

# Strong convergence for the alternating Halpern-Mann iteration in CAT(0) spaces \*

Bruno Dinis<sup>a</sup> and Pedro Pinto<sup>b</sup>

<sup>a</sup> Escola de Ciências e Tecnologia da Universidade de Évora,  
Rua Romão Ramalho, 59, 7000-671 Évora, Portugal  
E-mail: [bruno.dinis@uevora.pt](mailto:bruno.dinis@uevora.pt)

<sup>b</sup> Department of Mathematics, Technische Universität Darmstadt,  
Schlossgartenstrasse 7, 64289 Darmstadt, Germany  
E-mail: [pinto@mathematik.tu-darmstadt.de](mailto:pinto@mathematik.tu-darmstadt.de)

March 18, 2023

## Abstract

In this paper we consider, in the general context of CAT(0) spaces, an iterative schema which alternates between Halpern and Krasnoselskii-Mann style iterations. We prove, under suitable conditions, the strong convergence of this algorithm, benefiting from ideas from the proof mining program. We give quantitative information in the form of effective rates of asymptotic regularity and of metastability (in the sense of Tao). Motivated by these results we are also able to obtain strongly convergent versions of the forward-backward and the Douglas-Rachford algorithms. Our results generalize recent work by Boţ, Csetnek and Meier, and Cheval and Leuştean.

## 1 Introduction

In the context of CAT(0) spaces, we show the strong convergence of an iterative schema which alternates between Halpern and Krasnoselskii-Mann style iterations, while also obtaining quantitative information in the form of rates of asymptotic regularity and rates of metastability.

Let  $H$  be a Hilbert space and  $C$  a nonempty closed convex subset. A mapping  $U : C \rightarrow C$  is said to be *nonexpansive* on  $C$  if for all  $x, y \in C$  one has  $\|U(x) - U(y)\| \leq \|x - y\|$ . A well-known method for finding fixed points of a nonexpansive mapping  $U : C \rightarrow C$  is

$$x_{n+1} = (1 - \beta_n)U(x_n) + \beta_n x_n, \quad (\text{KM})$$

with  $x_0 \in C$  a starting point and  $(\beta_n) \subset [0, 1]$ . This iteration was introduced by Mann [46] and Krasnoselskii [41]. Even though this iteration has the propitious property of being Fejér monotone with respect to the fixed point set, in general it only converges weakly to a fixed point (see e.g. [51, Theorem 2] and [5, Corollary 5.2 and Remark 5.3]). This motivated several modifications of (KM) in order to ensure strong convergence. One such version was introduced by Yao, Zho and Liou [63], and rediscovered recently by Boţ, Csetnek, and Meier [7], where strong convergence is guaranteed by introducing Tikhonov regularization terms  $(\gamma_n) \subset [0, 1]$ ,

$$x_{n+1} = (1 - \beta_n)U(\gamma_n x_n) + \beta_n(\gamma_n x_n). \quad (\text{T-KM})$$

In [7, Theorem 3] Boţ *et al.* showed that  $(x_n)$  generated by (T-KM) converges strongly to the fixed point of  $U$  with minimum norm  $P_{\text{Fix}(U)}(0)$ , provided that the following conditions hold

- (i)  $(\gamma_n) \subset (0, 1]$ ,  $\lim \gamma_n = 1$ ,  $\sum_{n \geq 0} (1 - \gamma_n) = \infty$ ,  $\sum_{n \geq 0} |\gamma_{n+1} - \gamma_n| < \infty$ ,
- (ii)  $(\beta_n) \subset [0, 1)$ ,  $\sum_{n \geq 0} |\beta_{n+1} - \beta_n| < \infty$ , and  $\limsup \beta_n < 1$ .

---

\*2010 Mathematics Subject Classification: 47J25, 47H09, 47H10, 03F10, 47H05. Keywords: Strong convergence, CAT(0) spaces, metastability, asymptotic regularity.

From this result, Bot *et al.* introduced strongly convergent versions of the forward-backward and the Douglas-Rachford algorithms [7, Theorems 7 and 10], which are well-known splitting methods that weakly approximate zeros of a sum of monotone operators.

A different iteration was introduced by Halpern in [27]

$$x_{n+1} = (1 - \alpha_n)T(x_n) + \alpha_n u, \quad (\text{H})$$

where  $T : C \rightarrow C$  is a nonexpansive map,  $x_0, u \in C$  and  $(\alpha_n) \subset [0, 1]$ . Halpern showed the strong convergence of (H) to the metric projection onto the set  $\text{Fix}(T)$  of the anchor point  $u$ , in Hilbert spaces. The conditions considered by Halpern prevented the natural choice  $\alpha_n = \frac{1}{n+1}$ , which was later overcome by Wittmann [60]. The strong convergence of (H) was first extended to the setting of uniformly smooth Banach spaces by Reich in [52]. In [61, 62], Xu established strong convergence, also in uniformly smooth Banach spaces, under conditions which are incomparable to those considered by Wittmann but which still allowed  $\alpha_n = \frac{1}{n+1}$ .

As observed by Cheval and Leuştean in [14], the iteration (T-KM) can be seen as alternating between a Mann style construction and a special case of the Halpern schema. Indeed, the iteration considered by Cheval and Leuştean can be written in the following way

$$\begin{cases} x_{2n+1} &= (1 - \alpha_n)x_{2n} + \alpha_n u \\ x_{2n+2} &= (1 - \beta_n)U(x_{2n+1}) + \beta_n x_{2n+1}, \end{cases} \quad (\text{CL})$$

which corresponds to (T-KM) when  $u = 0$ . They considered this iteration in the general setting of hyperbolic spaces and obtained rates of asymptotic regularity. Yet no strong convergence results were established. Moreover, it was shown recently in [13] that this iteration is essentially a modified Halpern iteration, as introduced in [28], and studied in the setting of CAT(0) spaces in [15]. Rates of asymptotic regularity and of metastability for the modified Halpern iteration were obtained in [55], and shown to be transferable to the iteration (CL) (and *vice-versa*) in [13].

Let  $X$  be a CAT(0) space and  $C$  a nonempty convex closed subset. In this paper we will consider an iteration  $(x_n)$  generated by the recursive schema

$$\begin{cases} x_{2n+1} &= (1 - \alpha_n)T(x_{2n}) \oplus \alpha_n u \\ x_{2n+2} &= (1 - \beta_n)U(x_{2n+1}) \oplus \beta_n x_{2n+1}, \end{cases} \quad (\text{HM})$$

where  $x_0, u \in C$ ,  $(\alpha_n), (\beta_n) \subset [0, 1]$  are sequences of real numbers and  $T, U : C \rightarrow C$  are nonexpansive mappings. The symbol “ $\oplus$ ” denotes the application of the convex operator  $W$  (see Section 2 for details), which is the usual linear convex combination in the case of normed spaces. We show the following result:

**Theorem 1.1.** *Let  $X$  be a complete CAT(0) space and  $C \subseteq X$  a nonempty convex closed subset. Consider nonexpansive mappings  $T, U : C \rightarrow C$  such that  $F := \text{Fix}(T) \cap \text{Fix}(U) \neq \emptyset$  and  $u, x_0 \in C$ . Assume that  $(\alpha_n) \subset [0, 1]$ ,  $(\beta_n) \subset (0, 1)$  are sequences of real numbers satisfying*

$$\begin{aligned} (i) \quad & \lim \alpha_n = 0; & (ii) \quad & \sum_{n \geq 0} \alpha_n = \infty; & (iii) \quad & \sum_{n \geq 0} |\alpha_{n+1} - \alpha_n| < \infty; \\ (iv) \quad & \sum_{n \geq 0} |\beta_{n+1} - \beta_n| < \infty; & (v) \quad & 0 < \liminf \beta_n \leq \limsup \beta_n < 1. \end{aligned}$$

*Then  $(x_n)$  generated by (HM) converges strongly to  $P_F(u)$ .*

Theorem 1.1 is a twofold generalization of [7, Theorem 3]. On the one hand, our iteration has (T-KM) as a particular case. On the other hand, our strong convergence result is established in CAT(0) spaces – frequently considered the non-linear generalization of Hilbert spaces. The (HM) iteration also extends the one in [14] – which is the particular case when  $T = \text{Id}_C$ . Even though the context of [14] is more general than that of CAT(0) spaces, Theorem 1.1 proves strong convergence while [14, Theorems 4.1 and 4.2] only establish asymptotic regularity for the iteration. This greater generality is achieved at the cost of additionally assuming the mild condition  $0 < \liminf \beta_n$ . Similarly to the strongly convergent versions of the forward-backward and the Douglas-Rachford algorithms introduced in [7], our generalized iteration is also used to define extended versions of these algorithms.

As an initial step into proving our main theorem we first establish quantitative results in the spirit of the proof mining program [31, 34]. In this framework one looks to obtain quantitative information, such as rates of convergence and rates of metastability, guided by proof-theoretical techniques. For example, while in general it is not possible to obtain computable information on the value of  $n$  in the Cauchy property (see e.g. [47]),

$$\forall \varepsilon > 0 \exists n \in \mathbb{N} \forall i, j \geq n (d(x_i, x_j) \leq \varepsilon),$$

in many cases, logical results guarantee quantitative information for the equivalent finitary version

$$\forall \varepsilon > 0 \forall f : \mathbb{N} \rightarrow \mathbb{N} \exists n \forall i, j \in [n, f(n)] (d(x_i, x_j) \leq \varepsilon), \quad (\dagger)$$

where  $[n, f(n)]$  denotes the set  $\{n, n+1, \dots, f(n)\}$ . The property  $(\dagger)$  was popularized by Tao [58, 59] under the name *metastability*. Quantitative information for  $(\dagger)$  takes the form of a *rate of metastability*, i.e. a computable functional  $\mu : (0, +\infty) \times \mathbb{N}^{\mathbb{N}} \rightarrow \mathbb{N}$  satisfying

$$\forall \varepsilon > 0 \forall f : \mathbb{N} \rightarrow \mathbb{N} \exists n \leq \mu(\varepsilon, f) \forall i, j \in [n, f(n)] (d(x_i, x_j) \leq \varepsilon).$$

Note that a rate of metastability for  $(\dagger)$  does not entail computable information for the equivalent Cauchy property. Even though these two notions are equivalent, this equivalence is non-effective as the proof is by contradiction. Still, metastability results have a far reaching scope as can be seen in several recent results, see for example [29, 2, 25, 36, 58]. Another striking feature of proof mining is the possibility to identify precisely the required conditions needed to prove a result. This frequently allows to single out unused hypotheses and thus to obtain generalized results. In this sense, Theorem 1.1 follows from a generalization to a non-linear setting of a corresponding quantitative result in Hilbert spaces. In Hilbert spaces the result can be established via a sequential weak compactness argument (see Section 8). It is not clear how to carry out such an argument in CAT(0) spaces. Nevertheless, the technique developed in [22] shows that, through a quantitative treatment, it is possible to bypass this argument in Hilbert spaces (as was done e.g. in [19, 20]). Extending such quantitative results to the setting of CAT(0) spaces allows to conclude the main theorem. These quantitative results may also be seen as a natural continuation of previous quantitative analyses (see e.g. [19, 37, 18]). Even though our results and proofs are inspired by proof theoretical techniques, these are only used as an intermediate step and are not visible in the final product. As such we do not presuppose any particular knowledge of logical tools.

The structure of the paper is the following. Some relevant terminology and useful lemmas are recalled in Section 2. The central results are obtained in Sections 3 and 5 where we show that a sequence  $(x_n)$  generated by (HM) is asymptotically regular (Section 3) and has the metastability property (Section 5), while also obtaining the corresponding quantitative information. In Section 4 we show a quantitative version of the metric projection in CAT(0) spaces which is used to establish the metastability of the sequence. We show how Theorem 1.1 follows from the metastability property in Section 6. Afterwards, in Section 7, we study generalized versions of the forward-backward and the Douglas-Rachford algorithms (now in the context of Hilbert spaces). Some final remarks are left for Section 8.

## 2 Preliminaries

### 2.1 Hyperbolic and CAT(0) spaces

Consider a triple  $(X, d, W)$  where  $(X, d)$  is a metric space and  $W : X \times X \times [0, 1] \rightarrow X$  is a function satisfying, for all  $x, y, z, w \in X$  and  $\lambda, \lambda' \in [0, 1]$ ,

$$(W1) \quad d(z, W(x, y, \lambda)) \leq (1 - \lambda)d(z, x) + \lambda d(z, y)$$

$$(W2) \quad d(W(x, y, \lambda), W(x, y, \lambda')) = |\lambda - \lambda'|d(x, y)$$

$$(W3) \quad W(x, y, \lambda) = W(y, x, 1 - \lambda)$$

$$(W4) \quad d(W(x, y, \lambda), W(z, w, \lambda)) \leq (1 - \lambda)d(x, z) + \lambda d(y, w).$$

A triple in the conditions above is called a *hyperbolic space*. This formulation was introduced by Kohlenbach in [30] under the name of  $W$ -hyperbolic spaces. The convexity function  $W$  was first considered by Takahashi in [57], where a triple  $(X, d, W)$  satisfying (W1) was called a convex metric space. The notion of hyperbolic space considered here is more general than the hyperbolic spaces in the sense of Reich and Shafrir [53], and slightly more restrictive than the notion of space of hyperbolic type by Goebel and Kirk [23].

If  $x, y \in X$  and  $\lambda \in [0, 1]$ , we use the notation  $(1 - \lambda)x \oplus \lambda y$  to denote  $W(x, y, \lambda)$ . It is easy to see using (W1) that

$$d(x, (1 - \lambda)x \oplus \lambda y) = \lambda d(x, y) \quad \text{and} \quad d(y, (1 - \lambda)x \oplus \lambda y) = (1 - \lambda)d(x, y). \quad (1)$$

The metric segment with endpoints  $x, y \in X$  is the set  $\{(1 - \lambda)x \oplus \lambda y : \lambda \in [0, 1]\}$  and is denoted by  $[x, y]$ . A nonempty subset  $C \subseteq X$  is said to be *convex* if  $\forall x, y \in C ([x, y] \subseteq C)$ .

The class of hyperbolic spaces includes the normed spaces and their convex subsets (with  $(1 - \lambda)x \oplus \lambda y = (1 - \lambda)x + \lambda y$ , the usual convex linear combination), the Hilbert ball [24] and the CAT(0) spaces. The important class of CAT(0) spaces (introduced by Alexandrov in [1], and named as such by Gromov in [26]; see [8] for a detailed treatment) is characterized as the hyperbolic spaces that satisfy the property  $CN^-$  (which in the

presence of the other axioms is equivalent to the Bruhat-Tits CN-inequality [11], but contrary to the latter is purely universal<sup>1</sup>):

$$\forall x, y, z \in X \left( d^2 \left( z, \frac{1}{2}x \oplus \frac{1}{2}y \right) \leq \frac{1}{2}d^2(z, x) + \frac{1}{2}d^2(z, y) - \frac{1}{4}d^2(x, y) \right) \quad (\text{CN}^-)$$

This inequality actually extends beyond midpoints – see e.g. [17, Lemma 2.5]: for all  $x, y, z \in X$  and  $\lambda \in [0, 1]$

$$d^2(z, (1 - \lambda)x \oplus \lambda y) \leq (1 - \lambda)d^2(z, x) + \lambda d^2(z, y) - \lambda(1 - \lambda)d^2(x, y). \quad (\text{CN}^+)$$

In the same way as hyperbolic spaces are considered the non-linear counterpart of normed spaces, CAT(0) spaces are the non-linear generalizations of Hilbert spaces.

As shown by Leuştean [42], CAT(0) spaces are uniformly convex with a quadratic modulus of uniform convexity:

**Lemma 2.1** ([42, Proposition 8]). *Every CAT(0) space is a uniformly convex space and  $\eta(\varepsilon) = \frac{\varepsilon^2}{8}$  is a modulus of uniform convexity, i.e. for all  $\varepsilon \in (0, 2]$ ,  $r > 0$ , and  $x, y, a \in X$*

$$\left. \begin{array}{l} d(x, a) \leq r \\ d(y, a) \leq r \\ d(x, y) \geq \varepsilon r \end{array} \right\} \rightarrow d \left( \frac{1}{2}x \oplus \frac{1}{2}y, a \right) \leq (1 - \eta(\varepsilon))r.$$

In [6, Proposition 14] it was shown that in every metric space there exists a unique function  $\langle \cdot, \cdot \rangle : X^2 \times X^2 \rightarrow \mathbb{R}$  satisfying

- (i)  $\langle \overrightarrow{xy}, \overrightarrow{xy} \rangle = d^2(x, y)$
- (ii)  $\langle \overrightarrow{xy}, \overrightarrow{uv} \rangle = \langle \overrightarrow{uv}, \overrightarrow{xy} \rangle$
- (iii)  $\langle \overrightarrow{yx}, \overrightarrow{uv} \rangle = -\langle \overrightarrow{xy}, \overrightarrow{uv} \rangle$
- (iv)  $\langle \overrightarrow{xy}, \overrightarrow{uv} \rangle + \langle \overrightarrow{xy}, \overrightarrow{vw} \rangle = \langle \overrightarrow{xy}, \overrightarrow{uw} \rangle$ ,

where  $\overrightarrow{xy}$  denotes the pair  $(x, y)$ . This unique function, called the *quasi-linearization function*, is defined for any  $(x, y), (u, v) \in X^2$  by

$$\langle \overrightarrow{xy}, \overrightarrow{uv} \rangle := \frac{1}{2} (d^2(x, v) + d^2(y, u) - d^2(x, u) - d^2(y, v)).$$

In CAT(0) spaces the quasi-linearization function satisfies the Cauchy-Schwarz inequality,

$$\langle \overrightarrow{xy}, \overrightarrow{uv} \rangle \leq d(x, y)d(u, v). \quad (2)$$

We also have the following useful inequality.

**Lemma 2.2.** *Let  $X$  be a CAT(0) space. For every  $x, y, z \in X$  and  $t \in [0, 1]$*

$$d^2((1 - t)x \oplus ty, z) \leq (1 - t)^2 d^2(x, z) + 2t(1 - t)\langle \overrightarrow{xz}, \overrightarrow{yz} \rangle + t^2 d^2(y, z).$$

*Proof.* Using (CN<sup>+</sup>) we derive

$$\begin{aligned} d^2((1 - t)x \oplus ty, z) &\leq (1 - t)d^2(x, z) + td^2(y, z) - t(1 - t)d^2(x, y) \\ &= (1 - t)^2 d^2(x, z) + t(1 - t)(d^2(x, z) + d^2(y, z) - d^2(x, y)) + t^2 d^2(y, z) \\ &= (1 - t)^2 d^2(x, z) + 2t(1 - t)\langle \overrightarrow{xz}, \overrightarrow{yz} \rangle + t^2 d^2(y, z). \end{aligned} \quad \square$$

## 2.2 Quantitative notions

**Definition 2.3.** Let  $(a_n)$  be a sequence of real numbers.

- (i) A *rate of convergence* for  $\lim a_n = 0$  is a function  $\gamma : (0, +\infty) \rightarrow \mathbb{N}$  such that

$$\forall \varepsilon > 0 \forall n \geq \gamma(\varepsilon) (|a_n| \leq \varepsilon).$$

- (ii) A *rate of divergence* for  $\lim a_n = +\infty$  is a function  $\gamma : \mathbb{N} \rightarrow \mathbb{N}$  such that

$$\forall k \in \mathbb{N} \forall n \geq \gamma(k) (a_n \geq k).$$

---

<sup>1</sup>I.e. it can be formally restated as  $\forall x_1 \dots \forall x_n A_0(x_1, \dots, x_n)$ , where  $A_0$  is quantifier-free/decidable.

(iii) A *Cauchy rate* for  $(a_n)$  is a function  $\gamma : (0, +\infty) \rightarrow \mathbb{N}$  such that

$$\forall \varepsilon > 0 \forall i, j \geq \gamma(\varepsilon) (|a_i - a_j| \leq \varepsilon).$$

When  $a_n = \sum_{i=0}^n b_i$ , for some sequence  $(b_n) \subset [0, \infty)$ ,  $\gamma$  is a *Cauchy rate* for  $(a_n)$  if

$$\forall \varepsilon > 0 \forall n \in \mathbb{N} \left( \sum_{i=\gamma(\varepsilon)+1}^n b_i \leq \varepsilon \right).$$

**Definition 2.4.** Let  $(x_n)$  be a sequence in a metric space  $(X, d)$  and  $x \in X$ .

(i) A *rate of convergence* for  $\lim x_n = x$  is a rate of convergence for  $\lim d(x_n, x) = 0$ .

(ii) A *Cauchy rate* for  $(x_n)$  is a function  $\gamma : (0, +\infty) \rightarrow \mathbb{N}$  such that

$$\forall \varepsilon > 0 \forall i, j \geq \gamma(\varepsilon) (d(x_i, x_j) \leq \varepsilon).$$

(iii) A *rate of metastability* for  $(x_n)$  is a functional  $\Gamma : (0, +\infty) \times \mathbb{N}^{\mathbb{N}} \rightarrow \mathbb{N}$  such that

$$\forall \varepsilon > 0 \forall f : \mathbb{N} \rightarrow \mathbb{N} \exists n \leq \Gamma(\varepsilon, f) \forall i, j \in [n, f(n)] (d(x_i, x_j) \leq \varepsilon).$$

We say that a function  $f : \mathbb{N} \rightarrow \mathbb{N}$  is *monotone* if

$$k \leq k' \rightarrow f(k) \leq f(k').$$

Without loss of generality, we can always assume to have this property, since if needed we can replace the function  $f$  with the function  $f^{\max} : \mathbb{N} \rightarrow \mathbb{N}$  defined by  $f^{\max}(k) := \max\{f(k') : k' \leq k\}$  for all  $k \in \mathbb{N}$ . We say that a function  $g : (0, +\infty) \rightarrow \mathbb{N}$  is *monotone* if

$$0 \leq \varepsilon \leq \varepsilon' \rightarrow g(\varepsilon) \geq g(\varepsilon').$$

This notion follows from a proof mining treatment where one works with a function  $g' : \mathbb{N} \rightarrow \mathbb{N}$  satisfying

$$g(\varepsilon) = g'(k) \text{ with } k = \lceil \varepsilon^{-1} \rceil.$$

**Lemma 2.5.** Let  $(X, d)$  be a metric space and  $(x_n)$  be a sequence in  $X$ . Then  $(x_n)$  has the metastability property, i.e.

$$\forall \varepsilon > 0 \forall f : \mathbb{N} \rightarrow \mathbb{N} \exists n \in \mathbb{N} \forall i, j \in [n, f(n)] (d(x_i, x_j) \leq \varepsilon)$$

if and only if  $(x_n)$  is a Cauchy sequence.

*Proof.* If  $(x_n)$  is a Cauchy sequence then it clearly has the metastability property. Assume that  $(x_n)$  is not a Cauchy sequence. Then, there exists  $\varepsilon_0 > 0$  such that for every  $n \in \mathbb{N}$

$$\exists i_n > j_n \geq n (d(x_{i_n}, x_{j_n}) > \varepsilon_0).$$

Hence, the metastability property fails for  $\varepsilon_0$  and the function  $f$  defined by  $f(n) := i_n$ , for all  $n \in \mathbb{N}$ .  $\square$

Note that the proof that metastability implies the Cauchy property is non-effective. As a consequence, the existence of a computable rate of metastability does not entail a computable Cauchy rate. On the other hand, a function  $\gamma$  is a Cauchy rate if and only if  $\Gamma(\varepsilon, f) := \gamma(\varepsilon)$  is a rate of metastability (see e.g. [39, Proposition 2.6]). We now recall the notion of asymptotic regularity (cf. [10]), and corresponding quantitative notions of rate of asymptotic regularity.

**Definition 2.6.** Let  $(x_n)$  be a sequence in a metric space  $(X, d)$  and consider a mapping  $T : X \rightarrow X$ .

(i) The sequence  $(x_n)$  is *asymptotically regular* if  $\lim d(x_{n+1}, x_n) = 0$ . A *rate of asymptotic regularity* for  $(x_n)$  is a rate of convergence for  $\lim d(x_{n+1}, x_n) = 0$ .

(ii) The sequence  $(x_n)$  is *asymptotically regular* with respect to  $T$  if  $\lim d(T(x_n), x_n) = 0$ . A *rate of asymptotic regularity* for  $(x_n)$  with respect to  $T$  is a rate of convergence for  $\lim d(T(x_n), x_n) = 0$ .

## 2.3 Useful lemmas

We recall the following well-known result due to Xu.

**Lemma 2.7** ([61]). *Let  $(a_n) \subset (0, 1)$  and  $(r_n), (v_n)$  be real sequences such that*

$$(i) \sum a_n = \infty; \quad (ii) \limsup r_n \leq 0; \quad (iii) \sum v_n < \infty.$$

*Let  $(s_n)$  be a non-negative real sequence satisfying  $s_{n+1} \leq (1 - a_n)s_n + a_nr_n + v_n$ , for all  $n \in \mathbb{N}$ . Then  $\lim s_n = 0$ .*

Xu's lemma has received several quantitative analyses. The next two results are (essentially) from [43, 49].

**Notation 2.8.**

1. Throughout this paper  $[x]$  is defined as  $\max\{0, [x]\}$  with the usual definition of  $[\cdot]$  in the latter.
2. Consider a function  $\varphi$  on tuples of variables  $\bar{x}, \bar{y}$ . If we wish to consider the variables  $\bar{x}$  as parameters we write  $\varphi[\bar{x}](\bar{y})$ . For simplicity of notation we may then even omit the parameters and simply write  $\varphi(\bar{y})$ .

**Lemma 2.9.** *Let  $(s_n)$  be a bounded sequence of non-negative real numbers and  $D \in \mathbb{N} \setminus \{0\}$  an upper bound on  $(s_n)$ . Consider sequences of real numbers  $(a_n) \subset [0, 1]$ ,  $(r_n) \subset \mathbb{R}$  and  $(v_n) \subset \mathbb{R}_0^+$  and functions  $A : \mathbb{N} \rightarrow \mathbb{N}$  and  $R, V : (0, +\infty) \rightarrow \mathbb{N}$  such that*

- (i)  $A$  is a rate of divergence for  $(\sum a_n)$ ,
- (ii)  $R$  is such that  $\forall \varepsilon > 0 \forall n \geq R(\varepsilon) (r_n \leq \varepsilon)$ ,
- (iii)  $V$  is a Cauchy rate for  $(\sum v_n)$ .

*Assume that for all  $n \in \mathbb{N}$ ,  $s_{n+1} \leq (1 - a_n)s_n + a_nr_n + v_n$ . Then  $\lim s_n = 0$  with rate of convergence*

$$\theta(\varepsilon) := \theta[A, R, V, D](\varepsilon) := A \left( K + \left\lceil \ln \left( \frac{3D}{\varepsilon} \right) \right\rceil \right) + 1, \quad \text{with } K := \max \left\{ R \left( \frac{\varepsilon}{3} \right), V \left( \frac{\varepsilon}{3} \right) + 1 \right\}.$$

Moreover, the following hold:

- (1) If  $r_n \equiv 0$ , then the function  $\theta$  is simplified to

$$\theta(\varepsilon) := \hat{\theta}[A, V, D](\varepsilon) := A \left( V \left( \frac{\varepsilon}{2} \right) + \left\lceil \ln \left( \frac{2D}{\varepsilon} \right) \right\rceil + 1 \right) + 1.$$

- (2) If  $v_n \equiv 0$ , then the function  $\theta$  is simplified to

$$\theta(\varepsilon) := \check{\theta}[A, R, D](\varepsilon) := A \left( R \left( \frac{\varepsilon}{2} \right) + \left\lceil \ln \left( \frac{2D}{\varepsilon} \right) \right\rceil \right) + 1.$$

Instead of considering  $\sum a_n = \infty$ , one can work with the following equivalent condition

$$\forall m \in \mathbb{N} \left( \prod_{i=m}^{\infty} (1 - a_i) = 0 \right).$$

Hence, it makes sense to also consider a corresponding quantitative hypothesis:

$A' : \mathbb{N} \times (0, +\infty) \rightarrow \mathbb{N}$  is a function satisfying

$$\forall \varepsilon > 0 \forall m \in \mathbb{N} \left( \prod_{i=m}^{A'(m, \varepsilon)} (1 - a_i) \leq \varepsilon \right), \quad (Q^*)$$

i.e.  $A'(m, \cdot)$  is a rate of convergence towards zero for the sequence  $(\prod_{i=m}^n (1 - a_i))_n$ .

Next we state a quantitative version of Lemma 2.7 which relies on the condition  $(Q^*)$  (see also [40, Lemma 2.4] and [43]).

**Lemma 2.10.** *Let  $(s_n)$  be a bounded sequence of non-negative real numbers and  $D \in \mathbb{N} \setminus \{0\}$  an upper bound on  $(s_n)$ . Consider sequences of real numbers  $(a_n) \subset [0, 1]$ ,  $(r_n) \subset \mathbb{R}$  and  $(v_n) \subset \mathbb{R}_0^+$ , and functions  $A' : \mathbb{N} \times (0, +\infty) \rightarrow \mathbb{N}$  and  $R, V : (0, +\infty) \rightarrow \mathbb{N}$  such that*

- (i)  $A'$  satisfies  $(Q^*)$ ,

(ii)  $R$  is such that  $\forall \varepsilon > 0 \forall n \geq R(\varepsilon) (r_n \leq \varepsilon)$ ,

(iii)  $V$  is a Cauchy rate for  $(\sum v_n)$ .

Assume that for all  $n \in \mathbb{N}$ ,  $s_{n+1} \leq (1 - a_n)s_n + a_nr_n + v_n$ . Then  $\lim s_n = 0$  with rate of convergence

$$\theta'[A', R, V, D](\varepsilon) := A' \left( K, \frac{\varepsilon}{3D} \right) + 1, \text{ with } K \text{ as in Lemma 2.9.}$$

Moreover, the following hold:

(1) If  $r_n \equiv 0$ , then the function  $\theta'$  is simplified to

$$\theta'(\varepsilon) := \hat{\theta}'[A, V, D](\varepsilon) := A' \left( V \left( \frac{\varepsilon}{2} \right) + 1, \frac{\varepsilon}{2D} \right) + 1.$$

(2) If  $v_n \equiv 0$ , then the function  $\theta'$  is simplified to

$$\theta'(\varepsilon) := \check{\theta}'[A, R, D](\varepsilon) := A' \left( R \left( \frac{\varepsilon}{2} \right), \frac{\varepsilon}{3D} \right) + 1.$$

We will also require particular instances of Lemma 2.7 that allow for a error term. The next two results are trivial variants of [37, Lemmas 5.2 and 5.3]. For completeness, we include the proofs nevertheless.

**Lemma 2.11.** Let  $(s_n)$  be a bounded sequence of non-negative real numbers and  $D \in \mathbb{N} \setminus \{0\}$  an upper bound on  $(s_n)$ . Consider sequences of real numbers  $(a_n) \subset [0, 1]$ ,  $(r_n) \subset \mathbb{R}$  and assume that  $\sum a_n = \infty$  with a rate of divergence  $A$ . Let  $\varepsilon > 0$ ,  $K, P \in \mathbb{N}$  be given. If for all  $n \in [K, P]$

$$(i) \quad s_{n+1} \leq (1 - a_n)s_n + a_nr_n + \mathcal{E}, \quad (ii) \quad r_n \leq \frac{\varepsilon}{3}, \quad (iii) \quad \mathcal{E} \leq \frac{\varepsilon}{3(P+1)},$$

then  $\forall n \in [\sigma, P] (s_n \leq \varepsilon)$ , where

$$\sigma := \sigma[A, D](\varepsilon, K) := A \left( K + \left\lceil \ln \left( \frac{3D}{\varepsilon} \right) \right\rceil \right) + 1.$$

*Proof.* First note that if  $\sigma > P$ , then the result is trivial. Hence we may assume that  $\sigma \leq P$  (which in particular implies that  $K \leq P$ ). Next, we may assume that  $\varepsilon < D$ , otherwise the result also trivially holds. Now, by induction one easily sees that for all  $m \leq P - K$ ,

$$s_{K+m+1} \leq \left( \prod_{i=K}^{K+m} (1 - a_i) \right) s_K + \left( 1 - \prod_{i=K}^{K+m} (1 - a_i) \right) \frac{\varepsilon}{3} + (m+1)\mathcal{E}.$$

Hence, for all  $m \leq P - K$ ,

$$s_{K+m+1} \leq \left( \prod_{i=K}^{K+m} (1 - a_i) \right) s_K + \frac{\varepsilon}{3} + (P+1)\mathcal{E} \leq \left( \prod_{i=K}^{K+m} (1 - a_i) \right) D + \frac{2\varepsilon}{3}. \quad (3)$$

Since  $j+1 \leq \sum_{i=0}^{A(j+1)} a_i \leq A(j+1) + 1$ , we have  $A(j+1) \geq j$  for all  $j \in \mathbb{N}$ . By the assumption on  $\varepsilon$ , we have  $3D/\varepsilon \geq 3$  and so  $\lceil \ln(3D/\varepsilon) \rceil \geq 1$ . Hence  $A(K + \lceil \ln(3D/\varepsilon) \rceil) \geq K$ , and can consider the natural number  $M := A(K + \lceil \ln(3D/\varepsilon) \rceil) - K$ .

For all  $m \geq M$ , we have

$$\sum_{i=0}^{K+m} a_i \geq \sum_{i=0}^{K+M} a_i = \sum_{i=0}^{A(K + \lceil \ln(3D/\varepsilon) \rceil)} a_i \geq K + \ln \left( \frac{3D}{\varepsilon} \right) \geq \sum_{i=0}^{K-1} a_i + \ln \left( \frac{3D}{\varepsilon} \right),$$

which entails  $\sum_{i=K}^{K+m} a_i \geq \ln(3D/\varepsilon)$ . Using the inequality  $\forall x \in \mathbb{R}_0^+ (1 - x \leq \exp(-x))$ , we derive for all  $m \geq M$ ,

$$\left( \prod_{i=K}^{K+m} (1 - a_i) \right) D \leq \exp \left( - \sum_{i=K}^{K+m} a_i \right) D \leq \frac{\varepsilon}{3D} D = \frac{\varepsilon}{3}.$$

Together with (3), we have thus concluded that for  $m \in [M, P - K]$ ,

$$s_{K+m+1} \leq \left( \prod_{i=K}^{K+m} (1 - a_i) \right) D + \frac{2\varepsilon}{3} \leq \varepsilon,$$

which entails the result.  $\square$

As before, instead of a rate of divergence  $A$ , we can establish an analogue of the previous result with a function  $A'$  satisfying  $(Q^*)$ .

**Lemma 2.12.** *Let  $(s_n)$  be a bounded sequence of non-negative real numbers and  $D \in \mathbb{N} \setminus \{0\}$  an upper bound on  $(s_n)$ . Consider sequences of real numbers  $(a_n) \subset [0, 1]$ ,  $(r_n) \subset \mathbb{R}$  assume that  $\sum a_n = \infty$  with a function  $A'$  satisfying  $(Q^*)$ . Let  $\varepsilon > 0$ ,  $K, P \in \mathbb{N}$  be given. If for all  $n \in [K, P]$*

$$(i) \quad s_{n+1} \leq (1 - a_n)s_n + a_n r_n + \mathcal{E}, \quad (ii) \quad r_n \leq \frac{\varepsilon}{3}, \quad (iii) \quad \mathcal{E} \leq \frac{\varepsilon}{3(P+1)},$$

then  $\forall n \in [\sigma', P] (s_n \leq \varepsilon)$ , where

$$\sigma' := \sigma'[A', D](\varepsilon, K) := A' \left( K, \frac{\varepsilon}{3D} \right) + 1.$$

*Proof.* Following the proof of Lemma 2.11, we have for all  $m \leq P - K$ ,

$$s_{K+m+1} \leq \left( \prod_{i=K}^{K+m} (1 - a_i) \right) D + \frac{2\varepsilon}{3}.$$

Again from the assumption on  $\varepsilon$ , we may consider the natural number  $M := A'(K, \varepsilon/3D) - K$ . For all  $m \geq M$ ,

$$\left( \prod_{i=K}^{K+m} (1 - a_i) \right) D \leq \left( \prod_{i=K}^{K+M} (1 - a_i) \right) D = \left( \prod_{i=K}^{A'(K, \varepsilon/3D)} (1 - a_i) \right) D \leq \frac{\varepsilon}{3},$$

which as before entails the result.  $\square$

### 3 Asymptotic regularity

In this section we prove that a sequence  $(x_n)$  generated by (HM) is an asymptotically regular sequence of approximate common fixed points of  $T$  and  $U$ . We will assume that the conditions (i) – (v) of Theorem 1.1 hold and that their corresponding quantitative information is given explicitly. Namely, given  $(\alpha_n), (\beta_n) \subseteq [0, 1]$ , we consider the following conditions:

(Q<sub>1</sub>)  $\Gamma_1 : (0, +\infty) \rightarrow \mathbb{N}$  is a rate of convergence for  $\lim \alpha_n = 0$ , i.e.

$$\forall \varepsilon > 0 \forall n \geq \Gamma_1(\varepsilon) (\alpha_n \leq \varepsilon).$$

(Q<sub>2</sub>)  $\Gamma_2 : \mathbb{N} \rightarrow \mathbb{N}$  is a rate of divergence for  $(\sum_{n \geq 0} \alpha_n)$ , i.e.

$$\forall k \in \mathbb{N} \left( \sum_{i=0}^{\Gamma_2(k)} \alpha_i \geq k \right).$$

(Q<sub>3</sub>)  $\Gamma_3 : (0, +\infty) \rightarrow \mathbb{N}$  is a Cauchy rate for  $(\sum_{n \geq 0} |\alpha_{n+1} - \alpha_n|)$ , i.e.

$$\forall \varepsilon > 0 \forall n \in \mathbb{N} \left( \sum_{i=\Gamma_3(\varepsilon)+1}^{\Gamma_3(\varepsilon)+n} |\alpha_{i+1} - \alpha_i| \leq \varepsilon \right).$$

(Q<sub>4</sub>)  $\Gamma_4 : (0, +\infty) \rightarrow \mathbb{N}$  is a Cauchy rate for  $(\sum_{n \geq 0} |\beta_{n+1} - \beta_n|)$ .

(Q<sub>5</sub>)  $\gamma \in (0, 1/2]$  is such that  $\gamma \leq \beta_n \leq 1 - \gamma$ , for all  $n \in \mathbb{N}$ .

We will assume furthermore that the functions above satisfy the natural monotonicity conditions:

1. if  $\varepsilon \leq \varepsilon'$  then  $\Gamma_i(\varepsilon) \geq \Gamma_i(\varepsilon')$ , for  $i = 1, 3, 4$
2. if  $k \leq k'$  then  $\Gamma_2(k) \leq \Gamma_2(k')$ .

**Example 3.1.** Consider the sequences defined by  $\alpha_n := \frac{1}{n+1}$ ,  $\beta_n := \beta \in (0, 1)$  for all  $n \in \mathbb{N}$ . Then the conditions (Q<sub>1</sub>) – (Q<sub>5</sub>) are satisfied with

$$\Gamma_1(\varepsilon) = \Gamma_3(\varepsilon) := \lfloor \varepsilon^{-1} \rfloor, \quad \Gamma_2(k) := \lfloor e^k \rfloor, \quad \Gamma_4 \equiv 0, \quad \text{and} \quad \gamma := \min\{\beta, 1 - \beta\}.$$



We start by showing that  $(x_n)$  is bounded.

**Lemma 3.2.** *Consider a natural number  $N \in \mathbb{N} \setminus \{0\}$  such that  $N \geq \max\{d(x_0, p), d(u, p)\}$ , for some  $p \in F$ . We have that*

$$d(x_n, p) \leq N \quad (4)$$

and, in particular, the sequence  $(x_n)$  is bounded.

*Proof.* Since  $U$  is nonexpansive and  $U(p) = p$ , using (W1) we have

$$d(x_{2n+2}, p) \leq (1 - \alpha_n)d(U(x_{2n+1}), p) + \alpha_n d(x_{2n+1}, p) = d(x_{2n+1}, p). \quad (5)$$

By induction, we have that  $\forall n \in \mathbb{N} (d(x_{2n}, p) \leq N)$ . The base case is trivial. If the induction hypothesis holds for  $n$ , using the fact that  $T$  is nonexpansive,  $T(p) = p$ , and again (W1), we get

$$d(x_{2n+1}, p) \leq (1 - \alpha_n)d(x_{2n}, p) + \alpha_n d(u, p) \leq (1 - \alpha_n)N + \alpha_n N = N. \quad (6)$$

By (5), the induction step follows. Moreover, from (6) it is clear that the bound is also valid for the odd terms, and we conclude the result.  $\square$

The next lemma allows to prove the asymptotically regularity of the sequence  $(x_n)$ .

**Lemma 3.3.** *Consider a natural number  $N \in \mathbb{N} \setminus \{0\}$  such that  $N \geq \max\{d(x_0, p), d(u, p)\}$ , for some  $p \in F$ . If conditions  $(Q_1) - (Q_5)$  hold, then*

- (i)  $\theta_1$  is a rate of convergence for  $\lim d(x_{2n+2}, x_{2n}) = 0$ ,
- (ii)  $\theta_2$  is a rate of convergence for  $\lim d(x_{2n+3}, x_{2n+1}) = 0$ ,
- (iii)  $\theta_3$  is a rate of convergence for both  $\lim d(U(x_{2n+1}), x_{2n+1}) = 0$  and for  $\lim d(x_{2n+2}, x_{2n+1}) = 0$ ,
- (iv)  $\theta_4$  is a rate of convergence for  $\lim d(x_{2n+1}, x_{2n}) = 0$ .

where

$$\begin{aligned} \theta_1(\varepsilon) &:= \hat{\theta}[A, V, D], \text{ and } \hat{\theta} \text{ is as in Lemma 2.9(1), with parameters} \\ &\bullet A(k) := \Gamma_2(k+1), \text{ for all } k \in \mathbb{N} \\ &\bullet V(\varepsilon) := \max \left\{ \Gamma_3 \left( \frac{\varepsilon}{4N} \right), \Gamma_4 \left( \frac{\varepsilon}{4N} \right) \right\}, \text{ for all } \varepsilon > 0 \\ &\bullet D := 2N - \text{which by (4) is a bound on the sequence } (d(x_{2n+2}, x_{2n})) \\ \theta_2(\varepsilon) &:= \max \left\{ \theta_1 \left( \frac{\varepsilon}{2} \right), \Gamma_3 \left( \frac{\varepsilon}{4N} \right) + 1 \right\} \\ \theta_3(\varepsilon) &:= \max \left\{ \theta_1 \left( \frac{(\gamma\varepsilon)^2}{8N} \right), \Gamma_1 \left( \frac{(\gamma\varepsilon)^2}{2N^2} \right) \right\} \\ \theta_4(\varepsilon) &:= \max \left\{ \theta_1 \left( \frac{\varepsilon}{2} \right), \theta_3 \left( \frac{\varepsilon}{2} \right) \right\}. \end{aligned}$$

*Proof.* Using (W2) and (W4) we have

$$\begin{aligned} d(x_{2n+3}, x_{2n+1}) &\leq d(x_{2n+3}, (1 - \alpha_{n+1})T(x_{2n}) \oplus \alpha_{n+1}u) + d((1 - \alpha_{n+1})T(x_{2n}) \oplus \alpha_{n+1}u, x_{2n+1}) \\ &\leq (1 - \alpha_{n+1})d(T(x_{2n+2}), T(x_{2n})) + |\alpha_{n+1} - \alpha_n|d(T(x_{2n}), u) \\ &\leq (1 - \alpha_{n+1})d(x_{2n+2}, x_{2n}) + |\alpha_{n+1} - \alpha_n|d(T(x_{2n}), u). \end{aligned}$$

Hence, using (4), for all  $n \in \mathbb{N}$

$$d(x_{2n+3}, x_{2n+1}) \leq (1 - \alpha_{n+1})d(x_{2n+2}, x_{2n}) + 2N|\alpha_{n+1} - \alpha_n|. \quad (7)$$

Similarly,

$$\begin{aligned} d(x_{2n+4}, x_{2n+2}) &\leq d(x_{2n+2}, (1 - \beta_{n+1})U(x_{2n+1}) \oplus \beta_{n+1}x_{2n+1}) \\ &\quad + d((1 - \beta_{n+1})U(x_{2n+1}) \oplus \beta_{n+1}x_{2n+1}, x_{2n+2}) \\ &\leq (1 - \beta_{n+1})d(U(x_{2n+3}), U(x_{2n+1})) + \beta_{n+1}d(x_{2n+3}, x_{2n+1}) \\ &\quad + |\beta_{n+1} - \beta_n|d(U(x_{2n+1}), x_{2n+1}) \\ &\leq d(x_{2n+3}, x_{2n+1}) + |\beta_{n+1} - \beta_n|d(U(x_{2n+1}), x_{2n+1}). \end{aligned}$$

Hence, using (4), for all  $n \in \mathbb{N}$

$$d(x_{2n+4}, x_{2n+2}) \leq d(x_{2n+3}, x_{2n+1}) + 2N|\beta_{n+1} - \beta_n|. \quad (8)$$

From (7) and (8) we conclude that for all  $n \in \mathbb{N}$

$$d(x_{2n+4}, x_{2n+2}) \leq (1 - \alpha_{n+1})d(x_{2n+2}, x_{2n}) + 2N(|\alpha_{n+1} - \alpha_n| + |\beta_{n+1} - \beta_n|).$$

By  $(Q_2)$ ,  $(Q_3)$  and  $(Q_4)$ , we may apply Lemma 2.9 with  $s_n = d(x_{2n+2}, x_{2n})$ ,  $a_n = \alpha_{n+1}$ ,  $r_n \equiv 0$ , and  $v_n = 2N(|\alpha_{n+1} - \alpha_n| + |\beta_{n+1} - \beta_n|)$  to conclude that

$$\lim d(x_{2n+2}, x_{2n}) = 0, \text{ with rate of convergence } \theta_1. \quad (9)$$

Indeed, for all  $k \in \mathbb{N}$ ,

$$\sum_{i=0}^{A(k)} a_i = \sum_{i=0}^{\Gamma_2(k+1)} \alpha_{i+1} \geq \sum_{i=0}^{\Gamma_2(k+1)} \alpha_i - \alpha_0 \geq k$$

and, for all  $\varepsilon > 0$  and  $n \in \mathbb{N}$ ,

$$\begin{aligned} \sum_{i=V(\varepsilon)+1}^{V(\varepsilon)+n} v_i &= 2N \left( \sum_{i=V(\varepsilon)+1}^{V(\varepsilon)+n} |\alpha_{i+1} - \alpha_i| + \sum_{i=V(\varepsilon)+1}^{V(\varepsilon)+n} |\beta_{i+1} - \beta_i| \right) \\ &= 2N \left( \sum_{i=\Gamma_3(\frac{\varepsilon}{4N})+j+1}^{\Gamma_3(\frac{\varepsilon}{4N})+j+n} |\alpha_{i+1} - \alpha_i| + \sum_{i=\Gamma_4(\frac{\varepsilon}{4N})+j'+1}^{\Gamma_4(\frac{\varepsilon}{4N})+j'+n} |\beta_{i+1} - \beta_i| \right) \\ &\leq 2N \left( \sum_{i=\Gamma_3(\frac{\varepsilon}{4N})+1}^{\Gamma_3(\frac{\varepsilon}{4N})+j+n} |\alpha_{i+1} - \alpha_i| + \sum_{i=\Gamma_4(\frac{\varepsilon}{4N})+1}^{\Gamma_4(\frac{\varepsilon}{4N})+j'+n} |\beta_{i+1} - \beta_i| \right) \\ &\leq 2N \left( \frac{\varepsilon}{4N} + \frac{\varepsilon}{4N} \right) = \varepsilon. \end{aligned}$$

From  $(Q_3)$ , for any  $\varepsilon > 0$  and  $n \geq \Gamma_3(\varepsilon/4N) + 1$ , we have  $|\alpha_{n+1} - \alpha_n| \leq \varepsilon/4N$ . Then, by (7),

$$\lim d(x_{2n+3}, x_{2n+1}) = 0, \text{ with rate of convergence } \theta_2. \quad (10)$$

With  $p \in F$  as in the hypothesis over  $N$ , using  $(\text{CN}^+)$  and the fact that  $U$  is nonexpansive, we have

$$\begin{aligned} d^2(x_{2n+2}, p) &\leq (1 - \beta_n)d^2(U(x_{2n+1}), p) + \beta_n d^2(x_{2n+1}, p) - \beta_n(1 - \beta_n)d^2(U(x_{2n+1}), x_{2n+1}) \\ &\leq d^2(x_{2n+1}, p) - \beta_n(1 - \beta_n)d^2(U(x_{2n+1}), x_{2n+1}) \end{aligned}$$

and similarly, now using the fact that  $T$  is nonexpansive,

$$\begin{aligned} d^2(x_{2n+1}, p) &\leq (1 - \alpha_n)d^2(T(x_{2n}), p) + \alpha_n d^2(u, p) - \alpha_n(1 - \alpha_n)d^2(T(x_{2n}), u) \\ &\leq d^2(x_{2n}, p) + \alpha_n d^2(u, p). \end{aligned}$$

Then

$$d^2(x_{2n+2}, p) \leq d^2(x_{2n}, p) + \alpha_n d^2(u, p) - \beta_n(1 - \beta_n)d^2(U(x_{2n+1}), x_{2n+1}),$$

which implies

$$\begin{aligned} \beta_n(1 - \beta_n)d^2(U(x_{2n+1}), x_{2n+1}) &\leq d^2(x_{2n}, p) - d^2(x_{2n+2}, p) + \alpha_n d^2(u, p) \\ &\leq d(x_{2n+2}, x_{2n})(d(x_{2n+2}, x_{2n}) + 2d(x_{2n+2}, p)) + \alpha_n d^2(u, p). \end{aligned}$$

Now, condition  $(Q_5)$  entails

$$(\gamma \cdot d(U(x_{2n+1}), x_{2n+1}))^2 \leq d(x_{2n+2}, x_{2n})(d(x_{2n+2}, x_{2n}) + 2d(x_{2n+2}, p)) + \alpha_n d^2(u, p). \quad (11)$$

For  $\varepsilon > 0$ , using (4) and (9), we conclude that for  $n \geq \theta_1((\gamma\varepsilon)^2/8N)$ ,

$$d(x_{2n+2}, x_{2n})(d(x_{2n+2}, x_{2n}) + 2d(x_{2n+2}, p)) \leq d(x_{2n+2}, x_{2n})(2N + 2N) \leq \frac{(\gamma\varepsilon)^2}{2}.$$

Moreover, from condition  $(Q_1)$ , we get for  $n \geq \Gamma_1((\gamma\varepsilon)^2/2N^2)$

$$\alpha_n d^2(u, p) \leq \alpha_n N^2 \leq \frac{(\gamma\varepsilon)^2}{2}.$$

Hence, from (11) we conclude that

$$\lim d(U(x_{2n+1}), x_{2n+1}) = 0, \text{ with rate of convergence } \theta_3. \quad (12)$$

Since  $d(x_{2n+2}, x_{2n+1}) = (1 - \beta_n)d(U(x_{2n+1}), x_{2n+1})$  by (1), from (12) we also get that  $\lim d(x_{2n+2}, x_{2n+1}) = 0$  with rate of convergence  $\theta_3$ . On the other hand, since  $d(x_{2n+1}, x_{2n}) \leq d(x_{2n+2}, x_{2n+1}) + d(x_{2n+2}, x_{2n})$  it follows that  $\lim d(x_{2n+1}, x_{2n}) = 0$  with rate of convergence  $\theta_4$ .  $\square$

**Example 3.4.** With  $(\alpha_n), (\beta_n)$  as in Example 3.1, it is possible to verify that we have the following rates of convergence

$$\begin{aligned} \varepsilon &\mapsto \left\lceil \exp\left(\frac{12N}{\varepsilon} + 2\right) \right\rceil + 1 \text{ for } \lim d(x_{2n+2}, x_{2n}) = 0, \\ \varepsilon &\mapsto \left\lceil \exp\left(\frac{24N}{\varepsilon} + 2\right) \right\rceil + 1 \text{ for } \lim d(x_{2n+3}, x_{2n+1}) = 0, \\ \varepsilon &\mapsto \left\lceil \exp\left(\left(\frac{14N}{\gamma\varepsilon}\right)^2 + 2\right) \right\rceil + 1 \text{ for both } \lim d(U(x_{2n+1}), x_{2n+1}) = 0 \text{ and } \lim d(x_{2n+2}, x_{2n+1}) = 0, \\ \varepsilon &\mapsto \left\lceil \exp\left(\left(\frac{20N}{\gamma\varepsilon}\right)^2 + 2\right) \right\rceil + 1 \text{ for } \lim d(x_{2n+1}, x_{2n}) = 0. \end{aligned}$$

**Remark 3.5.** Lemma 3.3, and Proposition 3.6 below, still hold if instead of a function  $\Gamma_2$  satisfying  $(Q_2)$ , we have a monotone<sup>2</sup> function  $\Gamma'_2 : \mathbb{N} \times (0, +\infty) \rightarrow \mathbb{N}$  such that for all  $m \in \mathbb{N}$ ,  $\Gamma'_2(m, \cdot)$  is a rate of convergence towards zero for the sequence  $(\prod_{i=m}^n (1 - \alpha_i))$ . In this case, one defines  $\theta_1$  using Lemma 2.10(1) instead of Lemma 2.9(1), namely  $\theta_1 = \hat{\theta}'[A', V, 2N]$ , where  $A'(m, \varepsilon) := \Gamma'_2(m + 1, \varepsilon)$ , and the remaining functions are the same as before.

We now compute rates of convergence for the asymptotic properties of the sequence  $(x_n)$  using Lemma 3.3.

**Proposition 3.6.** *Under the conditions of Lemma 3.3, we have that*

$$\lim d(x_{n+1}, x_n) = \lim d(U(x_n), x_n) = \lim d(T(x_n), x_n) = 0,$$

with monotone rates of convergence, respectively,

$$\begin{aligned} (i) \quad \rho_1(\varepsilon) &:= \max\{2\theta_3(\varepsilon) + 1, 2\theta_4(\varepsilon)\} \\ (ii) \quad \rho_2(\varepsilon) &:= 2\theta_3\left(\frac{\varepsilon}{3}\right) + 2 \\ (iii) \quad \rho_3(\varepsilon) &:= 2 \max\left\{\theta_4\left(\frac{\varepsilon}{6}\right), \Gamma_1\left(\frac{\varepsilon}{4N}\right)\right\} + 1, \end{aligned}$$

where  $\theta_1, \theta_2, \theta_3$  and  $\theta_4$  are as in Lemma 3.3.

*Proof.* Let  $\varepsilon > 0$  be given. For part (i), take  $n \geq \rho_1(\varepsilon)$ . If  $n = 2n' + 1$ , then  $n' \geq \theta_3(\varepsilon)$  and the result follows from Lemma 3.3(iii). Similarly, if  $n = 2n'$  we get  $n' \geq \theta_4(\varepsilon)$  and the result now follows from Lemma 3.3(iv). For part (ii), using the fact that  $U$  is nonexpansive we have

$$\begin{aligned} d(U(x_{2n+2}), x_{2n+2}) &\leq d(U(x_{2n+2}), U(x_{2n+1})) + d(U(x_{2n+1}), x_{2n+1}) + d(x_{2n+2}, x_{2n+1}) \\ &\leq 2d(x_{2n+2}, x_{2n+1}) + d(U(x_{2n+1}), x_{2n+1}) \leq 3d(U(x_{2n+1}), x_{2n+1}), \end{aligned}$$

which from Lemma 3.3(iii) implies that  $\lim d(U(x_{2n+2}), x_{2n+2}) = 0$  with rate of convergence  $\varepsilon \mapsto \theta_3(\varepsilon/3)$ , and consequently

$$\lim d(U(x_n), x_n) = 0, \text{ with rate of convergence } \tilde{\rho}_2(\varepsilon) := \max\left\{2\theta_3(\varepsilon) + 1, 2\theta_3\left(\frac{\varepsilon}{3}\right) + 2\right\}.$$

Note that the assumption of monotonicity on the functions  $\Gamma_i$  (and the definition of the function  $\theta$  in Lemma 2.9), entail that the functions  $\theta_i$  (and  $\rho_i$ ) are also monotone. In particular, the monotonicity of  $\theta_3$  entails that  $\theta_3(\varepsilon/3) \geq \theta_3(\varepsilon)$ , and so  $\tilde{\rho}_2$  coincides with  $\rho_2$ .

Finally for part (iii), using (W1), the fact that  $T$  is a nonexpansive map and (4), we obtain

$$\begin{aligned} d(T(x_{2n+1}), x_{2n+1}) &\leq (1 - \alpha_n)d(T(x_{2n+1}), T(x_{2n})) + \alpha_n d(T(x_{2n+1}), u) \\ &\leq d(x_{2n+1}, x_{2n}) + 2N\alpha_n \end{aligned}$$

---

<sup>2</sup>I.e.  $\Gamma'_2$  satisfies  $\forall \varepsilon_1, \varepsilon_2 > 0 \forall m_1, m_2 \in \mathbb{N} (\varepsilon_1 \leq \varepsilon_2 \wedge m_1 \leq m_2 \rightarrow \Gamma'_2(m_1, \varepsilon_2) \leq \Gamma'_2(m_2, \varepsilon_1))$ .

and

$$\begin{aligned} d(T(x_{2n}), x_{2n}) &\leq d(T(x_{2n+1}), T(x_{2n})) + d(T(x_{2n+1}), x_{2n+1}) + d(x_{2n+1}, x_{2n}) \\ &\leq 2d(x_{2n+1}, x_{2n}) + d(T(x_{2n+1}), x_{2n+1}) \\ &\leq 3d(x_{2n+1}, x_{2n}) + 2N\alpha_n. \end{aligned}$$

By  $(Q_1)$  and Lemma 3.3(iv), we conclude that

$$\begin{aligned} \varepsilon &\mapsto \max \left\{ \theta_4 \left( \frac{\varepsilon}{2} \right), \Gamma_1 \left( \frac{\varepsilon}{4N} \right) \right\} \text{ is a rate of convergence for } \lim d(T(x_{2n+1}), x_{2n+1}) = 0 \\ \varepsilon &\mapsto \max \left\{ \theta_4 \left( \frac{\varepsilon}{6} \right), \Gamma_1 \left( \frac{\varepsilon}{4N} \right) \right\} \text{ is a rate of convergence for } \lim d(T(x_{2n}), x_{2n}) = 0. \end{aligned}$$

Using the monotonicity of  $\theta_4$ , we conclude part (iii).  $\square$

**Example 3.7.** With  $(\alpha_n), (\beta_n)$  as in Example 3.1, it is possible to verify that for any  $\varepsilon > 0$ ,

$$\begin{aligned} n &\geq \left\lceil \exp \left( \left( \frac{20N}{\gamma\varepsilon} \right)^2 + 3 \right) \right\rceil + 2 \rightarrow d(x_{n+1}, x_n) \leq \varepsilon, \\ n &\geq \left\lceil \exp \left( \left( \frac{14N}{\gamma\varepsilon} \right)^2 + 3 \right) \right\rceil + 3 \rightarrow d(U(x_n), x_n) \leq \varepsilon, \\ n &\geq \left\lceil \exp \left( \left( \frac{20N}{\gamma\varepsilon} \right)^2 + 3 \right) \right\rceil + 3 \rightarrow d(T(x_n), x_n) \leq \varepsilon. \end{aligned}$$

We have the following trivial consequence of Proposition 3.6.

**Lemma 3.8.** For any  $\varepsilon > 0$  define  $\rho(\varepsilon) := \max\{\rho_2(\varepsilon), \rho_3(\varepsilon)\}$ , where  $\rho_2, \rho_3$  are as in Proposition 3.6. Then  $\rho$  is a monotone rate of asymptotic regularity for both  $U$  and  $T$ , i.e.

$$\forall \varepsilon > 0 \forall n \geq \rho(\varepsilon) (d(x_n, T(x_n)), d(x_n, U(x_n))) \leq \varepsilon).$$

## 4 Metric Projection in CAT(0) spaces

Let  $S$  be a nonempty closed convex subset of a complete CAT(0) space  $X$ . The *metric projection* onto  $S$  is the mapping  $P_S : X \rightarrow S$  defined, for each  $u \in X$ , as the unique point  $P_S u$  satisfying  $d(u, P_S u) = \min_{x \in S} d(u, x)$ , [8]. In this section, we show a useful quantitative version regarding the metric projection. The proof follows the arguments from [22, 49] (see also [32]). We point out that our analysis only requires an  $\varepsilon$ -weakening of the metric projection (similarly to [32]), and thus the results in this section are valid even without assuming the completeness of the space.

Let  $T, U : C \rightarrow C$  be nonexpansive maps on a nonempty and convex subset of a CAT(0) space  $X$  and assume that  $F := \text{Fix}(T) \cap \text{Fix}(U) \neq \emptyset$ . Given  $\eta > 0$  and  $a \in X$  we write

$$F_N(a, \eta) := \{x \in C : d(x, T(x)), d(x, U(x)) \leq \eta\} \cap \overline{B}_N(a).$$

**Proposition 4.1.** Given  $u \in C$ , let  $N \in \mathbb{N} \setminus \{0\}$  be such that  $N \geq 2d(u, p)$ , for some  $p \in F$ . Then, for every  $\varepsilon > 0$  and  $\delta : (0, 1] \rightarrow (0, 1]$  there exists  $\eta \in [\varphi(\varepsilon, \delta), 1]$  and  $x \in F_N(p, \delta(\eta))$  such that

$$\forall y \in F_N(p, \eta) (d^2(u, x) \leq d^2(u, y) + \varepsilon),$$

where  $\varphi(\varepsilon, \delta) := \varphi[N](\varepsilon, \delta) := \min\{\delta^{(i)}(1) : i \leq r\}$  with  $r := r(N, \varepsilon) := \lceil \frac{N^2}{4\varepsilon} \rceil$ .

Moreover,  $\eta$  is witnessed by  $\delta^{(i)}(1)$ , for some  $i \leq r$ .

*Proof.* Suppose that for some  $\varepsilon > 0$  and  $\delta : (0, 1] \rightarrow (0, 1]$  we have

$$\forall \eta \in [\varphi(\varepsilon, \delta), 1] \forall x \in F_N(p, \delta(\eta)) \exists y \in F_N(p, \eta) (d^2(u, y) < d^2(u, x) - \varepsilon). \quad (13)$$

We construct a finite sequence  $x_0, \dots, x_r$  of elements of  $C$  as follows. We define  $x_0 := p$ . In particular,  $x_0 \in F_N(p, \delta^{(r)}(1))$ . Assume that, for  $j \leq r - 1$  we have  $x_j \in F_N(p, \delta^{(r-j)}(1))$ . Then, by (13) there exists  $y_0 \in F_N(p, \delta^{(r-j-1)}(1))$  such that

$$d^2(u, y_0) < d^2(u, x_j) - \varepsilon,$$

and we define  $x_{j+1} := y_0$ .

By construction, for all  $j \leq r$  we have  $d^2(u, x_{j+1}) < d^2(u, x_j) - \varepsilon$  and so

$$d^2(u, x_r) < d^2(u, x_0) - r\varepsilon \leq \frac{N^2}{4} - \frac{N^2}{4\varepsilon}\varepsilon = 0,$$

which is a contradiction.  $\square$

The following lemmas are a generalization to CAT(0) spaces of Lemmas 2.3 and 2.7 of [32]. The proofs are similar but we include them nevertheless for completeness.

**Lemma 4.2.** *Let  $X$  be a CAT(0) space,  $C \subseteq X$  a convex bounded subset and  $b \in \mathbb{N} \setminus \{0\}$ , a bound on the diameter of  $C$ . Consider  $T : C \rightarrow C$  a nonexpansive map on  $C$ . Then*

$$\forall \varepsilon > 0 \forall x_1, x_2 \in C \left( \bigwedge_{i=1}^2 d(x_i, T(x_i)) < \frac{\varepsilon^2}{12b} \rightarrow \forall t \in [0, 1] (d(w_t, T(w_t)) < \varepsilon) \right),$$

where  $w_t := (1-t)x_1 \oplus tx_2$ .

*Proof.* Let  $x_1, x_2 \in C$  be such that

$$\bigwedge_{i=1}^2 d(x_i, T(x_i)) < \frac{\varepsilon^2}{12b}. \quad (14)$$

For any  $t \in [0, 1]$  we have  $K := \max\{d(x_1, w_t), d(x_1, T(w_t))\} \leq b$ . If  $K < \varepsilon/2$ , then  $d(w_t, T(w_t)) \leq d(x_1, w_t) + d(x_1, T(w_t)) \leq \varepsilon$ . So, we may then assume that  $K \geq \frac{\varepsilon}{2}$ . Since  $T$  is nonexpansive, using (14) we derive

$$\begin{aligned} d\left(x_2, \frac{1}{2}w_t \oplus \frac{1}{2}T(w_t)\right) &\stackrel{(W1)}{\leq} \frac{1}{2}d(x_2, w_t) + \frac{1}{2}d(x_2, T(w_t)) \\ &\leq \frac{1}{2}d(x_2, w_t) + \frac{1}{2}d(T(x_2), T(w_t)) + \frac{1}{2}d(x_2, T(x_2)) \\ &< d(x_2, w_t) + \frac{\varepsilon^2}{24b} \\ &\leq d(x_2, w_t) + \frac{\varepsilon^2}{24K}. \end{aligned}$$

We have that

$$d\left(x_1, \frac{1}{2}w_t \oplus \frac{1}{2}T(w_t)\right) > d(x_1, w_t) - \frac{\varepsilon^2}{24K}, \quad (15)$$

otherwise we would obtain the following contradiction

$$\begin{aligned} d(x_1, x_2) &\leq d\left(x_1, \frac{1}{2}w_t \oplus \frac{1}{2}T(w_t)\right) + d\left(x_2, \frac{1}{2}w_t \oplus \frac{1}{2}T(w_t)\right) \\ &< d(x_1, w_t) - \frac{\varepsilon^2}{24K} + d(x_2, w_t) + \frac{\varepsilon^2}{24K} \\ &= td(x_1, x_2) + (1-t)d(x_1, x_2) = d(x_1, x_2). \end{aligned}$$

Since  $d(x_1, T(w_t)) < d(x_1, w_t) + \frac{\varepsilon^2}{12K}$ , from (15) we obtain

$$d\left(x_1, \frac{1}{2}w_t \oplus \frac{1}{2}T(w_t)\right) > d(x_1, w_t) - \frac{\varepsilon^2}{24K} > K - \frac{\varepsilon^2}{12K} - \frac{\varepsilon^2}{24K} = K \left(1 - \frac{(\varepsilon/K)^2}{8}\right).$$

From Lemma 2.1 with  $a = x_1$ ,  $x = w_t$ ,  $y = T(w_t)$ ,  $r = K$  and  $\varepsilon = \frac{\varepsilon}{K} \in (0, 2]$  we obtain

$$d(w_t, T(w_t)) < K \frac{\varepsilon}{K} = \varepsilon. \quad \square$$

**Lemma 4.3.** *Let  $X$  be a CAT(0) space and  $x, y, u \in X, t \in [0, 1]$  and define  $w_t := (1-t)x \oplus ty$ . Then*

$$\forall \varepsilon \in (0, b^2] \left( \forall t \in [0, 1] \left( d^2(u, x) \leq d^2(u, w_t) + \frac{\varepsilon^2}{b^2} \right) \rightarrow \langle \overrightarrow{ux}, \overrightarrow{yx} \rangle \leq \varepsilon \right),$$

where  $b \geq d(x, y)$ .

*Proof.* We have, for any  $t \in [0, 1]$

$$\begin{aligned} d^2(u, w_t) &\leq (1-t)d^2(u, x) + td^2(u, y) - t(1-t)d^2(x, y) \\ &= d^2(u, x) + 2t \left( \frac{1}{2} (d^2(u, y) + d^2(x, x) - d^2(u, x) - d^2(x, y)) \right) + t^2 d^2(x, y) \\ &= d^2(u, x) + 2t \langle \overrightarrow{ux}, \overrightarrow{xy} \rangle + t^2 d^2(x, y). \end{aligned}$$

Assume that  $d^2(u, x) \leq d^2(u, w_t) + \frac{\varepsilon^2}{b^2}$ . Hence, for  $t = \frac{\varepsilon}{b^2} \in [0, 1]$  we have

$$-\frac{\varepsilon^2}{b^2} \leq 2 \frac{\varepsilon}{b^2} \langle \overrightarrow{ux}, \overrightarrow{xy} \rangle + \left( \frac{\varepsilon}{b^2} \right)^2 d^2(x, y) \leq 2 \frac{\varepsilon}{b^2} \langle \overrightarrow{ux}, \overrightarrow{xy} \rangle + \left( \frac{\varepsilon}{b^2} \right)^2 b^2 = -2 \frac{\varepsilon}{b^2} \langle \overrightarrow{ux}, \overrightarrow{yx} \rangle + \frac{\varepsilon^2}{b^2},$$

from which we derive that  $\langle \overrightarrow{ux}, \overrightarrow{yx} \rangle \leq \varepsilon$ .  $\square$

By combining the previous results we obtain a quantitative version of the characterization of the metric projection in CAT(0) spaces (cf. Lemma 6.1).

**Proposition 4.4.** *Given  $u \in C$ , let  $N \in \mathbb{N} \setminus \{0\}$  be such that  $N \geq 2d(u, p)$  for some  $p \in F$ . For any  $\varepsilon > 0$  and function  $\delta : (0, 1] \rightarrow (0, 1]$ , there exist  $\eta \in [\Phi(\varepsilon, \delta), 1]$  and  $x \in F_N(p, \delta(\eta))$  such that*

$$\forall y \in F_N(p, \eta) \left( \langle \overline{ux}, \overline{yx} \rangle \leq \varepsilon \right),$$

where  $\Phi(\varepsilon, \delta) := \Phi[N](\varepsilon, \delta) := \frac{\varphi(\tilde{\varepsilon}, \tilde{\delta})^2}{24N}$ , with  $\tilde{\varepsilon} := \frac{\varepsilon^2}{4N^2}$ ,  $\tilde{\delta}(\xi) := \min \left\{ \delta \left( \frac{\xi^2}{24N} \right), \frac{\xi^2}{24N} \right\}$ , for any  $\xi \in (0, 1]$ , and  $\varphi := \varphi[N]$  is as in Proposition 4.1.

*Proof.* Let  $\varepsilon > 0$  and  $\delta : (0, 1] \rightarrow (0, 1]$  be given. We may assume that  $\varepsilon \leq 4N^2$ , otherwise the result follows easily using the Cauchy-Schwarz inequality. By Proposition 4.1, there exist  $\eta_0 \in [\varphi(\tilde{\varepsilon}, \tilde{\delta}), 1]$  and  $x \in F_N(p, \tilde{\delta}(\eta_0))$  such that

$$\forall y \in F_N(p, \eta_0) \left( d^2(u, x) \leq d^2(u, y) + \tilde{\varepsilon} \right). \quad (16)$$

Define  $\eta_1 := \frac{\eta_0^2}{24N} \in [\Phi(\varepsilon, \delta), 1]$ . Since  $x \in F_N(p, \eta_1)$ , for any  $y \in F_N(p, \eta_1)$ , by Lemma 4.2 (applied to  $T$  and  $U$  restricted to  $C \cap \overline{B}_N(p)$  and  $b = 2N$ )

$$\forall t \in [0, 1] \left( w_t \in F_N(p, \eta_0) \right),$$

with  $w_t := (1 - t)x \oplus ty$ . Hence, by (16)

$$\forall t \in [0, 1] \left( d^2(u, x) \leq d^2(u, w_t) + \frac{\varepsilon^2}{4N^2} \right)$$

and the result follows from Lemma 4.3 (with  $b = 2N$ ) and the fact that  $d(x, T(x)), d(x, U(x)) \leq \delta(\eta_1)$ .  $\square$

We now apply, in the setting of CAT(0) spaces, a technique which corresponds to the quantitative removal of sequential weak compactness in Hilbert spaces – see [22].

**Proposition 4.5.** *Given  $u \in C$ , let  $N \in \mathbb{N} \setminus \{0\}$  be such that  $N \geq \max\{d(x_0, p), 2d(u, p)\}$  for some  $p \in F$ . Let  $\rho$  be a monotone rate of asymptotic regularity for both  $U$  and  $T$ . For any  $\varepsilon > 0$  and function  $\Delta : \mathbb{N} \rightarrow (0, 1]$ , there exist  $n \leq \Psi[N, \rho](\varepsilon, \Delta)$  and  $x \in F_N(p, \Delta(n))$  such that*

$$\forall m \geq n \left( \langle \overline{ux}, \overline{x_m x} \rangle \leq \varepsilon \right),$$

where  $\Psi[N, \rho](\varepsilon, \Delta) := \rho(\Phi(\varepsilon, \bar{\Delta}))$ , with  $\bar{\Delta}(\eta) := \Delta(\rho(\eta))$  for all  $\eta \in (0, 1]$ , and  $\Phi := \Phi[N]$  is as in Proposition 4.4.

*Proof.* Let  $\varepsilon > 0$  and  $\Delta : \mathbb{N} \rightarrow (0, 1]$  be given. By Proposition 4.4, there exist  $\eta_0 \in [\Phi(\varepsilon, \bar{\Delta}), 1]$  and  $x \in F_N(p, \bar{\Delta}(\eta_0))$  such that

$$\forall y \in F_N(p, \eta_0) \left( \langle \overline{ux}, \overline{yx} \rangle \leq \varepsilon \right). \quad (17)$$

Let  $n := \rho(\eta_0)$ . Observe that  $d(x, T(x)), d(x, U(x)) \leq \bar{\Delta}(\eta_0) = \Delta(n)$  and using the monotonicity of  $\rho$ , we have  $n \leq \Psi[N, \rho](\varepsilon, \Delta)$ . Since  $d(x_n, p) \leq N$  by Lemma 3.2, and  $\forall m \geq n \left( d(x_n, T(x_n)), d(x_n, U(x_n)) \leq \eta_0 \right)$ , the result follows from (17).  $\square$

## 5 Metastability

In this section we prove metastability results pertaining to the iteration  $(x_n)$  and extract some consequences of that property. We also consider an iteration  $(\mathbf{HM}_e)$  which is a generalization of the iteration  $(\mathbf{HM})$  with error terms. We show that rates of metastability for a sequence generated by  $(\mathbf{HM}_e)$  follow from the metastability of  $(x_n)$ .

**Theorem 5.1.** *Let  $X$  be a CAT(0) space and  $C \subseteq X$  a nonempty convex subset. Consider nonexpansive mappings  $T, U : C \rightarrow C$  such that  $F := \text{Fix}(T) \cap \text{Fix}(U) \neq \emptyset$ . Let  $(\alpha_n), (\beta_n) \subset [0, 1]$  and  $u, x_0 \in C$ . Assume that we have monotone functions  $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$  and a constant  $\gamma \in (0, 1)$  satisfying  $(Q_1) - (Q_5)$ , respectively. Let  $N \in \mathbb{N} \setminus \{0\}$  be a natural number such that  $N \geq \max\{d(x_0, p), 2d(u, p)\}$ , for some  $p \in F$ .*

*Then  $(x_n)$  generated by  $(\mathbf{HM})$  is a Cauchy sequence with rate of metastability*

$$\mu(\varepsilon, f) := 2 \max \left\{ \Sigma \left( \Psi \left( \frac{\tilde{\varepsilon}}{24}, \Delta_{\varepsilon, f} \right) \right), \theta_4 \left( \frac{\varepsilon}{4} \right) \right\} + 1,$$

where  $\mu := \mu[N, \gamma, \Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4]$  and

$$\tilde{\varepsilon} := \frac{\varepsilon^2}{16},$$

$$\Delta_{\varepsilon, f}(n) := \min \left\{ \frac{\tilde{\varepsilon}}{30N(P(n)+1)}, 1 \right\},$$

$$P(n) := f \left( 2 \max \left\{ \Sigma(n), \theta_4 \left( \frac{\varepsilon}{4} \right) \right\} + 1 \right),$$

$$\Sigma(n) := \sigma \left[ \Gamma_2, 4N^2 \right] (\tilde{\varepsilon}, K(n)), \text{ with } \sigma \text{ as defined in Lemma 2.11,}$$

$$K(n) := \max \left\{ n, \rho_3 \left( \frac{\tilde{\varepsilon}}{36N} \right) \right\},$$

$$\theta_4 \text{ as in Lemma 3.3, } \rho_3 \text{ as in Proposition 3.6,}$$

$$\Psi = \Psi[N, \rho] \text{ as in Proposition 4.5, with } \rho \text{ as in Lemma 3.8.}$$

*Proof.* Let  $\varepsilon > 0$  and a function  $f : \mathbb{N} \rightarrow \mathbb{N}$  be given. By Proposition 4.5 we may consider  $n_0 \leq \Psi(\tilde{\varepsilon}/24, \Delta_{\varepsilon, f})$  and  $\tilde{x} \in F_N(p, \Delta_{\varepsilon, f}(n_0))$  such that

$$\forall n \geq n_0 \left( \langle \overrightarrow{u\tilde{x}}, \overrightarrow{x_n\tilde{x}} \rangle \leq \frac{\tilde{\varepsilon}}{24} \right).$$

Using (W1) and the nonexpansivity of  $U$ , we have for all  $n \in \mathbb{N}$

$$d(x_{2n+2}, \tilde{x}) \leq (1 - \beta_n)d(U(x_{2n+1}), \tilde{x}) + \beta_n d(x_{2n+1}, \tilde{x}) \leq d(x_{2n+1}, \tilde{x}) + d(U(\tilde{x}), \tilde{x}).$$

So, using Lemma 2.2, we have

$$\begin{aligned} d^2(x_{2n+2}, \tilde{x}) &\leq (d(x_{2n+1}, \tilde{x}) + d(U(\tilde{x}), \tilde{x}))^2 \\ &= d^2(x_{2n+1}, \tilde{x}) + d(U(\tilde{x}), \tilde{x}) (d(U(\tilde{x}), \tilde{x}) + 2d(x_{2n+1}, \tilde{x})) \\ &\leq (1 - \alpha_n)^2 d^2(T(x_{2n}), \tilde{x}) + 2\alpha_n(1 - \alpha_n) \langle \overrightarrow{T(x_{2n})\tilde{x}}, \overrightarrow{u\tilde{x}} \rangle + \alpha_n^2 d^2(u, \tilde{x}) + 5N\Delta_{\varepsilon, f}(n_0) \\ &\leq (1 - \alpha_n)d^2(x_{2n}, \tilde{x}) + \alpha_n \left( 2\langle \overrightarrow{u\tilde{x}}, \overrightarrow{T(x_{2n})\tilde{x}} \rangle + 3N^2\alpha_n \right) + 10N\Delta_{\varepsilon, f}(n_0), \end{aligned} \quad (18)$$

since  $2d^2(u, \tilde{x}) \leq (d(u, p) + d(\tilde{x}, p))^2 \leq (3N/2)^2 \leq 3N^2$  and

$$\begin{aligned} d^2(T(x_{2n}), \tilde{x}) &\leq d^2(T(x_{2n}), T(\tilde{x})) + d(T(\tilde{x}), \tilde{x}) (d(T(\tilde{x}), \tilde{x}) + 2d(T(x_{2n}), T(\tilde{x}))) \\ &\leq d^2(x_{2n}, \tilde{x}) + \Delta_{\varepsilon, f}(n_0)(1 + 2 \cdot 2N) \leq d^2(x_{2n}, \tilde{x}) + 5N\Delta_{\varepsilon, f}(n_0). \end{aligned}$$

Thus, the inequality (18) can be written in the compact form

$$s_{n+1} \leq (1 - \alpha_n)s_n + \alpha_n r_n + \mathcal{E},$$

where  $s_n := d^2(x_{2n}, \tilde{x})$ ,  $r_n := 2\langle \overrightarrow{u\tilde{x}}, \overrightarrow{T(x_{2n})\tilde{x}} \rangle + 3N^2\alpha_n$  and  $\mathcal{E} := 10N\Delta_{\varepsilon, f}(n_0)$ . Using Cauchy-Schwarz, for  $n \geq \max\{n_0, \rho_3(\frac{\tilde{\varepsilon}}{36N})\}$ , we have

$$\begin{aligned} 2\langle \overrightarrow{u\tilde{x}}, \overrightarrow{T(x_{2n})\tilde{x}} \rangle &= 2\langle \overrightarrow{u\tilde{x}}, \overrightarrow{T(x_{2n})x_{2n}} \rangle + 2\langle \overrightarrow{u\tilde{x}}, \overrightarrow{x_{2n}\tilde{x}} \rangle \\ &\leq 2d(u, \tilde{x})d(T(x_{2n}), x_{2n}) + 2\langle \overrightarrow{u\tilde{x}}, \overrightarrow{x_{2n}\tilde{x}} \rangle \\ &\leq 3N\frac{\tilde{\varepsilon}}{36N} + \frac{2\tilde{\varepsilon}}{24} = \frac{\tilde{\varepsilon}}{6}. \end{aligned}$$

By the definition of  $\rho_3$ , note that  $\max\{n_0, \rho_3(\frac{\tilde{\varepsilon}}{36N})\} \geq \Gamma_1(\frac{\tilde{\varepsilon}}{18N^2})$  and so

$$3N^2\alpha_n \leq \frac{\tilde{\varepsilon}}{6}.$$

Hence, for all  $n \geq K(n_0)$ ,  $r_n \leq \frac{\tilde{\varepsilon}}{3}$ . Since

$$\mathcal{E} \leq \frac{10N\tilde{\varepsilon}}{30N(P(n_0)+1)} = \frac{\tilde{\varepsilon}}{3(P(n_0)+1)},$$

we can apply Lemma 2.11 to conclude

$$\forall n \in [\Sigma(n_0), P(n_0)] \left( d(x_{2n}, \tilde{x}) \leq \frac{\varepsilon}{4} \right). \quad (19)$$

Since  $d(x_{2n+1}, \tilde{x}) \leq d(x_{2n+1}, x_{2n}) + d(x_{2n}, \tilde{x})$ , with  $n_1 := \max\{\Sigma(n_0), \theta_4(\varepsilon/4)\}$  we obtain

$$\forall n \in [n_1, P(n_0)] \left( d(x_{2n}, \tilde{x}), d(x_{2n+1}, \tilde{x}) \leq \frac{\varepsilon}{2} \right).$$

The latter entails that for  $n \in [2n_1 + 1, P(n_0)]$ ,  $d(x_n, \tilde{x}) \leq \frac{\varepsilon}{2}$ . One easily sees that the function  $\Sigma$  is monotone, and therefore  $2n_1 + 1 \leq \mu(\varepsilon, f)$ . Finally, noticing that  $P(n_0) = f(2n_1 + 1)$ , by the triangle inequality we conclude that  $\mu$  is a rate of metastability for  $(x_n)$ .  $\square$

**Remark 5.2.** 1. If  $\Gamma_2$  is replaced by a monotone function  $\Gamma'_2 : \mathbb{N} \times (0, +\infty) \rightarrow \mathbb{N}$  satisfying

$$\forall \varepsilon > 0 \forall m \in \mathbb{N} \left( \prod_{i=m}^{\Gamma'_2(m, \varepsilon)} (1 - \alpha_i) \leq \varepsilon \right),$$

then one can compute a rate of metastability as before but applying Lemma 2.12 instead of Lemma 2.11. Indeed the only difference would be in the definition of the number function  $\Sigma$ , which now would use  $\sigma'$  instead of  $\sigma$ ,

$$\Sigma(n) := \sigma'[\Gamma'_2, 4N^2](\tilde{\varepsilon}, K(n)).$$

2. The function  $\rho_3$  was only required as a rate of convergence towards zero for  $(d(T(x_{2n}), x_{2n}))$ . It is then clear that  $\rho_3$  could be replaced by  $\varepsilon \mapsto \max\{\theta_4(\frac{\varepsilon}{6}), \Gamma_1(\frac{\varepsilon}{4N})\}$  (cf. the end of the proof of Proposition 3.6).

As a corollary to the proof of Theorem 5.1, we can derive quantitative information on the Halpern iteration. Previous rates of metastability were obtained by Kohlenbach and Leuştean in [37] using different methods. Their results follow from a quantitative analysis of a proof by Saejung [54] and rely on a technique to eliminate the use of Banach limits needed in the original proof.

**Corollary 5.3.** Let  $T : C \rightarrow C$  be a nonexpansive mapping,  $(\alpha_n) \subset [0, 1]$  and  $y_0, u \in C$ . Let  $(y_n)$  be a sequence generated by the Halpern schema (H),

$$y_{n+1} := (1 - \alpha_n)T(y_n) \oplus \alpha_n u.$$

Assume that we have monotone functions  $\Gamma_1, \Gamma_2, \Gamma_3$  satisfying  $(Q_1) - (Q_3)$ , respectively. Let  $N \in \mathbb{N} \setminus \{0\}$  be a natural number such that  $N \geq \max\{d(y_0, p), 2d(u, p)\}$ , for some  $p \in \text{Fix}(T)$ . Then

- (i)  $(y_n)$  is asymptotically regular, and we have the following rates of asymptotic regularity

$$\lim d(y_{n+1}, y_n) = 0 \text{ with rate of convergence } \hat{\theta}[A, V, D],$$

$$\lim d(T(y_n), y_n) = 0 \text{ with rate of convergence } \tilde{\rho}(\varepsilon) := \max\left\{\hat{\theta}[A, V, D](\frac{\varepsilon}{2}), \Gamma_1\left(\frac{\varepsilon}{3N}\right)\right\},$$

where  $\hat{\theta}$  is as in Lemma 2.9(1), with parameters  $A(k) := \Gamma_2(k+1)$ ,  $V(\varepsilon) := \Gamma_3(\frac{2\varepsilon}{3N})$ , and  $D := 2N$ .

- (ii)  $(y_n)$  is a Cauchy sequence with rate of metastability

$$\zeta(\varepsilon, f) := \tilde{\Sigma}\left(\Psi\left(\frac{\tilde{\varepsilon}}{24}, \tilde{\Delta}_{\varepsilon, f}\right)\right),$$

where

$$\tilde{\varepsilon} := \frac{\varepsilon^2}{4},$$

$$\tilde{\Delta}_{\varepsilon, f} := \min\left\{\frac{\tilde{\varepsilon}}{30N\left(f\left(\tilde{\Sigma}(n)\right) + 1\right)}, 1\right\},$$

$$\tilde{\Sigma}(n) := \sigma[\Gamma_3, 4N^2](\tilde{\varepsilon}, \tilde{K}(n)), \text{ with } \sigma \text{ as defined in Lemma 2.11,}$$

$$\tilde{K}(n) := \max\left\{n, \tilde{\rho}\left(\frac{\tilde{\varepsilon}}{36N}\right)\right\},$$

$$\Psi = \Psi[N, \tilde{\rho}] \text{ as in Proposition 4.5.}$$

*Proof.* To prove part (i), we simplify the arguments in the proof of Lemma 3.3. As before, one easily proves by induction that  $d(y_n, p) \leq N$ , for all  $n \in \mathbb{N}$ . Using (W2) and (W4) one shows that for all  $n \in \mathbb{N}$

$$d(y_{n+2}, y_{n+1}) \leq (1 - \alpha_{n+1})d(y_{n+1}, y_n) + \frac{3N}{2}|\alpha_{n+1} - \alpha_n|.$$

By an application of Lemma 2.9(1) we have that  $\lim d(y_{n+1}, y_n) = 0$  with rate of convergence  $\hat{\theta}[A, V, D]$ . Since, using (W1) and the assumption on  $N$ , for all  $n \in \mathbb{N}$

$$d(T(y_n), y_n) \leq \frac{3N}{2}\alpha_n + d(y_{n+1}, y_n),$$

we see that  $\tilde{\rho}$  is a rate of convergence for  $\lim d(T(y_n), y_n) = 0$ , and conclude the proof of part (i).



We now argue the metastability property of  $(y_n)$ . First consider  $(x_n)$  generated by (HM) with parameters  $(\alpha_n, T, \beta_n, U, x_0, u) := (\alpha_n, T, 1/2, \text{Id}_C, y_0, u)$ . By an easy induction one sees that  $y_n = x_{2n}$ , for all  $n \in \mathbb{N}$ . In light of Remark 5.2(2) and the fact that  $\tilde{\rho}$  is a rate of convergence towards zero for  $(d(T(y_n), y_n)) \equiv (d(T(x_{2n}), x_{2n}))$ , we can follow the proof of Theorem 5.1 (with  $\tilde{\varepsilon} := \varepsilon^2/4$ ) until (19), to conclude that

$$\forall n \in \left[ \tilde{\Sigma}(n_0), f\left(\tilde{\Sigma}(n_0)\right) \right] \left( d(y_n, \tilde{x}) \leq \frac{\varepsilon}{2} \right).$$

The result then follows using the triangular inequality.  $\square$

The next results take into account the presence of error terms in the inductive construction of the iteration (HM). We start with an easy lemma which is well-known (see e.g. [43, Proposition 7]). We include its short proof for the sake of completeness.

**Lemma 5.4.** *Consider two sequences  $(w_n)$  and  $(z_n)$  in a metric space. If  $(w_n)$  is a Cauchy sequence with rate of metastability  $\tau$  and  $\lim d(w_n, z_n) = 0$  with rate of convergence  $\nu$ , then  $(z_n)$  is also a Cauchy sequence with rate of metastability*

$$\tau_\nu(\varepsilon, f) := \max \left\{ \tau \left( \frac{\varepsilon}{3}, f_{\varepsilon, \nu} \right), \nu \left( \frac{\varepsilon}{3} \right) \right\},$$

where  $f_{\varepsilon, \nu} := f(\max\{n, \nu(\varepsilon/3)\})$ , for all  $n \in \mathbb{N}$ .

*Proof.* Let  $\varepsilon > 0$  and  $f : \mathbb{N} \rightarrow \mathbb{N}$  be given. By the assumption on  $\tau$ , consider  $n_0 \leq \tau(\varepsilon/3, f_{\varepsilon, \nu})$  such that  $d(w_i, w_j) \leq \varepsilon/3$ , for all  $i, j \in [n_0, f_{\varepsilon, \nu}(n_0)]$ . Define  $n := \max\{n_0, \nu(\varepsilon/3)\} \leq \tau_\nu(\varepsilon, f)$ . Then, for all  $i, j \in [n, f_{\varepsilon, \nu}(n_0)] = [n, f(n)]$  we have

$$d(z_i, z_j) \leq d(z_i, w_i) + d(w_i, w_j) + d(w_j, z_j) \leq \varepsilon,$$

using the assumption on  $\nu$  and the fact that  $n \geq n_0$ .  $\square$

**Theorem 5.5.** *Let  $T, U : C \rightarrow C$  be nonexpansive mappings and  $(\alpha_n), (\beta_n) \subset [0, 1]$ . Let  $(x_n)$  be generated by (HM) and assume that  $(Q_2)$  holds with rate of divergence  $\Gamma_2$ . Consider a sequence  $(x'_n)$  satisfying  $x'_0 = x_0$  and*

$$\begin{cases} d(x'_{2n+1}, (1 - \alpha_n)T(x'_{2n}) \oplus \alpha_n u) & \leq \delta_{2n} \\ d(x'_{2n+2}, (1 - \beta_n)U(x'_{2n+1}) \oplus \beta_n x'_{2n+1}) & \leq \delta_{2n+1}, \end{cases} \quad (\text{HM}_e)$$

where  $(\delta_n) \subset [0, +\infty)$ . If  $\sum \delta_n < \infty$ , or  $\lim \frac{\delta_{2n} + \delta_{2n+1}}{\alpha_n} = 0$  (in the case  $\alpha_n > 0$ ), then  $\lim d(x'_n, x_n) = 0$ . Moreover, the following hold:

(1) If  $\chi_1$  is a Cauchy rate for  $(\sum \delta_n)$ , then  $\hat{\nu}$  is a rate of convergence for  $\lim d(x'_n, x_n) = 0$ , where

$$\hat{\nu}(\varepsilon) := 2 \max \left\{ \theta \left( \frac{\varepsilon}{2} \right), \chi_1 \left( \frac{\varepsilon}{2} \right) \right\} + 3,$$

with  $\theta := \hat{\theta}[\Gamma_2, \chi_1, [\sum_{i=0}^{\chi_1(1)} \delta_i] + 1]$ , where  $\hat{\theta}$  is as in Lemma 2.9(1).

(2) If  $\chi_2$  is a rate of convergence for  $\lim \frac{\delta_{2n} + \delta_{2n+1}}{\alpha_n} = 0$ , then  $\check{\nu}$  is a rate of convergence for  $\lim d(x'_n, x_n) = 0$ , where

$$\check{\nu}(\varepsilon) := 2 \max \left\{ \theta \left( \frac{\varepsilon}{2} \right), \chi_2 \left( \frac{\varepsilon}{2} \right) \right\} + 1,$$

with  $\theta := \check{\theta}[\Gamma_2, \chi_2, [\max_{i \leq \chi_2(1)} \left\{ \frac{\delta_{2i} + \delta_{2i+1}}{\alpha_i}, 1 \right\}]]$ , where  $\check{\theta}$  is as in Lemma 2.9(2).

*Proof.* By the nonexpansivity of  $T$  and (W4), we have for all  $n \in \mathbb{N}$

$$\begin{aligned} d(x'_{2n+1}, x_{2n+1}) &\leq d(x'_{2n+1}, (1 - \alpha_n)T(x'_{2n}) \oplus \alpha_n u) \\ &\quad + d((1 - \alpha_n)T(x'_{2n}) \oplus \alpha_n u, x_{2n+1}) \\ &\leq (1 - \alpha_n)d(x'_{2n}, x_{2n}) + \delta_{2n}. \end{aligned} \quad (20)$$

Similarly

$$\begin{aligned} d(x'_{2n+2}, x_{2n+2}) &\leq d(x'_{2n+2}, (1 - \beta_n)U(x'_{2n+1}) \oplus \beta_n x'_{2n+1}) \\ &\quad + d((1 - \beta_n)U(x'_{2n+1}) \oplus \beta_n x'_{2n+1}, x_{2n+2}) \\ &\leq d(x'_{2n+1}, x_{2n+1}) + \delta_{2n+1} \\ &\leq (1 - \alpha_n)d(x'_{2n}, x_{2n}) + \delta_{2n} + \delta_{2n+1}. \end{aligned} \quad (21)$$

We then conclude that  $\lim d(x'_{2n}, x_{2n}) = 0$  by applying Lemma 2.7 with  $r_n := 0$  and  $v_n := \delta_{2n} + \delta_{2n+1}$ , if  $\sum \delta_n < \infty$ , or with  $r_n := \frac{\delta_{2n} + \delta_{2n+1}}{\alpha_n}$  and  $v_n := 0$ , if  $\lim \frac{\delta_{2n} + \delta_{2n+1}}{\alpha_n} = 0$ . Then by (20), since  $\delta_n \rightarrow 0$ , also  $\lim d(x'_{2n+1}, x_{2n+1}) = 0$ . Hence  $\lim d(x'_n, x_n) = 0$ .

Let us show part (1). Clearly, (21) entails that  $d(x'_{2n}, x_{2n}) \leq \sum_{i=0}^{2n-1} \delta_i$ . With  $D := [\sum_{i=0}^{\chi_1(1)} \delta_i] + 1$  it is easy to see that the assumption on  $\chi_1$  implies that  $D$  is an upper bound on  $(\sum \delta_n)$ , and so also on  $(d(x'_{2n}, x_{2n}))$ . The function  $\chi_1$  is also a Cauchy rate for  $(\sum v_n)$ , where  $v_n := \delta_{2n} + \delta_{2n+1}$ . Indeed, for all  $\varepsilon > 0$  and  $n \in \mathbb{N}$

$$\sum_{i=\chi_1(\varepsilon)+1}^{\chi_1(\varepsilon)+n} v_n = \sum_{i=\chi_1(\varepsilon)+1}^{\chi_1(\varepsilon)+n} \delta_{2i} + \delta_{2i+1} = \sum_{i=2\chi_1(\varepsilon)+2}^{2\chi_1(\varepsilon)+2n+1} \delta_i \leq \sum_{i=\chi_1(\varepsilon)+1}^{2\chi_1(\varepsilon)+2n+1} \delta_i \leq \varepsilon.$$

Now, by Lemma 2.9(1) we have that  $\theta = \hat{\theta}[\Gamma_2, \chi_1, D]$  is a rate of convergence for  $\lim d(x'_{2n}, x_{2n}) = 0$ . Since  $\chi_1$  is a Cauchy rate for  $(\sum \delta_n)$ , we have  $\delta_{2n} \rightarrow 0$  with rate of convergence  $\chi_1 + 1$ . Using (20) we conclude that

$$\forall n \geq \max \left\{ \theta \left( \frac{\varepsilon}{2} \right), \chi_1 \left( \frac{\varepsilon}{2} \right) + 1 \right\} (d(x'_{2n}, x_{2n}), d(x'_{2n+1}, x_{2n+1}) \leq \varepsilon),$$

which entails part (1).

We now turn our attention to part (2). From the assumption on  $\chi_2$ , we have that  $\max_{i \leq \chi_2(1)} \left\{ \frac{\delta_{2i} + \delta_{2i+1}}{\alpha_i}, 1 \right\}$  is an upper bound on the sequence  $\left( \frac{\delta_{2n} + \delta_{2n+1}}{\alpha_n} \right)$ . Using (21), one shows by induction that  $(d(x'_{2n}, x_{2n}))$  is bounded by  $D$ , where  $D := \left\lceil \max_{i \leq \chi_2(1)} \left\{ \frac{\delta_{2i} + \delta_{2i+1}}{\alpha_i}, 1 \right\} \right\rceil$ . Now, by Lemma 2.9(2) we have that  $\theta = \check{\theta}[\Gamma_2, \chi_2, D]$  is a rate of convergence for  $\lim d(x'_{2n}, x_{2n}) = 0$ . Since  $\chi_2$  is a rate of convergence for  $\lim \frac{\delta_{2n} + \delta_{2n+1}}{\alpha_n} = 0$ , we have  $\delta_{2n} \rightarrow 0$  with the same rate of convergence. Using (20) we conclude that

$$\forall n \geq \max \left\{ \theta \left( \frac{\varepsilon}{2} \right), \chi_2 \left( \frac{\varepsilon}{2} \right) \right\} (d(x'_{2n}, x_{2n}), d(x'_{2n+1}, x_{2n+1}) \leq \varepsilon),$$

which entails the result.  $\square$

A rate of metastability for the generalized sequence with error terms  $(\text{HM}_e)$  can now be immediately obtained from Theorem 5.1 using Lemma 5.4 and Theorem 5.5. Additionally, it is clear that  $(x'_n)$  is asymptotically regular and rates of asymptotic regularity can be trivially computed using the rates in Proposition 3.6 together with the rates of convergence in Theorem 5.5.

**Corollary 5.6.** Consider sequences  $(\alpha_n), (\beta_n) \subset [0, 1]$ , a sequence of error terms  $(\delta_n) \subset [0, +\infty)$  satisfying  $\sum \delta_n < \infty$ , or  $\lim \frac{\delta_{2n} + \delta_{2n+1}}{\alpha_n} = 0$  (if  $\alpha_n > 0$ ), and let  $(x'_n)$  be generated by  $(\text{HM}_e)$ . Then, under the conditions of Theorem 5.1, the sequence  $(x'_n)$  is a Cauchy sequence. Moreover,

- (1) if  $\chi_1$  is a Cauchy rate for  $(\sum \delta_n)$ , then  $\mu_{\hat{\nu}}$  is a rate of metastability for  $(x'_n)$ ;
- (2) if  $\chi_2$  is a rate of convergence for  $\lim \frac{\delta_{2n} + \delta_{2n+1}}{\alpha_n} = 0$ , then  $\mu_{\check{\nu}}$  is a rate of metastability for  $(x'_n)$ ;

where  $\mu$  is as in Theorem 5.1,  $\hat{\nu}, \check{\nu}$  are as in Theorem 5.5 and the construction  $\tau_{\nu}$  is as in Lemma 5.4.

## 6 Strong Convergence

We will now argue how the metastability result from the previous section actually entails the strong convergence of the iteration  $(x_n)$ . We begin by recalling a well-known characterization of the metric projection in terms of the quasi-linearization function.

**Lemma 6.1** ([16]). *Let  $S$  be a nonempty convex closed subset of a complete CAT(0) space  $X$ . For any  $u \in X$ , let  $P_S(u)$  denote the metric projection of  $u \in X$  onto  $S$ . Then,*

$$\forall y \in S \left( \langle \overrightarrow{uP_S(u)}, \overrightarrow{yP_S(u)} \rangle \leq 0 \right).$$

We are now ready to give the proof of the main theorem.

**Proof of Theorem 1.1.** From Theorem 5.1 and Lemma 2.5,  $(x_n)$  is a Cauchy sequence and therefore convergent since  $X$  is complete. Let  $z = \lim x_n$ , which is an element of  $C$ , because  $C$  is closed. By Proposition 3.6

and the continuity of  $T$  and  $U$ , we conclude that  $z \in F$ . From Lemma 6.1 (with  $S = F$ ), we conclude that  $\langle \overrightarrow{uP_F(u)}, \overrightarrow{zP_F(u)} \rangle \leq 0$ . Using Lemma 2.2, with  $s_n := d^2(x_{2n}, P_F(u))$  we have

$$\begin{aligned} s_{n+1} &\leq d^2(x_{2n+1}, P_F(u)) \\ &\leq (1 - \alpha_n)^2 d^2(T(x_{2n}), P_F(u)) + 2\alpha_n(1 - \alpha_n) \langle \overrightarrow{uP_F(u)}, \overrightarrow{T(x_{2n})P_F(u)} \rangle + \alpha_n^2 d^2(u, P_F(u)) \\ &\leq (1 - \alpha_n) d^2(x_{2n}, P_F(u)) + \alpha_n \left( 2 \langle \overrightarrow{uP_F(u)}, \overrightarrow{T(x_{2n})P_F(u)} \rangle + \alpha_n d^2(u, P_F(u)) \right) \\ &= (1 - \alpha_n) s_n + \alpha_n r_n, \end{aligned} \tag{22}$$

where  $r_n = 2 \langle \overrightarrow{uP_F(u)}, \overrightarrow{T(x_{2n})P_F(u)} \rangle + \alpha_n d^2(u, P_F(u))$ . Since

$$\begin{aligned} \langle \overrightarrow{uP_F(u)}, \overrightarrow{T(x_{2n})P_F(u)} \rangle &= \langle \overrightarrow{uP_F(u)}, \overrightarrow{T(x_{2n})x_{2n}} \rangle + \langle \overrightarrow{uP_F(u)}, \overrightarrow{x_{2n}z} \rangle + \langle \overrightarrow{uP_F(u)}, \overrightarrow{zP_F(u)} \rangle \\ &\leq d(u, P_F(u)) (d(T(x_{2n}), x_{2n}) + d(x_{2n}, z)) \rightarrow 0, \end{aligned}$$

and  $\alpha_n \rightarrow 0$ , we conclude that  $\limsup r_n \leq 0$ . Since  $\sum \alpha_n = \infty$ , we can apply Lemma 2.7 to conclude that  $\lim s_n = 0$ . This means that  $\lim x_{2n} = P_F(u)$ , and so  $z = P_F(u)$ .  $\square$

As discussed in the Introduction, Theorem 1.1 is a twofold generalization of the recent strong convergence result, established in the setting of Hilbert spaces, by Boř, Csetnek, and Meier [7, Theorem 3]. On the one hand, our iteration is more general than the one considered by Boř *et al.* and, on the other hand, the strong convergence is established in CAT(0) spaces, which are frequently considered the non-linear generalization of Hilbert spaces. Our schema is also a generalization of the algorithm considered by Leuştean and Cheval in [14] in the setting of hyperbolic spaces. Even though the latter context is more general than that of CAT(0) spaces, our Theorem 1.1 succeeds in proving strong convergence while [14, Theorems 4.1 and 4.2] only show the asymptotic regularity of the iteration. Observe that not only the restriction to CAT(0) spaces was required to prove strong convergence, but we also needed to consider the condition  $0 < \liminf \beta_n$  (which has no corresponding condition in [7] nor in [14]) in order to obtain asymptotic regularity for the more general iteration (HM). In light of the new results in [13] connecting the iteration (CL) with the modified Halpern iteration, our results also provides quantitative information on the modified Halpern iteration (first obtained in [55]).

Since the Halpern iteration follows from a particular case of (HM) – cf. the proof of Corollary 5.3 –, Theorem 1.1 also allows to recover Saejung’s strong convergence of the Halpern iteration in CAT(0) spaces from [54, Theorem 2.3]. We would like to point out that, contrary to Saejung’s proof, we do not require the heavy machinery of Banach limits. Finally, we point out that Theorem 5.5, also extends the strong convergence result to the iterative schema (HM<sub>e</sub>).

## 7 Forward-backward and Douglas-Rachford algorithms

In this section we work in a Hilbert space  $H$  and apply our main results to obtain strongly convergent variants of the Forward-backward and Douglas-Rachford algorithms.

We recall that a multi-valued operator  $\mathsf{T} : H \rightrightarrows H$  is *monotone* if whenever  $(x, y)$  and  $(x', y')$  are elements of the graph of  $\mathsf{T}$ , it holds that  $\langle x - x', y - y' \rangle \geq 0$ . A monotone operator  $\mathsf{T}$  is said to be *maximal monotone* if the graph of  $\mathsf{T}$  is not properly contained in the graph of any other monotone operator on  $H$ .

We use  $J_{\mathsf{T}}$  to denote the *resolvent function* of  $\mathsf{T}$ , i.e. the single-valued function defined by  $J_{\mathsf{T}} = (I + \mathsf{T})^{-1}$  and  $R_{\mathsf{T}}$  to denote the nonexpansive *reflected resolvent function* defined by  $R_{\mathsf{T}} := 2J_{\mathsf{T}} - \text{Id}$ .

**Definition 7.1.** A mapping  $U : H \rightarrow H$  is called *firmly nonexpansive* if

$$\forall x, y \in H \left( \|U(x) - U(y)\|^2 \leq \|x - y\|^2 - \|(\text{Id} - U)(x) - (\text{Id} - U)(y)\|^2 \right).$$

Clearly, if  $U$  is firmly nonexpansive then it is nonexpansive. For  $c > 0$ , the resolvent function  $J_{c\mathsf{T}}$  is firmly nonexpansive and the set of fixed points of  $J_{c\mathsf{T}}$  coincides with the set of all zeros of  $\mathsf{T}$ . More information on firmly nonexpansive mappings can be found in [50, 24].

For  $\alpha \in (0, 1]$ , a function  $U : H \rightarrow H$  is called  $\alpha$ -*averaged*<sup>3</sup> if there exists a nonexpansive operator  $U' : H \rightarrow H$  such that  $U = (1 - \alpha)\text{Id} + \alpha U'$ . The  $\alpha$ -averaged operators are always nonexpansive. Moreover, the  $\frac{1}{2}$ -averaged operators coincide with the firmly nonexpansive operators.

In the following corollaries we assume that we have functions  $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$  satisfying conditions  $(Q_1) - (Q_4)$ .

<sup>3</sup>The standard definition [3] asks for  $\alpha \in (0, 1)$ . With this extension, 1-averaged is just another way of saying nonexpansive.

**Corollary 7.2.** Let  $\alpha \in (0, 1]$  and  $U : H \rightarrow H$  be  $\alpha$ -averaged. Given  $(\alpha_n) \subset [0, 1]$ ,  $(\beta_n) \subset [1 - \frac{1}{\alpha}, 1]$ , and  $x_0, u \in H$ , consider  $(x_n)$  generated by

$$\begin{cases} x_{2n+1} &= (1 - \alpha_n)x_{2n} + \alpha_n u \\ x_{2n+2} &= (1 - \beta_n)U(x_{2n+1}) + \beta_n x_{2n+1}. \end{cases} \quad (\text{HIdM})$$

Let  $\sigma \in (0, 1)$  be such that  $\alpha \geq \sigma$ . Assume that there exist  $N \in \mathbb{N} \setminus \{0\}$  such that  $N \geq \max\{\|x_0 - p\|, 2\|u - p\|\}$ , for some  $p \in \text{Fix } U$  and  $\gamma \in (0, \frac{1}{2\sigma}]$  satisfying  $1 - \frac{1}{\alpha} + \gamma \leq \beta_n \leq 1 - \gamma$ , for all  $n \in \mathbb{N}$ .

Then  $(x_n)$  converges strongly to  $P_{\text{Fix}(U)}(u)$  and

$$\mu_1[N, \sigma, \gamma, \Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4] := \mu[N, \sigma\gamma, \Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4], \text{ with } \mu \text{ as in Theorem 5.1,}$$

is a rate of metastability for  $(x_n)$ .

*Proof.* Since  $U$  is  $\alpha$ -averaged, there exists  $U'$  nonexpansive and such that  $U = (1 - \alpha)\text{Id} + \alpha U'$ . For all  $n \in \mathbb{N}$

$$\begin{aligned} x_{2n+2} &= (1 - \beta_n)U(x_{2n+1}) + \beta_n x_{2n+1} \\ &= (1 - \beta_n)[(1 - \alpha)\text{Id} + \alpha U'](x_{2n+1}) + \beta_n x_{2n+1} \\ &= (1 - \tilde{\beta}_n)U'(x_{2n+1}) + \tilde{\beta}_n x_{2n+1}, \end{aligned}$$

with  $\tilde{\beta}_n := 1 - \alpha + \alpha\beta_n$ . Hence  $(x_n)$  is generated by (HM), using the sequences  $(\alpha_n), (\tilde{\beta}_n) \subset [0, 1]$ ,  $T = \text{Id}$  and  $U = U'$ . It is easy to see that  $\sigma\gamma$  satisfies condition (Q<sub>5</sub>) for the sequence  $(\tilde{\beta}_n)$ . Note that  $\text{Fix } U = \text{Fix } U' = F$ . Since  $\alpha \in (0, 1]$ , for all  $\varepsilon > 0$  and  $n \in \mathbb{N}$  we have

$$\sum_{i=\Gamma_4(\varepsilon)+1}^{\Gamma_4(\varepsilon)+n} |\tilde{\beta}_{i+1} - \tilde{\beta}_i| \leq \sum_{i=\Gamma_4(\varepsilon)+1}^{\Gamma_4(\varepsilon)+n} |\beta_{i+1} - \beta_i| \leq \varepsilon,$$

and so condition (Q<sub>4</sub>) still holds with  $\Gamma_4$  for the sequence  $(\tilde{\beta}_n)$ . Hence, by Theorem 1.1 the sequence  $(x_n)$  converges strongly to  $P_{\text{Fix}(U)}(u)$ . The rate of metastability follows from an application of Theorem 5.1 with  $\gamma := \sigma\gamma$ .  $\square$

For  $\delta > 0$ , a function  $U : H \rightarrow H$  is said to be  $\delta$ -cocoercive if for all  $x, y \in H$ ,

$$\langle x - y, U(x) - U(y) \rangle \geq \delta \|U(x) - U(y)\|^2,$$

which is equivalent to say that  $\delta U$  is firmly nonexpansive. In the following result we give a strongly convergent version of the forward-backward algorithm which extends [7, Theorem 7] (as well as the quantitative analysis from [19, Corollary 2]). The algorithm (GFB) below can be seen as a generalization in the sense that for  $\alpha_n = \beta_n \equiv 0$  the sequence of even terms  $(x_{2n})$  is indeed the original weakly convergent forward-backward algorithm by Lions and Mercier [45].

**Corollary 7.3.** Let  $U_1 : H \rightrightarrows H$  be maximal monotone and  $U_2 : H \rightarrow H$  be  $\delta$ -cocoercive, for some  $\delta > 0$ . Let  $c \in (0, 2\delta]$ . Given  $(\alpha_n) \subset [0, 1]$ ,  $(\beta_n) \subset [1 - \frac{4\delta-c}{2\delta}, 1]$ , and  $x_0, u \in H$ , consider  $(x_n)$  generated by

$$\begin{cases} x_{2n+1} &= (1 - \alpha_n)x_{2n} + \alpha_n u \\ x_{2n+2} &= (1 - \beta_n)J_{cU_1}(x_{2n+1} - cU_2(x_{2n+1})) + \beta_n x_{2n+1}. \end{cases} \quad (\text{GFB})$$

Assume that there exist  $N \in \mathbb{N} \setminus \{0\}$  such that  $N \geq \max\{\|x_0 - p\|, 2\|u - p\|\}$ , for some  $p \in \text{zer}(U_1 + U_2)$  and  $\gamma \in (0, 1]$  satisfying  $1 - \frac{4\delta-c}{2\delta} + \gamma \leq \beta_n \leq 1 - \gamma$ , for all  $n \in \mathbb{N}$ . Then  $(x_n)$  converges strongly to  $P_{\text{zer}(U_1+U_2)}(u)$  and

$$\mu_2 := \mu_2[N, \gamma, \Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4] := \mu[N, \gamma/2, \Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4], \text{ with } \mu \text{ as in Theorem 5.1,}$$

is a rate of metastability for  $(x_n)$ .

*Proof.* It is straightforward to see that the iteration  $(x_n)$  is generated by (HIdM) with  $U := J_{cU_1} \circ (\text{Id} - cU_2)$ . The resolvent function  $J_{cU_1}$  is firmly nonexpansive, i.e.  $\frac{1}{2}$ -averaged. Furthermore, we have  $\frac{c}{2\delta} \in (0, 1]$  and so it follows from [4, Proposition 4.39] that  $(\text{Id} - cU_2)$  is  $\frac{c}{2\delta}$ -averaged. If  $c < 2\delta$ , then we use [48, Theorem 3(b)] to conclude that  $U$  is  $\frac{2\delta}{4\delta-c}$ -averaged. If  $c = 2\delta$ , the fact that  $U_2$  is  $\delta$ -cocoercive entails that  $\text{Id} - cU_2$  is nonexpansive. Hence  $U$  is nonexpansive, or equivalently (following the extended notion of 1-averaged),  $U$  is  $\frac{2\delta}{4\delta-c}$ -averaged since  $\frac{2\delta}{4\delta-c} = 1$ . Since  $\text{Fix } U = \text{zer}(U_1 + U_2)$  [4, Proposition 26.1(iv)(a)], we have  $N \geq \max\{\|x_0 - p\|, \|u - p\|\}$ , for some  $p \in \text{Fix } U$ . From the fact that  $\frac{2\delta}{4\delta-c} \geq \frac{1}{2}$ , we may apply Corollary 7.2 with  $\sigma := \frac{1}{2}$  to conclude the result.  $\square$

**Remark 7.4.** The fact that no quantitative information regarding ‘ $c > 0$ ’ is present in the rate of metastability reflects the fact that this condition is only needed to make sense of the definition of the resolvent function. In fact, if in Corollary 7.3 the resolvent function  $J_{cU_1}$  is replaced by an arbitrary firmly nonexpansive mapping  $J$ , and  $p$  is some point in  $\text{Fix}(J \circ (\text{Id} - cU_2))$ , then the result holds also for  $c = 0$  (immediately by Corollary 7.2 with  $\alpha = 1/2$ ). In such case, the sequence  $(\beta_n)$  is allowed to vary in the interval  $[-1, 1]$ .

In the following result we give a strongly convergent version of the Douglas–Rachford algorithm which extends [7, Theorem 10] (as well as the quantitative analysis from [19, Corollary 3]). The algorithm (GDR) below can be seen as a generalization in the sense that for  $\alpha_n \equiv 0$  and  $\beta_n \equiv -1$  the sequence of even terms  $(x_{2n})$  is indeed the original weakly convergent Douglas–Rachford algorithm [21].

**Corollary 7.5.** Let  $U_1, U_2 : H \rightrightarrows H$  be two maximal monotone operators and  $c > 0$ . Given  $(\alpha_n) \subset [0, 1]$ ,  $(\beta_n) \subset [-1, 1]$ , and  $x_0, u \in H$ , consider  $(x_n)$  generated by

$$\begin{cases} x_{2n+1} &= (1 - \alpha_n)x_{2n} + \alpha_n u \\ y_n &= J_{cU_2}(x_{2n+1}) \\ z_n &= J_{cU_1}(2y_n - x_{2n+1}) \\ x_{2n+2} &= x_{2n+1} + (1 - \beta_n)(z_n - y_n). \end{cases} \quad (\text{GDR})$$

Assume that there exist  $N \in \mathbb{N} \setminus \{0\}$  such that  $N \geq \max\{\|x_0 - p\|, 2\|u - p\|\}$  for some  $p \in \text{Fix}(R_{cU_1} \circ R_{cU_2})$ , and  $\gamma \in (0, 1]$  satisfying  $-1 + \gamma \leq \beta_n \leq 1 - \gamma$ , for all  $n \in \mathbb{N}$ . Then,

(i)  $(x_n)$  converges strongly to  $\bar{x} = P_{\text{Fix}(R_{cU_1} \circ R_{cU_2})}(u)$  and

$$\mu_3 := \mu_3[N, \gamma, \Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4] := \mu[N, \gamma/2, \Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4], \quad \text{with } \mu \text{ as in Theorem 5.1,}$$

is a rate of metastability for  $(x_n)$ ;

(ii)  $(y_n)$  and  $(z_n)$  converge strongly to  $J_{cU_2}(\bar{x}) \in \text{zer}(U_1 + U_2)$  and

$$\mu_4(\varepsilon, f) := \mu_3(\varepsilon, 2f + 1), \quad \mu_5(\varepsilon, f) := \mu_3(\varepsilon/3, 2f + 1),$$

are rates of metastability for  $(y_n)$ ,  $(z_n)$ , respectively.

*Proof.* It is easy to see that the iteration  $(x_n)$  is generated by (H<sub>Id</sub>M) using the sequences  $(\alpha_n), (\tilde{\beta}_n) \subset [0, 1]$ , where  $\tilde{\beta}_n := \frac{1+\beta_n}{2}$  for all  $n \in \mathbb{N}$ , and with the nonexpansive map  $U := R_{cU_1} \circ R_{cU_2}$ . It is clear that we are still in the conditions of Theorem 1.1, and so we conclude that  $(x_n)$  converges strongly to  $\bar{x} = P_{\text{Fix}(R_{cU_1} \circ R_{cU_2})}(u)$ .

Since the condition (Q<sub>4</sub>) still holds for  $(\tilde{\beta}_n)$  with the function  $\Gamma_4$ , and we have for all  $n \in \mathbb{N}$ ,  $\frac{\gamma}{2} \leq \tilde{\beta}_n \leq 1 - \frac{\gamma}{2}$ , we can apply Theorem 5.1 to obtain the desired rate of metastability. This concludes part (i).

We now prove part (ii). From part (i), the definition of  $(y_n)$  and the continuity of  $J_{cU_2}$ , we conclude that  $(y_n)$  converges strongly to  $J_{cU_2}(\bar{x})$ . By [4, Proposition 26.1(iii)(b)], we have  $\text{zer}(U_1 + U_2) = J_{cU_2}[\text{Fix}(U)]$ , hence  $J_{cU_2}(\bar{x}) \in \text{zer}(U_1 + U_2)$ . From the recursive definition of  $x_{2n+2}$ , we see that

$$\|z_n - y_n\| = \frac{1}{1 - \beta_n} \|x_{2n+2} - x_{2n+1}\| \leq \frac{1}{\gamma} \|x_{2n+2} - x_{2n+1}\| \rightarrow 0,$$

which entails that  $(z_n)$  also converges strongly to  $J_{cU_2}(\bar{x})$ . We now argue the rates of metastability for  $(y_n)$  and  $(z_n)$ . For all  $i, j \in \mathbb{N}$ , we have

$$\|y_i - y_j\| = \|J_{cU_2}(x_{2i+1}) - J_{cU_2}(x_{2j+1})\| \leq \|x_{2i+1} - x_{2j+1}\|$$

and

$$\begin{aligned} \|z_i - z_j\| &= \|J_{cU_1}(2y_i - x_{2i+1}) - J_{cU_1}(2y_j - x_{2j+1})\| \\ &\leq \|2y_i - x_{2i+1} - (2y_j - x_{2j+1})\| \\ &\leq 2\|y_i - y_j\| + \|x_{2i+1} - x_{2j+1}\| \\ &\leq 3\|x_{2i+1} - x_{2j+1}\|. \end{aligned}$$

The result now follows from part (i) and the fact that if  $i \in [n, f(n)]$ , then  $2i + 1 \in [n, 2f(n) + 1]$ .  $\square$

The algorithms considered by Boţ, Csetnek and Meier [7] are the particular case of the algorithms considered in this section when  $u = 0$ .

## 8 Final remarks

We consider in this paper an algorithm which mixes Halpern and Krasnoselskii-Mann iterative definitions in an alternating way. Under appropriate conditions we show asymptotic regularity for any sequence  $(x_n)$  generated by this algorithm, and construct effective rates of asymptotic regularity. We prove that  $(x_n)$  is a Cauchy sequence and obtain a rate of metastability in the sense of Tao [58, 59]. Using this result, we show that our iterative method strongly approximates a common fixed point of two nonexpansive mappings. This strong convergence result allowed us to define strongly convergent versions of the forward-backward and the Douglas-Rachford algorithms. Our results generalize recent work by Boţ, Csetnek and Meier [7], and Cheval and Leuştean [14].

In fact, the proof mining program enables, in certain instances, to obtain generalized versions of mathematical results. We briefly comment on three recent examples. The first example concerns the analysis, carried out in [38], of a proof of the “Lion-Man” game whose convergence crucially uses the compactness of the metric space together with its betweenness property. Proof mining allowed to weaken the compactness assumption to a boundedness condition, if *betweenness* is upgraded to *uniform betweenness* (which coincides with betweenness in the compact case anyway), resulting in a striking generalization since now the convergence holds in all bounded subsets of uniformly convex Banach spaces, CAT( $\kappa$ ) spaces ( $\kappa > 0$ , where one anyhow always has a boundedness assumption), etc. The second example comes from [39] where the authors analysed a proof that the strong convergence of the so-called viscosity generalizations of Browder- and Halpern-type algorithms can be reduced to the convergence of the original Browder/Halpern algorithms in the setting of Banach spaces. The quantitative analysis allowed the authors to generalize this reduction to the non-linear setting of hyperbolic spaces, which in turn allowed for several applications of previously known rates of metastability for these algorithms. Our final example can be found in [56] where the author adapts the analysis of the strong convergence of the Halpern type Proximal Point Algorithm, given in [35], from the setting of Banach spaces to CAT(0) spaces. This gave rise to new qualitative convergence results. The key observation is that instead of *strong nonexpansivity* and corresponding (SNE)-moduli, one can also work with *strong quasi-nonexpansivity* and corresponding (SQNE)-moduli (introduced in [33]). More information on SQNE operators can be found in [12, Section 2.1].

Finally, let us explain how the proof of Theorem 1.1 was obtained. A proof of strong convergence can be carried out with the usual arguments in the setting of Hilbert spaces, as explained below. However, it is not clear if such arguments can be generalized to a non-linear setting. Nevertheless, a passage through quantitative results made it possible to overcome this problem.

Let us elaborate on this matter. A proof in Hilbert spaces is guided by the following steps:

- (1)  $(x_n)$  is bounded: Follow the arguments of Lemma 3.2.
- (2) Asymptotic regularity of  $(x_n)$ : Follow the arguments of Lemma 3.3 and Proposition 3.6.
- (3) Projection argument: With  $\tilde{x}$  the projection point onto  $F$  of  $u$ , we have  $\forall y \in F$  ( $\langle u - \tilde{x}, y - \tilde{x} \rangle \leq 0$ ).
- (4) Sequential weak compactness and demiclosedness: Pick a subsequence  $(x_{n_j})$  of  $(x_n)$  such that

$$\limsup \langle u - \tilde{x}, x_{2n} - \tilde{x} \rangle = \lim_{j \rightarrow \infty} \langle u - \tilde{x}, x_{n_j} - \tilde{x} \rangle,$$

and simultaneously  $(x_{n_j})$  converges weakly to some  $y \in F$ . Here we are using (twice) the following demiclosedness principle.

**Lemma 8.1** (Demiclosedness principle [9]). *Let  $C$  be a closed convex subset of  $H$  and let  $f : C \rightarrow C$  be a nonexpansive mapping such that  $\text{Fix}(f) \neq \emptyset$ . Assume that  $(x_n)$  is a sequence in  $C$  such that  $(x_n)$  weakly converges to  $x \in C$  and  $((\text{Id}_C - f)(x_n))$  converges strongly to  $y \in H$ . Then  $(\text{Id}_C - f)(x) = y$ .*

By step (3) it follows that  $\limsup \langle u - \tilde{x}, x_n - \tilde{x} \rangle \leq 0$ .

- (5) Main combinatorial part: Following the inequalities culminating in (22) (in the proof of Theorem 1.1), we are then able to apply Lemma 2.7 and conclude that  $x_n \rightarrow \tilde{x}$ .

As shown in Section 3, steps (1) and (2) are valid in CAT(0) spaces (actually, even in general UCW spaces [44]). It is however not clear how the sequential weak compactness argument could be carried out in this more general context.

On the other hand, the quantitative analysis of this proof in Hilbert spaces is known to be possible. The only roadblocks to the analysis are the projection and the sequential weak compactness arguments. Yet, recent developments in proof mining provide a method to eliminate such arguments. The solution is twofold. For eliminating the full logical strength of the projection argument one relies on the crucial observation by Kohlenbach [32] that a weaker  $\varepsilon$ -version already suffices to obtain quantitative results. This holds true in both Hilbert and CAT(0) spaces. As for sequential weak compactness one makes use of the general principle developed in [22]. This macro was developed in the general setting of metric spaces aiming to bypass sequential weak



compactness arguments in Hilbert spaces. The fundamental observation that the macro can be applied in the setting of CAT(0) spaces (no longer with concerns about weak compactness) rests on establishing some version of [22, Proposition 4.3]. This can be seen to correspond to our treatment of the metric projection via the quasi-linearization function characterization carried out in Section 4 and culminating in Proposition 4.5 (which is the conclusion of [22, Proposition 4.3], adapted to two maps, with  $\varphi(x, y) := \langle \bar{u}\bar{x}, \bar{u}\bar{y} \rangle$ ). The argument in Section 6 brings us back to a “qualitative” statement, giving rise to our proof of Theorem 1.1.

Finally, we would like to mention some very recently obtained results on asymptotic regularity for particular choices of parameters. For the particular case of the (CL) iteration (and also the modified Halpern iteration), linear rates of asymptotic regularity were obtained in [13]. The asymptotic regularity of the general iteration (HM) was studied in the context of UCW spaces in [44]. For a particular choice of parameters it was shown to admit quadratic rates in CAT(0) spaces.

## Acknowledgements

The first author acknowledges the support of FCT - Fundação para a Ciência e Tecnologia under the projects: UIDP/04561/2020 and UIDP/04674/2020, and the research centers CMAFcIO – Centro de Matemática, Aplicações Fundamentais e Investigação Operacional and CIMA – Centro de Investigação em Matemática e Aplicações. The second author was supported by the German Science Foundation (DFG Project KO 1737/6-2). This work benefited from discussions with Ulrich Kohlenbach.

## References

- [1] A. D. Alexandrov. A theorem on triangles in a metric space and some of its applications. *Trudy Mat. Inst. Steklov.*, 38:5–23, 1951.
- [2] J. Avigad, P. Gerhardy, and H. Towsner. Local stability of ergodic averages. *Trans. Amer. Math. Soc.*, 362(1):261–288, 2010.
- [3] J. B. Baillon, R. E. Bruck, and S. Reich. On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces. *Houston J. Math.*, 4(1):1–9, 1978.
- [4] H. Bauschke and P. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, Cham, second edition, 2017.
- [5] H. Bauschke, E. Matoušková, and S. Reich. Projection and proximal point methods: convergence results and counterexamples. *Nonlinear Anal.*, 56(5):715–738, 2004.
- [6] I. D. Berg and I. Nikolaev. Quasilinearization and curvature of Aleksandrov spaces. *Geom. Dedicata*, 133:195–218, 2008.
- [7] R. I. Boţ, E. R. Csetnek, and D. Meier. Inducing strong convergence into the asymptotic behaviour of proximal splitting algorithms in Hilbert spaces. *Optim. Methods Softw.*, 34(3):489–514, 2019.
- [8] M. R. Bridson and A. Haefliger. *Metric spaces of non-positive curvature*, volume 319. Springer Science & Business Media, 2013.
- [9] F. E. Browder. Nonexpansive nonlinear operators in a Banach space. *Proc. Nat. Acad. Sci. U.S.A.*, 54:1041–1044, 1965.
- [10] F. E. Browder and W. V. Petryshyn. The solution by iteration of nonlinear functional equations in Banach spaces. *Bull. Amer. Math. Soc.*, 72:571–575, 1966.
- [11] F. Bruhat and J. Tits. Groupes réductifs sur un corps local. *Publications Mathématiques de l’Institut des Hautes Études Scientifiques*, 41(1):5–251, 1972.
- [12] A. Cegielski, S. Reich, and R. Zalas. Regular sequences of quasi-nonexpansive operators and their applications. *SIAM J. Optim.*, 28(2):1508–1532, 2018.
- [13] H. Cheval, U. Kohlenbach, and L. Leuştean. On modified Halpern and Tikhonov-Mann iterations. Final version to appear in: *Journal of Optimization Theory and Applications*. arXiv:2203.11003, 2023.
- [14] H. Cheval and L. Leuştean. Quadratic rates of asymptotic regularity for the Tikhonov–Mann iteration. *Optimization Methods and Software*, pages 1–16, 2022.

- [15] A. Cuntavepanit and B. Panyanak. Strong convergence of modified Halpern iterations in  $CAT(0)$  spaces. *Fixed Point Theory Appl.*, pages Art. ID 869458, 11, 2011.
- [16] H. Dehghan and J. Rooin. A Characterization of Metric Projection in  $CAT(0)$  Spaces. In *International Conference on Functional Equation*, pages 41–43, 2012. arXiv:1311.4174.
- [17] S. Dhompangsa and B. Panyanak. On  $\Delta$ -convergence theorems in  $CAT(0)$  spaces. *Computers & Mathematics with Applications*, 56(10):2572–2579, 2008.
- [18] B. Dinis and P. Pinto. Effective metastability for a method of alternating resolvents. Final version to appear in: *Fixed Point Theory*. arXiv:2101.12675.
- [19] B. Dinis and P. Pinto. On the convergence of algorithms with Tikhonov regularization terms. *Optimization Letters*, 15(4):1263–1276, 2021.
- [20] B. Dinis and P. Pinto. Quantitative results on the multi-parameters proximal point algorithm. *Journal of Convex Analysis*, 28(3), 2021.
- [21] J. Douglas, Jr. and H. H. Rachford, Jr. On the numerical solution of heat conduction problems in two and three space variables. *Trans. Amer. Math. Soc.*, 82:421–439, 1956.
- [22] F. Ferreira, L. Leuştean, and P. Pinto. On the removal of weak compactness arguments in proof mining. *Advances in Mathematics*, 354:106728, 2019.
- [23] K. Goebel and W. A. Kirk. Iteration processes for nonexpansive mappings. *Contemp. Math*, 21:115–123, 1983.
- [24] K. Goebel and S. Reich. Uniform convexity, hyperbolic geometry, and nonexpansive mappings. In *Series of Monographs and Textbooks in Pure and Applied Mathematics 83*. Dekker, 1984.
- [25] B. Green and T. Tao. The primes contain arbitrarily long arithmetic progressions. *Ann. of Math. (2)*, 167(2):481–547, 2008.
- [26] M. Gromov. Hyperbolic groups. In *Essays in group theory*, pages 75–263. Springer, 1987.
- [27] B. Halpern. Fixed points of nonexpanding maps. *Bull. Amer. Math. Soc.*, 73:957–961, 1967.
- [28] T.-H. Kim and H.-K. Xu. Strong convergence of modified Mann iterations. *Nonlinear Anal.*, 61(1-2):51–60, 2005.
- [29] U. Kohlenbach. Some computational aspects of metric fixed-point theory. *Nonlinear Anal.*, 61(5):823–837, 2005.
- [30] U. Kohlenbach. Some logical metatheorems with applications in functional analysis. *Transactions of the American Mathematical Society*, 357(1):89–128, 2005.
- [31] U. Kohlenbach. *Applied proof theory: proof interpretations and their use in mathematics*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2008.
- [32] U. Kohlenbach. On quantitative versions of theorems due to F.E. Browder and R. Wittmann. *Advances in Mathematics*, 226(3):2764–2795, 2011.
- [33] U. Kohlenbach. On the quantitative asymptotic behavior of strongly nonexpansive mappings in Banach and geodesic spaces. *Israel J. Math.*, 216(1):215–246, 2016.
- [34] U. Kohlenbach. Proof-theoretic methods in nonlinear analysis. In M. V. B. Sirakov, P. Ney de Souza, editor, *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. II. Invited lectures*, pages 61–82. World Sci. Publ., Hackensack, NJ, 2019.
- [35] U. Kohlenbach. Quantitative analysis of a Halpern-type proximal point algorithm for accretive operators in Banach spaces. *J. Nonlinear Convex Anal.*, 21(9):2125–2138, 2020.
- [36] U. Kohlenbach and L. Leuştean. A quantitative mean ergodic theorem for uniformly convex Banach spaces. *Ergodic Theory Dynam. Systems*, 29(6):1907–1915, 2009.
- [37] U. Kohlenbach and L. Leuştean. Effective metastability of Halpern iterates in  $CAT(0)$  spaces. *Advances in Mathematics*, 231(5):2526–2556, 2012.



- [38] U. Kohlenbach, G. López-Acedo, and A. Nicolae. A uniform betweenness property in metric spaces and its role in the quantitative analysis of the “Lion-Man” game. *Pacific J. Math.*, 310(1):181–212, 2021.
- [39] U. Kohlenbach and P. Pinto. Quantitative translations for viscosity approximation methods in hyperbolic spaces. *J. Math. Anal. Appl.*, 507(2):125823, 2022.
- [40] D. Körnlein. Quantitative results for Halpern iterations of nonexpansive mappings. *Journal of Mathematical Analysis and Applications*, 428(2):1161–1172, 2015.
- [41] M. A. Krasnoselskii. Two remarks on the method of successive approximations. *Uspehi Mat. Nauk (N.S.)*, 10(1(63)):123–127, 1955.
- [42] L. Leuştean. A quadratic rate of asymptotic regularity for CAT(0)-spaces. *Journal of Mathematical Analysis and Applications*, 325(1):386–399, 2007.
- [43] L. Leuştean and P. Pinto. Quantitative results on a Halpern-type proximal point algorithm. *Computational Optimization and Applications*, 79(1):101–125, 2021.
- [44] L. Leustean and P. Pinto. Rates of asymptotic regularity for the alternating Halpern-Mann iteration. arXiv:2206.02226.
- [45] P.-L. Lions and B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.*, 16(6):964–979, 1979.
- [46] W. R. Mann. Mean value methods in iteration. *Proc. Amer. Math. Soc.*, 4:506–510, 1953.
- [47] E. Neumann. Computational problems in metric fixed point theory and their Weihrauch degrees. *Log. Methods Comput. Sci.*, 11(4):4:20, 44, 2015.
- [48] N. Ogura and I. Yamada. Non-strictly convex minimization over the fixed point set of an asymptotically shrinking nonexpansive mapping. *Numer. Funct. Anal. Optim.*, 23(1-2):113–137, 2002.
- [49] P. Pinto. A rate of metastability for the Halpern type Proximal Point Algorithm. *Numerical Functional Analysis and Optimization*, 42(3):320–343, 2021.
- [50] S. Reich. Extension problems for accretive sets in Banach spaces. *J. Functional Analysis*, 26(4):378–395, 1977.
- [51] S. Reich. Weak convergence theorems for nonexpansive mappings in Banach spaces. *J. Math. Anal. Appl.*, 67(2):274–276, 1979.
- [52] S. Reich. Strong convergence theorems for resolvents of accretive operators in Banach spaces. *J. Math. Anal. Appl.*, 75(1):287–292, 1980.
- [53] S. Reich and I. Shafrir. Nonexpansive iterations in hyperbolic spaces. *Nonlinear analysis: theory, methods & applications*, 15(6):537–558, 1990.
- [54] S. Saejung. Halpern’s iteration in CAT(0) spaces. *Fixed Point Theory Appl.*, pages Art. ID 471781, 13, 2010.
- [55] K. Schade and U. Kohlenbach. Effective metastability for modified Halpern iterations in CAT(0) spaces. *Fixed Point Theory Appl.*, pages 2012:191, 19, 2012.
- [56] A. Sipoş. Abstract strongly convergent variants of the proximal point algorithm. *Comput. Optim. Appl.*, 83(1):349–380, 2022.
- [57] W. Takahashi. A convexity in metric space and nonexpansive mappings, I. In *Kodai Mathematical Seminar Reports*, volume 22, pages 142–149. Department of Mathematics, Tokyo Institute of Technology, 1970.
- [58] T. Tao. Norm convergence of multiple ergodic averages for commuting transformations. *Ergodic Theory Dynam. Systems*, 28(2):657–688, 2008.
- [59] T. Tao. *Structure and randomness*. American Mathematical Society, Providence, RI, 2008. Pages from year one of a mathematical blog.
- [60] R. Wittmann. Approximation of fixed points of nonexpansive mappings. *Arch. Math. (Basel)*, 58(5):486–491, 1992.
- [61] H.-K. Xu. Iterative algorithms for nonlinear operators. *J. London Math. Soc. (2)*, 66(1):240–256, 2002.

- [62] H.-K. Xu. Viscosity approximation methods for nonexpansive mappings. *J. Math. Anal. Appl.*, 298(1):279–291, 2004.
- [63] Y. Yao, H. Zhou, and Y.-C. Liou. Strong convergence of a modified Krasnoselski-Mann iterative algorithm for non-expansive mappings. *J. Appl. Math. Comput.*, 29(1-2):383–389, 2009.