{sr}'[A, B](\overline{x}) \le \frac{2m}{1-c}$.*

Proof. Let $\rho > 0$ be so small that any alternating projections sequence starting in $A \cap \mathbb{B}_\rho(\overline{x}) \cap \Lambda$ converges R-linearly to a point in S with rate c. Take any $x_0 \in A \cap \mathbb{B}_\rho(\overline{x}) \cap \Lambda$ and generate an alternating projections sequence $(x_k)_{k\in\mathbb{N}}$. By Lemma 1, there is a subsequence $(x_{k_n})_{n\in\mathbb{N}}$ linearly monotone with respect to S at rate c. Then

$$\|x_{k_1} - x_{k_0}\| \ge \operatorname{dist}(x_{k_0}, S) - \operatorname{dist}(x_{k_1}, S) \ge (1 - c) \operatorname{dist}(x_{k_0}, S).$$

Let $b_0 \in P_B x_0$ be such that $x_1 \in P_A b_0$ and note that $x_{k_0} = x_0$. By the definition of the projection and $k_1 \le m$ it follows that

$$\begin{aligned} 2m \operatorname{dist}(x_0, B) &= 2m\|b_0 - x_0\| \\ &\ge \|x_{k_1} - x_0\| \ge (1 - c) \operatorname{dist}(x_0, S) \\ &\ge (1 - c) \operatorname{dist}(x_0, A \cap B \cap \Lambda). \end{aligned}$$

Hence

$$\operatorname{dist}(x_0, A \cap B) \le \frac{2m}{1 - c} \operatorname{dist}(x_0, B) \quad \forall x_0 \in A \cap \mathbb{B}_\rho(\overline{x}) \cap \Lambda.$$

This yields subtransversality of $\{A, B\}$ at $\overline{x}$ relative to Λ and $\operatorname{sr}'[A, B](\overline{x}) \le \frac{2m}{1-c}$ as claimed. □

The joining alternating projections sequence $(z_k)_{k\in\mathbb{N}}$ given by (18) often plays a role as an intermediate step in the analysis of alternating projections. As we shall see, property of linear extendability itself can also be of interest when dealing with the alternating projections algorithm, especially for non-convex setting. This observation can be seen for example in [7, 11, 23, 24, 28].

Theorem 6 (subtransversality is necessary for linear extendability of subsequences) *Let Λ, A, and B be closed subsets of $\mathbb{E}$, let $\overline{x} \in A\cap B\cap\Lambda$, and let $1 \le n \in \mathbb{N}$ and $c \in [0,1)$ be fixed. Suppose that every alternating projections sequence $(x_k)_{k\in\mathbb{N}}$ starting in $A\cap\Lambda$ and sufficiently close to $\overline{x}$ has a subsequence of the form $(x_{j+nk})_{k\in\mathbb{N}}$ for some $j \in \{0, 1, \dots, n-1\}$ such that the joining sequence $(z_k)_{k\in\mathbb{N}}$ given by (18) is a linear extension of (x_{j+nk}) on Λ with frequency $2n$ and rate c. Then the collection of sets $\{A, B\}$ is subtransversal at $\overline{x}$ relative to Λ with constant $\mathrm{sr}'[A, B](\overline{x}) \le \frac{2(2n-1-c(n-1))}{1-c}$.*

Proof. Let $\rho > 0$ be so small that any alternating projections sequence starting in $A \cap \mathbb{B}_\rho(\overline{x}) \cap \Lambda$ has a subsequence of the described form which admits the joining sequence as a linear extension on Λ with frequency $2n$ and rate c. Take any $x_0 \in A \cap \mathbb{B}_\rho(\overline{x}) \cap \Lambda$. Let us consider any alternating projections sequence $(x_k)_{k\in\mathbb{N}}$ starting at x_0, the corresponding joining sequence $(z_k)_{k\in\mathbb{N}}$ and the subsequence $(x_{j+nk})_{k\in\mathbb{N}}$. Let $\widetilde{x} \in \Lambda$ be the limit of $(z_k)_{k\in\mathbb{N}}$ as verified in Proposition 2.

On one hand,

$$\begin{aligned}
\mathrm{dist}(x_j, A\cap B\cap\Lambda) &\le \|x_j - \widetilde{x}\| = \|z_{2j} - \widetilde{x}\| \\
&\le \sum_{i=2j}^{\infty} \|z_i - z_{i+1}\| \le \frac{2n}{1-c}\|z_{2j} - z_{2j+1}\| \\
&= \frac{2n}{1-c}\,\mathrm{dist}(z_{2j}, B) = \frac{2n}{1-c}\,\mathrm{dist}(x_j, B) \\
&\le \frac{2n}{1-c}\,\mathrm{dist}(x_0, B),
\end{aligned} \tag{29}$$

where the last estimate follows from the nature of alternating projections.

On the other hand,

$$\begin{aligned}
\mathrm{dist}(x_j, A\cap B\cap\Lambda) &\ge \mathrm{dist}(x_0, A\cap B\cap\Lambda) - \|x_0 - x_j\| \\
&\ge \mathrm{dist}(x_0, A\cap B\cap\Lambda) - 2(n-1)\,\mathrm{dist}(x_0, B),
\end{aligned} \tag{30}$$

where the last estimate holds true since

$$\|x_0 - x_j\| \le 2j\,\mathrm{dist}(x_0, B) \le 2(n-1)\,\mathrm{dist}(x_0, B).$$

A combination of (29) and (30) then implies

$$\mathrm{dist}(x_0, A\cap B\cap\Lambda) \le \frac{2(2n-1-c(n-1))}{1-c}\,\mathrm{dist}(x_0, B) \quad \forall x_0 \in A\cap\mathbb{B}_\rho(\overline{x})\cap\Lambda,$$

which yields subtransversality of $\{A, B\}$ at $\overline{x}$ relative to Λ and $\mathrm{sr}'[A, B](\overline{x}) \le \frac{2(2n-1-c(n-1))}{1-c}$ as claimed. □

In general, subtransversality is not a sufficient condition for an alternating projections sequence to converge to a point in the intersection of the sets. For example, let us define the function $f : [0,1] \to \mathbb{R}$ by $f(0) = 0$ and on each interval of form $(1/2^{n+1}, 1/2^n]$,

$$f(t) = \begin{cases} -t + 1/2^{n+1}, & \text{if } t \in (1/2^{n+1}, 3/2^{n+2}], \\ t - 1/2^n, & \text{if } t \in (3/2^{n+2}, 1/2^n], \end{cases} \qquad (\forall n \in \mathbb{N})$$

and consider the sets: $A = \operatorname{gph} f$ and $B = \{(t, t/3) \mid t \in [0,1]\}$ and the point $\bar{x} = (0,0) \in A \cap B$ in $\mathbb{R}^2$. Then it can be verified that the collection of sets $\{A, B\}$ is subtransversal at $\bar{x}$ while the alternating projections method gets stuck at points $(1/2^n, 0) \notin A \cap B$.

To conclude this section, we show that the property of subtransversality of the collection of sets has been imposed either explicitly or implicitly in all existing linear convergence criteria for the method of alternating projections that we are aware of. The next proposition catalogs existing linear convergence criteria for alternating projections which complement Proposition 11.

Proposition 12 (R-linear convergence of nonconvex alternating projections) *Let A and B be closed and $\bar{x} \in A \cap B$. The collection of sets is denoted $\{A, B\}$. All alternating projections iterations starting sufficiently close to $\bar{x}$ converge R-linearly to some point in $A \cap B$ if one of the following conditions holds.*

(i) *[23, Theorem 4.3] A and B are smooth manifolds around $\bar{x}$ and $\{A, B\}$ is transversal at $\bar{x}$.*
(ii) *[11, Theorem 6.1] $\{A, B\}$ is intrinsically transversal at $\bar{x}$.*
(iii) *[24, Theorem 5.16] A is super-regular at $\bar{x}$ and $\{A, B\}$ is transversal at $\bar{x}$.*
(iv) *[6, Theorem 3.17] A is (B, ε, δ)-regular at $\bar{x}$ and the (A, B)-qualification condition holds at $\bar{x}$.*
(v) *[28, Theorem 2] A is 0-Hölder regular relative to B at $\bar{x}$ and $\{A, B\}$ intersects separably at $\bar{x}$.*

It can be recognized without much effort that under any item of Proposition 12, the sequences generated by alternating projections starting sufficiently close to $\bar{x}$ are actually linearly extendible.

Proposition 13 (ubiquity of subtransversality in linear convergence criteria) *Suppose than one of the conditions* (i)–(v) *of Proposition 12 is satisfied. Then for any alternating projections sequence $(x_k)_{k\in\mathbb{N}}$ starting sufficiently close to $\bar{x}$, the corresponding joining sequence $(z_k)_{k\in\mathbb{N}}$ given by* (18) *is a linear extension of $(x_k)_{k\in\mathbb{N}}$ with frequency* 2 *and rate $c \in [0,1)$.*

Proof. The statement can be observed directly from the key estimates that were used in proving the corresponding convergence criterion. In fact, all the criteria listed in Proposition 12 essentially were obtained from the same fundamental estimate which we named *linear extendability* in this paper. □

Taking Theorem 6 into account we conclude that subtransversality of the collection of sets $\{A, B\}$ at $\overline{x}$ is a consequence of each item listed in Proposition 12. This observation gives some insights about relationships between various regularity notions of collections of sets and has been formulated partly in [11, Theorem 6.2] and [22, Theorem 4]. Hence, the subtransversality property lies at the foundation of all linear convergence criteria for the method of alternating projections for both convex and nonconvex sets appearing in the literature to this point.

5 Application: alternating projections with convexity

In the convex setting, statements with sharper convergence rate estimates are possible. This is the main goal of the present section. Note that a convex set is elementally regular at all points in the set for all normal vectors with constant $\varepsilon = 0$ and the neighborhood $\mathbb{E}$ [22, Proposition 4(vii)]. We can thus, without loss of generality, remove the restriction to the subset Λ that is omnipresent in the nonconvex setting. We also write $P_A x$ and $P_B x$ for the projections since the projectors are single-valued.

The next technical lemma is fundamental for the subsequent analysis.

Lemma 2 (non-decreasing rate) *Let A and B be two closed convex sets in $\mathbb{E}$. We have*

$$\|P_B P_A P_B x - P_A P_B x\| \cdot \|P_B x - x\| \geq \|P_A P_B x - P_B x\|^2 \quad \forall x \in A. \tag{31}$$

Proof. Using the basic facts of the projection operators on a closed and convex sets, we obtain

$$\begin{aligned}
&\|P_A P_B x - P_B x\|^2 \leq \langle x - P_B x,\ P_A P_B x - P_B x\rangle \\
&\quad = \langle x - P_B x,\ P_A P_B x - P_B P_A P_B x\rangle + \langle x - P_B x,\ P_B P_A P_B x - P_B x\rangle \\
&\quad \leq \|x - P_B x\| \cdot \|P_A P_B x - P_B P_A P_B x\| .
\end{aligned}$$

The last estimate holds true since the second term on the previous line is non-positive. □

Lemma 2 implies that for any sequence $(x_k)_{k\in\mathbb{N}}$ of alternating projections for convex sets, the rate $\frac{\|x_{k+1}-x_k\|}{\|x_k-x_{k-1}\|}$ is nondecreasing when k increases. This allows us to deduce the following fact about the algorithm.

Theorem 7 (lower bound of complexity) *Consider the alternating projections algorithm for two closed convex sets A and B with a nonempty intersection. Then one of the following statements holds true.*

(i) *The alternating projections method finds a solution after one iteration.*
(ii) *Alternating projections will not reach a solution after any finite number of iterations.*

Proof. If the starting point is actually in $A \cap B$, the proof becomes trivial. Let us consider the case that $x_0 \in A \setminus B$. Suppose that the alternating projections method does not find a solution after one iterate, that is, $x_1 = P_A P_B x_0 \notin A \cap B$. In other words, we suppose that scenario (i) does not occur and prove the validity of scenario (ii).

In this case, it holds that $P_B x_0 \in B \setminus A$ as $x_1 = P_A P_B x_0 \notin A \cap B$. As a result, $\|x_1 - P_B x_0\| > 0$. We also claim that $\|x_1 - P_B x_0\| < \|P_B x_0 - x_0\|$. Indeed, suppose otherwise that $\|x_1 - P_B x_0\| = \|P_B x_0 - x_0\|$ (note that $\|x_1 - P_B x_0\| \leq \|P_B x_0 - x_0\|$ by definition of projection). Then $\|P_B x_0 - x_0\| = \operatorname{dist}(P_B x_0, A)$, which implies that x_0 is a fixed point of $P_A P_B$, $x_0 = P_A P_B x_0$. This contradicts the fact that $x_0 \in A \setminus B$ and $A \cap B \neq \emptyset$ since any alternating projections sequence for convex sets with nonempty intersection will converge to a point in the intersection [4]. Hence, we have checked that

$$0 < \|x_1 - P_B x_0\| < \|P_B x_0 - x_0\|.$$

Then the following constant is well defined:

$$\sqrt{c} := \frac{\|x_1 - P_B x_0\|}{\|P_B x_0 - x_0\|} \in (0, 1). \tag{32}$$

Using Lemma 2 we get

$$\frac{\operatorname{dist}(x_1, B)}{\|x_1 - P_B x_0\|} = \frac{\|P_B x_1 - x_1\|}{\|x_1 - P_B x_0\|} \geq \frac{\|x_1 - P_B x_0\|}{\|P_B x_0 - x_0\|} = \sqrt{c}.$$

Hence

$$\operatorname{dist}(x_1, B) \geq \sqrt{c}\|x_1 - P_B x_0\| = c\|P_B x_0 - x_0\| = c \operatorname{dist}(x_0, B) > 0.$$

Applying Lemma 2 consecutively, we obtain

$$\operatorname{dist}(x_k, B) \geq c^k \operatorname{dist}(x_0, B) > 0 \quad \forall k \in \mathbb{N}.$$

This particularly implies that $x_k \notin A \cap B$ for any natural number $k \in \mathbb{N}$, and the proof is complete. □

Remark 6 In contrast to Theorem 7 for convex sets, there are simple examples of nonconvex sets such that for any given number $n \in \mathbb{N}$, the alternating projections method will find a solution after exactly n iterates. For instance, let us consider a geometric sequence $z_k = \left(\frac{1}{3}\right)^k z_0$ where $0 \neq z_0 \in \mathbb{E}$. For any number $n \in \mathbb{N}$, one can construct the two finite sets by $A := \{z_{2k} \mid k = 0, 1, \ldots, n\}$ and $B := \{z_{2n}\} \cup \{z_{2k+1} \mid k = 0, 1, \ldots, n-1\}$. Then the alternating projections method starting at z_0 will find the unique solution z_{2n} after exactly n iterates.

Theorem 8 (necessary and sufficient condition: local version) *Let A and B be closed convex sets and $\overline{x} \in A \cap B$. If the collection of sets $\{A, B\}$ is subtransversal at $\overline{x}$ with constant $\mathrm{sr}'[A, B](\overline{x}) < +\infty$, then for any number $c \in (1-\mathrm{sr}'[A, B](\overline{x})^{-2}, 1)$, all alternating projections sequences starting sufficiently close to $\overline{x}$ are linearly monotone with respect to $A \cap B$ with rate not greater than c.*

Conversely, if there exists a number $c \in [0, 1)$ such that every alternating projections iteration starting sufficiently close to $\overline{x}$ converges R-linearly to some point in $A \cap B$ with rate not greater than c, then the collection of sets $\{A, B\}$ is subtransversal at $\overline{x}$ with constant $\mathrm{sr}'[A, B](\overline{x}) \le \frac{1}{1-c}$.

Proof. The first implication is an adaption of [13, Corollary 3.13(c)] to the terminology of this paper.

We now prove the converse implication. Let $\rho > 0$ be so small that every alternating projections iteration starting in $B_\rho(\overline{x})$ converges R-linearly to a point in $A \cap B$ with rate not greater than c. Take any $x_0 \in A \cap B_\rho(\overline{x})$. Let us consider the alternating projections sequence $(x_k)_{k\in\mathbb{N}}$ starting at x_0 and converging R-linearly to $\widetilde{x} \in A \cap B$ with rate not greater than c. By definition of R-linear convergence, there is a number $\gamma > 0$ such that

$$\|x_k - \widetilde{x}\| \le \gamma c^k \quad \forall k \in \mathbb{N}. \tag{33}$$

Taking Theorem 7 into account, we consider the two possible cases as follows.

Case 1. The alternating projections method finds a solution after one iterate, $x_1 = P_A P_B x_0 \in A \cap B$. Lemma 2 yields

$$\|x_1 - P_B x_0\|^2 \le \|P_B x_1 - x_1\| \cdot \|P_B x_0 - x_0\| = 0.$$

This implies that $P_B x_0 = x_1 \in A \cap B$ and as a result,

$$\mathrm{dist}(x_0, A \cap B) \le \|x_0 - P_B x_0\| = \mathrm{dist}(x_0, B). \tag{34}$$

Case 2. The alternating projections do not reach a solution after any finite number of iterates. We will make use of the joining sequence $(z_k)_{k\in\mathbb{N}}$ given by (18). Since P_A and P_B are firmly nonexpansive, the sequence $(z_k)_{k\in\mathbb{N}}$ is Fejér monotone with respect to $A \cap B$. Then it follows that

$$\|z_{k+1} - z_k\| = \max\{\mathrm{dist}(z_k, A), \mathrm{dist}(z_k, B)\} \le \|z_k - \widetilde{x}\| \le \frac{\gamma}{\sqrt{c}} \sqrt{c}^k \quad \forall k \in \mathbb{N}. \tag{35}$$

We claim that

$$\|z_{k+1} - z_k\| \le \sqrt{c}\|z_k - z_{k-1}\| \quad \forall k \in \mathbb{N}. \tag{36}$$

Suppose to the contrary that there exists a natural number $p \ge 1$ such that $\|z_{p+1} - z_p\| > \sqrt{c}\|z_p - z_{p-1}\|$. Choose a number $\theta > \sqrt{c}$ such that $\|z_{p+1} - z_p\| \ge \theta\|z_p - z_{p-1}\|$. Then applying Lemma 2, we get

$$\|z_{k+1} - z_k\| \ge \theta^{k-p+1}\|z_p - z_{p-1}\| \quad \forall k \ge p,\ k \in \mathbb{N}.$$

This together with (35) implies that for all natural number $k \geq p$,

$$
\begin{aligned}
&\frac{\gamma}{\sqrt{c}}\sqrt{c}^k \geq \theta^{k-p+1}\|z_p - z_{p-1}\| \\
\Leftrightarrow &\frac{\gamma}{\sqrt{c}}\sqrt{c}^k \geq \sqrt{c}^{k-p+1}\left(\frac{\theta}{\sqrt{c}}\right)^{k-p+1}\|z_p - z_{p-1}\| \\
\Leftrightarrow &\frac{\gamma}{\sqrt{c}}\sqrt{c}^{p-1} \geq \left(\frac{\theta}{\sqrt{c}}\right)^{k-p+1}\|z_p - z_{p-1}\|.
\end{aligned}
$$

Letting $k \to +\infty$, the last inequality leads to a contradiction since $\frac{\theta}{\sqrt{c}} > 1$. Hence, (36) has been proved.

Now, using (36) and the firm nonexpansiveness of P_A and P_B, we obtain that

$$
\begin{aligned}
\operatorname{dist}(x_0, A \cap B) \leq &\|x_0 - \widetilde{x}\| \leq \sum_{j=0}^{\infty}\|z_{2j+2} - z_{2j}\| \leq \sum_{j=0}^{\infty}\|z_{2j+1} - z_{2j}\| \\
\leq &\sum_{j=0}^{\infty}\|z_1 - z_0\|\sqrt{c}^{2j} \leq \frac{1}{1-c}\|z_1 - z_0\| = \frac{1}{1-c}\operatorname{dist}(x_0, B).
\end{aligned}
\tag{37}
$$

A combination of (34) and (37), which respectively correspond to the two cases, implies that

$$
\operatorname{dist}(x_0, A \cap B) \leq \frac{1}{1-c}\operatorname{dist}(x_0, B) \quad \forall x_0 \in A \cap B_\rho(\overline{x}).
$$

Hence $\{A, B\}$ is subtransversal at $\overline{x}$ and the constant $\operatorname{sr}'[A, B](\overline{x}) \leq \frac{1}{1-c}$ as claimed. □

The next theorem provides a global version of Theorem 8.

Theorem 9 (necessary and sufficient condition: global version) *Let A and B be closed convex sets with nonempty intersection. If the collection of sets $\{A, B\}$ is subtransversal at every point of (the boundary of) $A \cap B$ with constants bounded from above by $\kappa < +\infty$, then for any number $c \in (1 - \kappa^{-2}, 1)$, every alternating projections iteration converges R-linearly to a point in $A \cap B$ with rate not greater than c.*

Conversely, if there exists a number $c \in [0, 1)$ such that every alternating projections sequence eventually converges R-linearly to a point in $A \cap B$ with rate not greater than c, then the collection of sets $\{A, B\}$ is globally subtransversal with constant $\kappa \leq \frac{1}{1-c}$, that is,

$$
\operatorname{dist}(x, A \cap B) \leq \frac{1}{1-c}\operatorname{dist}(x, B) \quad \forall x \in A. \tag{38}
$$

Proof. We prove the first implication. Let us take any point $x_0 \in \mathbb{E}$ and consider the alternating projections sequence $(x_k)_{k\in\mathbb{N}}$ starting at x_0. It suffices to consider only the case that the alternating projections method does not find a solution after one iterate. It is well known that $(x_k)_{k\in\mathbb{N}}$ converges to some point $\widetilde{x} \in \mathrm{bd}\,(A\cap B)$ [4]. Hence, after a finite number, say p, of iterates, the iterate x_p must be sufficiently close to $\widetilde{x}$. Using the assumption that $\{A,B\}$ is subtransversal at $\widetilde{x}$ with constant $\mathrm{sr}'[A,B](\widetilde{x}) \le \kappa$ and applying Theorem 8, we deduce that the alternating projections sequence starting from x_p converges R-linearly to $\widetilde{x}$ with rate not greater than c. On one hand, using (36) for the alternating projections sequence starting from x_p and the corresponding joining sequence, we get

$$\|z_{k+1}-z_k\| \le \sqrt{c}\|z_k - z_{k-1}\| \quad \forall k \ge p+1, k\in\mathbb{N}. \tag{39}$$

On the other hand, applying Lemma 2, we get

$$\frac{\|z_{k+1}-z_k\|}{\|z_k-z_{k-1}\|} \le \frac{\|z_{p+2}-z_{p+1}\|}{\|z_{p+1}-z_p\|} \le \sqrt{c} \quad \forall k \le p, k\in\mathbb{N}. \tag{40}$$

A combination of (39) and (40) yields the estimate (36). Proposition 4 then ensures that the joining sequence $(z_k)_{k\in\mathbb{N}}$ converges R-linearly to $\widetilde{x}$ with rate not greater than $\sqrt{c}$. This implies that the sequence $(x_k)_{k\in\mathbb{N}}$ converges R-linearly to $\widetilde{x}$ with rate not greater than c as claimed.

We now prove the converse implication. Suppose that every sequence of alternating projections eventually converges R-linearly to a point in $A\cap B$ with rate not greater than c. We need to verify (38). Note that the estimate (38) is trivial for $x\in A\cap B$. Let us take an arbitrary $x\in A\setminus B$ and consider the alternating projections sequence $(x_k)_{k\in\mathbb{N}}$ starting at $x_0 = x$. We consider the two possible cases as stated in Theorem 7.

Case 1. The alternating projections method finds a solution after one iterate. The argument for *Case 1* of the proof of Theorem 8 yields (38).

Case 2. The alternating projections method does not find a solution after any finite number of iterates. Since $(x_k)_{k\in\mathbb{N}}$ eventually converges R-linearly to a point $\widetilde{x}\in A\cap B$ with rate not greater than c, there exists a natural number $p\in\mathbb{N}$ and a constant $\gamma'>0$ such that

$$\|x_k-\widetilde{x}\| \le \gamma' c^{k-p} = \frac{\gamma'}{c^p}c^k \quad \forall k\ge p. \tag{41}$$

Let us define the number

$$\gamma := \max\left\{\frac{\gamma'}{c^p}, \frac{\|x_k-\widetilde{x}\|}{c^k} \,:\; k=0,1,\ldots,p\right\} > 0. \tag{42}$$

Combining (41) and (42) yields

$$\|x_k-\widetilde{x}\| \le \gamma c^k \quad \forall k\in\mathbb{N}.$$

The argument for *Case 2* in the proof of Theorem 8 implies that the sequence $(z_k)_{k\in\mathbb{N}}$ defined at (18) satisfies

$$\|z_{k+1}-z_k\| \le \sqrt{c}\|z_k-z_{k-1}\| \quad \forall k\in\mathbb{N}.$$

From this condition, the estimate (38) is obtained by using the estimates at (37).

The proof is complete. □

It is clear that Theorem 8 does not cover Theorem 9. The following example also rules out the inverse implication.

Example 3 (Theorem 9 does not cover Theorem 8) *Consider the convex function $f : \mathbb{R} \to \mathbb{R}$ given by*

$$f(t) = \begin{cases} t^2, & \text{if } t \in [0, \infty), \\ 0, & \text{if } t \in [-1, 0), \\ -t-1, & \text{if } t \in (-\infty, -1). \end{cases}$$

In $\mathbb{R}^2$, we define two closed convex sets $A := \operatorname{epi} f$ and $B := \mathbb{R} \times \mathbb{R}_-$ and a point $\overline{x} = (-1, 0) \in A \cap B$. Then the two equivalent properties (namely, transversality of $\{A, B\}$ at $\overline{x}$ and local linear convergence of T_{AP} around $\overline{x}$) involved in Theorem 8 hold true while the two global ones involved in Theorem 9 do not.

To establish global convergence of a fixed point iteration, one normally needs some kind of global regularity behavior of the fixed point set. In Theorem 9, we formally impose only subtransversality in order to deduce global R-linear convergence and vice versa. Beside the global behavior of convexity, the hidden reason behind this seemingly contradictory phenomenon is a well known fact about subtransversality of collections of convex sets. We next deduce this result from the convergence analysis above. The proof is given for completeness.

Corollary 3 *Let A and B be closed and convex subsets of $\mathbb{E}$ with nonempty intersection. The collection of sets $\{A, B\}$ is globally subtransversal, that is, there is a constant $\kappa < +\infty$ such that*

$$\operatorname{dist}(x, A \cap B) \leq \kappa \operatorname{dist}(x, B) \quad \forall x \in A, \tag{43}$$

if and only if $\{A, B\}$ is subtransversal at every point in $\operatorname{bd}(A \cap B)$ with constants bounded from above by some $\overline{\kappa} < +\infty$.

Proof. ($\Rightarrow$) This implication is trivial with $\overline{\kappa} = \kappa$.
($\Leftarrow$) Note that the estimate (43) is trivial for $x \in A \cap B$. Let us take an arbitrary $x \in A \setminus B$ and consider the alternating projections sequence $(x_k)_{k \in \mathbb{N}}$ starting at $x_0 = x$. Take any number $c \in \left(1 - \overline{\kappa}^{-2}, 1\right)$ The argument in the first part of Theorem 9 implies that $(x_k)_{k \in \mathbb{N}}$ converges R-linearly to some $\widetilde{x} \in A \cap B$ with rate c and

$$\operatorname{dist}(x, A \cap B) \leq \frac{1}{1-c} \operatorname{dist}(x, B).$$

By letting $c \downarrow 1 - \overline{\kappa}^{-2}$ in the above inequality, we obtain (43) with $\kappa = \overline{\kappa}^2$.

The proof is complete. □

The convergence counterpart of Corollary 3 can also be of interest.

Corollary 4 *Let $(x_k)_{k\in\mathbb{N}}$ be an alternating projections sequence for two closed convex subsets of $\mathbb{E}$ with nonempty intersection and $c \in [0,1)$. If there exists a natural number $p \in \mathbb{N}$ such that $\|x_k - \widetilde{x}\| \le \gamma c^k$ for all $k \ge p$, then $\|x_k - \widetilde{x}\| \le \gamma c^k$ for all $k \in \mathbb{N}$.*

We emphasize that the two statements in Corollary 4 are always equivalent (by the argument for the second part of Theorem 9) if the constant γ is not required to be the same. However, this requirement becomes important when one wants to estimate global rate of convergence via the local rate of convergence. The next statement can easily be observed as a by-product via the proof of Theorem 8.

Proposition 14 (equivalence of linear monotonicity and R-linear convergence) *For sequences of alternating projections between convex sets, R-linear convergence and linear monotonicity of the sequence of iterates are equivalent.*

The next statement can serve as a motivation for Definition 3.

Proposition 15 (Q-linear convergence implies linear extendability) *Let $(x_k)_{k\in\mathbb{N}}$ be a sequence of alternating projections for two closed convex sets $A, B \subset \mathbb{E}$ with nonempty intersection. If $(x_k)_{k\in\mathbb{N}}$ converges Q-linearly to a point $\widetilde{x} \in A\cap B$ with rate $c \in [0,1)$, then $(x_k)_{k\in\mathbb{N}}$ is linearly extendible with frequency 2 and rate c, and the corresponding joining sequence $(z_k)_{k\in\mathbb{N}}$ is such a linear extension sequence.*

Before proving this, we first establish the following technical fact.

Lemma 3 *Let A and B be two closed convex sets in $\mathbb{E}$ with nonempty intersection. We have*

$$\|P_Ba - x\|\|P_Ba - a\| \ge \|a - x\|\|P_AP_Ba - P_Ba\| \quad \forall a \in A,\ \forall x \in A\cap B. \tag{44}$$

Proof. [of Lemma 3.] Denote $b = P_Ba$ and $a_+ = P_AP_Ba$. It suffices to consider the two cases as follows.

Case 1. $\|b - x\| = 0$ or $\|b - a\| = 0$. This implies that $b \in A\cap B$, which in turn implies that $a_+ = b$. Hence, inequality (44) is satisfied.

Case 2. Both sides of (44) are strictly positive. Let a' be the projection of b on the line (segment, equivalently since $\langle x - b,\ a - b\rangle \le 0$) joining x and a. The elementary geometry for triangles Δxba and $\Delta xba'$, respectively, yields

$$\begin{aligned}\|b - a\| &\ge \|a - x\|\sin\angle(b - x, a - x) > 0,\\ \|a' - b\| &= \|b - x\|\sin\angle(b - x, a - x) > 0.\end{aligned}$$

This implies

$$\|b - x\|\|b - a\| \ge \|b - x\|\|a - x\|\sin\angle(b - x, a - x) = \|a - x\|\|a' - b\|.$$

Inequality (44) now follows since, by convexity of A, $a' \in A$, and by definition of the projector,

$$\|a' - b\| \geq \operatorname{dist}(b, A) = \|a_+ - b\|.$$

This proves Lemma 3. □

We conclude with the proof of Proposition 15.

Proof. [of Proposition 15.] It suffices to prove that the sequence $(z_k)_{k\in\mathbb{N}}$ given in (18) satisfies

$$\|z_{k+2} - z_{k+1}\| \leq \sqrt{c}\|z_{k+1} - z_k\| \quad \forall k \in \mathbb{N}.$$

We will prove this by way of contradiction. Suppose otherwise that there exists some $p \in \mathbb{N}$ such that

$$\|z_{p+2} - z_{p+1}\| > \sqrt{c}\|z_{p+1} - z_p\|.$$

We can assume $p = 2k$ without loss of generality. By Lemma 2 we get

$$\frac{\|z_{2k+3} - z_{2k+2}\|}{\|z_{2k+2} - z_{2k+1}\|} \geq \frac{\|z_{2k+2} - z_{2k+1}\|}{\|z_{2k+1} - z_{2k}\|} > \sqrt{c}.$$

Lemma 3 then implies

$$\begin{aligned} \frac{\|x_{k+1} - \widetilde{x}\|}{\|x_k - \widetilde{x}\|} &= \frac{\|z_{2k+2} - \widetilde{x}\|}{\|z_{2k} - \widetilde{x}\|} = \frac{\|z_{2k+2} - \widetilde{x}\|}{\|z_{2k+1} - \widetilde{x}\|} \frac{\|z_{2k+1} - \widetilde{x}\|}{\|z_{2k} - \widetilde{x}\|} \\ &\geq \frac{\|z_{2k+3} - z_{2k+2}\|}{\|z_{2k+2} - z_{2k+1}\|} \frac{\|z_{2k+2} - z_{2k+1}\|}{\|z_{2k+1} - z_{2k}\|} > c. \end{aligned}$$

This contradicts Q-linear convergence of $(x_k)_{k\in\mathbb{N}}$ to $\widetilde{x}$ with rate c, and the proof is complete. □

References

1. Aragón Artacho FJ, Mordukhovich BS (2011) Enhanced metric regularity and Lipschitzian properties of variational systems. J Global Optim 50(1):145–167, DOI 10.1007/s10898-011-9698-x, URL http://dx.doi.org/10.1007/s10898-011-9698-x
2. Aspelmeier T, Charitha C, Luke DR (2016) Local linear convergence of the ADMM/Douglas–Rachford algorithms without strong convexity and application to statistical imaging. SIAM J Imaging Sci 9(2):842–868
3. Bauschke HH, Borwein JM (1993) On the convergence of von Neumann's alternating projection algorithm for two sets. Set-Valued Anal 1(2):185–212
4. Bauschke HH, Borwein JM (1996) On projection algorithms for solving convex feasibility problems. SIAM Rev 38(3):367–426
5. Bauschke HH, Combettes PL (2011) Convex Analysis and Monotone Operator Theory in Hilbert Spaces. CMS Books Math./Ouvrages Math. SMC, Springer, New York
6. Bauschke HH, Luke DR, Phan HM, Wang X (2013a) Restricted Normal Cones and the Method of Alternating Projections: Applications. Set-Valued Var Anal 21:475–501, DOI 10.1007/s11228-013-0238-3, URL http://dx.doi.org/10.1007/s11228-013-0238-3
7. Bauschke HH, Luke DR, Phan HM, Wang X (2013b) Restricted Normal Cones and the Method of Alternating Projections: Theory. Set-Valued Var Anal 21:431–473, DOI 10.1007/s11228-013-0239-2, URL http://dx.doi.org/10.1007/s11228-013-0239-2

8. Bolte J, Daniilidis A, Ley O, Mazet L (2010) Characterizations of Łojasiewicz inequalities: subgradient flows, talweg, convexity. Trans Am Math Soc 362(6):3319–3363, DOI 10.1090/S0002-9947-09-05048-X
9. Bunt LNH (1934) Bitdrage tot de theorie der konvekse puntverzamelingen. PhD thesis, Univ. of Groningen, Amsterdam
10. Dontchev AL, Rockafellar RT (2009) Implicit Functions and Solution Mappings. Monographs in Mathematics, Springer, New York
11. Drusvyatskiy D, Ioffe AD, Lewis AS (2015) Transversality and alternating projections for nonconvex sets. Found Comput Math 15(6):1637–1651, DOI 10.1007/s10208-015-9279-3, URL http://dx.doi.org/10.1007/s10208-015-9279-3
12. Giselsson P (2015) Tight global linear convergence rate bounds for Douglas–Rachford splitting, arXiv:1506.01556
13. Hesse R, Luke DR (2013) Nonconvex notions of regularity and convergence of fundamental algorithms for feasibility problems. SIAM J Optim 23(4):2397–2419
14. Ioffe AD (1989) Approximate subdifferentials and applications: III. Mathematika 36(71):1–38
15. Ioffe AD (2000) Metric regularity and subdifferential calculus. Russian Mathematical Surveys 55(3):501, URL http://stacks.iop.org/0036-0279/55/i=3/a=R03
16. Ioffe AD (2011) Regularity on a fixed set. SIAM J Optim 21(4):1345–1370, DOI 10.1137/110820981, URL http://10.1137/110820981
17. Ioffe AD (2013) Nonlinear regularity models. Math Program 139(1-2):223–242, DOI 10.1007/s10107-013-0670-z, URL http://dx.doi.org/10.1007/s10107-013-0670-z
18. Kruger AY (2006) About regularity of collections of sets. Set-Valued Anal 14:187–206
19. Kruger AY (2009) About stationarity and regularity in variational analysis. Taiwanese J Math 13(6A):1737–1785
20. Kruger AY, Thao NH (2015) Quantitative characterizations of regularity properties of collections of sets. J Optim Theory and Appl 164:41–67, DOI 10.1007/s10957-014-0556-0
21. Kruger AY, Luke DR, Thao NH (2017) About subtransversality of collections of sets. Set-Valued and Variational Analysis 25(4):701–729
22. Kruger AY, Luke DR, Thao NH (2018) Set regularities and feasibility problems. Math Program B 168:279–311, DOI 10.1007/s10107-016-1039-x, arXiv:1602.04935
23. Lewis AS, Malick J (2008) Alternating projections on manifolds. Math Oper Res 33:216–234
24. Lewis AS, Luke DR, Malick J (2009) Local linear convergence of alternating and averaged projections. Found Comput Math 9(4):485–513
25. Luke DR, Thao NH, Tam MK (2018) Quantitative convergence analysis of iterated expansive, set-valued mappings. Math Oper Res DOI 10.1287/moor.2017.0898, URL https://doi.org/10.1287/moor.2017.0898, http://arxiv.org/abs/1605.05725
26. Mordukhovich B (2006) Variational Analysis and Generalized Differentiation, I: Basic Theory; II: Applications. Grundlehren Math. Wiss., Springer-Verlag, New York
27. Ngai HV, Théra M (2001) Metric inequality, subdifferential calculus and applications. Set-Valued Anal 9:187–216, URL http://dx.doi.org/10.1023/A:1011291608129, 10.1023/A:1011291608129
28. Noll D, Rondepierre A (2016) On local convergence of the method of alternating projections. Found Comput Math 16(2):425–455, DOI 10.1007/s10208-015-9253-0, URL http://dx.doi.org/10.1007/s10208-015-9253-0
29. Ortega JM, Rheinboldt WC (1970) Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York
30. Penot JP (2013) Calculus Without Derivatives. Springer, New York
31. Phan H (2016) Linear convergence of the Douglas–Rachford method for two closed sets. Optimization 65:369–385
32. Poliquin RA, Rockafellar RT, Thibault L (2000) Local differentiability of distance functions. Trans Amer Math Soc 352(11):5231–5249
33. Rockafellar RT, Wets RJ (1998) Variational Analysis. Grundlehren Math. Wiss., Springer-Verlag, Berlin
34. Zheng XY, Ng KF (2008) Linear regularity for a collection of subsmooth sets in Banach spaces. SIAM J Optim 19(1):62–76