On the convergence of the Krasnoselskij iteration for strictly pseudocontractive operators

Diego Deplano, Sergio Grammatico, Mauro Franceschelli

***Abstract*— We study the convergence of the nonlinear Krasnoselskij iteration $x(k+1) = (1-\theta)x(k) + \theta\mathsf{T}(x(k))$ in real vector spaces of finite dimension equipped with a p-norm, which is relevant for stability analysis and distributed computation in several discrete-time dynamical systems. Specifically, we provide sufficient conditions for the convergence of the Krasnoselskij iteration, derived via implications between the strict pseudocontractivity of the operator T and the nonexpansiveness of $(1-\theta)\mathsf{Id} + \theta\mathsf{T}$. Interestingly, it turns out that strict pseudocontractivity of T is necessary for the Euclidean norm ($p = 2$) only; not necessary for non-Euclidean norms ($p \neq 2$); sufficient for any finite norm $p \in (1, \infty)$; not sufficient for the taxi-cab norm ($p = 1$) and the supremum norm ($p = \infty$). We numerically verify the above results in the context of recurrent neural networks and multi-agent systems with nonlinear Laplacian dynamics.**

I. Introduction

Consider the Banach-Picard iteration [1, Eq. (1.69)] in the form of discrete-time dynamical system:

$$\boldsymbol{x}(k+1) = \mathsf{T}_\theta(\boldsymbol{x}(k)) = (1-\theta)\boldsymbol{x}(k) + \theta\mathsf{T}(\boldsymbol{x}(k)), k \in \mathbb{N}, \quad (1)$$

where $\theta \in (0, 1)$ and $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ such that $\text{fix}(\mathsf{T}) \neq \emptyset$.

One of the first convergence results dates back to 1955 and it is due to Krasnoselskii [2][3, Theorem 6.4.1], who proved convergence of $\boldsymbol{x}(k)$ to a fixed point when T is nonexpansive and $\theta = \frac{1}{2}$ for *uniformly convex* spaces [4, Definition 1.8]. More than 10 years later, Edelstein in [5] extended this result to $\theta \in (0, 1)$ and *strictly convex* spaces[1] [4, Definition 1.10]. In 1976, the convergence results for the Banach-Picard iteration in (1) in uniformly/strictly convex spaces were extended to general Banach spaces by Ishikawa [6, Theorem 1], see also [3, Theorem 6.4.3]. By limiting their analysis to Hilbert spaces, Marino and Xu in [7] proved that the iteration in (1) converges also when the map T is κ-strictly pseudocontractive and $\theta < 1 - \kappa$. Moreover, for linear maps in Hilbert spaces, it has been recently proven that κ-strictly pseudocontractivity of T is both necessary and sufficient for the convergence of the Krasnoselkij iteration, given $\theta < 1 - \kappa$ [8, Theorem 1]. Marino and Xu in [7] also posed the currently open question: "Does this result hold also in Banach spaces which are uniformly convex?". Since then, many authors have provided different answers to this question by considering several iteration schemes and sets of assumptions [9]–[15]. From a general mathematical perspective, the convergence problem is a fixed-point problem [4], or equivalently, a zero finding problem [1]. For example, consensus in nonlinear multi-agent systems is equivalent to finding a collective state in the kernel of the nonlinear Laplacian operator [16]–[18]. Variations of the Krasnoselskij fixed-point iteration have also been adopted to design distributed algorithms for computing fixed-points in networks [19]–[23], splitting methods in distributed convex optimization [24]–[28], aggregative game theory [29], [30], monotone dynamical systems [31]–[35], and so on.

The main contribution of this paper is showing that, in real Banach spaces $\mathcal{S}_p = (\mathbb{R}^n, \|\cdot\|_p)$ of finite-dimension n equipped with a p-norm for $p \in (1, \infty)$, the Krasnoselskij iteration converges if $\theta^{r-1} < (1-\kappa)/c_p$ (see Theorems 1-2) where $r = \min\{p, 2\}$ and $c_p \geq 1$ is a constant that depends on p, whose best (smallest) value is characterized in Lemma 5. We apply this result to verify alternative set of assumptions on which some of the state-of-the-art results are built upon. It turns out that some of these assumptions can not actually hold, as they imply that the constant c_p takes smaller values than those provided in Lemma 5, which is not possible. Finally, we discuss two examples of application, namely recurrent neural networks and consensus via nonlinear Laplacian dynamics.

II. Notation and Preliminaries

The set of real and integer numbers are denoted by $\mathbb{R}$ and $\mathbb{Z}$, and their restriction to nonnegative and positive values are denoted with $\mathbb{R}_{\geq 0}$, $\mathbb{N}$ and $\mathbb{R}_{>0}$, $\mathbb{N}_+$, respectively. Matrices $M \in \mathbb{R}^{n\times n}$ are denoted by uppercase letters, vectors $\boldsymbol{v} \in \mathbb{R}^n$ by bold letters, scalars $s \in \mathbb{R}$ by lowercase letters, while sets and spaces $\mathcal{S}$ are denoted by uppercase calligraphic letters. We denote by $\mathbf{0}_n$ and $\mathbf{1}_n$ the vector of zeros and ones of dimension n, respectively. Mappings $\mathsf{T} : \mathcal{X}_1 \to \mathcal{X}_2$ between two spaces $\mathcal{X}_1, \mathcal{X}_2$ are usually denoted with block capital letters; for instance, the linear operator associated to the identity matrix I is defined by $\mathsf{Id} : \boldsymbol{x} \mapsto I\boldsymbol{x}$. When $\mathcal{X}_2 \equiv \mathbb{R}$, block lowercase letters are used instead, e.g., $\mathsf{t} : \mathcal{X} \to \mathbb{R}$. Given a self-mapping $\mathsf{T} : \mathcal{X} \to \mathcal{X}$, $\text{fix}(\mathsf{T}) = \{\boldsymbol{x} \in \mathcal{X} \mid \mathsf{T}(\boldsymbol{x}) = \boldsymbol{x}\}$ denotes the set of its fixed points and $\text{zer}(\mathsf{T}) = \{\boldsymbol{x} \in \mathcal{X} \mid \mathsf{T}(\boldsymbol{x}) = \mathbf{0}\}$ denotes the set of its zeros.

Work supported by project e.INS- Ecosystem of Innovation for Next Generation Sardinia (cod. ECS 00000038) funded by the Italian Ministry for Research and Education (MUR) under the National Recovery and Resilience Plan (NRRP) - MISSION 4 COMPONENT 2, "From research to business" INVESTMENT 1.5, "Creation and strengthening of Ecosystems of innovation" and construction of "Territorial R& D Leaders".

Diego Deplano and Mauro Franceschelli are with the DIEE, University of Cagliari, 09123, Italy. Emails: {diego.deplano, mauro.franceschelli}@unica.it.

Sergio Grammatico is with the Delft Center for Systems and Control, TU Delft, The Netherlands. Email: s.grammatico@tudelft.nl.

A. Operator-Theoretic definitions in real Banach spaces

A *normed vector space* is a pair $(\mathcal{X}, \|\cdot\|)$ where $\mathcal{X}$ is a vector space and $\|\cdot\|$ is a norm on $\mathcal{X}$, which induces in the natural way a metric, i.e., a notion of distance: the distance between two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ is given by $\|\boldsymbol{x}-\boldsymbol{y}\|$. We focus on the real vector space $\mathcal{X} = \mathbb{R}^n$ of finite dimension $n \in \mathbb{N}$ equipped with a p-norm $\|\cdot\|_p$, for $p \in [1, \infty]$. We denote these spaces with $\mathcal{S}_p = (\mathbb{R}^n, \|\cdot\|_p)$, which are Banach spaces since every finite-dimensional normed vector space is complete as in [4, Def. 1.5 and Rem. 2 on page 7]. The only Hilbert space is for $p = 2$, for which the inner product is well defined by $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = \boldsymbol{x}^\top \boldsymbol{x} = \|\boldsymbol{x}\|_2^2$. We now introduce some duality concepts of real Banach spaces.

Definition 1. *[4, Def. 1.11] The dual of $\mathcal{S}_p$ is denoted by $\mathcal{S}_p^* = ((\mathbb{R}^n)^*, \|\cdot\|_p^*)$ and it is defined as follows:*

- *The dual space $(\mathbb{R}^n)^*$ is the set of all continuous linear mappings $\mathsf{L}_{\boldsymbol{z}} : \mathbb{R}^n \to \mathbb{R}$ uniquely defined by a vector $\boldsymbol{z} \in \mathbb{R}^n$, such that $\mathsf{L}_{\boldsymbol{z}}(\boldsymbol{x}) = \boldsymbol{z}^\top \boldsymbol{x}$;*
- *The dual norm is defined by $\|\mathsf{L}_{\boldsymbol{z}}\|_p^* = \sup_{\|\boldsymbol{x}\|_p \le 1} \left|\boldsymbol{z}^\top \boldsymbol{x}\right|$.*

The concept of a duality mapping was introduced by Beurling and Livingston in [36]. We define its generalized form in the case of spaces $\mathcal{S}_p$ by means of the Holder's conjugate numbers.

Definition 2. *Two elements $p, q \in [1, \infty]$ are Holder's conjugate if $\frac{1}{p} + \frac{1}{q} = 1$ where, by convention, $1/\infty = 0$.*

Definition 3. *[9, Page 1, Paragraph 2] The generalized duality mapping $\mathsf{J}_r : \mathcal{S}_p \to 2^{\mathcal{S}_p^*}$ with $\mathcal{S}_p = (\mathbb{R}^n, \|\cdot\|_p)$ is defined[2] for any $\boldsymbol{x} \in \mathbb{R}^n$ by*

$$\mathsf{J}_r(\boldsymbol{x}) = \{\mathsf{L}_{\boldsymbol{z}} : \mathbb{R}^n \to \mathbb{R} \mid \boldsymbol{x}^\top \boldsymbol{z} = \|\boldsymbol{x}\|_p^r, \, \|\boldsymbol{x}\|_p^{r-1} = \|\boldsymbol{z}\|_q, \, \boldsymbol{z} \in \mathbb{R}^n\},$$

where $r \in [1, 2]$ and $p, q \in [1, \infty]$ are Holder's conjugate.

We now recall some useful results in Lemmas 1-2-3, a proof of which can be found in the appendix.

Lemma 1. *Let $p, q \in [1, \infty]$ be Holder's conjugate, then the dual norm $\|\cdot\|_p^*$ is given by $\|\mathsf{L}_{\boldsymbol{z}}\|_p^* = \|\boldsymbol{z}\|_q$.*

Lemma 2. *Let $\mathcal{S}_p = (\mathbb{R}^n, \|\cdot\|_p)$ with $p \in [1, \infty]$. Given $r = \min\{p, 2\}$, the generalized duality mapping is not empty $\mathsf{J}_r(\boldsymbol{x}) \neq \emptyset$ for any $\boldsymbol{x} \in \mathcal{X}$ and consists of (at least) one linear mapping $\mathsf{L}_{\mathsf{j}_r(\boldsymbol{x})} \in \mathsf{J}_r(\boldsymbol{x})$ with $\mathsf{j}_r(\boldsymbol{x}) \in \mathbb{R}^n$ given by*

$$\mathsf{j}_r(\boldsymbol{x}) := \begin{cases} \operatorname{sign}(\boldsymbol{x}) \circ |\boldsymbol{x}|^{p-1} / \|\boldsymbol{x}\|_p^{p-r} & \text{if} \quad p \in [1, \infty) \\ \boldsymbol{x} \circ \boldsymbol{x}_\infty / \mathbf{1}^\mathsf{T} \boldsymbol{x}_\infty & \text{if} \quad p = \infty \end{cases}, \quad (2)$$

where $\circ$ denotes the Hadamard product and where

$$\boldsymbol{x}_\infty = [\cdots, x_{\infty,i} \cdots]^\top, \quad x_{\infty,i} = \begin{cases} 1 & \text{if} \quad |x_i| = \max |\boldsymbol{x}|, \\ 0 & \text{otherwise}. \end{cases}$$

Lemma 3. *Consider a Banach space $\mathcal{S}_p = (\mathbb{R}^n, \|\cdot\|_p)$ with $p \in [1, \infty]$. Given $r = \min\{p, 2\}$, the generalized duality mapping is single-valued $\mathsf{J}_r(\boldsymbol{x}) = \{\mathsf{L}_{\mathsf{j}_r(\boldsymbol{x})}\}$ for any $\boldsymbol{x} \in \mathcal{X}$ if and only if $p \in (1, \infty)$.*

[2]Special case of [9] for spaces $\mathcal{S}_p$, where $\|\mathsf{L}_{\boldsymbol{z}}\|_p^* = \|\boldsymbol{z}\|_q$ by Lemma 1.

Among nonlinear mappings, the classes of nonexpansive mappings and pseudocontractions play a pivotal role. Let us define these properties in the context of Banach spaces $\mathcal{S}_p$.

Definition 4. *[37] Consider the space $\mathcal{S}_p = (\mathbb{R}^n, \|\cdot\|_p)$ and a mapping $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$. If for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ it holds*

$$\|\mathsf{T}(\boldsymbol{x}) - \mathsf{T}(\boldsymbol{y})\|_p \le \ell \, \|\boldsymbol{x} - \boldsymbol{y}\|_p, \quad (3)$$

then the mapping is is called:

- *ℓ-contractive (ℓ-C) if $\ell \in (0, 1)$;*
- *nonexpansive (NE) if $\ell = 1$.*

Definition 5. *[9] Consider the space $\mathcal{S}_p = (\mathbb{R}^n, \|\cdot\|_p)$ and a mapping $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$. If for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ there exists $\mathsf{L}_r := \mathsf{L}_{\mathsf{j}_r(\boldsymbol{x}-\boldsymbol{y})} \in \mathsf{J}_r(\boldsymbol{x}-\boldsymbol{y})$ with $r = \min\{p, 2\}$ such that*

$$\begin{aligned} \mathsf{L}_r(\mathsf{T}(\boldsymbol{x}) - \mathsf{T}(\boldsymbol{y})) \le {} & \|\boldsymbol{x}-\boldsymbol{y}\|_p^r \\ & - \frac{1-\kappa}{r} \|\boldsymbol{x} - \boldsymbol{y} - (\mathsf{T}(\boldsymbol{x}) - \mathsf{T}(\boldsymbol{y}))\|_p^r, \end{aligned} \quad (4)$$

then the mapping is called:

- *κ-strictly pseudocontractive (κ-sPC) if $\kappa \in (0, 1)$;*
- *pseudocontractive (PC) if $\kappa = 1$.*

We note that it holds: ℓ-C $\Rightarrow$ NE $\Rightarrow$ κ-sPC $\Rightarrow$ PC.

III. Main results

Our first main result in Theorem 1 characterizes the relation between the nonexpansiveness of the Krasnoselkij iteration operator and the strict pseudocontractivity of the corresponding mapping, which is instrumental to obtain sufficient conditions for its convergence, our second main result, Theorem 2. More precisely, given a κ-sPC mapping T, Theorems 1-2 ensure that the Krasnoselksji iteration T_θ in (1) converges for

$$\theta^{r-1} < \frac{1-\kappa}{c_p}, \qquad \forall p \in (1, \infty) \quad (5)$$

where $r = \min\{p, 2\}$ and c_p is a constant that depends on the space $\mathcal{S}_p$ of interest. The existence of such constant c_p follows by the Reich's inequality [38], which is given in the following Lemma 4 in the special case of $\mathcal{S}_p$ spaces, by exploiting the results of Honh-Kun Xu in [39]. We then prove in Lemma 5 that the value of the constant c_p provided by Lemma 4 is the best possible.

Lemma 4. *[38][39, Eqs. (3.5)' and (3.8)' in Corollary 2] Consider the Banach space $\mathcal{S}_p$ with $p \in (1, \infty)$. Given two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and the (unique) dual linear mapping $\mathsf{L}_{\mathsf{j}_r(\boldsymbol{x})} \in \mathsf{J}_r(\boldsymbol{x})$, then it holds that*

$$\|\boldsymbol{x} + \boldsymbol{y}\|_p^r \le \|\boldsymbol{x}\|_p^r + r \mathsf{L}_{\mathsf{j}_r(\boldsymbol{x})}(\boldsymbol{y}) + c_p \|\boldsymbol{y}\|_p^r, \quad (6)$$

where $r = \min\{p, 2\}$ and

$$c_p = \begin{cases} p - 1 & \text{if } p \ge 2 \\ (1 + t_p^{p-1})(1 + t_p)^{1-p} & \text{if } p \in (1, 2) \end{cases}, \quad (7)$$

with t_p being the unique solution of the following equation

$$(p-2)t^{p-1} + (p-1)t^{p-2} = 1.$$

In [39, Remark1] is stated that the constant c_p as give in Lemma 4 is the best possible, i.e., there are not smaller values of c_p for which inequality in eq. (6) holds for any pair of points $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. However, a proof of this claim is not provided in [39], and also in the references therein. We now provide such proof, which – to the best of our knowledge – is a novel technical result.

Lemma 5. *The constant c_p in* (7) *is the best possible and is such that $c_p > 1$ for all $p \neq 2$.*

Proof: The proof consists in providing pair of points $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ such that the inequality in eq. (6), reported next

$$\|\boldsymbol{x}+\boldsymbol{y}\|_p^r \leq \|\boldsymbol{x}\|_p^r + r\mathsf{L}_{\mathsf{j}_r(\boldsymbol{x})}(\boldsymbol{y}) + c_p\|\boldsymbol{y}\|_p^r,$$

holds strictly for c_p as in eq. (7). Let us first consider the case of $p \in [2, \infty)$ and the points: $\boldsymbol{x} = [a,\ a]^\top$, $\boldsymbol{y} = [-1,\ 1]^\top$. Since $r = \min\{p, 2\} = 2$, one can compute the generalized duality mapping $\mathsf{J}_r = \mathsf{J}_2$ of $\boldsymbol{x}$ by means of Lemma 2, which is unique by Lemma 3, namely

$$\mathsf{J}_2(\boldsymbol{x}) = \left\{\mathsf{L}_{\mathsf{j}_2(\boldsymbol{x})} : \mathbb{R}^n \to \mathbb{R} \mid \mathsf{j}_2(\boldsymbol{x}) = 2^{\frac{p}{p-2}} a \begin{bmatrix}1 & 1\end{bmatrix}^\top\right\}.$$

One can verify that the above is correct by $\boldsymbol{x}^\top \mathsf{j}_2(\boldsymbol{x}) = \|\boldsymbol{x}\|_p^2 = \|\mathsf{j}_2(\boldsymbol{x})\|_q^2 = \sqrt[p]{4}a^2$, where $q = p/(p-1)$ is the Holder's conjiugate of p according to Definition 2. Let us now compute the following norms $\|\boldsymbol{x}\|_p = \sqrt[p]{2}a$, $\|\boldsymbol{x}+\boldsymbol{y}\|_p = \sqrt[p]{(a+1)^p + (a-1)^p}$, $\|\boldsymbol{y}\|_p = \sqrt[p]{2}$, and also

$$\mathsf{L}_{\mathsf{j}_2(\boldsymbol{x})}(\boldsymbol{y}) = \boldsymbol{y}^\top \mathsf{j}_2(\boldsymbol{x}) = 2^{\frac{p}{p-2}} a[-1,\ 1][1,\ 1]^\top = 0.$$

Substituting the above into inequality (6) of Lemma 4 yields $(\sqrt[p]{(a+1)^p + (a-1)^p})^2 \leq \sqrt[p]{4}a^2 + c_p\sqrt[p]{4}$ and therefore

$$c_p \geq f_p(a) := \sqrt[p]{4^{-1}}\left((\sqrt[p]{(a+1)^p + (a-1)^p})^2 - \sqrt[p]{4}a^2\right),$$

for all $p \geq 2$. The maximum of the function $f_p(a)$ provides a lower bound to the minimum value of c_p. It can be verified that $f_p(a)$ is monotonically increasing w.r.t. a, and thus the maximum is attained in the limit of a to infinity[1]:

$$c_p \geq \lim_{a\to\infty} f_p(a) = p - 1, \quad \forall p \geq 2.$$

[1]Verifiable by symbolic analysis tools, e.g., Wolfram - Alpha Pro engine.

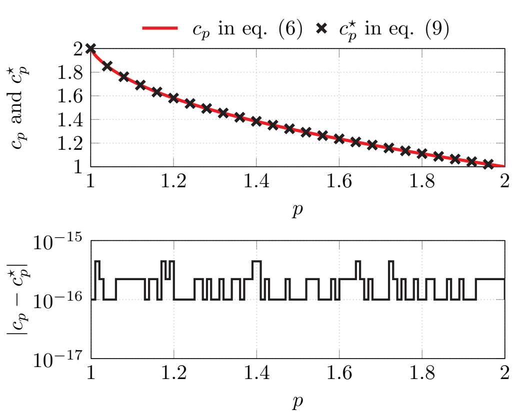

Fig. 1. Comparison between the value of c_p provided by Lemma 4 with the lower bound $c_p^\star$ provided within the proof of Lemma 5.

Let us now consider the case of $p \in (1, 2)$ and the points: $\boldsymbol{x} = [1,\ -1]^\top$, $\boldsymbol{y} = [0,\ a]^\top$. Since $r = \min\{p, 2\} = p$, one can compute the generalized duality mapping $\mathsf{J}_r = \mathsf{J}_p$ of $\boldsymbol{x}$ by means of Lemma 2, which is unique by Lemma 3, namely

$$\mathsf{J}_p(\boldsymbol{x}) = \left\{\mathsf{L}_{\mathsf{j}_p(\boldsymbol{x})} : \mathbb{R}^n \to \mathbb{R} \mid \mathsf{j}_p(\boldsymbol{x}) = \begin{bmatrix}1 & -1\end{bmatrix}^\top\right\}.$$

One can verify that the above is correct by $\boldsymbol{x}^\top \mathsf{j}_p(\boldsymbol{x}) = \|\boldsymbol{x}\|_p^p = \|\mathsf{j}_2(\boldsymbol{x})\|_q^{p/(p-1)} = 2$, where $q = p/(p-1)$ is the Holder's conjiugate of p according to Definition 2. Let us now compute the following norms: $\|\boldsymbol{x}\|_p = \sqrt[p]{2}$, $\|\boldsymbol{x}+\boldsymbol{y}\|_p = \sqrt[p]{1 + (a-1)^p}$, $\|\boldsymbol{y}\|_p = a$, and also

$$\mathsf{L}_{\mathsf{j}_2(\boldsymbol{x})}(\boldsymbol{y}) = \boldsymbol{y}^\top \mathsf{j}_2(\boldsymbol{x}) = [a,\ 0][1,\ -1] = -a.$$

Substituting the above into inequality (6) of Lemma 4 yields $1 + (a-1)^p \leq 2 - pa + c_p a^p$ and therefore

$$c_p \geq f_p(a) := a^{-p}\left((a-1)^p - 1 + pa\right), \quad \forall p \in (1, 2).$$

The maximum of the function $f_p(a)$ provides a lower bound $c_p^\star$ to the minimum value of c_p, which occurs at the (unique) value of a_p that makes the derivative zero, i.e.,

$$a_p: \quad f_p'(a_p) = 0, \qquad \Rightarrow \qquad c_p^\star := f_p(a_p). \tag{8}$$

Since there is not a closed form solution of a_p, we computed it numerically for $p \in (1, 2)$ and compared the lower-bound $c_p^\star$ in eq. (8) with the value of c_p given in eq. (7) of Lemma 4. The results are displayed in Figure 1, which demonstrates that the lower bound $c_p^\star$ is equal to c_p (up to numerical precision). Since $c_p > 1$, $\forall p \neq 2$, the thesis holds. ■

Our main result about the relation between the pseudo-contractivity of a mapping and the nonexpansiveness of the Krasnoselkij iteration map is given next.

Theorem 1. *Consider a Banach space $\mathcal{S}_p$ with $p \in [1, \infty]$, a mapping $\mathsf{T} : \mathcal{X} \to \mathcal{X}$ and the following properties:*

(a) T *is* κ-sPC *for some* $\kappa \in (0, 1)$*;*
(b) $\mathsf{T}_\theta = (1-\theta)\mathsf{Id} + \theta\mathsf{T}$ *is* NE *for some* $\theta \in (0, 1]$*;*

Given $r = \min\{p, 2\}$, the following statements hold:

$(s1)$ $(a) \Leftarrow (b)$ *with $\theta \geq 1 - \kappa$ if and only if $p = 2$;*
$(s2)$ $(a) \Rightarrow (b)$ *with $\theta^{r-1} \leq (1-\kappa)/c_p$ if $p \in (1, \infty)$;*
$(s3)$ $(a) \not\Rightarrow (b)$ *for any $\theta \in (0, 1)$ if $p \in \{1, \infty\}$.*

where c_p is given in eq. (7) *of Lemma 4.*

Proof: We recall that a continuous linear mapping $\mathsf{L}_r \in \mathsf{J}_r(\boldsymbol{x})$ is such that the following properties hold:

- $\mathsf{L}_r(\theta\boldsymbol{x}) = \theta\mathsf{L}_r(\boldsymbol{x})$ for all $\theta \in \mathbb{R}$ and $\boldsymbol{x} \in \mathbb{R}^n$;
- $\mathsf{L}_r(\boldsymbol{x} \pm \boldsymbol{y}) = \mathsf{L}_r(\boldsymbol{x}) \pm \mathsf{L}_r(\boldsymbol{y})$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$.

We first prove statement $(s1)$. For each $p \in (1, \infty)$, let $r = \min\{p, 2\}$ and $\mathsf{L}_r \in \mathsf{J}_r(\boldsymbol{x} - \boldsymbol{y})$, then it holds:

$$\begin{aligned}
\mathsf{L}_r(\mathsf{T}(\boldsymbol{x}) - \mathsf{T}(\boldsymbol{y})) &= \frac{1}{\theta}\mathsf{L}_r\big(\mathsf{T}_\theta(\boldsymbol{x}) - (1-\theta)\boldsymbol{x} - \mathsf{T}_\theta(\boldsymbol{y}) + (1-\theta)\boldsymbol{y}\big) \\
&\overset{(i)}{=} \frac{1}{\theta}\Big[\mathsf{L}_r\big(\mathsf{T}_\theta(\boldsymbol{x}) - \mathsf{T}_\theta(\boldsymbol{y})\big) - (1-\theta)\mathsf{L}_r(\boldsymbol{x} - \boldsymbol{y})\Big] \\
&\overset{(ii)}{=} -\frac{1}{\theta}\Big[\mathsf{L}_r\big(\mathsf{T}_\theta(\boldsymbol{y}) - \mathsf{T}_\theta(\boldsymbol{x})\big) + (1-\theta)\|\boldsymbol{x} - \boldsymbol{y}\|_p^r\Big]
\end{aligned}$$

$$\overset{(iii)}{\le} -\frac{1}{\theta}\Big[\frac{1}{r}\Big(\|\boldsymbol{x}-\boldsymbol{y}-(\mathsf{T}_\theta(\boldsymbol{x})-\mathsf{T}_\theta(\boldsymbol{y}))\|_p^r - \|\boldsymbol{x}-\boldsymbol{y}\|_p^r - c_p\|\mathsf{T}_\theta(\boldsymbol{x})-\mathsf{T}_\theta(\boldsymbol{y})\|_p^r\Big) + (1-\theta)\|\boldsymbol{x}-\boldsymbol{y}\|_p^r\Big]$$
$$\overset{(iv)}{\le} -\frac{1}{\theta}\Big[(1-\theta-\frac{1+c_p}{r})\|\boldsymbol{x}-\boldsymbol{y}\|_p^r + \frac{\theta^r}{r}\|\boldsymbol{x}-\boldsymbol{y}-(\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y}))\|_p^r$$
$$=(1-\frac{1}{\theta}+\frac{1+c_p}{\theta r})\|\boldsymbol{x}-\boldsymbol{y}\|_p^r - \frac{\theta^{r-1}}{r}\|\boldsymbol{x}-\boldsymbol{y}-(\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y}))\|_p^r$$
$$\overset{(v)}{\le}\|\boldsymbol{x}-\boldsymbol{y}\|_p^r - \frac{\theta^{r-1}}{r}\|\boldsymbol{x}-\boldsymbol{y}-(\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y}))\|_p^r$$
$$\overset{(vi)}{\le}\|\boldsymbol{x}-\boldsymbol{y}\|_p^r - \frac{1-\kappa}{r}\|\boldsymbol{x}-\boldsymbol{y}-(\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y}))\|_p^r,$$

where (i) holds by linearity of L; (ii) follows by Definition 3; (iii) follows by the tight bound eq. (6) in Lemmas 4-5, which implies $-\mathsf{L}_{\mathrm{j}_r(\boldsymbol{a})}(\boldsymbol{b}) \le -\frac{1}{r}\left(\|\boldsymbol{a}+\boldsymbol{b}\|_p^r - \|\boldsymbol{a}\|_p^r - c_p\|\boldsymbol{b}\|_p^r\right)$ for any $\boldsymbol{a},\boldsymbol{b}\in\mathbb{R}^n$; (iv) follows by nonexpansiveness of T_θ; (v) holds if and only if $c_p \le r-1 \le 1$, which holds if and only if $p=r=2$ by Lemma 5 (otherwise, $c_p > 1$ for $p \neq 2$); (vi) holds for $\theta^{r-1} \ge 1-\kappa$. This proves statement $(s1)$.

We now prove statement $(s2)$ by letting $\mathsf{L}_r(\boldsymbol{x}-\boldsymbol{y})$:

$$||\mathsf{T}_\theta(\boldsymbol{x})-\mathsf{T}_\theta(\boldsymbol{y})||_p^r = \|(1-\theta)(\boldsymbol{x}-\boldsymbol{y})+\theta(\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y}))\|_p^r$$
$$=\|\boldsymbol{x}-\boldsymbol{y}-\theta(\boldsymbol{x}-\boldsymbol{y}-(\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y})))\|_p^r$$
$$\overset{(i)}{\le}\|\boldsymbol{x}-\boldsymbol{y}\|_p^r - r\theta\mathsf{L}_r(\boldsymbol{x}-\boldsymbol{y}-(\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y}))) + c_p\|\theta(\boldsymbol{x}-\boldsymbol{y}-(\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y})))\|_p^r$$
$$=(1-r\theta)\|\boldsymbol{x}-\boldsymbol{y}\|_p^r + r\theta\mathsf{L}_r(\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y})) + \theta^r c_p\|\boldsymbol{x}-\boldsymbol{y}-(\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y}))\|_p^r$$
$$\overset{(ii)}{\le}\|\boldsymbol{x}-\boldsymbol{y}\|_p^r - \theta(1-\kappa)\|\boldsymbol{x}-\boldsymbol{y}-(\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y}))\|_p^r + \theta^r c_p\|\boldsymbol{x}-\boldsymbol{y}-(\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y}))\|_p^r$$
$$\overset{(iii)}{\le}\|\boldsymbol{x}-\boldsymbol{y}\|_p^r$$

where (i) holds by Lemma 4; (ii) holds since map T is κ-sPC; (iii) holds for $-\theta(1-\kappa)+\theta^r c_p \le 0$ which implies $\theta^{r-1} \le (1-\kappa)/c_p$. This proves statement $(s2)$. On the other hand: in the limit of $p \to 1^+$, it holds that $r=p$ and $c_p \to 2^-$ (see Figure 1), and therefore $1 \le (1-\kappa)/2$, which is in contrast with $\kappa \ge 0$; in the limit of $p\to\infty$, it holds that $r=2$ and $c_p\to\infty$ (see eq. (7)), and therefore $\theta \le 0$, which is a contradiction. This proves statement $(s3)$. ■

Corollary 1. *For the Euclidean norm, i.e. for $p=2$, statements $(s1)$ and $(s2)$ of Theorem 1 implythat $(a) \Leftrightarrow (b)$ with $\theta = 1-\kappa$.*

A result related to the above corollary has been recently proven for linear maps in the Hilbert space $(\mathbb{R}^n, \|\cdot\|_{2,P})$ where $\|\cdot\|_{2,P} = \sqrt{x^\top P x}$ and where P is a symmetric and positive definite matrix [8, Lemma 5]. We are now in the position to prove our second main result for the convergence of the Krasnoselsij iteration. In particular, the tight bound on the constant c_p ensured by Lemma 5 allows to determine the largest value of θ anticipated in eq. (5) ensuring convergence of the iteration.

Theorem 2. *Consider a Banach space $\mathcal{S}_p$ and a map $\mathsf{T}:\mathbb{R}^n\to\mathbb{R}^n$ such that $\mathrm{fix}(\mathsf{T})\neq\emptyset$. Given $r=\min\{p,2\}$, the following statements hold:*

- *[6] The iteration in* (1) *converges if* T *is* NE *for any* $p\in[1,\infty]$ *and* $\theta<1$*;*
- *The iteration in* (1) *converges if* T *is* κ-sPC *for any* $p\in(1,\infty)$ *and* $\theta^{r-1} < (1-\kappa)/c_p$*, with* c_p *as in* (7)*.*

Proof: The first statement is due to [6, Theorem 1]. By our Theorem 1, for $p\in(1,\infty)$ it holds that for $\theta^* = ((1-\kappa)/c_p)^{1/(r-1)} \in (0,1)$ the map $\mathsf{T}_{\theta^*} = (1-\theta^*)\mathsf{Id}+\theta^*\mathsf{T}$ is NE. Consequently, map T_θ ruling the iteration in (1) can be equivalently written as $\mathsf{T}_\theta = (1-\frac{\theta}{\theta^*})\mathsf{Id} + \frac{\theta}{\theta^*}\mathsf{T}_{\theta^*}$. Thus, T_θ can be seen as the Krasnoselskij iteration of the nonexpansive map T_{θ^*} with coefficient θ/θ^*, which is known to converge for $\theta/\theta^* < 1$ by [6, Theorem 1], i.e.,

$$\frac{\theta}{\theta^*} < 1 \Rightarrow \frac{\theta}{((1-\kappa)/c_p)^{1/(r-1)}} < 1 \Rightarrow \frac{\theta^{r-1}}{(1-\kappa)/c_p} < 1,$$

completing the proof. ■

Remark 1. *Theorem 2 provides an upper bound to the value of θ, namely $\theta < (1-\kappa)/c_p$, which limits the convergence speed of the Krasnoselkij iteration. Higher values of the strict pseudocontracivity constant κ imply smaller admissible values of θ, i.e., slower convergence. The same holds for the constant c_p, whose best (smallest) value is characterized in Lemmas 4-5. In particular, the minimum value $c_p = 1$ (yielding faster convergence) holds only for $p=2$, while for any other $p\neq 2$ then $c_p > 1$ (yielding slower convergence).*

The following corollary directly follows from Theorem 2 and [1, Theorem 5.14(iii)], which ensures the convergence of the Krasnoselksij-Mann iteration given by

$$\boldsymbol{x}(k+1) = (1-\theta_k)\boldsymbol{x}(k) + \theta_k\mathsf{T}(\boldsymbol{x}(k)), \tag{9}$$

when T is sPC and where the sequence $(\theta_k)_{k\in\mathbb{N}}$ is such that $0\le\theta_k\le\theta_{\mathrm{MAX}}<\infty$ for all $k\in\mathbb{N}$, for some θ_{MAX}, and such that $\lim_{k\to\infty}\theta_k = 0$ with $\sum_{k=0}^\infty \theta_k = \infty$.

Corollary 2. *Consider a Banach space $\mathcal{S}_p$ and a map $\mathsf{T}:\mathbb{R}^n\to\mathbb{R}^n$ such that $\mathrm{fix}(\mathsf{T})\neq\emptyset$. Given $r=\min\{p,2\}$, the following statements hold:*

- *Iteration* (9) *converges if* T *is* NE *and* $p\in[1,\infty]$*;*
- *Iteration* (9) *converges if* T *is* κ-sPC *and* $p\in(1,\infty)$.

IV. Comparison with the state-of-the-art and applications to dynamical systems

In this section, we discuss the applicability of the results in [13], [14], [40] limited to finite spaces $\mathcal{S}_p = (\mathbb{R}^n,\|\cdot\|_p)$ with $p\in[2,\infty)$. Let us recall the general form of the Reich inequality in (6) in Lemma 4, as originally formulated in [38] and then recalled in [13, Lemma 2.3] and [14, Lemma 1.5]:

$$\|\boldsymbol{x}+\boldsymbol{y}\|_p^2 \le \|\boldsymbol{x}\|_p^2 + 2\mathsf{L}_{\mathrm{j}_r(\boldsymbol{x})}(\boldsymbol{y}) + \max\{\|\boldsymbol{x}\|_p, 1\}\|\boldsymbol{y}\|_p\beta(\|\boldsymbol{y}\|_p),$$

where $\beta:\mathbb{R}_{>0}\to\mathbb{R}_{>0}$ is a continuous function such that

$$\lim_{t\to 0^+}\beta(t) = 0, \qquad \beta(ct)\le c\beta(t), \quad \forall c\ge 1.$$

Chidume and Su in [13, Lemma 3.2] and also Sahu and Petrusel [40] have based their result on the assumption "$\beta(t) \leq t$", while Cholamjiak and Suntai in [14] made the following assumption "$\beta(t) \leq 2t$". Under these assumptions, for vectors $\|\boldsymbol{x}\|_p \leq 1$ the Reich inequality becomes

$$\beta(t) \leq t \quad \Rightarrow \quad \|\boldsymbol{x}+\boldsymbol{y}\|_p^2 \leq \|\boldsymbol{x}\|_p^2 + 2\mathsf{L}_{\mathsf{j}_r(\boldsymbol{x})}(\boldsymbol{y}) + \|\boldsymbol{y}\|_p^2, \tag{10}$$

$$\beta(t) \leq 2t \quad \Rightarrow \quad \|\boldsymbol{x}+\boldsymbol{y}\|_p^2 \leq \|\boldsymbol{x}\|_p^2 + 2\mathsf{L}_{\mathsf{j}_r(\boldsymbol{x})}(\boldsymbol{y}) + 2\|\boldsymbol{y}\|_p^2. \tag{11}$$

Comparing the above with the bound in eq. (6) in Lemma 4, which is the best possible by Lemma 5, one can verify that:

- Assumption in (10) of [13], [40] holds only for $p = 2$;
- Assumption in (11) of [14] holds only for $p \in [2, 3]$;

We provide an example corroborating these statements, let

$$\boldsymbol{x} = \tfrac{1}{\sqrt[p]{2}} \begin{bmatrix} 1 & 1 \end{bmatrix}^\top, \quad \boldsymbol{y} = \tfrac{p^{-p}}{\sqrt[p]{2}} \begin{bmatrix} -1 & 1 \end{bmatrix}^\top, \tag{12}$$

and compute $\|\boldsymbol{x}\|_p^2 = 1$, $\|\boldsymbol{y}\|_p^2 = p^{-2p}$, $\mathsf{L}_{\mathsf{j}_2(\boldsymbol{x})}(\boldsymbol{y}) = 0$, $\|\boldsymbol{x}+\boldsymbol{y}\|_p^2 = (\sqrt[p]{(p^{-p}+1)^p + (p^{-p}-1)^p})^2 / \sqrt[p]{4}$. For this pair of points, assumptions in eqs. (10)-(11) do not hold for $p > 2$ and $p > 3$, respectively, as it is shown in Figure 2. The plot on the right of Figure 2 displays the sign between the distance of $\|\boldsymbol{x}+\boldsymbol{y}\|_p^2$ to the bounds, which takes value of -1 when the bound holds and $+1$ when it does not. On the other hand, our results are consistent with those of Zhang in [9] where, however, the explicit values of the constant c_p determining the upper bound on θ are not given.

A. Application example: Recurrent Neural Networks

Consider the following continuous-time recurrent neural-network, usually referred to as the *firing-rate* model:

$$\dot{\boldsymbol{x}}(t) = -\mathsf{F}(\boldsymbol{x}(t)) := -\boldsymbol{x}(t) + \underbrace{\Phi(A\boldsymbol{x}(t) + \boldsymbol{b})}_{\mathsf{T}(\boldsymbol{x}(t))} \tag{13}$$

where $\boldsymbol{x}, \boldsymbol{b} \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times n}$ and $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ is an activation mapping applied entrywise, i.e., $\Phi(\boldsymbol{x}) = (\phi(x_1), \dots, \phi(x_n))$ with $\phi : \mathbb{R} \to \mathbb{R}$. In this example, we consider the case that ϕ is a LeakyReLU activation function, i.e., $\phi(x) = \max\{x, ax\}$, where we select $a = 0.1$. Stationary point of F are also fixed points of T, i.e., $\operatorname{zer}(\mathsf{F}) \equiv \operatorname{fix}(\mathsf{T})$. In order to find a stationary point, one can apply the forward step method $\boldsymbol{x}(k+1) = (\mathsf{Id} - \theta\mathsf{F})\boldsymbol{x}(k)$, which leads to the Krasnoselkij iteration in (1):

$$\boldsymbol{x}(k+1) = \mathsf{T}_\theta \boldsymbol{x}(k) = (1-\theta)\boldsymbol{x}(k) + \theta\mathsf{T}(\boldsymbol{x}(k)).$$

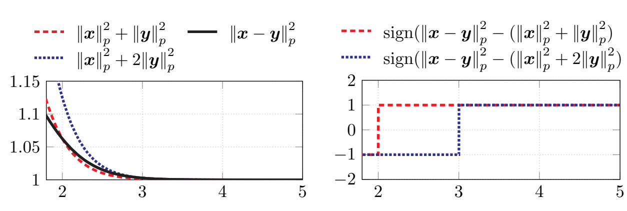

Fig. 2. Empirical validation for the pair of points in eq. (12) of the assumption of Chidume and Su in eq. (10) [13] (red dashed curves) and of Cholamjiak and Suntai in eq. (11) [14] (blue dotted curves).

Now, let us consider the following matrix:

$$A = \begin{bmatrix} -0.12 & -0.63 & -0.33 & +0.21 \\ +0.12 & +0.15 & +0.03 & +0.09 \\ -0.63 & -0.30 & +0.36 & +1.65 \\ -0.90 & -5.79 & +0.45 & -6.39 \end{bmatrix}.$$

Note that for $\theta = 1$ the iteration surely does not converge because the matrix A has an eigenvalue outside the unitary disk, that is $\lambda = -6.3912$. For $p = 2$, one can verify that the operator T is not κ-strictly pseudocontractive w.r.t. $\|\cdot\|_2$ for any $\kappa \in (0,1)$, as exemplified by following choice of vectors $\boldsymbol{x} = [3.34, -4.82, 4.87, 1.05]$, $\boldsymbol{y} = [3.42, -1.86, 0.18, -1.25]$. Instead, for $p = 4$ we have empirically verified that the operator T is κ-strictly pseudocontractive w.r.t. $\|\cdot\|_4$ with $\kappa \approx 0.972$. Thus, the forward step method converges for $\theta < (1-\kappa)/(p-1) \approx 0.0093$ according to Theorem 2.

B. Application example: Nonlinear Laplacian dynamics

Consider a network of n agents with discrete-time dynamics seeking consensus via the following nonlinear protocol

$$\boldsymbol{x}(k+1) = (1-\theta_k)\boldsymbol{x}(k) + \theta_k \underbrace{(-\boldsymbol{x}(k) + f(L\boldsymbol{x}(k)))}_{\mathsf{T}(\boldsymbol{x}(k))},$$

where $L \in \mathbb{R}^{n \times n}$ is the Laplacian matrix associated to the graph describing the interactions among the agents, and $f : \mathbb{R}^n \to \mathbb{R}^n$ is a nonlinear operator such that $f = [\dots, f_i, \dots]$ with $f_i : \mathbb{R} \to \mathbb{R}$. In this example, we consider the case all f_i for $i = 1, \dots, n$ are the same saturating function: $f_i(x) = (1 - e^{-mx})/(1 + e^{-mx})$ with $m \geq 0$. Note that for $m = 2$ the above reduces to the hyperbolic tangent function and for $m \to \infty$ it approximates the sign function; from now on we consider $m = 10$. For $p = 2$, one can verify that the operator T is not κ-strictly pseudocontractive w.r.t. $\|\cdot\|_2$ for any $\kappa \in (0,1)$ as exemplified by the following choice of vectors: $\boldsymbol{x} = [0.2, -1.0, 1.0, 0.3]$, $yb = [-0.1, -1.3, -0.6, -0.6]$. Instead, for $p > 5$ we have empirically verified that the operator T is κ-strictly pseudocontractive w.r.t. $\|\cdot\|_p$ for some $\kappa \in (0,1)$. Thus, the agents could employ a vanishing time-varying sequence θ_k as in eq. (9) and converge to a consensus according to Corollary 2. In this case, it is not necessary that the agents know the constant κ of strict pseudocontractivity of the operator T, i.e., they do not need global information about the system to guarantee convergence.

V. Discussion and future directions

The class of κ-strictly pseudocontractive operators has attracted attention because it leads to generalized convergence results of fixed-point iterations. This work provides the tightest condition for the convergence of the Krasnoselskij iteration on strict pseudocontractive operators for finite and real Banach spaces $\mathcal{S}_p = (\mathbb{R}^n, \|\cdot\|_p)$ with $p \in [1, \infty]$. Notably, for $p = 2$, κ-strict pseudocontractivity of a mapping T is necessary and sufficient for nonexpansiveness of the averaged mapping $\mathsf{T}_\theta = (1-\theta)\mathsf{Id} + \theta\mathsf{T}$ with $\theta = 1 - \kappa$. In contrast, for $p \neq 2$ this is not the case, but there exists a sufficiently small θ such that T_θ becomes nonexpansive given that T is strict pseudocontractive. Finally, for $p \in \{1, \infty\}$, there exists no such a θ. The weakening of the link between strict pseudocontractivity of T and nonexpansiveness of T_θ when transitioning from the Hilbert space $\mathcal{S}_2$ to Banach spaces $\mathcal{S}_p$ with $p \neq 2$ – and its complete loss for $p \in \{1, \infty\}$ – motivates two research directions: 1) consider weighted norms $\|\cdot\|_{p,P}$ for symmetric, positive semidefinite matrices P, and 2) search for alternative properties to strict pseudocontractivity.

REFERENCES

[1] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, 2nd ed. Cham, Switzerland: Springer, 2017.

[2] M. A. Krasnosel'skii, "Two comments on the method of successive approximations," *Usp. Math. Nauk*, vol. 10, pp. 123–127, 1955.

[3] R. P. Agarwal, D. O'Regan, and D. Sahu, *Fixed point theory for Lipschitzian-type mappings with applications*. Springer, 2009, vol. 6.

[4] V. Berinde and F. Takens, *Iterative approximation of fixed points*. Springer, 2007, vol. 1912.

[5] M. Edelstein, "A remark on a theorem of M.A. Krasnoselski," *Amer. Math. Monthly*, vol. 73, pp. 509–510, 1966.

[6] S. Ishikawa, "Fixed points and iteration of a nonexpansive mapping in a Banach space," *Proceedings of the American Mathematical Society*, vol. 59, no. 1, pp. 65–71, 1976.

[7] G. Marino and H.-K. Xu, "Weak and strong convergence theorems for strict pseudo-contractions in hilbert spaces," *Journal of Mathematical Analysis and Applications*, vol. 329, no. 1, pp. 336–346, 2007.

[8] G. Belgioioso, F. Fabiani, F. Blanchini, and S. Grammatico, "On the convergence of discrete-time linear systems: A linear time-varying mann iteration converges iff its operator is strictly pseudocontractive," *IEEE Control Systems Letters*, vol. 2, no. 3, pp. 453–458, 2018.

[9] H. Zhang and Y. Su, "Convergence theorems for strict pseudo-contractions in q-uniformly smooth Banach spaces," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 71, no. 10, pp. 4572–4580, 2009.

[10] H. Zhang and Y. Su, "Strong convergence theorems for strict pseudo-contractions in q-uniformly smooth Banach spaces," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 70, no. 9, pp. 3236–3242, 2009.

[11] G. Cai and C. song Hu, "Strong convergence theorems of a general iterative process for a finite family of λ-strict pseudo-contractions in q-uniformly smooth Banach spaces," *Computers & Mathematics with Applications*, vol. 59, no. 1, pp. 149–160, 2010.

[12] H. Y. Zhou, "Convergence theorems for λ-strict pseudo-contractions in q-uniformly smooth Banach spaces," *Acta Mathematica Sinica, English Series*, vol. 26, no. 4, pp. 743–758, 2010.

[13] C. Chidume and N. Shahzad, "Weak convergence theorems for a finite family of strict pseudocontractions," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 72, no. 3-4, pp. 1257–1265, 2010.

[14] P. Cholamjiak and S. Suantai, "Weak convergence theorems for a countable family of strict pseudocontractions in Banach spaces," *Fixed Point Theory and Applications*, pp. 1–16, 2010.

[15] P. Cholamjiak and S. Suantai, "Strong convergence for a countable family of strict pseudocontractions in q-uniformly smooth Banach spaces," *Computers & Mathematics with Applications*, vol. 62, no. 2, pp. 787–796, 2011.

[16] D. Deplano, M. Franceschelli, and A. Giua, "Lyapunov-free analysis for consensus of nonlinear discrete- time multi-agent systems," in *IEEE Conference on Decision and Control (CDC)*, 2018, pp. 2525–2530.

[17] D. Deplano, M. Franceschelli, and A. Giua, "A nonlinear perron–frobenius approach for stability and consensus of discrete-time multi-agent systems," *Automatica*, vol. 118, p. 109 025, 2020.

[18] R. Bonetto and H. J. Kojakhmetov, "Nonlinear laplacian dynamics: Symmetries, perturbations, and consensus," *arXiv preprint arXiv:2206.04442*, 2022.

[19] X. Li and L. Xie, "Distributed algorithms for computing a fixed point of multi-agent nonexpansive operators," *Automatica*, vol. 122, p. 109 286, 2020.

[20] N. Bastianello, R. Carli, L. Schenato, and M. Todescato, "Asynchronous distributed optimization over lossy networks via relaxed ADMM: Stability and linear convergence," *IEEE Transactions on Automatic Control*, vol. 66, no. 6, pp. 2620–2635, 2020.

[21] F. Andrade, M. A. Figueiredo, and J. Xavier, "Distributed Banach–Picard iteration for locally contractive maps," *IEEE Transactions on Automatic Control*, vol. 68, no. 2, pp. 1275–1280, 2022.

[22] X. Li, M. Meng, and L. Xie, "Dot and dop: Linearly convergent algorithms for finding fixed points of multi-agent operators," *IEEE Transactions on Automatic Control*, 2023.

[23] X. Nian, D. Liu, and F. Li, "Continuous-time distributed algorithm for seeking fixed points of multi-agent quasi-nonexpansive operators," *IEEE Transactions on Control of Network Systems*, 2023.

[24] P. Giselsson and S. Boyd, "Linear convergence and metric selection for douglas-rachford splitting and ADMM," *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 532–544, 2016.

[25] L. Pavel, "Distributed gne seeking under partial-decision information over networks via a doubly-augmented operator splitting approach," *IEEE Transactions on Automatic Control*, vol. 65, no. 4, pp. 1584–1597, 2019.

[26] E. Dall'Anese, A. Simonetto, and A. Bernstein, "On the convergence of the inexact running Krasnosel'skij–mann method," *IEEE Control Systems Letters*, vol. 3, no. 3, pp. 613–618, 2019.

[27] L. Guo, X. Shi, S. Yang, and J. Cao, "Disa: A dual inexact splitting algorithm for distributed convex composite optimization," *IEEE Transactions on Automatic Control*, 2023.

[28] N. Bastianello, D. Deplano, M. Franceschelli, and K. H. Johansson, "Robust online learning over networks," *IEEE Transactions on Automatic Control*, vol. 70, no. 2, pp. 933–946, 2025.

[29] S. Grammatico, F. Parise, M. Colombino, and J. Lygeros, "Decentralized convergence to nash equilibria in constrained deterministic mean field control," *IEEE Transactions on Automatic Control*, vol. 61, no. 11, pp. 3315–3329, 2015.

[30] S. Grammatico, "Dynamic control of agents playing aggregative games with coupling constraints," *IEEE Transactions on Automatic Control*, vol. 62, no. 9, pp. 4537–4548, 2017.

[31] S. Jafarpour, A. Davydov, and F. Bullo, "Non-euclidean contraction theory for monotone and positive systems," *IEEE Transactions on Automatic Control*, 2022.

[32] D. Deplano, M. Franceschelli, and A. Giua, "Novel stability conditions for nonlinear monotone systems and consensus in multiagent networks," *IEEE Transactions on Automatic Control*, vol. 68, no. 12, pp. 7028–7040, 2023.

[33] D. Deplano, M. Franceschelli, and A. Giua, "Stability of nonexpansive monotone systems and application to recurrent neural networks," *IEEE Control Systems Letters*, vol. 8, pp. 1793–1798, 2024.

[34] Y. Kawano and Y. Hosoe, "Contraction analysis of differentially positive discrete-time stochastic systems," *IEEE Transactions on Automatic Control*, 2025.

[35] R. Parasnis, M. Franceschetti, and B. Touri, "A perron-frobenius theorem for strongly aperiodic stochastic chains," *IEEE Transactions on Automatic Control*, 2025.

[36] A. Beurling and A. Livingston, "A theorem on duality mappings in Banach spaces," *Arkiv för Matematik*, vol. 4, no. 5, pp. 405–411, 1962.

[37] F. E. Browder, "Nonexpansive nonlinear operators in a banach space," *Proceedings of the National Academy of Sciences*, vol. 54, no. 4, pp. 1041–1044, 1965.

[38] S. Reich, "An iterative procedure for constructing zeros of accretive sets in Banach spaces," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 2, no. 1, pp. 85–92, 1978.

[39] H.-K. Xu, "Inequalities in Banach spaces with applications," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 16, no. 12, pp. 1127–1138, 1991.

[40] D. Sahu and A. Petruşel, "Strong convergence of iterative methods by strictly pseudocontractive mappings in Banach spaces," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 74, no. 17, pp. 6012–6023, 2011.

Appendix

A. Proof of Lemma 1

By means of the Hölder's inequality it holds,

$$\left|\boldsymbol{y}^\top \boldsymbol{x}\right| = \left|\sum_i^n x_i y_i\right| \leq \sum_i^n |x_i y_i| \leq \|\boldsymbol{x}\|_p \|\boldsymbol{y}\|_q,$$

which leads to an upper bound to the norm,

$$\|\mathsf{L}_{\boldsymbol{y}}\|_p^* = \sup_{\|\boldsymbol{x}\|_p \leq 1} |\mathsf{L}_{\boldsymbol{y}}(\boldsymbol{x})| \leq \sup_{\|\boldsymbol{x}\|_p \leq 1} \|\boldsymbol{x}\|_p \|\boldsymbol{y}\|_q \leq \|\boldsymbol{y}\|_q.$$

To prove the converse inequality, we should estimate the supremum from below. We start by considering $p \in (1, \infty)$ and the vector $\tilde{\boldsymbol{x}} = [\tilde{x}_1, \ldots, \tilde{x}_n]^\mathsf{T}$ defined component-wise by $\tilde{x}_i = |y_i|^{q-2} y_i \|\boldsymbol{y}\|_q^{1-q}$. By simple manipulation one can verify that $\|\tilde{\boldsymbol{x}}\|_p = 1$, i.e., $\tilde{\boldsymbol{x}}$ belongs to the constraint set $\|\boldsymbol{x}\|_p \leq 1$ of the supremum function, and thus we can write

$$\|\mathsf{L}_{\boldsymbol{y}}\|_p^* \geq \left|\boldsymbol{y}^\top \tilde{\boldsymbol{x}}\right| = \left|\sum_{i=1}^n \tilde{x}_i y_i\right| = \frac{\sum_{i=1}^n |y_i|^q}{\|\boldsymbol{y}\|_q^{q-1}} = \frac{\|\boldsymbol{y}\|_q^q}{\|\boldsymbol{y}\|_q^{q-1}} = \|\boldsymbol{y}\|_q.$$

Whereas, for $p = \infty$ we consider the vector of ones $\tilde{\boldsymbol{x}} = \mathbf{1}$ which is such that $\|\tilde{\boldsymbol{x}}\|_\infty = 1$. Thus, we get

$$\|\mathsf{L}_{\boldsymbol{y}}\|_p^* \geq \left|\boldsymbol{y}^\top \tilde{\boldsymbol{x}}\right| = \left|\sum_{i=1}^n \tilde{x}_i y_i\right| = \sum_{i=1}^n \tilde{x}_i |y_i| = \left|\sum_{i=1}^n y_i\right| = \|\boldsymbol{y}\|_1.$$

Finally, for $p = 1$ we let I be the set of indexes such that $|x_i| = \max|\boldsymbol{x}|$, i.e., $I = \{i \in [1, n] : |y_i| = \max|\boldsymbol{y}|\}$, we let $|I|$ be the cardinality of the set and consider the vector $\tilde{\boldsymbol{x}}$ defined component-wise by

$$\tilde{x}_i = \begin{cases} 1/|I| & \text{if} \quad i \in I \\ 0 & \text{otherwise} \end{cases},$$

which clearly satisfies $\|\tilde{\boldsymbol{x}}\|_1 = 1$. Thus, we get

$$\|\mathsf{L}_{\boldsymbol{y}}\|_p^* \geq \left|\boldsymbol{y}^\top \tilde{\boldsymbol{x}}\right| = \left|\sum_{i=1}^n \tilde{x}_i y_i\right| = \left|\sum_{i \in I} \frac{y_i}{|I|}\right| = \sum_{i \in I} \frac{|y_i|}{|I|} = \|\boldsymbol{y}\|_\infty.$$

Since $\|\boldsymbol{y}\|_q \leq \|\mathsf{L}_{\boldsymbol{y}}\|_p^* \leq \|\boldsymbol{y}\|_q$, for all p, the proof is completed.

B. Proof of Lemma 2

Given $r = \min\{p, 2\}$, we are going to prove that $\boldsymbol{y} = \mathsf{j}_r(\boldsymbol{x})$ as in (2) belongs to $\mathsf{J}_r(\boldsymbol{x})$ for any $\boldsymbol{x} \in \mathcal{X}$. To do so, we need to verify the following two conditions:

a) $\|\boldsymbol{x}\|_p^r = \boldsymbol{x}^\mathsf{T} \boldsymbol{y}$;
b) $\|\boldsymbol{x}\|_p^{r-1} = \|\boldsymbol{y}\|_q$.

We go through the proof case by case:

- *Condition a) for* $p \in [1, \infty)]$:

$$\boldsymbol{x}^\mathsf{T} \boldsymbol{y} = \frac{\boldsymbol{x}^\mathsf{T}(\operatorname{sign}(\boldsymbol{x}) \circ |\boldsymbol{x}|^{p-1})}{\|\boldsymbol{x}\|_p^{p-r}} = \frac{|\boldsymbol{x}|^\mathsf{T} |\boldsymbol{x}|^{p-1}}{\|\boldsymbol{x}\|_p^{p-r}} = \\ = \frac{\sum_{i=1}^n |x_i|^p}{\|\boldsymbol{x}\|_p^{p-r}} = \frac{\|\boldsymbol{x}\|_p^p}{\|\boldsymbol{x}\|_p^{p-r}} = \|\boldsymbol{x}\|_p^r.$$

- *Condition b) for* $p \in [1, \infty)$:

$$\|\boldsymbol{y}\|_q^r = \left\|\frac{\operatorname{sign}(\boldsymbol{x}) \circ |\boldsymbol{x}|^{p-1}}{\|\boldsymbol{x}\|_p^{p-r}}\right\|_q^r = \\ = \left[\frac{\left\|\boldsymbol{x}^{p-1}\right\|_q}{\|\boldsymbol{x}\|_p^{p-r}}\right]^r = \left[\frac{\left(\sum_{i=1}^n |x_i|^{(p-1)q}\right)^{1/q}}{\|\boldsymbol{x}\|_p^{p-r}}\right]^r = \\ = \left[\frac{\left(\sum_{i=1}^n |x_i|^p\right)^{\frac{p-1}{p}}}{\|\boldsymbol{x}\|_p^{p-r}}\right]^r = \left[\frac{\|\boldsymbol{x}\|_p^{p-1}}{\|\boldsymbol{x}\|_p^{p-r}}\right]^r = \|\boldsymbol{x}\|_p^r.$$

- *Condition a) for* $p = \infty$ such that $r = 2$: let n_∞ be the number of entries of $\boldsymbol{x}$ such that $|x_i| = \max|\boldsymbol{x}|$, then it holds

$$\boldsymbol{x}^\mathsf{T} \boldsymbol{y} = \frac{\boldsymbol{x}^\mathsf{T}(\boldsymbol{x} \circ \boldsymbol{x}_\infty)}{\mathbf{1}^\mathsf{T} \boldsymbol{x}_\infty} = \frac{(\boldsymbol{x} \circ \boldsymbol{x})^\mathsf{T} \boldsymbol{x}_\infty}{\mathbf{1}^\mathsf{T} \boldsymbol{x}_\infty} = \frac{(\boldsymbol{x}^2)^\mathsf{T} \boldsymbol{x}_\infty}{\mathbf{1}^\mathsf{T} \boldsymbol{x}_\infty} = \\ = \frac{n_\infty \max \boldsymbol{x}^2}{n_\infty} = \|\boldsymbol{x}\|_\infty^2,$$

- *Condition b) for* $p = \infty$ such that $r = 2$: let n_∞ be the number of entries of $\boldsymbol{x}$ such that $|x_i| = \max|\boldsymbol{x}|$, then

$$\|\boldsymbol{y}\|_1^2 = \left\|\frac{\boldsymbol{x} \circ \boldsymbol{x}_\infty}{\mathbf{1}^\mathsf{T} \boldsymbol{x}_\infty}\right\|_1^2 = \frac{1}{n_\infty^2} \|\boldsymbol{x} \circ \boldsymbol{x}_\infty\|_1^2 = \\ = \frac{1}{n_\infty^2} \left[\sum_{i=1}^n |x_i x_{\infty,i}|\right]^2 = \frac{n_\infty^2}{n_\infty^2} \max \boldsymbol{x}^2 = \|\boldsymbol{x}\|_\infty^2.$$

This completes the proof.

C. Proof of Lemma 3

We first note that for $p \in (1, \infty)$ the generalized duality mapping J_r mapping is in one-to-tone relation with the so-called normalized duality mapping J_2 since $\mathsf{J}_r(\boldsymbol{x}) = \|\boldsymbol{x}\|_p^{p-2} \mathsf{J}_2(\boldsymbol{x})$ (cfr. [39]). Thus, the sufficiency of the statement is due to the strict convexity of the Banach space S_p (cfr. with Definition 1.10 and Remark 1 on Page 9 of [4]). The necessity follows from the next two counter examples: *(Case* $p = 1$ *and* $r = 1$*)* Let $\boldsymbol{x} = [1, 0]^\mathsf{T} \in \mathbb{R}^2$, then all points $\boldsymbol{y} = [1, \theta]^\mathsf{T} \in \mathbb{R}^2$ with $|\theta| \leq 1$ belongs to $\mathsf{J}_1(\boldsymbol{x})$, i.e., $\boldsymbol{y}^\mathsf{T} \boldsymbol{x} = \|\boldsymbol{x}\|_1 = \|\boldsymbol{y}\|_\infty = 1$. *(Case* $p = \infty$ *and* $r = 2$*)* Let $\boldsymbol{x} = [2, 0, 2]^\mathsf{T} \in \mathbb{R}^2$, then all points $\boldsymbol{y} = [\theta, 0, 2 - \theta]^\mathsf{T} \in \mathbb{R}^2$ with $\theta \in [0, 2]$ belongs to $\mathsf{J}_2(\boldsymbol{x})$, i.e., $\boldsymbol{y}^\mathsf{T} \boldsymbol{x} = \|\boldsymbol{x}\|_\infty^2 = \|\boldsymbol{y}\|_1^2 = 4$.