Implicit error bounds for Picard iterations on Hilbert spaces

Luke, D.R.; Nguyen, Thao; Tam, M.K.

DOI
10.1007/s10013-018-0279-x

Publication date
2018

Document Version
Final published version

Published in
Vietnam Journal of Mathematics

Vietnam J. Math. (2018) 46:243–258
https://doi.org/10.1007/s10013-018-0279-x

Implicit Error Bounds for Picard Iterations on Hilbert Spaces

D. Russell Luke[1] · **Nguyen H. Thao**[2] ·
Matthew K. Tam[1]

Received: 27 April 2017 / Accepted: 27 November 2017 / Published online: 21 February 2018

Abstract We investigate the role of error bounds, or metric subregularity, in the convergence of Picard iterations of nonexpansive maps in Hilbert spaces. Our main results show, on one hand, that the existence of an error bound is sufficient for strong convergence and, on the other hand, that an error bound exists on bounded sets for nonexpansive mappings possessing a fixed point whenever the space is finite dimensional. In the Hilbert space setting, we show that a monotonicity property of the distances of the Picard iterations is all that is needed to guarantee the existence of an error bound. The same monotonicity assumption turns out also to guarantee that the distance of Picard iterates to the fixed point set converges to zero. Our results provide a quantitative characterization of strong convergence as well as new criteria for when strong, as opposed to just weak, convergence holds.

Keywords Averaged operators · Error bounds · Strong convergence · Fixed points · Picard iteration · Metric regularity · Metric subregularity · Nonexpansiveness

Mathematics Subject Classification (2010) 49J53 · 65K10 · 49M05 · 49M27 · 65K05 · 90C30

This paper is dedicated to Professor Michel Théra on his 70th birthday.

✉ D. Russell Luke
r.luke@math.uni-goettingen.de

Nguyen H. Thao
h.t.nguyen-3@tudelft.nl

Matthew K. Tam
m.tam@math.uni-goettingen.de

[1] Institut für Numerische und Angewandte Mathematik, Universität Göttingen, 37083 Göttingen, Germany

[2] Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands

1 Introduction

Consider a simple Picard iteration, $x^{k+1} = Tx^k$, on a Hilbert space $\mathcal{H}$. The self-mapping $T : \mathcal{H} \to \mathcal{H}$ can represent a variety of sophisticated algorithms for finding the zeros of an operator, from accelerated first-order splitting methods to Newton's method, where step-lengths and dependencies on previous points are deterministically prescribed. While much of fixed point theory is concerned with existence of fixed points, we take this for granted and focus instead on characterizing convergence.

A celebrated theorem of Opial [25] shows that any fixed point iteration of a nonexpansive, asymptotically regular operator converges weakly to a fixed point (provided that said operator possesses at least one fixed point). Such a setting has wide-ranging applications including all *averaged* operators possessing fixed points. In a general (infinite) dimensional space, the conclusion of weak convergence in Opial's theorem cannot be strengthened to strong convergence without additional structure even in the simplest infinite dimensional setting of $\ell_2(\mathbb{N})$. Counter-examples to strong convergence can be found for a firmly nonexpansive operator in [13] and for an instance of the alternating projections algorithm in [15]. To ensure strong convergence in infinite dimensions, it suffices to assume so-called *linear* [2, 6, 20] or *Hölder regularity* properties [8, 9] of the underlying operator T – more precisely, *metric (Hölder) subregularity* of the operator $(\mathrm{Id} - T)$ at fixed points of T for 0 where Id denotes the identity mapping. Moreover, such properties allow one to deduce convergence rates for the fixed point sequence.

In [22], a program of analysis was proposed for quantifying the convergence of fixed point iterations in finite dimensions based on two properties of the mapping T: *pointwise (almost) nonexpansiveness* and *(gauge) metric subregularity* at certain points of interest (fixed points, for instance, but not exclusively). As was discussed above, when mere convergence is all that one is after, Opial-type results show that the requirement of metric subregularity can be dropped in the setting in which T is nonexpansive. On the other hand, in finite dimensions, it has been shown that strict monotonicity of a Picard iteration with respect to the set of fixed points, together with (almost) averagedness of the fixed point mapping, implies metric subregularity of $\mathrm{Id} - T$ [23, Corollary 3.12]. Other authors have approached this analysis from different angles (see for instance [7, 10, 11, 21, 24]). In contrast to these finite dimensional facts, in infinite dimensions, the critical distinction is not between quantitative convergence estimates and just convergence, but rather between strong and weak convergence; in some cases, the former is associated with faster rates of convergence [14]. The question we address in this note is whether metric subregularity is necessary for strong convergence, both in finite and infinite dimensional settings.

Continuing along these lines, our first main result, Theorem 1 in Section 3, shows that a general *functional subregularity* property suffices to deduce strong convergence of fixed point iterations. The type of subregularity is neither metric, nor of gauge-type because the associated function quantifying convergence need not be monotone. While alone this result is not surprising, it does unify many results in the literature, each concerned with a different rate of convergence for the same algorithm. The rest of Section 3 refines this result for averaged mappings which are gauge-metrically subregular at fixed points and explores the role of error bounds in deducing the rate of convergence of fixed point iterations (see Theorem 2). In Section 4, we study a converse implication, namely necessary conditions for the existence of a gauge-type subregularity property – what we refer to as an *implicit error bound*. Our second main result, Theorem 4, shows that, for self-mappings possessing fixed points, uniform monotonicity of the distance of Picard iterates to the set of fixed points, from any starting point on bounded sets, implies existence of an error bound. The same uniform monotonicity on bounded sets of the distance of Picard iterates of self-mappings to their fixed points is shown in Proposition 3 to imply that the distance of the iterates to the set of fixed points converges to zero. In finite dimensions, we show in Theorem 3 that it suffices to assume that the mapping is nonexpansive and possesses fixed points in order to guarantee the existence of error bounds on bounded sets. We begin in Section 2 with preliminary results and definitions.

2 Preliminaries

We limit our attention to self-mappings on a Hilbert space, $T : \mathcal{H} \to \mathcal{H}$. The domain of such mappings, denoted dom T is the set of points x where $Tx \neq \emptyset$. The distance of a point x to a set $C \subset \mathcal{H}$ is denoted $\operatorname{dist}(x, C)$ and defined by

$$\operatorname{dist}(x, C) := \inf_{y \in C}\{\|x - y\|\}.$$

We use the convention that the distance to the empty set is $+\infty$. We use the *excess* to characterize the distance between sets: for two sets C_1 and C_2

$$\operatorname{exc}(C_1, C_2) := \sup\{\operatorname{dist}(x, C_2) : \ x \in C_1\}.$$

This is finite whenever C_2 is nonempty and C_1 is bounded and nonempty.

Definition 1 Let D be a nonempty subset of $\mathcal{H}$ and let $T : D \to \mathcal{H}$. We say that T is

(i) *nonexpansive* if

$$\|Tx - Ty\| \leq \|x - y\| \quad \forall x, y \in D;$$

(ii) *firmly nonexpansive* if

$$\|Tx - Ty\|^2 + \|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 \leq \|x - y\|^2 \quad \forall x, y \in D;$$

(iii) *averaged with constant* $\gamma > 0$ if

$$\|Tx - Ty\|^2 + \gamma\|(\mathrm{Id} - T)x - (\mathrm{Id} - T)y\|^2 \leq \|x - y\|^2 \quad \forall x, y \in D;$$

(iv) *asymptotically regular* if

$$(\mathrm{Id} - T)T^n x \to 0 \text{ as } n \to \infty \quad \forall x \in D.$$

Fact 1 (Properties of nonexpansive operators) *Let D be a nonempty closed convex set and $T : D \to D$.*

(i) *If T is averaged and* $\mathsf{Fix}\, T \neq \emptyset$, *then T is asymptotically regular.*
(ii) *If T is nonexpansive and* $\mathsf{Fix}\, T \neq \emptyset$, *then* $\mathsf{Fix}\, T$ *is a closed convex set.*

Proof (i): [1, Theorem 2.1] or [17]. (ii): [3, Propositions 4.13 and 4.14]. □

Definition 2 Let C be a nonempty subset of $\mathcal{H}$ and (x_n) be a sequence in $\mathcal{H}$. Then, (x_n) is *Fejér monotone* with respect to C if, for all $x \in C$, it holds that

$$\|x_{n+1} - x\| \leq \|x_n - x\| \quad \forall n \in \mathbb{N}.$$

Fact 2 (Properties of Fejér monotone sequences) *Let C be a nonempty closed convex subset of $\mathcal{H}$ and suppose $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to C. The following assertions hold.*

(i) $(\mathrm{dist}(x_n, C))_{n\in\mathbb{N}}$ *is nonincreasing and convergent.*
(ii) $(x_n)_{n\in\mathbb{N}}$ *converges strongly to a point in C, say x, if and only if $d(x_n, C) \to 0$. In this case,*

$$\|x_n - x\| \le 2\mathrm{dist}(x_n, C) \quad \forall n \in \mathbb{N}.$$

Proof (i): [3, Proposition 5.4]. (ii): [3, Theorem 5.11 and (5.8)]. □

For the remainder of the note, we consider the sequence of Picard iterates of T, that is, a sequence (x_n) with

$$x_0 \in \mathcal{H}, \quad x_{n+1} = Tx_n \quad \forall n \in \mathbb{N}.$$

3 Sufficient Conditions for Strong Convergence

We begin this section by showing that, in the setting of Opial's theorem, the existence of an error bound is sufficient for strong convergence of Picard iterates.

Theorem 1 (Strong convergence for nonexpansive maps) *Let D be a nonempty closed convex subset of $\mathcal{H}$ and let $T : D \to D$ be nonexpansive and asymptotically regular with $\mathsf{Fix}\, T \neq \emptyset$. Suppose that, on each bounded subset U of D, there exists a function $\kappa : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\lim_{t\searrow 0} \kappa(t) = 0$ and*

$$\mathrm{dist}(x, \mathsf{Fix}\, T) \le \kappa(\|x - Tx\|) \quad \forall x \in U. \tag{1}$$

Then, for any $x_0 \in D$, the sequence $(x_n)_{n\in\mathbb{N}}$ defined by $x_{n+1} := Tx_n$ for all $n \in \mathbb{N}$ converges strongly to a point in $\mathsf{Fix}\, T$.

Proof Since T is nonexpansive $\mathsf{Fix}\, T$ is closed and convex and, for any $z \in \mathsf{Fix}\, T$, it holds that

$$\|x_{n+1} - z\| \le \|x_n - z\| \quad \forall n \in \mathbb{N}.$$

In other words, $(x_n)_{n\in\mathbb{N}}$ is Fejér monotone with respect to $\mathsf{Fix}\, T$. By Fact 2(i), the sequence (x_n) is bounded and so there is a bounded set U containing (x_n). By assumption, there is a function κ which satisfies (1), hence

$$\mathrm{dist}(x_n, \mathsf{Fix}\, T) \le \kappa(\|x_n - x_{n+1}\|).$$

But since T is asymptotically regular $\kappa(\|x_n - x_{n+1}\|) \to 0$ and so by Fact 2(ii), we deduce that (x_n) is strongly convergent to a point in $\mathsf{Fix}\, T$. □

In particular, the previous theorem applies to any averaged operator with fixed points for which the error bound in (1) holds.

Corollary 1 (Strong convergence for averaged maps) *Let D be a nonempty closed convex subset of $\mathcal{H}$ and let $T : D \to D$ be averaged with $\mathsf{Fix}\, T \neq \emptyset$. Suppose that, on each bounded subset U of D, there exists a function $\kappa : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\lim_{t\searrow 0} \kappa(t) = 0$ and condition (1) is satisfied. For any $x_0 \in D$, define $x_{n+1} := Tx_n$ for all $n \in \mathbb{N}$. Then, (x_n) converges strongly to a point in $\mathsf{Fix}\, T$.*

Proof Since an averaged operator with a fixed point is asymptotically regular by Fact 1(i), the result follows from Theorem 1. □

We note that, in order to deduce strong convergence in Theorem 1, it is not necessary to assume that the function κ is a *gauge function*. Indeed, only continuity at the origin is required (rather than on all of $\mathbb{R}_+$) and, moreover, κ need not be strictly increasing. However, in the case where κ is a gauge function, we have the following refinement.

Definition 3 (Gauge function [16, § 2]) A function $\kappa : \mathbb{R}_+ \to \mathbb{R}_+$ is said to be a *gauge function* if it is continuous, strictly increasing, $\kappa(0) = 0$ and $\kappa(t) \to +\infty$ as $t \to +\infty$.

In what follows, the n-fold composition of a function $\varphi : \mathbb{R} \to \mathbb{R}$ is denoted by

$$\varphi^n := \underbrace{\varphi \circ \cdots \varphi \circ \varphi}_{n \text{ times}}.$$

Theorem 2 (Error bound estimate for convergence rate) *Let D be a nonempty closed convex subset of $\mathcal{H}$ and let $T : D \to D$ be averaged with $\operatorname{Fix} T \neq \emptyset$. Suppose that, on each bounded subset U of D, there exists a gauge function $\kappa : \mathbb{R}_+ \to \mathbb{R}_+$ such that condition (1) is satisfied and*

$$\lim_{n\to\infty} \varphi^n(t) = 0 \quad \forall t \geq 0 \qquad \textit{where} \quad \varphi(t) := \sqrt{t^2 - \gamma\kappa^{-1}(t)^2}. \tag{2}$$

For any $x_0 \in D$, define $x_{n+1} := Tx_n$ for all $n \in \mathbb{N}$. Then $x_n \to x^* \in \operatorname{Fix} T$ and

$$\|x_n - x^*\| \leq 2\varphi^n \left(\operatorname{dist}(x_0, \operatorname{Fix} T)\right) \to 0 \ \textit{ as } \ n \to \infty. \tag{3}$$

In other words, (x_n) converges strongly to x^* with rate no worse than the rate at which $\varphi^n(\operatorname{dist}(x_0, \operatorname{Fix} T)) \searrow 0$.

Proof For convenience, denote $d_n := \operatorname{dist}(x_n, \operatorname{Fix} T)$ for all $n \in \mathbb{N}$. As T is averaged, in particular, T is also nonexpansive and hence, (x_n) is Fejér monotone with respect to $\operatorname{Fix} T$. By Fact 2(i), the sequence (x_n) with $x_0 \in D$ is bounded and so we may let U denote a bounded subset of D containing (x_n). By assumption, there is a gauge function κ which satisfies (1) and, moreover, its inverse κ^{-1} is also a gauge function with

$$\kappa^{-1} \left(\operatorname{dist}(x, \operatorname{Fix} T)\right) \leq \|x - Tx\| \quad \forall x \in U.$$

Since T is averaged, it holds that

$$d_{n+1}^2 + \gamma\|x_n - x_{n+1}\|^2 \leq \|x_{n+1} - P_{\operatorname{Fix} T} x_n\|^2 + \gamma\|x_n - x_{n+1}\|^2 \leq d_n^2.$$

Combining the last two inequalities gives $d_{n+1}^2 \leq d_n^2 - \gamma[\kappa^{-1}(d_n)]^2$, or equivalently, $d_{n+1} \leq \varphi(d_n)$ for all $n \in \mathbb{N}$. Applying Fact 2(ii), there exists a point $x^* \in \operatorname{Fix} T$ such that

$$\frac{1}{2}\|x_n - x^*\| \leq d_n \leq \varphi^n(d_0) \to 0,$$

where the limit tends to zero by assumption (2). This completes the proof. □

Remark 1 We discuss some important special cases of Theorem 2.

(i) (Linear regularity). The setting in which κ is linear (i.e., $\kappa(t) = Kt$ for some $K > 0$) corresponds to *bounded linear regularity* of T as discussed in [6, 20]. In this case, $\kappa^{-1}(t) = t/K$ and so

$$\varphi(t) = \sqrt{t^2 - \gamma\,\frac{t^2}{K^2}} = t\sqrt{1 - \frac{\gamma}{K^2}} \quad \Longrightarrow \quad \varphi^n(t) = t\left(\sqrt{1 - \frac{\gamma}{K^2}}\right)^n.$$

Theorem 2 implies R-linear convergence with rate no worse than $c := \sqrt{1 - \frac{\gamma}{K^2}} < 1$ which recovers the single operator specialization of [6].

(ii) (Hölder regularity). The case in which κ is a "Hölder-type function" (i.e., $\kappa(t) = Kt^\tau$ for constants $K > 0$ and $\tau \in (0, 1)$) corresponds to *bounded Hölder regularity* of T as was discussed in [8]. In this case, $\kappa^{-1}(t) = \sqrt[\tau]{t/K}$ and so

$$\varphi(t) = \sqrt{t^2 - \frac{\gamma}{K^{\frac{2}{\tau}}}\,t^{\frac{2}{\tau}}} = t\sqrt{1 - \frac{\gamma}{K^{\frac{2}{\tau}}}\,t^\alpha},$$

where $\alpha := 2/\tau - 2 = 2(1-\tau)/\tau > 0$. By [9, §4], this yields

$$\varphi^n(t) \leq \left(t^{-\alpha} + \alpha n \frac{\gamma}{K^{2/\tau}}\right)^{-\alpha} = O(n^{-1/\alpha}) = O\left(n^{-\frac{\tau}{2(1-\tau)}}\right).$$

Theorem 2 then implies convergence with order $O\left(n^{-\frac{\tau}{2(1-\tau)}}\right)$ which recovers [8, Proposition 3.1].

As the following example shows that, at least in principle, Theorem 2 opens the possibility of characterizing different convergence rates by choosing U appropriately.

Example 1 (Convergence rate by regions of a fixed point) Consider the *alternating projection operator* $T := P_A P_B$ for the two convex subsets A and B of $\mathbb{R}^2$ given by

$$A := \{(x_1, x_2) \in \mathbb{R}^2 : \ x_2 = 0\}, \quad B := \operatorname{epi}(f) \quad \text{where} \ \ f(t) = \begin{cases} t & \text{if } t \geq 0, \\ t^2 & \text{if } t < 0. \end{cases}$$

In this setting, we have $\mathsf{Fix}\, T = A \cap B = \{0\}$. The *alternating projections sequence* given by $x_{n+1} := Tx_n$ always converges to 0. However, the rate which it does so depends on the starting point $x_0 \in \mathbb{R}^2$. We consider two cases:

(i) Let $U_1 := \mathbb{R}_+ \times \mathbb{R}$. Then, the linear error bound condition is satisfied on U_1 and (x_n) converges linearly.

(ii) Let $U_2 := \mathbb{R}_- \times \mathbb{R}$. Then, there is a Hölder-type gauge function κ such that the error bound condition with gauge κ is satisfied on U_2 and (x_n) converges sublinearly.

4 Existence of Implicit Error Bounds

We now prove a kind of converse to Theorem 1. The next results show that nonexpansiveness alone is enough to guarantee the existence of an error bound. This is remarkable since, without asymptotic regularity, the fixed point iteration need not even converge. The next lemma will be referred to frequently in our development.

Lemma 1 *Let $T : \mathcal{H} \rightrightarrows \mathcal{H}$ satisfy $\mathsf{Fix}\, T \neq \emptyset$. Let $U \subset \mathcal{H}$ with $U \cap \mathsf{Fix}\, T \neq \emptyset$. Define the set-valued map $S : \mathbb{R}_+ \rightrightarrows \mathcal{H}$ by*

$$S(t) := \{y \in \mathcal{H} \ : \ \operatorname{dist}(y, Ty) \leq t\} \tag{4}$$

and define the function $\kappa : \mathbb{R}_+ \to \mathbb{R}_+ \cup \{+\infty\}$ by

$$\kappa(t) := \sup_{y \in S(t) \cap U} \{\operatorname{dist}(y, \mathsf{Fix}\, T)\}. \tag{5}$$

The following assertions hold.

(i) *The set $S(t)$ is a nonempty subset of* $\operatorname{dom} T$ *for all $t \geq 0$ and satisfies*

$$\emptyset \neq \mathsf{Fix}\, T = S(0) \subset S(s) \subset S(t) \quad \forall t \geq s \geq 0.$$

(ii) *The function κ is nonnegative, nondecreasing, $\kappa(0) = 0$ and satisfies*

$$\operatorname{dist}(x, \mathsf{Fix}\, T) \leq \kappa(\operatorname{dist}(x, Tx)) \quad \forall x \in U. \tag{6}$$

If any of the following hold, then κ is bounded:

(a) *there is a bounded set V with $S(t) \cap U \subset V$ for all t;*
(b) *the function* $\operatorname{dist}(\cdot, \mathsf{Fix}\, T)$ *is bounded on U.*

Proof (i): This is immediate from the definition of $S(t)$ in (4).
(ii): Since $U \cap \mathsf{Fix}\, T \neq \emptyset$, by (i) it follows that $S(t) \cap U \neq \emptyset$ for all $t \geq 0$ and

$$\emptyset \neq (S(0) \cap U) \subset (S(s) \cap U) \subset (S(t) \cap U) \quad \forall\, 0 \leq s \leq t.$$

Thus, κ is nondecreasing and $\kappa(0) = 0$. For any $x \in U \cap \operatorname{dom} T$, we have $x \in S(\|x - Tx\|) \cap U$, and hence, (6) holds. That κ is bounded if the function $\operatorname{dist}(\cdot, \mathsf{Fix}\, T)$ is bounded on U is clear. To show that κ is bounded if there is a bounded subset V with $S(t) \cap U \subset V$ for all $t \geq 0$, fix a point $z_0 \in \mathsf{Fix}\, T$ and let $M > 0$ be such that $\|u\| < M$ for all $u \in S(t) \cap U$ for all $t \geq 0$. Then, we have

$$\kappa(t) := \sup_{y \in S(t) \cap U} \{\operatorname{dist}(y, \mathsf{Fix}\, T)\} \leq \sup_{y \in S(t) \cap U} \{\|y - z_0\|\} < M + \|z_0\| \quad \forall t \geq 0.$$

That is, κ is bounded and the proof is complete. □

Theorem 3 (Error bounds in finite dimensions) *Let $\mathcal{H}$ be a finite dimensional Hilbert space. Suppose that $T : \mathcal{H} \to \mathcal{H}$ is nonexpansive with $\mathsf{Fix}\, T \neq \emptyset$. Then, for each bounded set U containing a fixed point of T, the nondecreasing function $\kappa : \mathbb{R}_+ \to \mathbb{R}_+$ defined by* (5) *is bounded, right-continuous at $t = 0$ with $\kappa(0) = 0$ and satisfies*

$$\operatorname{dist}(x, \mathsf{Fix}\, T) \leq \kappa(\|x - Tx\|) \quad \forall x \in U.$$

Proof The function κ defined by (5) is nonnegative, nondecreasing, bounded, $\kappa(0) = 0$ and satisfies (6) by Lemma 1(ii). It remains only to show that the function κ is right continuous at zero. The proof is by contradiction. Suppose there exists a sequence $t_n \searrow 0$ with $\lim_{n\to\infty} \kappa(t_n) = \alpha > 0$. Since κ is nondecreasing, there exists a sequence (y_n) such that

$$y_n \in S(t_n) \cap U \quad \text{and} \quad \frac{\alpha}{2} \leq \operatorname{dist}(y_n, \mathsf{Fix}\, T) \quad \forall n \in \mathbb{N}.$$

As (y_n) is contained in the bounded set U, it possesses a convergent subsequence $y_{n_k} \to y$ for some $y \in \mathcal{H}$. From the definition of $S(t_n)$, it holds that

$$\|(\mathrm{Id} - T)y_{n_k}\| = \|y_{n_k} - Ty_{n_k}\| \leq t_{n_k} \to 0.$$

As T is continuous, it follows that $y \in \mathsf{Fix}\, T$. Since $\mathsf{Fix}\, T$ is nonempty closed convex, $\mathrm{dist}(\cdot, \mathsf{Fix}\, T)$ is continuous and hence,

$$\frac{\alpha}{2} \leq \mathrm{dist}(y_{n_k}, \mathsf{Fix}\, T) \to \mathrm{dist}(y, \mathsf{Fix}\, T) = 0,$$

which contradicts the assumption on α and the proof is complete. □

Note that the proof of Theorem 3 is not valid in infinite dimensions, since in this case, the bounded sequence (y_n) need only contain a weakly convergent subsequence and the $\mathrm{dist}(\cdot, \mathsf{Fix}\, T)$ need not be weakly (sequentially) continuous.

Remark 2 (Infinite dimensional counterexamples) In general, the assumption of finite dimensionality of $\mathcal{H}$ in Theorem 3 cannot be dropped. Indeed, if $\mathcal{H}$ is infinite dimensional, then a concrete counterexample is provided by any averaged operator with a fixed point, T, for which there is a starting point, $x_0 \in \mathcal{H}$, such that the sequence $(T^n x_0)_{n=0}^{\infty}$ converges weakly but not strongly. The explicit constructions of such examples can be found, for instance, in [13, 15].

In order to shift our discussion to the infinite dimensional setting, we first make the following observation.

Lemma 2 *Let $\mathcal{H}$ be a Hilbert space, and let $T : \mathcal{H} \to \mathcal{H}$ be averaged with $\mathsf{Fix}\, T \neq \emptyset$. For each Picard iteration (x_n) generated by T from a starting point $x_0 \in \mathcal{H}$, let us define $d_0 := \mathrm{dist}(x_0, \mathsf{Fix}\, T)$ and $\underline{d} := \lim_{n\to\infty} \mathrm{dist}(x_n, \mathsf{Fix}\, T)$. Then, there exists a continuous and nondecreasing function $\mu : [\underline{d}, d_0] \to [\underline{d}, d_0]$ satisfying $\mu(t) < t$ for all $t \in (\underline{d}, d_0]$ such that*

$$\mathrm{dist}(x_{n+1}, \mathsf{Fix}\, T) = \mu(\mathrm{dist}(x_n, \mathsf{Fix}\, T)) \quad \forall n \in \mathbb{N}. \tag{7}$$

Proof Let us denote $d_n := \mathrm{dist}(x_n, \mathsf{Fix}\, T)$ for all $n \in \mathbb{N}$. We first claim that there exists a sequence $(c_n) \subset [0, 1)$, dependent on x_0, such that

$$d_{n+1} = c_n d_n \quad \forall n \in \mathbb{N}. \tag{8}$$

For any $N \in \mathbb{N}$, if $x_{N+1} \in \mathsf{Fix}\, T$, then one can take $c_n = 0$ for all $n > N$. Suppose, then, that $x_{n+1} \notin \mathsf{Fix}\, T$, hence $x_n \notin \mathsf{Fix}\, T$ and $x_n \neq x_{n+1}$. In particular, $\|x_n - x_{n+1}\| > 0$. Since T is averaged (Definition 1(iii)), there is a constant $\gamma > 0$ such that

$$d_{n+1}^2 \leq d_n^2 - \gamma\|x_n - x_{n+1}\|^2.$$

Consequently, we have $0 < d_{n+1} < d_n$ and it follows that

$$c_n := \frac{d_{n+1}}{d_n} \in (0, 1)$$

is well-defined and satisfies (8).

We next define the piecewise linear function, μ, on $[\underline{d}, d_0]$ such that

$$\mu(\underline{d}) := \underline{d}, \qquad \mu(d_n) := c_n d_n \quad \forall n \in \mathbb{N} \tag{9}$$

and, on each interval of the form $[d_{n+1}, d_n]$, the value of μ is given by a linear interpolation of its values defined by (9).

To complete the proof, we check that μ is nondecreasing on $[\underline{d}, d_0]$. By the construction of μ, the sequence $(\mu(d_n))$ is nonincreasing as $n \to \infty$. It suffices to check that μ is nondecreasing on each (nontrivial) interval $[d_{n+1}, d_n]$. Indeed, let $d_{n+1} \leq t_1 < t_2 \leq d_n$, then

$$\begin{aligned}\mu(t_1) &= \mu(d_{n+1}) + \frac{t_1 - d_{n+1}}{d_n - d_{n+1}} (\mu(d_n) - \mu(d_{n+1})) \\ &\leq \mu(d_{n+1}) + \frac{t_2 - d_{n+1}}{d_n - d_{n+1}} (\mu(d_n) - \mu(d_{n+1})) = \mu(t_2).\end{aligned}$$

□

Proposition 1 *Let $\mathcal{H}$ be a Hilbert space and consider an operator $T : \mathcal{H} \to \mathcal{H}$ with $\operatorname{Fix} T \neq \emptyset$. Let $(x_n)_{n\in\mathbb{N}}$ be a Picard sequence such that $\operatorname{dist}(x_n, \operatorname{Fix} T) \to 0$. Then, the function κ defined by (5) with $U := (x_n)_{n\in\mathbb{N}}$ is nonnegative, nondecreasing, bounded, $\kappa(0) = 0$ and satisfies*

$$\operatorname{dist}(x_n, \operatorname{Fix} T) \leq \kappa(\|x_n - Tx_n\|) \quad \forall n \in \mathbb{N}. \tag{10}$$

In addition, if T is averaged, then the sequence $(x_n)_{n\in\mathbb{N}}$ converges strongly to some point x in $\operatorname{Fix} T$ and the function κ is right continuous at 0.

Proof Note that, since $\operatorname{dist}(x_n, \operatorname{Fix} T) \to 0$, $\sup_{y\in U}\{\operatorname{dist}(y, \operatorname{Fix} T)\} < \infty$ where $U = (x_n)_{n\in\mathbb{N}}$. That κ defined by (5) with $U := (x_n)_{n\in\mathbb{N}}$ is nonnegative, nondecreasing, bounded, $\kappa(0) = 0$ and satisfies (6) then follows immediately from Lemma 1(ii). But in this case (6) is just (10).

If T is averaged, then (x_n) is Fejér monotone with respect to $\operatorname{Fix} T$ which is nonempty closed and convex. Since $\operatorname{dist}(x_n, \operatorname{Fix} T) \to 0$, Fact 2(ii) implies that (x_n) is strongly convergent to some point x in $\operatorname{Fix} T$. Continuity from the right of κ at 0 follows from a pattern similar to the proof of Theorem 3. Let $t_n \searrow 0$ as $n \to \infty$. If $\lim_{n\to\infty} \kappa(t_n) = \alpha > 0$, where $\kappa(t_n) := \sup_{S(t_n)\cap U}\{\operatorname{dist}(\cdot, \operatorname{Fix} T)\}$ for $U = (x_n)_{n\in\mathbb{N}}$, then, since $x_n \to x \in \operatorname{Fix} T$, and T is continuous, for all t_n, there must be an $n_k \geq n$ such that $\|x_{n_k} - Tx_{n_k}\| \leq t_n$ and $\operatorname{dist}(x_{n_k}, \operatorname{Fix} T) \geq \alpha/2$. But the assumption $\operatorname{dist}(x_n, \operatorname{Fix} T) \to 0$ also applies to this subsequence, which leads to a contradiction. □

It is clear from the above observation that, in order to obtain a meaningful error bound, a suitable function κ needs to be found for *all* possible starting points on a bounded set containing fixed points of T. Nevertheless, the sequence (c_n) given by Lemma 2 does characterize strong convergence of the corresponding iteration (x_n). More specifically, we have the following.

Proposition 2 (Equivalences) *Let $\mathcal{H}$ be a Hilbert space, let $T : \mathcal{H} \to \mathcal{H}$ be averaged with $\operatorname{Fix} T \neq \emptyset$ and let (x_n) be a Picard iteration generated by T with initial point $x_0 \in \mathcal{H}$. The following statements are equivalent.*

(i) *(x_n) converges strongly to a point x in $\mathcal{H}$.*
(ii) *(x_n) converges strongly to a point x in $\operatorname{Fix} T$.*
(iii) *$(\operatorname{dist}(x_n, \operatorname{Fix} T))$ converges to zero.*
(iv) *There exists a nondecreasing function $\mu : [0, d_0] \to [0, d_0]$ satisfying $\mu(t) < t$ for all $t \in [0, d_0]$ such that (7) holds and $\mu^n(\operatorname{dist}(x_0, \operatorname{Fix} T)) \to 0$ as $n \to \infty$.*

Proof (i) $\iff$ (ii): Suppose that $x_n \to x$. Then, since T is continuous, applying Fact 1(i) yields

$$\|x - Tx\| = \lim_{n\to\infty} \|x_n - Tx_n\| = 0.$$

The reverse implication is trivial.

(ii) $\iff$ (iii): Since T is averaged, in particular, it holds, for all $z \in \operatorname{Fix} T$, that

$$\|x_{n+1} - z\| \le \|x_n - z\| \quad \forall n \in \mathbb{N}.$$

That is, (x_n) is Fejér monotone with respect to $\operatorname{Fix} T$. The equivalence follows from Fact 2(ii).

(iii) $\iff$ (iv): Note that

$$\operatorname{dist}(x_n, \operatorname{Fix} T) = \mu^n(\operatorname{dist}(x_0, \operatorname{Fix} T)) \quad \forall n \in \mathbb{N}.$$

Lemma 2 yields the existence of a nondecreasing $\mu : [\underline{d}, d_0] \to [\underline{d}, d_0]$ satisfying $\mu(t) < t$ for all $t \in [0, d_0]$ such that (7) holds where $\underline{d} := \lim_{n\to\infty} \operatorname{dist}(x_n, \operatorname{Fix} T)$. So if $\mu^n(\operatorname{dist}(x_0, \operatorname{Fix} T)) \to 0$ as $n \to \infty$, then $\underline{d} = 0$ and (iii) holds. On the other hand, if (iii) holds, then $\underline{d} = 0$ and $\mu^n(\operatorname{dist}(x_0, \operatorname{Fix} T)) \to 0$ as $n \to \infty$. □

Remark 3 The function μ in Proposition 2(iv) characterizes the convergence rate of (x_n).

(i) When μ is majorized by a linear function with slope $c \in [0, 1)$ on some interval $[0, \tau)$ where $\tau > 0$, that is,

$$\mu(\operatorname{dist}(x_n, \operatorname{Fix} T)) \le c\operatorname{dist}(x_n, \operatorname{Fix} T) \quad \forall n \text{ sufficiently large}$$

– equivalently, the sequence (c_n) defined in (8) satisfies $c := \sup_{n\in\mathbb{N}} c_n < 1$ – then we have a *linearly monotone* sequence as defined in [23] and *R-linear* convergence as detailed in [3, Theorem 5.12].

(ii) When $\mu^n(\operatorname{dist}(x_0, \operatorname{Fix} T))$ tends to zero slower or faster than a linear rate, the sequence (x_n) is said to converge sublinearly or superlinearly, respectively. An example of sublinear convergence corresponding to $\mu(t) = \frac{t}{\sqrt{t^2+1}}$ for all $t \in [0, \operatorname{dist}(x_0, \operatorname{Fix} T)]$ is detailed in Example 3 below.

In order to deduce a uniform version of the previous results, a property which holds uniformly on U is needed.

Theorem 4 (Sufficient condition for an error bound) *Let $\mathcal{H}$ be a Hilbert space, let $T : \mathcal{H} \to \mathcal{H}$ with $\operatorname{Fix} T \neq \emptyset$, let U be a bounded subset of $\mathcal{H}$ containing a fixed point of T. Suppose that there exists a function $c : [0, \infty) \to [0, 1]$ which is upper semi-continuous on $(0, \operatorname{exc}(U, \operatorname{Fix} T)]$ and satisfies $c(t) < 1$ for all t in this interval such that*

$$\operatorname{dist}(Tx, \operatorname{Fix} T) \le c\,(\operatorname{dist}(x, \operatorname{Fix} T)) \operatorname{dist}(x, \operatorname{Fix} T) \quad \forall x \in U. \tag{11}$$

Then, the nonnegative, nondecreasing function $\kappa : \mathbb{R}_+ \to \mathbb{R}_+$ defined by (5) is bounded, right continuous at $t = 0$ and satisfies

$$\operatorname{dist}(x, \operatorname{Fix} T) \le \kappa(\|x - Tx\|) \quad \forall x \in U.$$

Proof The function κ defined by (5) is nonnegative, nondecreasing, bounded (because U is bounded), $\kappa(0) = 0$ and satisfies (6) by Lemma 1(ii). It remains only to show that the function κ is right continuous at zero. By way of contradiction, suppose there exists a sequence

$t_n \searrow 0$ with $\lim_{n\to\infty} \kappa(t_n) \geq 2\alpha > 0$. Since the function c is upper semi-continuous (by assumption) and takes values strictly less than 1 on the compact set $[\alpha, \mathrm{exc}(U, \mathsf{Fix}\, T)]$, there exists a maximizer t_α of c on this interval. Denote the corresponding maximal value by c_α, then it holds that

$$c_\alpha = \max\{c(t) : \alpha \leq t \leq \mathrm{exc}(U, \mathsf{Fix}\, T)\} = c(t_\alpha) < 1. \tag{12}$$

In particular, $(1 - c_\alpha)\alpha > 0$. We choose a number ε satisfying

$$0 < \varepsilon < (1 - c_\alpha)\alpha. \tag{13}$$

Since $t_n \searrow 0$, there is an index $N \in \mathbb{N}$ such that $t_N < \varepsilon$. Since $\kappa(t_n)$ is nonincreasing as n increases and $\lim_{n\to\infty} \kappa(t_n) \geq 2\alpha > 0$, it holds that $\kappa(t_N) \geq 2\alpha$. Recall that $\kappa(t_N)$ is the supremum of the function $\mathrm{dist}(\cdot, \mathsf{Fix}\, T)$ over the set $U \cap S(t_N)$ with $S(t_N)$ defined by (4). Then, there exists a point $x \in U \cap S(t_N)$ such that $\mathrm{dist}(x, \mathsf{Fix}\, T) + \alpha \geq \kappa(t_N)$. Since $\kappa(t_N) \geq 2\alpha$, this implies

$$\mathrm{dist}(x, \mathsf{Fix}\, T) \geq \alpha. \tag{14}$$

Also, since $x \in U \cap S(t_N)$ and $t_N < \varepsilon$, it holds that $x \in U \cap S(\varepsilon)$ by Lemma 1(i). In particular,

$$\|x - Tx\| \leq \varepsilon. \tag{15}$$

Now by (15), the triangle inequality, (11), (12), and (14), successively, we have

$$\begin{aligned}
\varepsilon &\geq \|x - Tx\| \\
&\geq \mathrm{dist}(x, \mathsf{Fix}\, T) - \mathrm{dist}(Tx, \mathsf{Fix}\, T) \\
&\geq \mathrm{dist}(x, \mathsf{Fix}\, T) - c(\mathrm{dist}(x, \mathsf{Fix}\, T))\mathrm{dist}(x, \mathsf{Fix}\, T) \\
&= (1 - c(\mathrm{dist}(x, \mathsf{Fix}\, T)))\, \mathrm{dist}(x, \mathsf{Fix}\, T) \\
&\geq (1 - c_\alpha)\mathrm{dist}(x, \mathsf{Fix}\, T) \\
&\geq (1 - c_\alpha)\alpha.
\end{aligned}$$

This contradicts (13) and hence, it must hold that $\alpha = 0$ and the proof is complete. □

Example 2 (Arbitrarily slow convergence) There are two things to note about the theorem above, both hinging on the choice of the subset U. The first point is that it is possible to choose U such that no c satisfying the requirements of the theorem exists. We demonstrate this when U is simply a ball. Such a phenomenon shows that *uniform* linear error bounds are not always possible. The second point, however, is that when an iteration converges it is always possible to choose a set U such that a function c satisfying the requirements of Theorem 4 exists, but the resulting error bound may not always be informative. We also show an example of this below.

To put the above results in context, consider the method of alternating projections for finding the intersection of two closed subspaces of a Hilbert space, call them A and B. The alternating projections fixed point mapping is $T := P_A P_B$ with $\mathsf{Fix}\, T = A \cap B$. Von Neumann showed that the iterates of the method of alternating projections converges strongly to the projection of the starting point onto the intersection [26]. In the mid 1950's a rate was established in terms of what is known as the *Friedrich's angle*[12][1] between the sets defined as the number in $[0, \frac{\pi}{2}]$ whose cosine is given by

$$c(A, B) := \sup \left\{ |\langle a, b\rangle| \;\middle|\; \begin{array}{ll} a \in A \cap (A \cap B)^\perp, & \|a\| \leq 1, \\ b \in B \cap (A \cap B)^\perp, & \|b\| \leq 1. \end{array} \right\}$$

[1]Though it is called the Friedrichs angle, the notion goes back at least to Jordan [18, Eq. 60, pp. 122–130].

It is straightforward to see that $c(A, B) \leq 1$. Moreover, $c(A, B) < 1$ if and only if $A + B$ is closed [2, Lemma 4.10]. In this case, a bound on the rate of convergence in terms of the Friedrichs angle follows from the fact that [19]

$$\|T^n - P_{A\cap B}\| = c(A, B)^{2n-1} \quad \forall n \in \mathbb{N}. \tag{16}$$

In the context of Theorem 4, if $A + B$ is closed, then the function $c : [0, \infty) \to [0, 1]$ can be simply chosen to be the cosine of the Friedrichs angle [5, Theorem 3.16].

If $A + B$ is not closed, then it was shown in [4] (i.e., $c(A, B) = 1$) that convergence can be arbitrarily slow in the sense that for any non-increasing sequence $\lambda_n \to 0$ with $\lambda_0 < 1$, there is a starting point x_λ such that

$$\|T^n x_\lambda - P_{A\cap B} x_\lambda\| \geq \lambda_n \quad \forall n \in \mathbb{N}.$$

In the context of Theorem 4, if $A + B$ is not closed, then no function $c : [0, \infty) \to [0, 1]$ satisfying Theorem 4 exists as soon as the bounded set U contains dilate of the sphere $S := \{x \in \mathcal{H} : \|x\| = 1\}$. To see this, suppose on the contrary, that there exists a function c satisfying Theorem 4. In particular, we have that $c(t) < 1$ for all t in the interval $(0, \operatorname{exc}(U, \operatorname{\mathsf{Fix}} T))$. Then, for any $x \in S \subseteq U$, we have

$$\begin{aligned}\|Tx - P_{A\cap B}x\| = \operatorname{dist}(Tx, \operatorname{\mathsf{Fix}} T) &\leq c(\operatorname{dist}(x, \operatorname{\mathsf{Fix}} T))\operatorname{dist}(x, \operatorname{\mathsf{Fix}} T)\\ &= c(\operatorname{dist}(x, \operatorname{\mathsf{Fix}} T))\|x - P_{A\cap B}x\|\\ &\leq c(\operatorname{dist}(x, \operatorname{\mathsf{Fix}} T))\|x\|.\end{aligned}$$

Dividing both sides of the inequality by $\|x\|$, taking the supremum over S, and substituting (16) gives

$$1 \leq \sup_{x\in S} c(\operatorname{dist}(x, \operatorname{\mathsf{Fix}} T)),$$

which contradicts the assumption that $c(t) < 1$ (as c satisfies Theorem 4). The choice of U to be a scaled ball is the natural choice when one is interested in uniform error bounds. This example shows that even for the simple alternating projections algorithm, such bounds are not always possible.

To the second point, if for the above example, instead of choosing U to be a ball, we restrict U to be the *iterates* (x_n) of the alternating projections sequence together with their limit x_∞ for a fixed x_0, then we can construct a function c satisfying the assumptions of Theorem 4. Indeed, choose $c(t)$ to be a linear interpolation of the points

$$c(t_n) := \frac{\|Tx_n - x_\infty\|}{\|x_n - x_\infty\|} \quad \text{for } \; t_n = \|x_n - x_\infty\| \text{ whenever } \|x_n - x_\infty\| > 0.$$

Such a function satisfies the requirements of Theorem 4 and hence guarantees the existence of an error bound. But this is not informative, because the error bound depends on the iteration itself, and hence the initial guess x_0. Returning to the fact that if $A + B$ is not closed the alternating projections algorithm exhibits arbitrarily slow convergence, then even though we have an error bound for a particular instance, we cannot say anything about uniform rates of convergence.

The following example illustrates the role of the function c satisfying condition (11) as in Theorem 4.

Example 3 Consider the *alternating projections operator* $T := P_A P_B$ for the two convex subsets A and B of $\mathbb{R}^2$ given by

$$A := \{(x_1, x_2) \in \mathbb{R}^2 : x_2 = 0\}, \quad B := \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + (x_2 - 1)^2 \leq 1\}.$$

Then, we have $\operatorname{Fix} T = A \cap B = \{0\}$ and the only set U of interest is $U = A$. For each $x \in U$, say $x = (t, 0)$, it holds $Tx = \left(\frac{t}{\sqrt{t^2+1}}, 0\right)$ and consequently

$$\operatorname{dist}(x, \operatorname{Fix} T) = |t|, \quad \operatorname{dist}(Tx, \operatorname{Fix} T) = \frac{|t|}{\sqrt{t^2+1}},$$

and

$$\|x - Tx\| = |t|\left(1 - \frac{1}{\sqrt{t^2+1}}\right).$$

In this setting, we now can directly check the following statements.

(i) The function c defined by

$$c(t) := \frac{1}{\sqrt{t^2+1}} \quad \forall t \in \mathbb{R}_+$$

satisfies all the assumptions of Theorem 4. It is worth emphasizing that for each $\alpha > 0$,

$$c_\alpha := \sup\{c(t) : t \geq \alpha\} = \frac{1}{\sqrt{\alpha^2+1}} < 1 \quad \text{while} \quad \sup\{c(t) : t \geq 0\} = 1.$$

(ii) The function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ defined by

$$\varphi(t) := t\left(1 - \frac{1}{\sqrt{t^2+1}}\right) \quad \forall t \in \mathbb{R}_+$$

is a gauge function (see Definition 3) and the desired function, κ, defined by (5) is the inverse function φ^{-1} which is also a gauge function.

(iii) This development is an extension of *μ-monotonicity* introduced in [23]. A sequence (x_k) on $\mathcal{H}$ is said to be μ-monotone with respect to Ω ($\emptyset \neq \Omega \subset \mathcal{H}$) if there exists a nonnegative function $\mu : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying $\mu(0) = 0$ and $\mu(t_1) < \mu(t_2) \leq t_2$ whenever $0 \leq t_1 < t_2$ such that

$$(\forall k \in \mathbb{N}) \qquad \operatorname{dist}(x_{k+1}, \Omega) \leq \mu(\operatorname{dist}(x_k, \Omega)). \tag{17}$$

In the present example, the sequence (x_n) generated by T is μ-monotone with respect to $\operatorname{Fix} T$, where $\mu : \mathbb{R}_+ \to \mathbb{R}_+$ is given by

$$\mu(t) := \frac{t}{\sqrt{t^2+1}} \quad \forall t \in \mathbb{R}_+.$$

The sequence is said to be *linearly monotone with respect to* Ω if (17) is satisfied for $\mu(t) = ct$ for all $t \in \mathbb{R}_+$ with constant $c \in [0, 1]$.

Remark 4 Condition (11) can be viewed as the functional extension of the linear result in [23, Theorem 3.1] where linear monotonicity (part (ii) of Example 3) was shown to be sufficient for the existence of linear error bounds. Indeed, (11) is a realization of the notion of *μ-monotonicity* in which the function μ has the form $\mu(t) := c(t) \cdot t$ for all $t \geq 0$. In particular, if $c(t) := c_0$ for some constant $c_0 \in [0, 1)$, Theorem 4 recovers [23, Theorem 3.1].

Note that in Theorem 4, condition (11) is the only assumption required to obtain the error bound. An implicit consequence of the condition is that the distance of Picard iterates to $\operatorname{Fix} T$ converges to zero as soon as T has a fixed point and that the initial point of the iteration is in a set U which satisfies $T(U) \subset U$.

Proposition 3 (Convergence to zero of the distance to fixed points) *Let $\mathcal{H}$ be a Hilbert space, let $T : \mathcal{H} \to \mathcal{H}$ with $\mathsf{Fix}\, T \neq \emptyset$, and let U be a bounded subset containing a fixed point of T and $T(U) \subset U$. Suppose that there exists a function $c : [0, \infty) \to [0, 1]$ being upper semi-continuous on $(0, \mathrm{exc}(U, \mathsf{Fix}\, T)]$ and satisfying $c(t) < 1$ for all t in this interval such that condition (11) is satisfied. Then, every Picard iteration (x_n) with $x_0 \in U$ generated by T satisfies $\mathrm{dist}(x_n, \mathsf{Fix}\, T) \to 0$ as $n \to \infty$.*

Proof Let us consider an arbitrary iteration $(x_n) \subset U$ generated by T. Condition (11) implies that $\mathrm{dist}(x_n, \mathsf{Fix}\, T)$ is non-increasing and bounded from below by 0. Hence, $\mathrm{dist}(x_n, \mathsf{Fix}\, T)$ converges to some $\alpha \geq 0$. We complete the proof by showing that $\alpha = 0$. Again, by way of contradiction, assume that $\alpha > 0$. Then, by the same argument as the one for obtaining (12), we get

$$c_\alpha := \max\{c(t) : \alpha \leq t \leq \mathrm{exc}(U, \mathsf{Fix}\, T)\} < 1.$$

Since $\alpha \leq \mathrm{dist}(x_n, \mathsf{Fix}\, T) \leq \mathrm{exc}(U, \mathsf{Fix}\, T)$ for all $n \in \mathbb{N}$, it holds that

$$c(\mathrm{dist}(x_n, \mathsf{Fix}\, T)) \leq c_\alpha \quad \forall n \in \mathbb{N}. \tag{18}$$

Combining (18) and (11) then yields, for all $n \in \mathbb{N}$,

$$\mathrm{dist}(x_{n+1}, \mathsf{Fix}\, T) \leq c(\mathrm{dist}(x_n, \mathsf{Fix}\, T))\mathrm{dist}(x_n, \mathsf{Fix}\, T) \leq c_\alpha \mathrm{dist}(x_n, \mathsf{Fix}\, T).$$

Inductively,

$$\mathrm{dist}(x_n, \mathsf{Fix}\, T) \leq c_\alpha^n \mathrm{dist}(x_0, \mathsf{Fix}\, T) \quad \forall n \in \mathbb{N}. \tag{19}$$

Letting $n \to \infty$ in (19) with noting that $c_\alpha < 1$ yields

$$0 < \alpha \leq \mathrm{dist}(x_n, \mathsf{Fix}\, T) \leq c_\alpha^n \mathrm{dist}(x_0, \mathsf{Fix}\, T) \to 0.$$

This is a contradiction and the proof is complete. □

In the light of Proposition 3, Theorem 4 can be viewed as a uniform version of Proposition 1.

To conclude, we finally discuss some insights of condition (11) in the averaged operator setting.

Remark 5 Let $T : \mathcal{H} \to \mathcal{H}$ be averaged with $\mathsf{Fix}\, T \neq \emptyset$.

(i) Lemma 2 implies that, for each $x \in \mathcal{H}$, there exists a number $c_x < 1$ such that

$$\mathrm{dist}(Tx, \mathsf{Fix}\, T) \leq c_x \mathrm{dist}(x, \mathsf{Fix}\, T).$$

Note that, the existence of a function c satisfying condition (11) would require that the supremum of all such numbers c_x taken over each level set $\mathcal{L}_t := \{x : \mathrm{dist}(x, \mathsf{Fix}\, T) = t\}$ exists and is less than 1. In this case, c can be any function which is upper semi-continuous on $(0, \mathrm{exc}(U, \mathsf{Fix}\, T)]$ and satisfies

$$\sup\{c_x : x \in \mathcal{L}_t\} \leq c(t) < 1 \quad \forall t > 0.$$

Note that the function $f : \mathcal{H} \to \mathbb{R}_+$ given by

$$f(x) := \begin{cases} \frac{\mathrm{dist}(Tx, \mathsf{Fix}\, T)}{\mathrm{dist}(x, \mathsf{Fix}\, T)} & \text{if } x \notin \mathsf{Fix}\, T, \\ 0 & \text{if otherwise} \end{cases}$$

is continuous at all points $x \notin \mathsf{Fix}\, T$ as a quotient of two continuous functions $\mathrm{dist}(\cdot, \mathsf{Fix}\, T)$ and $\mathrm{dist}(T(\cdot), \mathsf{Fix}\, T)$ (because T is averaged). Thus, in particular, if $\mathcal{H}$ is finite dimensional and $\mathsf{Fix}\, T$ is bounded, then $\mathcal{L}_t$ is compact and hence, for all $t > 0$,

$\sup\{c_x : x \in \mathcal{L}_t\}$ is trivially less than one. In other words, for an averaged operator in a finite dimensional space, condition (11) in Theorem 4 is superfluous and only upper semi-continuity of c need be assumed.

(ii) Condition (11) quantifies the rate of decrease of $\mathrm{dist}(\cdot, \mathsf{Fix}\, T)$ on each level set $\mathcal{L}_t$. More precisely, if $x_n \in \mathcal{L}_t$, then the distance to $\mathsf{Fix}\, T$ will decrease by a factor of at least $c(t)$ in the next iterate x_{n+1}. Furthermore, a closer look at the proof of Proposition 3 shows that condition (11) can actually provide an estimate of the rate at which $\mathrm{dist}(T^n x, \mathsf{Fix}\, T) \to 0$ even in the infinite dimensional setting.

(iii) On one hand, Theorem 4 can be viewed as an attempt to extend Theorem 3 to the infinite dimensional settings. On the other hand, it shows that an error bound in the form of (6) is a necessary condition for a certain type of μ-monotonicity (see Example 3 and Remark 4). More precisely, μ-monotonicity with μ of the form $\mu(t) = c(t) \cdot t$ for all $t \geq 0$ where c denotes the function in (11).

Funding Information D. Russell Luke was supported in part by German Israeli Foundation Grant G-1253-304.6 and Deutsche Forschungsgemeinschaft Research Training Grant RTG2088.

Nguyen H. Thao was supported by German Israeli Foundation Grant G-1253-304.6 and the European Research Council grant agreement no. 339681.

Matthew K. Tam was supported by Deutsche Forschungsgemeinschaft Research Training Grant 2088 and a Postdoctoral Fellowship from the Alexander von Humboldt Foundation.

References

1. Baillon, J.B., Bruck, R.E., Reich, S.: On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces. Houst. J. Math. **4**, 1–9 (1978)
2. Bauschke, H.H., Borwein, J.M.: On the convergence of von Neumann's alternating projection algorithm for two sets. Set-Valued Anal. **1**, 185–212 (1993)
3. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces. CMS Books Math./Ouvrages Math. SMC. Springer, New York (2011)
4. Bauschke, H.H., Deutsch, F., Hundal, H.: Characterizing arbitrarily slow convergence in the method of alternating projections. Int. Trans. Oper. Res. **16**, 413–425 (2009)
5. Bauschke, H.H., Deutsch, F., Hundal, H., Park, S.-H.: Accelerating the convergence of the method of alternating projections. Trans. Am. Math. Soc. **355**, 3433–3461 (2003)
6. Bauschke, H.H., Noll, D., Phan, H.M.: Linear and strong convergence of algorithms involving averaged nonexpansive operators. J. Math. Anal. Appl. **421**, 1–20 (2015)
7. Bolte, J., Nguyen, T.P., Peypouquet, J., Suter, B.W.: From error bounds to the complexity of first-order descent methods for convex functions. Math. Program. **165**, 471–507 (2017)
8. Borwein, J.M., Li, G., Tam, M.K.: Convergence rate analysis for averaged fixed point iterations in common fixed point problems. SIAM J. Optim. **27**, 1–33 (2017)
9. Borwein, J.M., Li, G., Yao, L.: Analysis of the convergence rate for the cyclic projection algorithm applied to basic semialgebraic convex sets. SIAM J. Optim. **24**, 498–527 (2014)
10. Boţ, R.I., Csetnek, E.R.: A dynamical system associated with the fixed points set of a nonexpansive operator. J. Dyn. Differ. Equ. **29**, 155–168 (2017)
11. Drusvyatskiy, D., Ioffe, A.D., Lewis, A.: Transversality and alternating projections for nonconvex sets. Found. Comput. Math. **15**, 1637–1651 (2015)
12. Friedrichs, K.: On certain inequalities and characteristic value problems for analytic functions and for functions of two variables. Trans. Am. Math. Soc. **41**, 321–364 (1937)
13. Genel, A., Lindenstrauss, J.: An example concerning fixed points. Isr. J. Math. **22**, 81–86 (1975)
14. Güler, O.: On the convergence of the proximal point algorithm for convex minimization. SIAM J. Control Optim. **29**, 403–419 (1991)
15. Hundal, H.: An alternating projection that does not converge in norm. Nonlinear Anal. **57**, 35–61 (2004)
16. Ioffe, A.D.: Nonlinear regularity models. Math. Program. Ser. B **139**, 223–242 (2013)
17. Ishikawa, S.: Fixed points and iteration of a nonexpansive mapping in a Banach space. Proc. Am. Math. Soc. **59**, 65–71 (1976)

18. Jordan, C.: Essai sur la géométrie á n dimensions. Bull. Soc. Math. Fr. **3**, 103–174 (1875)
19. Kayalar, S., Weinert, H.: Error bounds for the method of alternating projections. Math. Control Signals Syst. **1**, 43–59 (1988)
20. Leventhal, D.: Metric subregularity and the proximal point method. J. Math. Anal. Appl. **360**, 681–688 (2009)
21. Li, G., Nghia, T.T.A., Mordukhovich, B.S., Phạm, T.S.: Error bounds for parametric polynomial systems with applications to higher-order stability analysis and convergence rates. Math Program. https://doi.org/10.1007/s10107-016-1014-6 (2016)
22. Luke, D.R., Thao, N.H., Tam, M.K.: Quantitative convergence analysis of iterated expansive, set-valued mappings. Math. Oper. Res. (to appear)
23. Luke, D.R., Thao, N.H., Teboulle, M.: Necessary conditions for linear convergence of Picard iterations and application to alternating projections. arXiv:1704.08926 (2017)
24. Noll, D., Rondepierre, A.: On local convergence of the method of alternating projections. Found. Comput. Math. **16**, 425–455 (2016)
25. Opial, Z.: Weak convergence of the sequence of successive approximations for nonexpansive mappings. Bull. Am. Math. Soc. **73**, 591–597 (1967)
26. von Neumann, J.: Functional Operators, Vol. II. The Geometry of Orthogonal Spaces Annals of Mathematical Studies, vol. 22. Princeton University Press, Princeton (1950)