Convergence of proximal splitting algorithms in $\mathrm{CAT}(\kappa)$ spaces and beyond

Florian Lauster[1] and D. Russell Luke[1*]

*Correspondence:
r.luke@math.uni-goettingen.de
[1] Institute for Numerical and Applied
Mathematics, University of
Goettingen, 37083, Goettingen,
Germany

Abstract

In the setting of CAT(κ) spaces, common fixed point iterations built from prox mappings (e.g. prox-prox, Krasnoselsky–Mann relaxations, nonlinear projected-gradients) converge locally linearly under the assumption of linear metric subregularity. Linear metric subregularity is in any case necessary for linearly convergent fixed point sequences, so the result is tight. To show this, we develop a theory of fixed point mappings that violate the usual assumptions of nonexpansiveness and firm nonexpansiveness in p-uniformly convex spaces.

MSC: Primary 47H09; 47H10; 53C22; 53C21; secondary 53C23; 53C20; 49M27

Keywords: Averaged mappings; p-uniformly convex; CAT(k) space; Nonexpansive mappings; Firmly nonexpansive; Fixed point iteration; Proximal point algorithm

1 Fundamentals of nonlinear spaces

Following [1] we focus on *p-uniformly convex spaces with parameter c* [12]: for $p \in (1, \infty)$, a metric space (G, d) is p-uniformly convex with constant $c > 0$ whenever it is a geodesic space, and

$$\left(\forall t \in [0, 1]\right)\left(\forall x, y, z \in G\right)$$

$$d\left(z, (1-t)x \oplus ty\right)^p \leq (1-t)d(z, x)^p + td(z, y)^p - \frac{c}{2}t(1-t)d(x, y)^p. \tag{1}$$

Examples of p-uniformly convex spaces are L^p spaces, CAT(0) spaces ($p = c = 2$), Hadamard spaces (complete CAT(0) spaces), Hilbert spaces (linear Hadamard spaces). Of particular interest are CAT(κ) spaces since these serve as the model space for applications on manifolds with curvature bounded above.

Lemma 1 ([13], Proposition 3.1)　*A* CAT(κ) *space is locally 2-uniformly convex with parameter $c \nearrow 2$ as the diameter of the local neighborhood vanishes. In particular, for any* CAT(κ) *space (G, d) and any point $\overline{x} \in G$, for all $\delta \in (0, \pi/(4\sqrt{\kappa}))$, the subspace $(\mathbb{B}_\delta(\overline{x}), d|_{\mathbb{B}_\delta(\overline{x})})$ is a 2-uniformly convex space with constant $c_\delta = 4\delta\sqrt{\kappa}\tan(\pi/2 - 2\delta\sqrt{\kappa})$.*

Note the asymptotic behavior of the constants: as $\delta \searrow 0$, the constant $c \nearrow 2$.

Definition 2 Let (G, d) be a geodesic space and γ and η be two geodesics through p. Then γ is said to be perpendicular to η at point p denoted by $\gamma \perp_p \eta$ if

$$d(x, p) \leq d(x, y) \quad \forall x \in \gamma, y \in \eta.$$

A space is said to be symmetric perpendicular if for all geodesics γ and η with common point p we have

$$\gamma \perp_p \eta \quad \Leftrightarrow \quad \eta \perp_p \gamma.$$

Remark 3 Any CAT(κ) space (G, d) with $\text{diam}(G) < \frac{\pi}{2\sqrt{\kappa}}$ is symmetric perpendicular [9, Theorem 2.11].

In a complete p-uniformly convex space the p-proximal mapping of a proper and lower semicontinuous function f is defined by

$$\text{prox}^p_{f, \lambda}(x) := \underset{y \in G}{\text{argmin}} f(y) + \frac{1}{p\lambda^{p-1}} d(y, x)^p \quad (\lambda > 0). \tag{2}$$

The main dividend of this work is the following.

Theorem 4 (Convergence of proximal algorithms in CAT(κ) spaces) *Let (G, d) be a complete* CAT(κ) *space with $\kappa > 0$, and for $j = 1, 2, \ldots, N$, let f_j and $g : G \to \mathbb{R} \cap \{+\infty\}$ be proper, convex, and lower semicontinuous with $\text{argmin} \, g \cap (\bigcap_{j=1}^N \text{argmin} f_j) \neq \emptyset$. Let $T : D \to D$ where $D \subset G$ denotes one of the following:*
- *(i) $T := \text{prox}_{f_N, \lambda_N} \circ \text{prox}_{f_{N-1}, \lambda_{N-1}} \circ \cdots \circ \text{prox}_{f_1, \lambda_1}$;*
- *(ii) $T := \beta \, \text{prox}_{g, \lambda} \oplus (1 - \beta) \, \text{Id}$;*
- *(iii) $T := \text{prox}_{f_1, \lambda_1} \circ (\beta \, \text{prox}_{g, \lambda_2} \oplus (1 - \beta) \, \text{Id})$;*
- *(iv) $T := P_C \circ (\beta \, \text{prox}_{g, \lambda_1} \oplus (1 - \beta) \, \text{Id})$,*

where in the last case P_C is the metric projector onto the closed convex set $C \subset G$ (in other words, this is the specialization of part (iii) to the case where f_1 is the indicator function of a closed convex subset of G). If T satisfies Fix $T \neq \emptyset$ *and*

$$d(x, \text{Fix} \, T \cap D) \leq \mu d(x, Tx), \quad \forall x \in D \subset G,$$

with constant μ, then the fixed point sequence initialized from any starting point close enough to Fix T *is at least linearly convergent to a point in* Fix T.

A more precise statement of this theorem, with proof, is Theorem 25. The intervening sections prove the fundamental building blocks.

2 Almost α-firmly nonexpansive mappings

The regularity of a mapping $T : G \to G$ is characterized by the behavior of the images of pairs of points under T. A key tool is what has been called the *transport discrepancy* in [4]:

$$\psi_T^{(p,c)}(x, y) := \frac{c}{2} \left(d(Tx, x)^p + d(Ty, y)^p + d(Tx, Ty)^p + d(x, y)^p - d(Tx, y)^p - d(x, Ty)^p \right). \tag{3}$$

Definition 5 Let (G, d) be a p-uniformly convex metric space with constant c.

(i) The mapping $T : G \to G$ is pointwise almost nonexpansive at $y \in D \subset G$ on D with violation $\epsilon \geq 0$ whenever

$$\exists \epsilon \geq 0 : \quad d(Tx, Ty)^p \leq (1 + \epsilon) d(x, y)^p \quad \forall x \in D. \tag{4}$$

The smallest ϵ for which (4) holds is called the *violation*. If (4) holds with $\epsilon = 0$, then T is pointwise nonexpansive at $y \in D \subset G$ on D. If (4) holds at all $y \in D$, then T is said to be (almost) nonexpansive on D. If $D = G$, then the mapping T is simply said to be (almost) nonexpansive. If $D \supset \operatorname{Fix} T \neq \emptyset$ and (4) holds at all $y \in \operatorname{Fix} T$ with the same violation, then T is said to be almost quasi nonexpansive.

(ii) The mapping T is said to be *pointwise asymptotically nonexpansive at y* whenever

$$\forall \epsilon > 0, \exists D_\epsilon(y) \subset G : \quad d(Tx, Ty)^p \leq (1 + \epsilon) d(x, y)^p \quad \forall x \in D_\epsilon(y), \tag{5}$$

where $D_\epsilon(y)$ is a neighborhood of y in D.

(iii) $T : G \to G$ is said to be *quasi strictly nonexpansive* whenever

$$d(Tx, \bar{x}) < d(x, \bar{x}) \quad \forall x \in G \setminus \operatorname{Fix} T, \forall \bar{x} \in \operatorname{Fix} T. \tag{6}$$

(iv) The operator $T : G \to G$ is *pointwise almost α-firmly nonexpansive at $y \in D \subset G$ on D with violation at most $\epsilon > 0$* whenever

$$\exists \alpha \in (0, 1), \epsilon \geq 0 : \quad d(Tx, Ty)^p \leq (1 + \epsilon) d(x, y)^p - \frac{1 - \alpha}{\alpha} \psi_T^{(p,c)}(x, y). \tag{7}$$

If (7) holds with $\epsilon = 0$, then T is pointwise α-firmly nonexpansive at $y \in D \subset G$ on D. If (7) holds at all $y \in D$ with the same constant α, then T is said to be (almost) α-firmly nonexpansive on D. If $D = G$, then the mapping T is simply said to be (almost) α-firmly nonexpansive. If $D \supset \operatorname{Fix} T \neq \emptyset$ and (7) holds at all $y \in \operatorname{Fix} T$ with the same constant α, then T is said to be almost quasi α-firmly nonexpansive.

(v) The mapping T is said to be *pointwise asymptotically α-firmly nonexpansive at y with constant $\alpha < 1$* whenever

$$\forall \epsilon > 0, \exists D_\epsilon(y) \subset G :$$
$$d(Tx, Ty)^p \leq (1 + \epsilon) d(x, y)^p - \frac{1 - \alpha}{\alpha} \psi_T^{(p,c)}(x, y) \quad \forall x \in D_\epsilon(y), \tag{8}$$

where $D_\epsilon(y)$ is a neighborhood of y in D.

Proposition 6 (Characterizations) *Let (G, d) be a p-uniformly convex space with constant $c > 0$, and let $T : D \to G$ for $D \subset G$.*

(i)

$$\psi_T^{(p,c)}(x, y) = \frac{c}{2} d(Tx, x)^p \quad \text{whenever } y \in \operatorname{Fix} T. \tag{9}$$

For fixed $y \in \operatorname{Fix} T$, the function $\psi_T^{(p,c)}(x, y) \geq 0$ for all $x \in D$ and $\psi_T^{(p,c)}(x, y) = 0$ only when $x \in \operatorname{Fix} T$.

(ii) *Let $y \in \text{Fix}\, T$. T is pointwise almost α-firmly nonexpansive at y on D with violation at most $\epsilon > 0$ if and only if*

$$\exists \alpha \in [0,1): \quad d(Tx,y)^p \le (1+\epsilon)d(x,y)^p - \frac{1-\alpha}{\alpha}\frac{c}{2}d(Tx,x)^p \quad \forall x \in D. \tag{10}$$

In particular, T is almost quasi α-firmly nonexpansive on D whenever T possesses fixed points and (10) holds at all $y \in \text{Fix}\, T$ with the same constant $\alpha \in [0,1)$ and violation at most ϵ.

(iii) *If T is pointwise almost α-firmly nonexpansive at $y \in \text{Fix}\, T$ on D with constant $\underline{\alpha} \in [0,1)$ and violation at most ϵ, then it is pointwise almost α-firmly nonexpansive at y with the same upper bound on the violaton violation on D for all $\alpha \in [\underline{\alpha},1]$. In particular, if T is pointwise almost α-firmly nonexpansive at $y \in \text{Fix}\, T$ on D, then it is pointwise almost nonexpansive at y on D.*

Proof This is a slight extension of [4, Proposition 4], which was for pointwise α-firmly nonexpansive mappings. The proof for pointwise almost α-firmly nonexpansive mappings is the same. □

2.1 Composition of operators

Before continuing with pointwise almost α-firmly nonexpansive mappings, we make a brief but important observation about fixed points of compositions of quasi strictly nonexpansive mappings (6).

Lemma 7 *Let T_1 and T_2 be quasi strictly nonexpansive on (G,d) with $\text{Fix}\, T_1 \cap \text{Fix}\, T_2 \ne \emptyset$. Then $\text{Fix}\, T_1 \circ T_2 = \text{Fix}\, T_1 \cap \text{Fix}\, T_2$.*

Proof The inclusion $\text{Fix}\, T_1 \cap \text{Fix}\, T_2 \subset \text{Fix}(T_1 \circ T_2)$ is clear. Assume that there exists $y \in \text{Fix}(T_1 \circ T_2) \setminus (\text{Fix}\, T_1 \cap \text{Fix}\, T_2)$ and choose $x \in \text{Fix}\, T_1 \cap \text{Fix}\, T_2$. Then

$$d(y,x) \le d(T_1 \circ T_2 y, x) \le d(T_2 y, x) \le d(y,x)$$

with either $d(T_2 y, x) < d(y,x)$ or $d(T_1 \circ T_2 y, x) < d(T_2 y, x)$ as $y \notin \text{Fix}\, T_1 \cap \text{Fix}\, T_2$. This is a contradiction, so $\text{Fix}\, T_1 \circ T_2 = \text{Fix}\, T_1 \cap \text{Fix}\, T_2$. □

Remark 8 The sufficiency of strict quasi nonexpansivity for the analogous identity for convex combinations of mappings in a Hadamard space was recognized in [3, Remark 7.11].

Lemma 9 *Let (G,d) be a p-uniformly convex space with constant $c > 0$, and let $D \subset G$. Let $T_0 : D \to G$ be pointwise almost α-firmly nonexpansive at y on D with constant α_0 and violation ϵ_0, and let $T_1 : T_0(D) \to G$ be pointwise almost α-firmly nonexpansive at $T_0 y$ on $T_0(D)$ with constant α_1 and violation ϵ_1. Then the composition $\overline{T} := T_1 \circ T_0$ is pointwise almost α-firmly nonexpansive at y with constant $\overline{\alpha} \in (0,1)$ and violation at most $\overline{\epsilon} = \epsilon_0 + \epsilon_1 + \epsilon_0 \epsilon_1$ on D whenever*

$$\frac{1-\overline{\alpha}}{\overline{\alpha}}\psi_{\overline{T}}^{(p,c)}(x,y) \le (1+\epsilon_1)\frac{1-\alpha_0}{\alpha_0}\psi_{T_0}^{(p,c)}(x,y) + \frac{1-\alpha_1}{\alpha_1}\psi_{T_1}^{(p,c)}(T_0 x, T_0 y) \quad \forall x \in D. \tag{11}$$

Proof The proof is a slight extension of the same result for compositions of α-firmly non-expansive mappings in [4, Lemma 10]. Since T_1 is pointwise α-firmly nonexpansive at $T_0 y$ with violation ϵ_1 and constant α_1 on $T_0(D)$, we have

$$d(\overline{T}x, \overline{T}y)^p \leq (1 + \epsilon_1)d(T_0 x, T_0 y)^p - \frac{1 - \alpha_1}{\alpha_1}\psi_{T_1}^{(p,c)}(T_0 x, T_0 y) \quad \forall x \in D,$$

where $\psi_{T_1}^{(p,c)}$ is defined by (3). Since T_0 is α-firmly nonexpansive at y with constant α_0 with violation ϵ_0 on D, we have

$$d(\overline{T}x, \overline{T}y)^p$$

$$\leq (1 + \epsilon_0)(1 + \epsilon_1)d(x, y)^p - (1 + \epsilon_1)\frac{1 - \alpha_0}{\alpha_0}\psi_{T_0}^{(p,c)}(x, y) - \frac{1 - \alpha_1}{\alpha_1}\psi_{T_1}^{(p,c)}(T_0 x, T_0 y)$$

for all $x \in D$. Whenever (11) holds, we conclude that

$$\exists \overline{\alpha} \in (0, 1): \quad d(\overline{T}x, \overline{T}y)^p \leq (1 + \overline{\epsilon})d(x, y)^p - \frac{1 - \overline{\alpha}}{\overline{\alpha}}\psi_{\overline{T}}^{(p,c)}(x, y) \quad \forall x \in D,$$

where $\overline{\epsilon} = \epsilon_0 + \epsilon_1 + \epsilon_0 \epsilon_1$.   □

Proposition 10 (Compositions of pointwise almost α-firmly nonexpansive mappings)
Let (G, d) be a p-uniformly convex space with constant $c > 0$, and let $D \subset G$. Let $T_0 : D \to G$ be pointwise almost α-firmly nonexpansive at y on D with constant α_0 and violation ϵ_0, and let $T_1 : T_0(D) \to G$ be pointwise almost α-firmly nonexpansive at y on $T_0(D)$ with constant α_1 and violation ϵ_1. Let $y \in \text{Fix } T_0 \cap \text{Fix } T_1$. Then the composite operator $\overline{T} = T_1 \circ T_0$ is pointwise almost α-firmly nonexpansive at y on D with violation at most $\overline{\epsilon} = \epsilon_0 + \epsilon_1 + \epsilon_0 \epsilon_1$ and constant

$$\overline{\alpha} = \frac{\alpha_0 + \alpha_1 - 2\alpha_0\alpha_1}{\frac{c}{2}(1 - \alpha_0 - \alpha_1 + \alpha_0\alpha_1) + \alpha_0 + \alpha_1 - 2\alpha_0\alpha_1}. \tag{12}$$

Proof This is a minor extension of [4, Theorem 11]. By Lemma 9, it suffices to show (11) at all points $y \in \text{Fix } T_1 \cap \text{Fix } T_0$. First, note that $\text{Fix } \overline{T} \supset \text{Fix } T_1 \cap \text{Fix } T_0$, so by (9) we have $\psi_{T_0}^{(p,c)}(x, y) = \frac{c}{2}d(x, T_0 x)^p$, $\psi_{T_1}^{(p,c)}(T_0 x, T_0 y) = \frac{c}{2}d(T_0 x, \overline{T}x)^p$, and $\psi_{\overline{T}}^{(p,c)}(x, y) = \frac{c}{2}d(x, \overline{T}x)^p$ whenever $y \in \text{Fix } T_1 \cap \text{Fix } T_0$. Inequality (11) in this case simplifies to

$$\exists \overline{\kappa} > 0: \quad \kappa_0(1 + \epsilon_1)d(x, T_0 x)^p + \kappa_1 d(T_0 x, \overline{T}x)^p \geq \overline{\kappa}d(x, \overline{T}x)^p \quad \forall x \in D, \tag{13}$$

where $\kappa_0 := \frac{1 - \alpha_0}{\alpha_0}$, $\kappa_1 := \frac{1 - \alpha_1}{\alpha_1}$, and $\overline{\kappa} := \frac{1 - \overline{\alpha}}{\overline{\alpha}}$ with $\overline{\alpha} \in (0, 1)$. By (1), we have

$$\frac{c}{2}t(1 - t)d(x, \overline{T}x)^p \leq \frac{c}{2}t(1 - t)d(x, \overline{T}x)^p + d\left(T_0 x, (1 - t)x \oplus t\overline{T}x\right)^p$$

$$\leq (1 - t)d(T_0 x, x)^p + td(T_0 x, \overline{T}x)^p \quad \forall x \in G, \forall t \in (0, 1). \tag{14}$$

Letting $t = \frac{\kappa_1}{\kappa_0 + \kappa_1}$ yields $(1 - t) = \frac{\kappa_0}{\kappa_0 + \kappa_1}$, so that (14) becomes

$$\frac{c}{2}\frac{\kappa_0\kappa_1}{\kappa_0 + \kappa_1}d(x, \overline{T}x)^p \leq (1 + \epsilon_1)\kappa_0 d(T_0 x, x)^p + \kappa_1 d(T_0 x, \overline{T}x)^p \quad \forall x \in G. \tag{15}$$

It follows that (13) holds for any $\overline{\kappa} \in (0, \frac{c\kappa_0\kappa_1}{2(\kappa_0+\kappa_1)}]$. We conclude that the composition $\overline{T}$ is quasi α-firmly nonexpansive with constant

$$\overline{\alpha} = \frac{\kappa_0 + \kappa_1}{\frac{c}{2}\kappa_0\kappa_1 + \kappa_0 + \kappa_1}.$$

A short calculation shows that this is the same as (12), which completes the proof.　　□

2.2　Averages of operators

Let $\mathcal{B}(G)$ be the Borel algebra on (G, d), $\mathcal{P}$ be the family of probability measures on $(G, \mathcal{B}(G))$, and $\mathcal{P}^p(G)$ be the family of probability measures on G such that the pth moment exists i.e.

$$\mathcal{P}^p(G) := \left\{ v \in \mathcal{P}(G) \,\middle|\, \int_y d(x,y)^p v(dx) < \infty \; \forall x \in G \right\}.$$

For $\mu \in \mathcal{P}^p(G)$, the minimizer of

$$x \mapsto \int_G d(x,y)^p v(dy)$$

is called p-barycenter of v and denoted by $b_p(v)$ if it exists. The p-barycenter of v always exists if (G, d) is a proper geodesic space and $\mu \in \mathcal{P}^p(G)$ [9, Proposition 3.3].

Let $T_i : G \to G, i \in I$ be a collection of mappings where I is an index space. Assume that $(I, \mathcal{I})$ is a measurable space and $(x, i) \mapsto T_i x$ is measurable. Let η be a probability measure on I and define $b_p(T_i, \eta) : G \to G$ by

$$b_p(T_i, \eta)(x) = \operatorname*{argmin}_{y \in G} y \mapsto \int_I d(T_i x, y)^p \eta(di).$$

We use the notation $T_i x_* \eta$ for the push forward of η with respect to the mapping $i \mapsto T_i x$ for fixed x i.e.

$$T_i x_* \eta(A) = \eta\big(\{i \mid T_i x \in A\}\big) \tag{16}$$

for $A \in \mathcal{B}(G)$. Then, by definition, $b_p(T_i, \eta)(x) = b_p(T_i x_* \eta)$.

Theorem 11 *Let G be a proper, symmetric perpendicular, p-uniformly convex space with constant $c > 0$, and let T_i, $i \in I$ be a family of almost quasi α-firmly nonexpansive operators with violation ϵ_i and constant α_i respectively on G. Let η be a probability measure on I such that $T_i x_* \eta \in \mathcal{P}^p(G)$ for all $x \in G$. Then $\mathcal{T} = b_p(T_i, \eta)$ is a pointwise almost α-firm operator at any $y \in \bigcap_{i \in I} \operatorname{Fix} T_i$ on G with constant $\overline{\alpha} = \sup_{i \in I} \alpha_i$ and violation at most $\overline{\epsilon} = \sup_{i \in I} \epsilon_i$.*

Proof Let $x \in G$ be arbitrary, $y \in \bigcap_{i \in I} \operatorname{Fix} T_i \subset \operatorname{Fix} \mathcal{T}$, $v = T_i x_* \eta$ as defined in (16) and $(d(\cdot, y)^p)_* v$ the push forward of v. Then

$$d(\mathcal{T}x, \mathcal{T}y)^p = d\big(b_p(T_i x_* \eta), y\big)^p = d\big(b_p(v), y\big)^p \leq b_p\big((d(\cdot, y)^p)_* v\big)$$

by Jensen's inequality [9, Theorem 4.1] since $d(\cdot, y)^p$ is a convex function and the fact that $\mathbb{R}$ is a p-uniformly convex space with constant c. Now

$$
\begin{aligned}
b_p\big((d(\cdot,y)^p)_* \nu\big) &= \operatorname*{argmin}_{t \in \mathbb{R}} \int_z \big|t - d(z,y)^p\big|^p \nu(dz) \\[2mm]
&= \operatorname*{argmin}_{t \in \mathbb{R}} \int_i \big|t - d(T_i x, y)^p\big|^p \eta(di) \\[2mm]
&\leq \operatorname*{argmin}_{t \in \mathbb{R}} \int_i \left| t - \left[(1 + \epsilon_i) d(x,y)^p - \frac{1 - \alpha_i}{\alpha_i} \frac{c}{2} d(T_i x, x)^p \right] \right|^p \eta(di) \\[2mm]
&\leq (1 + \overline{\epsilon}) d(x,y)^p - \operatorname*{argmin}_{t \in \mathbb{R}} \int_i \left| t - \frac{1 - \overline{\alpha}}{\overline{\alpha}} \frac{c}{2} d(T_i x, x)^p \right|^p \eta(di) \\[2mm]
&= (1 + \overline{\epsilon}) d(x,y)^p - \frac{1 - \overline{\alpha}}{\overline{\alpha}} \frac{c}{2} \operatorname*{argmin}_{t \in \mathbb{R}} \int_z \big|t - d(z,x)^p\big|^p \nu(dz) \\[2mm]
&= (1 + \overline{\epsilon}) d(x,y)^p - \frac{1 - \overline{\alpha}}{\overline{\alpha}} \frac{c}{2} b_p\big((d(\cdot, y)^p)_* \nu\big).
\end{aligned}
$$

And again Jensen's inequality completes the proof

$$
\begin{aligned}
(1 + \overline{\epsilon}) d(x,y)^p &- b_p\left(\frac{1 - \overline{\alpha}}{\overline{\alpha}} \frac{c}{2} d(x, T_i x)_* \eta \right) \\[2mm]
&\leq (1 + \overline{\epsilon}) d(x,y)^p - \frac{1 - \overline{\alpha}}{\overline{\alpha}} \frac{c}{2} d\big(x, b_p(T_i x_* \eta)\big)^p. \qquad \square
\end{aligned}
$$

For a finite index set I, without loss of generality $I = \{1, \ldots, n\}$ and a probability measure η on I, we can define $\omega_i := \eta(i)$ for all i. Then

$$
b_p(T_i, \eta)(x) = \operatorname*{argmin}_{z \in G} \sum_{i=1}^n \omega_i d(z, T_i x)^p.
$$

In case of a Hilbert space G and $p = c = 2$ this reduces further to $b_p(T_i, \eta)(x) = \sum_{i=1}^n \omega_i T_i x$.

If the support of the measure ν consists of two discrete points x_1 and x_2 i.e. for $\omega \in [0,1]$, $\nu = \omega \delta_{x_1} + (1 - \omega) \delta_{x_2}$, then $b_p(\nu)$ can be calculated explicitly. It is obvious that $b_p(\nu)$ has to lie on the geodesic connecting x_1 and x_2. Hence $b_p(\nu) = \overline{t} x_1 \oplus (1 - \overline{t}) x_2$ for some $\overline{t} \in [0,1]$. Minimizing the function $t \mapsto \omega d(t x_1 \oplus (1 - t) x_2, x_1)^p + (1 - \omega) d(t x_1 \oplus (1 - t) x_2, x_1)^p$ leads to $\overline{t} = \frac{1}{\sqrt[p-1]{\frac{1-\omega}{\omega}} + 1}$. If $I = \{1, 2\}$, $T_1 = T$ and $T_2 = \text{Id}$, $\eta = \omega \delta_1(\cdot) + (1 - \omega) \delta_2(\cdot)$ for $\omega \in [0,1]$, then $T_\beta = b_p(T_i, \eta)$ is the Krasnoselsky–Mann relaxation of T

$$
T_\beta = \beta T \oplus (1 - \beta) \, \text{Id} \quad \text{with } \beta = \frac{1}{\sqrt[p-1]{\frac{1-\omega}{\omega}} + 1}.
$$

The next result shows that the convex combination of an almost nonexpansive mapping with the identity mapping can be made arbitrarily close to α-firmly nonexpansive (no violation) by choosing the averaging constant small enough—this can be interpreted as choosing an appropriately small step size.

Proposition 12 (Krasnoselsky–Mann relaxations) *Let (G, d) be a p-uniformly convex space and $T \colon G \to G$ be pointwise almost nonexpansive at all $y \in \text{Fix}\, T$ with violation*

ϵ. *Then* $T_\beta := \beta T \oplus (1 - \beta)$ Id *is pointwise almost α-firmly nonexpansive at all $y \in$ Fix T with constant*

$$\alpha_\beta = \frac{\alpha\beta^{p-1}}{\alpha\beta^{p-1} - \alpha\beta + 1}$$

and violation at most $\epsilon_\beta := \epsilon\beta$.

Proof Clearly Fix T = Fix T_β and $d(x, T_\beta x)^p = \beta^p d(x, Tx)^p$. Let $y \in FixT_\beta$, then

$$\begin{aligned}
d(y, T_\beta x)^p &= d(y, \beta Tx \oplus (1 - \beta)x)^p \\
&\leq \beta d(y, Tx)^p + (1 - \beta)d(y, x)^p - \frac{c}{2}\beta(1 - \beta)d(x, Tx)^p \\
&\leq (1 + \epsilon\beta)d(x, y)^p - \frac{c}{2}\beta^{1-p}\left(\frac{1-\alpha}{\alpha} + 1 - \beta\right)d(x, T_\beta x)^p.
\end{aligned}$$

Setting

$$\frac{1 - \alpha_\beta}{\alpha_\beta} = \beta^{1-p}\left(\frac{1-\alpha}{\alpha} + 1 - \beta\right)$$

and solving for α_β yield the result. □

2.3 Metric subregularity

Recall that $\mu : [0, \infty) \to [0, \infty)$ is a *gauge function* if μ is continuous, strictly increasing with $\mu(0) = 0$, and $\lim_{t\to\infty} \mu(t) = \infty$.

Definition 13 (Metric regularity on a set) Let (G_1, d_1) and (G_2, d_2) be metric spaces, and let $\Phi : G_1 \rightrightarrows G_2$, $U \subset G_1$, $V \subset G_2$. The mapping Φ is called *metrically regular with gauge μ on $U \times V$ relative to* $\Lambda \subset G_1$ if

$$\forall y \in V, \forall x \in U \cap \Lambda, \quad \mathrm{dist}(x, \Phi^{-1}(y) \cap \Lambda) \leq \mu(\mathrm{dist}(y, \Phi(x))). \tag{17}$$

When the set V consists of a single point, $V = \{\bar{y}\}$, then Φ is said to be *metrically subregular for $\bar{y}$ on U with gauge μ relative to* $\Lambda \subset G_1$.

When μ is a linear function (that is, $\mu(t) = \kappa t, \forall t \in [0, \infty)$), this special case is distinguished as *linear metric (sub)regularity with constant κ*. When $\Lambda = G_1$, the quantifier "relative to" is dropped. When μ is linear, the infimum of all constants κ for which (17) holds is called the *modulus* of metric regularity.

The next statement is obvious from the definition.

Proposition 14 *Let (G_1, d_1) and (G_2, d_2) be metric spaces, and let $\Phi : G_1 \rightrightarrows G_2$, $U \subset G_1$, $V \subset G_2$. If Φ is metrically subregular with gauge μ at y on U relative to $\Lambda \subset G_1$, then Φ is metrically subregular with the same gauge μ at y on all subsets $U' \subset U$ relative to $\Lambda \subset G_1$.*

3 Quantitative convergence

To obtain convergence of fixed point iterations under the assumption of metric subregularity, the gauge of metric subregularity μ is constructed implicitly from another nonnegative function $\theta : [0, \infty) \to [0, \infty)$ satisfying

$$\text{(i)} \quad \theta(0) = 0; \qquad \text{(ii)} \quad 0 < \theta(t) < t \quad \forall t > 0; \qquad \text{(iii)} \quad \sum_{j=1}^{\infty} \theta^{(j)}(t) < \infty \quad \forall t \geq 0. \quad (18)$$

For a p-uniformly convex space the operative gauge satisfies

$$\mu\left(\left(\frac{(1+\epsilon)t^p - (\theta(t))^p}{\tau}\right)^{1/p}\right) = t \quad \Longleftrightarrow \quad \theta(t) = \left((1+\epsilon)t^p - \tau\left(\mu^{-1}(t)\right)^p\right)^{1/p} \quad (19)$$

for $\tau > 0$ fixed and θ satisfying (18).

In the case of linear metric subregularity on a CAT(κ) space this becomes

$$t \mapsto \mu t \quad \Longleftrightarrow \quad \theta(t) = \left((1+\epsilon) - \frac{\tau}{\mu^2}\right)^{1/2} t \left(\mu \geq \sqrt{\frac{\tau}{(1+\epsilon)}}\right).$$

If (17) is satisfied for some $\mu' > 0$, then the condition $\mu \geq \sqrt{\frac{\tau}{(1+\epsilon)}}$ is satisfied for all $\mu \geq \mu'$ large enough. The conditions in (18) in this case simplify to $\theta(t) = \gamma t$, where

$$0 < \gamma := \sqrt{1 + \epsilon - \frac{\tau}{\mu^2}} < 1 \quad \Longleftrightarrow \quad \sqrt{\frac{\tau}{(1+\epsilon)}} < \mu < \sqrt{\frac{\tau}{\epsilon}}. \quad (20)$$

The next definition characterizes the quantitative convergence of sequences in terms of gauge functions.

Definition 15 (Gauge monotonicity [10]) Let (G, d) be a metric space, let $(x_k)_{k \in \mathbb{N}}$ be a sequence on G, let $D \subset G$ be nonempty, and let the continuous mapping $\mu : \mathbb{R}_+ \to \mathbb{R}_+$ satisfy $\mu(0) = 0$ and

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

(i) $(x_k)_{k \in \mathbb{N}}$ is said to be *gauge monotone with respect to D with rate* μ whenever

$$d(x_{k+1}, D) \leq \mu\left(d(x_k, D)\right) \quad \forall k \in \mathbb{N}. \quad (21)$$

(ii) $(x_k)_{k \in \mathbb{N}}$ is said to be *linearly monotone with respect to D* with rate c if (21) is satisfied for $\mu(t) = c \cdot t$ for all $t \in \mathbb{R}_+$ and some constant $c \in [0, 1]$.

A sequence $(x_k)_{k \in \mathbb{N}}$ is said to converge *gauge monotonically* to some element $x^* \in G$ with rate $s_k(t) := \sum_{j=k}^{\infty} \mu^{(j)}(t)$ whenever it is gauge monotone with gauge μ satisfying $\sum_{j=1}^{\infty} \mu^{(j)}(t) < \infty \; \forall t \geq 0$, and there exists a constant $a > 0$ such that $d(x_k, x^*) \leq a s_k(t)$ for all $k \in \mathbb{N}$.

All Fejér monotone sequences are linearly monotone (with constant $c = 1$) but the converse does not hold (see Proposition 1 and Example 1 of [10]). Gauge-monotonic convergence for a linear gauge in the definition above is just *R*-linear convergence.

Metric subregularity and pointwise (almost) nonexpansiveness are fundamentally connected through the surrogate mapping $\mathcal{T}_S : G \to \mathbb{R}_+ \cup \{+\infty\}$ defined by

$$\mathcal{T}_S(x) := \left| \left(\inf_{y \in S} \psi_T^{(p,c)}(x,y) \right)^{1/p} \right|, \tag{22}$$

where $\psi_T^{(p,c)}$ is defined by (3) and $S \subset G$. If $S = \emptyset$ then, by definition, $\mathcal{T}_S(x) := +\infty$ for all x. Hence, $\mathcal{T}_S$ is proper when S is nonempty. For our purposes, $S \subseteq \operatorname{Fix} T$, in which case by Proposition 6(i) we have $\psi_T^{(p,c)}(x,y) \geq 0$ for all $x \in D$ and all $y \in S$ and $\psi_T^{(p,c)}(x,y) = 0$ only when both $x, y \in \operatorname{Fix} T$. Hence $\mathcal{T}_S$ is nonnegative, takes the value 0 only on $\operatorname{Fix} T$, and has the simple representation

$$\mathcal{T}_S(x) = \sqrt[p]{\frac{2}{c}} d(Tx, x) > 0 \quad (S \neq \emptyset). \tag{23}$$

Theorem 16 (Necessary and sufficient conditions for convergence rates) *Let (G, d) be a complete p-uniformly convex space with constant c; let $D \subset G$ with (D, d), let $T : D \to D$ with $(T(D), d)$ compact on bounded subsets, and let $S := \operatorname{Fix} T \cap D$ be nonempty. Assume further that T is pointwise almost α-firmly nonexpansive at all points $y \in S$ with the same constant $\overline{\alpha}$ and violation at most ϵ on D.*

  (a) *(necessity) Suppose that all sequences $(x^k)_{k \in \mathbb{N}}$ defined by $x^{k+1} = Tx^k$ and initialized in D are gauge monotone relative to S with rate θ satisfying (18), and $(\operatorname{Id} - \theta)^{-1}(\cdot)$ is continuous on $\mathbb{R}_+$, strictly increasing, and $(\operatorname{Id} - \theta)^{-1}(0) = 0$. Then all sequences initialized on D converge gauge monotonically to some $\overline{x} \in S$ with rate $O(s_k(t_0))$ where $s_k(t) := \sum_{j=k}^{\infty} \theta^{(j)}(t)$ and $t_0 := d(x^0, \operatorname{Fix} T \cap D)$. Moreover, $\mathcal{T}_S$ defined by (22) is metrically subregular for 0 relative to D on D with gauge $\mu(\cdot) = (\operatorname{Id} - \theta)^{-1}(\cdot)$.*

  (b) *(sufficiency) Let T satisfy*

$$d(x, \operatorname{Fix} T \cap D) \leq \mu\big(d(x, Tx)\big), \quad \forall x \in D, \tag{24}$$

  with gauge μ given implicitly by (19) with θ satisfying (18) for $\tau = (1 - \overline{\alpha})/\overline{\alpha}$ and $\epsilon \geq 0$ an upper bound on the violation of pointwise α firmness of T on D. Then, for any $x^0 \in D$, the sequence $(x^k)_{k \in \mathbb{N}}$ defined by $x^{k+1} = Tx^k$ satisfies

$$d\big(x^{k+1}, \operatorname{Fix} T \cap D\big) \leq \theta\big(d\big(x^k, \operatorname{Fix} T \cap D\big)\big) \quad \forall k \in \mathbb{N}. \tag{25}$$

  Moreover, the sequence $(x^k)_{k \in \mathbb{N}}$ converges gauge monotonically to some $x^* \in \operatorname{Fix} T \cap D$ with rate $O(s_k(t_0))$ where $s_k(t) := \sum_{j=k}^{\infty} \theta^{(j)}(t)$ and $t_0 := d(x^0, \operatorname{Fix} T \cap D)$.

Before proving this theorem, we collect some intermediate results.

Lemma 17 (Gauge monotonicity and almost quasi α-firmness implies convergence to fixed points) *Let (G, d) be a complete p-uniformly convex space with constant c. Let $T : G \to G$ with $T(D) \subseteq D \subseteq G$ boundedly compact. Suppose that $\operatorname{Fix} T \cap D$ is nonempty and that T is pointwise almost α-firmly nonexpansive at all $y \in \operatorname{Fix} T \cap D$ with the same constant $\overline{\alpha}$ and violation at most ϵ on D. If the sequence $(x^k)_{k \in \mathbb{N}}$ defined by $x^{k+1} = Tx^k$ and initialized in D is gauge monotone relative to $\operatorname{Fix} T \cap D$ with rate θ satisfying (18), then*

$(x^k)_{k \in \mathbb{N}}$ *converges gauge monotonically to some* $x^* \in \operatorname{Fix} T \cap D$ *with rate* $O(s_k(t_0))$ *where* $s_k(t) := \sum_{j=k}^{\infty} \theta^{(j)}(t)$ *and* $t_0 := d(x^0, \operatorname{Fix} T \cap D)$.

Proof The assumption that T is pointwise almost α-firmly nonexpansive at all $y \in \operatorname{Fix} T \cap D$ with constant $\overline{\alpha}$ and violation at most ϵ on D yields

$$d(Tx, y)^p \le (1 + \epsilon) d(x, y)^p - \frac{c(1 - \overline{\alpha})}{2\overline{\alpha}} d(Tx, x)^p, \quad \forall x \in D.$$

Let $x^0 \in D$ and define the sequence $x^{k+1} = Tx^k$ for all $k \in \mathbb{N}$. Since T is pointwise almost α-firmly nonexpansive at all points in $\operatorname{Fix} T \cap D$ on D, $\operatorname{Fix} T \cap D$ is closed and $P_{\operatorname{Fix} T \cap D} x^k$ is nonempty (though possibly set-valued) for all k. Denote any selection by $\bar{x}^k \in P_{\operatorname{Fix} T \cap D} x^k$ for each $k \in \mathbb{N}$. Then

$$d(x^{k+1}, \bar{x}^k)^p \le (1 + \epsilon) d(x^k, \bar{x}^k)^p - \frac{c(1 - \overline{\alpha})}{2\overline{\alpha}} d(x^k, x^{k+1})^p, \quad \forall k \in \mathbb{N},$$

which implies that

$$d(x^k, x^{k+1}) \le \left(\frac{2\overline{\alpha}(1 + \epsilon)}{c(1 - \overline{\alpha})} \right)^{1/p} d(x^k, \bar{x}^k), \quad \forall k \in \mathbb{N}.$$

On the other hand $d(x^k, \bar{x}^k) = d(x^k, \operatorname{Fix} T \cap D) \le \theta(d(x^{k-1}, \operatorname{Fix} T \cap D))$ since $(x^k)_{k \in \mathbb{N}}$ is gauge monotone relative to $\operatorname{Fix} T \cap D$ with rate θ. Therefore an iterative application of gauge monotonicity yields

$$d(x^k, x^{k+1}) \le \left(\frac{2\overline{\alpha}(1 + \epsilon)}{c(1 - \overline{\alpha})} \right)^{1/p} \theta^{(k)}(d(x^0, \operatorname{Fix} T \cap D)), \quad \forall k \in \mathbb{N}.$$

Let $t_0 = d(x^0, \operatorname{Fix} T \cap D)$. For any given natural numbers k, l with $k < l$, an iterative application of the triangle inequality yields the upper estimate

$$\begin{aligned} d(x^k, x^l) &\le d(x^k, x^{k+1}) + d(x^{k+1}, x^{k+2}) + \cdots + d(x^{l-1}, x^l) \\ &\le \left(\frac{2\overline{\alpha}(1 + \epsilon)}{c(1 - \overline{\alpha})} \right)^{1/p} \left(\theta^{(k)}(t_0) + \theta^{(k+1)}(t_0) + \cdots + \theta^{(l-1)}(t_0) \right) \\ &< \left(\frac{2\overline{\alpha}(1 + \epsilon)}{c(1 - \overline{\alpha})} \right)^{1/p} s_k(t_0), \end{aligned}$$

where $s_k(t_0) := \sum_{j=k}^{\infty} \theta^{(j)}(t_0) < \infty$ for θ satisfying (18). Since $(\theta^{(k)}(t_0))_{k \in \mathbb{N}}$ is a summable sequence of nonnegative numbers, the sequence of partial sums $s_k(t_0)$ converges to zero monotonically as $k \to \infty$, and hence $(x^k)_{k \in \mathbb{N}}$ is a Cauchy sequence and $x^k \to x^*$ for some $x^* \in G$. Letting $l \to +\infty$ yields

$$\lim_{l \to +\infty} d(x^k, x^l) = d(x^k, x^*) \le a s_k(t_0), \qquad a := \left(\frac{2\overline{\alpha}(1 + \epsilon)}{c(1 - \overline{\alpha})} \right)^{1/p}.$$

Therefore $(x^k)_{k \in \mathbb{N}}$ converges gauge monotonically to x^* with rate $O(s_k(t_0))$.

It remains to show that $x^* \in \text{Fix } T \cap D$. Note that for each $k \in \mathbb{N}$

$$d(x^k, \bar{x}^k) = d(x^k, \text{Fix } T \cap D) \leq \theta^{(k)}(t_0),$$

which yields $\lim_k d(x^k, \bar{x}^k) = 0$. But by the triangle inequality

$$d(\bar{x}^k, x^*) \leq d(x^k, \bar{x}^k) + d(x^k, x^*),$$

so $\lim_k d(\bar{x}^k, x^*) = 0$. By construction $(\bar{x}^k)_{k \in \mathbb{N}} \subseteq \text{Fix } T \cap D$ and $\text{Fix } T \cap D$ is closed, hence $x^* \in \text{Fix } T \cap D$. □

Proposition 18 ([4], Theorem 32) *Let (G, d) be a p-uniformly convex metric space with constant c. Let $T : D \to D$ with $D \subseteq G$. Suppose that $S := \text{Fix } T \cap D$ is nonempty. Suppose that all sequences $(x^k)_{k \in \mathbb{N}}$ defined by $x^{k+1} = Tx^k$ and initialized in D are gauge monotone relative to S with rate θ satisfying (18). Suppose, in addition, that $(\text{Id} - \theta)^{-1}(\cdot)$ is continuous on $\mathbb{R}_+$, strictly increasing, and $(\text{Id} - \theta)^{-1}(0) = 0$. Then $\mathcal{T}_S$ defined by (22) is metrically subregular for 0 relative to D on D with gauge $\mu(\cdot) = (\text{Id} - \theta)^{-1}(\cdot)$.*

Proof of Theorem 16   Part (a). This follows immediately from Lemma 17 and Proposition 18.

Part (b). Our pattern of proof follows the same logic as the analogous result for set-valued mappings in a Euclidean space setting [11, Theorem 2.2]. Since $S = \text{Fix } T \cap D$, Proposition 6(i) establishes that $\psi(x, y) = \frac{c}{2} d(x, Tx)^p$ for all $y \in \text{Fix } T$, so $\mathcal{T}_S(x) = \frac{c}{2} d(x, Tx)$. Also by Proposition 6(i) $\mathcal{T}_S$ takes the value 0 only on $\text{Fix } T$, that is, $\mathcal{T}_S^{-1}(0) = \text{Fix } T$. So by assumption that $\mathcal{T}_S$ satisfies (24) with gauge μ given by (19) for $\tau = (1 - \overline{\alpha})/\overline{\alpha}$, together with the definition of metric subregularity (Definition 13), this yields

$$d(x, \text{Fix } T \cap D) = d\big(x, \mathcal{T}_S^{-1}(0) \cap D\big)$$
$$\leq \mu\big(\mathcal{T}_S(x)\big) = \mu\big(d(x, Tx)\big) \quad \forall x \in D.$$

In other words,

$$\frac{1 - \overline{\alpha}}{\overline{\alpha}} \big(\mu^{-1}\big(d(x, \text{Fix } T \cap D)\big)\big)^p \leq \frac{1 - \overline{\alpha}}{\overline{\alpha}} d(x, Tx)^p \quad \forall x \in D. \tag{26}$$

On the other hand, by the assumption that T is pointwise almost α-firmly nonexpansive at all points $y \in S$ with the same constant $\overline{\alpha}$ and violation at most ϵ on D we have

$$0 \leq \frac{1 - \overline{\alpha}}{\overline{\alpha}} d(x, Tx)^p$$
$$\leq (1 + \epsilon) d(x, y)^p - d(Tx, y)^p \quad \forall y \in \text{Fix } T \cap D, \forall x \in D. \tag{27}$$

Incorporating (26) into (27) and rearranging the inequality yields

$$d(Tx, y)^p \leq (1 + \epsilon) d(x, y)^p - \frac{1 - \overline{\alpha}}{\overline{\alpha}} \big(\mu^{-1}\big(d(x, \text{Fix } T \cap D)\big)\big)^p$$

$$\forall y \in \text{Fix } T \cap D, \forall x \in D.$$

Since this holds at *any* $x \in D$, it certainly holds at the iterates x^k with initial point $x^0 \in D$ since T is a self-mapping on D. Therefore for all $k \in \mathbb{N}$

$$d\big(x^{k+1}, y\big) \leq \sqrt[p]{(1 + \epsilon)d\big(x^k, y\big)^p - \frac{1 - \overline{\alpha}}{\overline{\alpha}}\big(\mu^{-1}\big(d\big(x^k, \operatorname{Fix} T \cap D\big)\big)\big)^p}$$

$$\forall y \in \operatorname{Fix} T \cap D. \tag{28}$$

Equation (28) simplifies as follows. Since the space $(T(D), d)$ is boundedly compact and Fix T is closed by continuity, for every $k \in \mathbb{N}$, the distance $d(x^k, \operatorname{Fix} T \cap D)$ is attained at some $y^k \in \operatorname{Fix} T \cap D$. This yields

$$d\big(x^{k+1}, y^{k+1}\big)^p \leq d\big(x^{k+1}, y^k\big)^p \leq (1 + \epsilon)d\big(x^k, y^k\big)^p - \frac{1 - \overline{\alpha}}{\overline{\alpha}}\big(\mu^{-1}\big(d\big(x^k, y^k\big)\big)\big)^p$$

$$\forall k \in \mathbb{N}. \tag{29}$$

Taking the pth root and recalling (19) yields (25).

This establishes also that the sequence $(x^k)_{k \in \mathbb{N}}$ is gauge monotone relative to Fix $T \cap D$ with rate θ satisfying Eq. (18). By Lemma 17 it follows that the sequence $(x^k)_{k \in \mathbb{N}}$ converges gauge monotonically to $x^* \in \operatorname{Fix} T \cap D$ with the rate $O(s_k(d(x^0, \operatorname{Fix} T \cap D)))$ where $s_k(t) := \sum_{j=k}^{\infty} \theta^{(j)}(t)$. □

Corollary 19 (Linear convergence)

(a) (*necessity*) *In the setting of Theorem* 16(*a*), *if all sequences* $(x^k)_{k \in \mathbb{N}}$ *defined by* $x^{k+1} = Tx^k$ *and initialized in D are linearly monotone relative to S with rate $\gamma < 1$, then all sequences initialized on D converge R-linearly to some $\overline{x} \in S$ with rate $O(\gamma^k)$. Moreover, $\mathcal{T}_S$ defined by* (22) *is linearly metrically subregular for* 0 *relative to D on D with gauge $\mu(t) = (1 - \gamma)^{-1}t$.*

(b) (*sufficiency*) *In the setting of Theorem* 16(*b*) *suppose that T satisfies*

$$d(x, \operatorname{Fix} T \cap D) \leq \mu d(x, Tx), \quad \forall x \in D, \tag{30}$$

with the scalar μ satisfying

$$\sqrt[p]{\frac{1 - \overline{\alpha}}{\overline{\alpha}(1 + \epsilon)}} < \mu < \sqrt[p]{\frac{1 - \overline{\alpha}}{\overline{\alpha}\epsilon}}.$$

Then, for any $x^0 \in D$, the sequence $(x^k)_{k \in \mathbb{N}}$ defined by $x^{k+1} = Tx^k$ is R-linearly convergent to a point in Fix $T \cap D$ *with rate $\gamma = \sqrt[p]{1 + \epsilon - \frac{1-\overline{\alpha}}{\overline{\alpha}\mu^p}}$.*

In the statements above, the upper bound on the violation of α-firm nonexpansiveness ϵ has to be compensated for by an equally strong gauge of metric subregularity with this value of ϵ explicitly accounted for in the gauge. The next result shows that these can be decoupled if T is pointwise asymptotically α-firmly nonexpansive at fixed points. In particular, if T is pointwise almost α-firmly nonexpansive at $y \in \operatorname{Fix} T$ with arbitrarily small violation ϵ, then whenever T is (gauge) metrically subregular at y, there is a neighborhood of y on which convergence of the fixed point iteration can be quantified by the said

gauge. In this situation it suffices to *qualitatively* determine metric subregularity—the exact value of the constants in relation to the violation of α-firmness is not needed in order to determine local convergence on the order of the gauge.

Proposition 20 *Let (G, d) be a complete p-uniformly convex space with constant c; let $D \subset G$, let $T : D \to D$ with $T(D)$ boundedly compact, and let $S := \operatorname{Fix} T \cap D$ be nonempty. Assume that T is a self-mapping on sufficiently small balls around points in S restricted to D, and that T is pointwise asymptotically α-firmly nonexpansive at all points $y \in S$ with constant $\overline{\alpha} \in (0, 1)$. Suppose further that T satisfies*

$$d(x, \operatorname{Fix} T \cap D) \le \mu\big(d(x, Tx)\big), \quad \forall x \in D, \tag{31}$$

with gauge μ given by (19) and $\tau = (1 - \overline{\alpha})/\overline{\alpha}$. Then, for any x^0 close enough to S, the sequence $(x^k)_{k \in \mathbb{N}}$ defined by $x^{k+1} = Tx^k$ converges gauge monotonically to some $x^* \in \operatorname{Fix} T \cap D$ with rate $O(s_k(t_0))$ where $s_k(t) := \sum_{j=k}^{\infty} \theta^{(j)}(t)$ and $t_0 := d(x^0, \operatorname{Fix} T)$ for θ given implicitly by (19) satisfying (18).

Proof Since T is a self-mapping on $\mathbb{B}_\delta(S) \cap D$ for δ small enough, and T is pointwise asymptotically α-firmly nonexpansive with constant $\overline{\alpha} \in (0, 1)$, the result follows immediately from Proposition 14 and Theorem 16 when the domain D is restricted to $\mathbb{B}_\delta(S)$ for δ sufficiently small.   □

4 Proximal mappings

We return now to the prox mapping (2). It was shown in [8, Proposition 2.7] that the argmin in (2) exists and is unique if f is proper, lsc, and convex. In general the prox mapping of a convex function is not α-firmly nonexpansive. However, it was shown in [4, Corollary 23] that it is *almost α-firmly nonexpansive*. This and other properties of prox mappings are collected in the following result.

Theorem 21 *Let (G, d) be a p-uniformly convex metric space with constant $c \in (0, 2]$, and let $f : G \to \mathbb{R}$ be proper, convex, and lower semicontinuous.*

(i) *If (G, d) is symmetric perpendicular, then $\operatorname{prox}_{f,\lambda}^p$ is quasi strictly nonexpansive, that is,*

$$d\big(\operatorname{prox}_{f,\lambda}^p(x), \overline{x}\big) < d(x, \overline{x}) \quad \forall x \in G, \forall \overline{x} \in \operatorname{argmin} f = \operatorname{Fix} \operatorname{prox}_{f,\lambda}^p.$$

(ii) *The prox mapping $\operatorname{prox}_{f,\lambda}^p$ is pointwise almost α-firmly nonexpansive at all $y \in \operatorname{argmin} f$ on G with constant*

$$\alpha_c = \frac{c(c-1)}{c(c-1)+2} \quad \text{and violation bounded above by} \quad \epsilon_c = \frac{2-c}{c-1}. \tag{32}$$

(iii) *If (G, d) is a $\operatorname{CAT}(\kappa)$ space, then $\operatorname{prox}_{f,\lambda}^p$ is pointwise asymptotically α-firmly nonexpansive at all $y \in \operatorname{argmin} f$ with constant $\overline{\alpha} = 1/2$.*

Proof (i). Let $x \in G$ be arbitrary and $y := \operatorname{prox}_{f,\lambda}^p(x)$. We prove by contradiction that the projection of x onto the geodesic $[\overline{x}, y]$ connecting $\overline{x}$ and y is y i.e. $P_{[\overline{x},y]}(x) = y$. Therefore

assume that $P_{[\overline{x},y]}(x) \neq y$ i.e. $P_{[\overline{x},y]}(x) = (1-t)y \oplus t\overline{x}$ for some $t \in (0,1]$. Then $f((1-t)y \oplus t\overline{x}) \leq (1-t)f(y) + tf(\overline{x}) \leq f(y)$ and $d((1-t)y \oplus t\overline{x}, x) < d(y,x)$. Now

$$f\big((1-t)y \oplus t\overline{x}\big) + \frac{1}{p\lambda^{p-1}}d\big((1-t)y \oplus t\overline{x}, x\big)^p < f(y) + \frac{1}{p\lambda^{p-1}}d(y,x)$$

contradicts $y = \mathrm{prox}_{f,\lambda}^p(x)$. Hence our assumption must be discarded and $P_{[\overline{x},y]}(x) = y$. In particular $[\overline{x},y] \perp_y [x,y]$, and hence by the symmetric perpendicular property $[x,y] \perp_y [\overline{x},y]$. Now $[x,y] \perp_y [\overline{x},y]$ in turn yields the claim $d(y,\overline{x}) \leq d(x,\overline{x})$.

If in addition (G,d) is a p-uniformly convex space

$$
\begin{aligned}
d(\overline{x},y)^p = d\big(\overline{x}, P_{[x,y]}(\overline{x})\big)^d &\leq d\left(\overline{x}, \frac{1}{2}x \oplus \frac{1}{2}y\right)^p \\
&\leq \frac{1}{2}d(\overline{x},x)^p + \frac{1}{2}d(\overline{x},y)^p - \frac{c}{2}\frac{1}{4}d(x,y)^p \\
&\leq d(\overline{x},x)^p - \frac{c}{2}\frac{1}{4}d(x,y)^p.
\end{aligned}
$$

This is only possible if either $x = y$ or $d(\overline{x},y) < d(\overline{x},x)$. In both cases $\mathrm{prox}_{f,\lambda}^p$ is quasi strictly nonexpansive.

(ii). This is [4, Corollary 23].

(iii). Let $\epsilon > 0$, $y \in \mathrm{argmin}f$, and $c = \frac{2+\epsilon}{1+\epsilon}$. Then $c \in (1,2)$ and there is a unique solution $t \in (0,\pi/2)$ to $\frac{\epsilon}{2} = t\tan(\pi/2 - t)$. Set $\delta = t/(2\sqrt{\kappa})$. Then $\delta \in (0, \pi/(4\sqrt{\kappa}))$ and $(\mathbb{B}_\delta(y), d|_{\mathbb{B}_\delta(y)})$ is a p-uniformly convex space with constant c by Lemma 1. Part (i) ensures that $\mathrm{prox}_{f,\lambda}^p$ is a self-mapping on $\mathbb{B}_\delta(y)$, and hence

$$\operatorname*{argmin}_{z \in G} f(z) + \frac{1}{p\lambda^{p-1}}d(z,x)^p = \operatorname*{argmin}_{z \in \mathbb{B}_\delta(y)} f(y) + \frac{1}{p\lambda^{p-1}}d(y,x)^p,$$

and we will use the same notation $\mathrm{prox}_{f,\lambda}^p$ for the operator restricted to the subspace $(\mathbb{B}_\delta(y), d|_{\mathbb{B}_\delta(y)})$. The operator $\mathrm{prox}_{f,\lambda}^p$ is pointwise almost α-firmly nonexpansive with constant $\alpha_c = \frac{c(c-1)}{c(c-1)+2}$ and violation bounded above by $\epsilon_c = \frac{2-c}{c-1} = \epsilon$ on $(\mathbb{B}_\delta(y), d|_{\mathbb{B}_\delta(y)})$ by part (ii). Note that $\alpha_c \nearrow 1/2$ as $c \nearrow 2$ and by Propositon 6(iii) $\mathrm{prox}_{f,\lambda}^p$ is pointwise almost α-firmly nonexpansive with constant $\overline{\alpha} = 1/2$ for all $c \in (0,2)$. Since $\epsilon_c \searrow 0$ as $c \nearrow 2$, this implies that $\mathrm{prox}_{f,\lambda}^p$ is pointwise asymptotically α-firmly nonexpansive at $y \in \mathrm{argmin}f$ with constant $\overline{\alpha} = 1/2$ and the neighborhood $\mathbb{B}_\delta(y)$ of y. □

Remark 22 In part (i) of Theorem 21, quasi nonexpansiveness comes from the symmetric perpendicular property. Strict quasi nonexpansiveness comes from the property of p-uniformly convex spaces.

The next result gathers properties of mappings built from prox mappings.

Proposition 23 *Let (G,d) be a p-uniformly convex space with constant $c \in (1,2]$, and let $f, g: G \to \mathbb{R} \cup \{+\infty\}$ be convex, proper, and lower semicontinuous. Assume that $\mathrm{argmin}f \cap \mathrm{argmin}g \neq \emptyset$.*

(i) *The operator $T := \mathrm{prox}_{f,\lambda_2}^p \circ \mathrm{prox}_{g,\lambda_1}^p$ $(\lambda_1, \lambda_2 > 0)$ is pointwise almost α-firmly nonexpansive at all points $y \in \Omega := \mathrm{argmin}f \cap \mathrm{argmin}g$ on G with constant*

$\alpha_\circ = \frac{2(c-1)}{2c-1}$ *and violation bounded above by* $\epsilon_\circ = \frac{1}{(c-1)^2} - 1$. *If* (G,d) *is a* $\mathrm{CAT}(\kappa)$ *space, then* T *is pointwise asymptotically* α-*firmly nonexpansive at all* $y \in \Omega$ *with constant* $\overline{\alpha}_\circ = 2/3$.

(ii) *Let* $T_0 : G \to G$ *be pointwise asymptotically* α-*firmly nonexpansive with constant* $\overline{\alpha}_0$ *at all* $y \in \Omega := \operatorname{argmin} f \cap \operatorname{Fix} T_0$. *The operator* $T := \operatorname{prox}_{f,\lambda}^p \circ T_0$ $(\lambda > 0)$ *is pointwise almost* α-*firmly nonexpansive at all points* $y \in \Omega$ *on* G *with constant and violation bounded above by*

$$\alpha = \frac{\overline{\alpha}_0 + \alpha_c - 2\overline{\alpha}_0\alpha_c}{\frac{c}{2}(1 - \overline{\alpha}_0 - \alpha_c + \overline{\alpha}_0\alpha_c) + \overline{\alpha}_0 + \alpha_c - 2\overline{\alpha}_0\alpha_c} \quad and \quad \epsilon = \epsilon_0 + \epsilon_c + \epsilon_0\epsilon_c, \qquad (33)$$

where ϵ_0 *is the violation of* α-*firm nonexpansiveness of* T_0 *on some neighborhood small enough. If* (G,d) *is a* $\mathrm{CAT}(\kappa)$ *space, then* T *is pointwise asymptotically* α-*firmly nonexpansive at all* $y \in \Omega$ *with constant*

$$\overline{\alpha} := \frac{1}{2 - \overline{\alpha}_0}. \qquad (34)$$

(iii) *The Krasnoselsky–Mann relaxation* $T := \beta \operatorname{prox}_{f,\lambda}^p \oplus (1 - \beta) \operatorname{Id}$ *is pointwise almost* α-*firmly nonexpansive at all points* $y \in \Omega := \operatorname{argmin} f$ *on* G *with constant*

$$\alpha_\beta = \frac{\alpha_c \beta^{p-1}}{\alpha_c \beta^{p-1} - \alpha_c \beta + 1}$$

and violation bounded above by $\epsilon_\beta = \beta\epsilon_c$, *where* α_c *and* ϵ_c *are given by* (32). *If* (G,d) *is a* $\mathrm{CAT}(\kappa)$ *space, then* T *is pointwise asymptotically* α-*firmly nonexpansive at all* $y \in \Omega$ *with constant*

$$\overline{\alpha}_\beta := \frac{\beta^{p-1}}{\beta^{p-1} - \beta + 2}.$$

(iv) *The composition* $T := \operatorname{prox}_{f,\lambda_2}^p \circ (\beta \operatorname{prox}_{g,\lambda_1}^p \oplus (1 - \beta) \operatorname{Id})$ *is pointwise almost* α-*firmly nonexpansive at all points* $y \in \Omega := \operatorname{argmin} f \cap \operatorname{argmin} g$ *on* G *with constant*

$$\widehat{\alpha} = \frac{\alpha_\beta + \alpha_c - 2\alpha_\beta\alpha_c}{\frac{c}{2}(1 - \alpha_\beta - \alpha_c + \alpha_\beta\alpha_c) + \alpha_\beta + \alpha_c - 2\alpha_\beta\alpha_c}$$

and violation bounded above by

$$\widehat{\epsilon} = (1 + \beta)\epsilon_c + \beta\epsilon_c^2,$$

(35)

where α_β, α_c, *and* ϵ_c *are the constants in part* (iii) *and Theorem* 21(ii) *respectively. If* (G,d) *is a* $\mathrm{CAT}(\kappa)$ *space, then* T *is pointwise asymptotically* α-*firmly nonexpansive at all* $y \in \Omega$ *with constant*

$$\widehat{\alpha} = \frac{1}{2 - \overline{\alpha}_\beta},$$

where $\overline{\alpha}_\beta$ *is the constant in part* (iii),

(v) *If (G, d) is symmetric perpendicular, the projected gradient operator*

$$T := P_C \circ \left(\beta \operatorname{prox}_{g,\lambda}^p \oplus (1 - \beta) \operatorname{Id} \right)$$

is pointwise almost α-firmly nonexpansive at all points $y \in \Omega := C \cap \operatorname{argmin} g$ on G with

$$\alpha_{PG} = \frac{1}{\frac{c}{2}(1 - \alpha_\beta) + 1} \quad \text{and violation bounded above by} \quad \epsilon_{PG} = \epsilon_c \beta. \qquad (36)$$

If (G, d) is a $\mathrm{CAT}(\kappa)$ space, then T is pointwise asymptotically α-firmly nonexpansive at all $y \in \Omega$ with constant

$$\alpha_{PG} = \frac{1}{2 - \overline{\alpha}_\beta},$$

where $\overline{\alpha}_\beta$ is the constant in part (iii),

Proof (i). By Theorem 21(ii), the operators $\operatorname{prox}_{f,\lambda}^p$ and $\operatorname{prox}_{g,\lambda}^p$ are almost quasi α-firmly nonexpansive with constants $\alpha_c = \frac{c(c-1)}{c(c-1)+2}$ and violation bounded above by $\epsilon_c = \frac{2-c}{c-1}$. The operator T is almost α-firmly nonexpansive at all points $y \in \Omega$ on G with

$$\alpha_\circ = \frac{2\alpha_c - 2\alpha_c^2}{\frac{c}{2}(1 - 2\alpha_c + \alpha_c^2) + 2\alpha_c - 2\alpha_c^2}$$

and violation satisfying $1 + \epsilon_\circ = (1 + \epsilon_c)^2$ by Proposition 10. A short calculation yields

$$\alpha_\circ = \frac{2(c-1)}{2c-1} \quad \text{and} \quad \epsilon_\circ = \frac{1}{(c-1)^2} - 1.$$

Taking the limit as $c \to 2$ from below yields the constant $\overline{\alpha}_\circ = 2/3$ with limiting violation $\overline{\epsilon}_\circ = 0$. The same argument as Proposition 21(iii) then shows that, when (G, d) is a $\mathrm{CAT}(\kappa)$ space, the composition of two prox mappings is pointwise asymptotically α-firmly nonexpansive at points in Ω with constant $2/3$.

(ii). Theorem 21(ii) and Proposition 10 yield pointwise almost α-firm nonexpansiveness of T at $y \in \Omega$ with constant and violation characterized by (33), where α_c and ϵ_c are given by (32), $\overline{\alpha}_0$ is the asymptotic constant of α-firm nonexpansiveness of T_0, and ϵ_0 is the upper bound of the violation on some neighborhood. (By Proposition 6(iii) if T_0 is pointwise almost α-firmly nonexpansive with constant $\alpha_0 < \overline{\alpha}_0$, then T_0 is also pointwise almost α-firmly nonexpansive with constant $\overline{\alpha}_0$.) By Theorem 21(iii) and the assumption that T_0 is pointwise asymptotically α-firmly nonexpansive at Fix T_0 with constant $\overline{\alpha}_0$, the same argument as Proposition 21(iii) establishes that, when (G, d) is a $\mathrm{CAT}(\kappa)$ space, T is pointwise asymptotically α-firmly nonexpansive at points in Ω with constant $\overline{\alpha}$ given by (34).

(iii). The first statement is an immediate application of Proposition 12 and Theorem 21(ii). When (G, d) is a $\mathrm{CAT}(\kappa)$ space, the same argument as Proposition 21(iii) establishes that Krasnoselsky–Mann relaxations of prox mappings are pointwise asymptotically α-firmly nonexpansive with constant $\overline{\alpha}_\beta$ as claimed.

(iv). This is an application of part (ii) to part (iii).

(v). This is a specialization of part (iv) when f is the indicator function of a convex set C and follows from the fact that, on symmetric perpendicular p-uniformly convex spaces, the projector is pointwise α-firmly nonexpansive at all points in C with constant $\alpha = 1/2$ (no violation) as shown in [4, Proposition 25].   □

Remark 24   Part (i) of Proposition 23 coincides with $\alpha = \frac{2}{3}$ and $\epsilon = 0$ in the classic setting with $c = 2$. In particular, the composition $P_A \circ P_B$ of two projections P_A and P_B onto convex sets A and B with $A \cap B \neq \emptyset$ is α-firmly nonexpansive at all $y \in A \cap B$ on G with $\alpha = \frac{2}{3}$ and violation $\epsilon = 0$. However, this result does not apply if the problem is infeasible i.e. $A \cap B = \emptyset$.

These properties allow us to prove the following fundamental result.

Theorem 25 (Convergence of proximal algorithms in CAT(κ) spaces) *Let (G, d) be a complete* CAT(κ) *space with $\kappa > 0$ and $f, g \colon G \to \mathbb{R} \cap \{+\infty\}$ be proper, convex, and lower semicontinuous with* $\operatorname{argmin} f \cap \operatorname{argmin} g \neq \emptyset$. *Let T denote any of the mappings in Proposition* 23. *If T satisfies*

$$d(x, \operatorname{Fix} T \cap D) \leq \mu d(x, Tx), \quad \forall x \in D, \tag{37}$$

with constant μ, then the fixed point sequence initialized from any starting point close enough to $\operatorname{Fix} T \cap D$ *is at least linearly convergent to a point in* $\operatorname{Fix} T \cap D$ *with rate $\gamma = \sqrt{1 + \epsilon - \frac{1-\alpha}{\alpha \mu^2}} < 1$, where α and ϵ are the respective constant and violation of pointwise α-firm nonexpansiveness of the fixed point mapping T as given in Proposition* 23. *The asymptotic rate of convergence is $\overline{\gamma} = \sqrt{1 - \frac{1-\overline{\alpha}}{\overline{\alpha} \mu^2}} < 1$, where $\overline{\alpha}$ is the respective constant of pointwise asymptotic α-firm nonexpansiveness of the fixed point mapping T.*

Proof As established in Theorem 21 and Proposition 23, all of the mappings covered in those results are pointwise asymptotically α-firmly nonexpansive at points in Ω with constants $\overline{\alpha} < 1$, where Ω is one of the following subsets corresponding to the respective mappings (i)–(v) in 23: (i) $\Omega = \operatorname{argmin} f \cap \operatorname{argmin} g$; (ii) $\Omega \subset \operatorname{Fix} T_0 \cap \operatorname{argmin} f$; (iii) $\Omega = \operatorname{argmin} f$; (iv) $\Omega = \operatorname{argmin} f \cap \operatorname{argmin} g$; (v) $\Omega = C \cap \operatorname{argmin} g$.

As noted in Remark 3, any CAT(κ) space is symmetric perpendicular locally, so by Theorem 21(i) and Lemma 7, in every case $\operatorname{Fix} T = \Omega$. Now, if T satisfies (37) at all points in $\operatorname{Fix} T$ on G with constant μ, it follows immediately from Proposition 20 that the fixed point iteration converges linearly to a point in $\operatorname{Fix} T \cap D$ with the given rate for all starting points close enough to $\operatorname{Fix} T \cap D$, as claimed.   □

5 Proximal splitting methods

The concrete examples provided here have been well studied for p-uniformly convex spaces with $p = c = 2$ i.e. CAT(0) spaces. The tools established in the previous sections open the door to applying these methods in CAT(κ) spaces, which is new. Since CAT(κ) spaces are p-uniformly convex with $p = 2$, to avoid confusion, we revert to the usual notation for proximal operators in the setting, namely $\operatorname{prox}_{f,\lambda}$, omitting the exponent.

Let (G, d) be a CAT(κ) space, $f_i \colon G \to G$ be proper lsc convex functions for $i = 1, 2, \dots N$. Consider the problem

$$\inf_{x \in G} \sum_{i=1}^{N} f_i(x). \tag{38}$$

Algorithm 1: Proximal splitting

Parameters   : Functions $f_1 \ldots, f_N$ and $\lambda_i > 0$ $(i = 1, 2, \ldots, N)$.

Initialization: Choose $x_0 \in G$.

for $k = 0, 1, 2, \ldots$ **do**

$$x_{k+1} = Tx_k := (\text{prox}_{f_N, \lambda_N} \circ \cdots \circ \text{prox}_{f_2, \lambda_2} \circ \text{prox}_{f_1, \lambda_1})(x_k)$$

Algorithm 2: Metric projected gradients

Parameters   : $f : G \to \mathbb{R}$, the closed set $C \subset G$, $\lambda > 0$ and $\beta \in (0, 1)$.

Initialization: Choose $x_0 \in G$.

for $k = 0, 1, 2, \ldots$ **do**

$$x_{k+1} = T_{PG}(x_k) := P_C\big((1 - \beta)\,\text{Id} \oplus \beta\,\text{prox}_{f, \lambda}\big)(x_k)$$

Applying *backward-backward splitting* to this problem yields Algorithm 1. Local linear convergence follows from Theorem 25 and the extension of Proposition 23(i) via part (ii) of the same proposition and induction, under the assumption that $\mathcal{T}_\Omega$ defined by (22)—which simplifies to (23)—is linearly metrically subregular for 0 on G with constant μ and $\Omega := \bigcap_j \text{argmin} f_j \neq \emptyset$. By Proposition 6(i), linear metric subregularity simplifies to

$$d(x, \Omega) \leq \mu d(Tx, x) \quad \forall x \in D. \tag{39}$$

Recall, in a p-uniformly convex space with $p = 2$, the Moreau–Yosida envelope of f is defined by

$$e_{f, \lambda}(x) := \inf_{y \in G}\left(f(y) + \frac{1}{2\lambda} d(x, y)^2\right).$$

The analogue to the direction of steepest descent for the Moreau–Yosida envelope in p-uniformly convex settings is

$$(1 - \beta)x \oplus \beta\,\text{prox}_{f, \lambda}(x). \tag{40}$$

Specializing problem (38) to the case $N = 2$ and $f_2 = \iota_C$, the indicator function of some closed convex set $C \subset G$ yields Algorithm 2, the analog to projected gradients in CAT(κ) space, which is the projected resolvent/projected prox iteration. Local linear convergence follows immediately from Theorem 25 under the assumption that T satisfies (30) and $\Omega := \text{argmin} f \cap C \neq \emptyset$.

Compositions of projectors in CAT(κ) spaces has been studied in [4] and [1]. We consider Algorithm 1 when the functions $f_i := \iota_{C_i}$, the indicator functions of closed convex sets

$C_i \subset G$, where (G, d) is a complete, symmetric perpendicular p-uniformly convex space with constant c. The p-proximal mapping of the indicator function is the metric projector, and so by [4, Proposition 25] these are pointwise α-firmly nonexpansive at all points in $\bigcap_i C_i$ (assuming, of course, that this is nonempty) and by [4, Lemma 10] the cyclic projections mapping

$$T_{CP} := P_{C_N} \cdot P_{C_2} P_{C_1} \tag{41}$$

is pointwise α-firmly nonexpansive at all points in $\bigcap_i C_i = \operatorname{Fix} T_{CP}$, when the intersection is nonempty, with constant $\overline{\alpha}_N = \frac{N-1}{N}$ on G. Δ- or weak convergence (no rate) to a point in $\bigcap_i C_i$ follows from [4, Theorem 27], with strong convergence whenever one of the sets is compact. If in addition $d(x, \bigcap_i C_i) \le \mu d(T_{CP} x, x)$ for all $x \in G$, where $\mu > 0$ is the rate of linear metric subregularity, then, by Theorem 25, the sequence (x_k) initialized anywhere in G converges linearly to some $x^* \in \bigcap_i C_i$.

6 Open problems

There are two obvious next steps for this work. First and foremost is to determine the requirements for quantitative convergence of proximal splitting methods for the case when the individual prox mappings do not have common fixed points—the so-called *inconsistent case*—since it is too limiting to require that the fixed points of the constituent elements of splitting methods coincide. The second item to explore is qualitative settings in which metric subregularity comes "for free". In linear settings, polyhedrality and isolated fixed points suffice to guarantee metric subregularity [7, Propositions 3I.1 and 3I.2], and this was successfully used to prove local linear convergence of the ADMM/Douglas–Rachford algorithms in a convex setting [2, Theorem 2.7]. In more general settings, the Kurdyka–Łojasiewicz (KL) property—which implies metric subregularity [6, Corollary 4 and Remark 5]—is satisfied by *semi-algebraic* functions [5]. Analogues to these properties for p-uniformly convex spaces would be very useful.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
The authors contributed equally to all results. All authors read and approved the final manuscript.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Ariza-Ruiz, D., López-Acedo, G., Nicolae, A.: The asymptotic behavior of the composition of firmly nonexpansive mappings. J. Optim. Theory Appl. **167**, 409–429 (2015)
2. Aspelmeier, T., Charitha, C., Luke, D.R.: Local linear convergence of the ADMM/Douglas–Rachford algorithms without strong convexity and application to statistical imaging. SIAM J. Imaging Sci. **9**(2), 842–868 (2016)
3. Bërdëllima, A.: Investigations in Hadamard Spaces. PhD thesis, Georg-August Universität Göttingen, Göttingen (2020)
4. Bërdëllima, A., Lauster, F., Luke, D.R.: α-firmly nonexpansive operators on metric spaces. arXiv:2104.11302 (2021)
5. Bolte, J., Daniilidis, A., Lewis, A.: The Lojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems. SIAM J. Optim. **17**, 1205–1223 (2006)
6. Bolte, J., Daniilidis, A., Ley, O., Mazet, L.: Characterizations of Lojasiewicz inequalities: subgradient flows, talweg, convexity. Trans. Am. Math. Soc. **362**(6), 3319–3363 (2010)
7. Dontchev, A.L., Rockafellar, R.T.: Implicit Functions and Solution Mappings, 2nd edn. Springer, Dordrecht (2014)
8. Izuchukwu, C., Ugwunnadi, G.C., Mewomo, O.T., Khan, A.R., Abbas, M.: Proximal-type algorithms for split minimization problem in P-uniformly convex metric spaces. Numer. Algorithms **82**(3), 909–935 (2019)
9. Kuwae, K.: Jensen's inequality on convex spaces. Calc. Var. Partial Differ. Equ. **49**(3–4), 1359–1378 (2014)
10. Luke, D.R., Teboulle, M., Thao, N.H.: Necessary conditions for linear convergence of iterated expansive, set-valued mappings. Math. Program. **180**, 1–31 (2018)
11. Luke, D.R., Thao, N.H., Tam, M.K.: Quantitative convergence analysis of iterated expansive, set-valued mappings. Math. Oper. Res. **43**(4), 1143–1176 (2018)
12. Naor, A., Silberman, L.: Poincaré inequalities, embeddings, and wild groups. Compos. Math. **147**(5), 1546–1572 (2011)
13. Ohta, S.: Convexities of metric spaces. Geom. Dedic. **125**, 225–250 (2007)