On the convergence of the Krasnoselskij iteration for strictly pseudocontractive operators

Diego Deplano†, **Sergio Grammatico**‡, **Mauro Franceschelli**†

†Department of Electrical and Electronic Engineering, University of Cagliari, Italy
‡Delft Center for Systems and Control, TU Delft, The Netherlands

IEEE European Control Conference 2025 – June 23-27 – Thessaloniki, Greece

Outline

Outline

The Krasnoselskij iteration

We study nonlinear discrete-time iterations of the following kind:

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

where:

- $\theta \in (0,1)$ is a step size;
- $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ is an operator.

What are the most general property for the operator T and the largest values of θ ensuring the convergence of the iteration?

The Krasnoselskij iteration

We study nonlinear discrete-time iterations of the following kind:

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

where:

- $\theta \in (0,1)$ is a step size;
- $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ is an operator.

What are the most general property for the operator T and the largest values of θ ensuring the convergence of the iteration?

Figure: Multi-Robot Systems

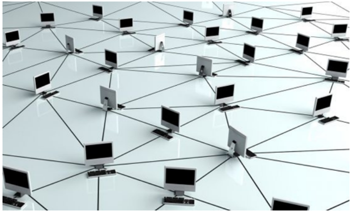


Figure: Distributed optimization

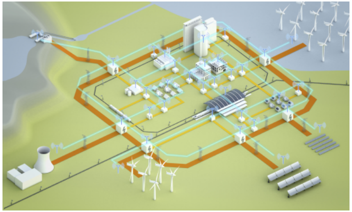


Figure: Smart grids

Figure: Neural Networks

Literature review with timeline on the iteration $x(k+1) = (1-\theta)x(k) + \theta \mathrm{T}(x(k))$

1955 - Krasnoselskij (*sufficient*):

- Uniformly convex Banach spaces
- T is nonexpansive and $\theta = 1/2$

M. A. Krasnoselskij (1955), "*Two remarks on the method of successive approximations*", Uspekhi matematicheskikh nauk.

1966 - Edelstein (*sufficient*):

- Strictly convex Banach spaces
- T is nonexpansive and $\theta \in (0,1)$

M. Edelstein (1966), "*A remark on a theorem of MA Krasnoselski*", Amer. Math. Monthly.

1976 - Ishikawa (*sufficient*):

- General Banach spaces (no convexity assumptions)
- T is nonexpansive and $\theta \in (0,1)$

S. Ishikawa (1976), "*Fixed points and iteration of a nonexpansive mapping in a Banach space*", Proc. Amer. Math. Soc.

Literature review with timeline on the iteration $x(k+1) = (1-\theta)x(k) + \theta \mathsf{T}(x(k))$

2007 - Marino and Xu (*sufficient*):

- Hilbert spaces
- T is κ-strictly pseudocontractive and $\theta \in (0, 1-\kappa)$

G. Marino and H.K. Xu (2007), "*Weak and strong convergence theorems for strict pseudo-contractions in Hilbert spaces*", Journal of Mathematical Analysis and Applications.

2018 - Belgioioso, Fabiami, Blanchini, Grammatico (*necessary and sufficient*):

- Hilbert spaces and linear maps
- T is κ-strictly pseudocontractive and $\theta \in (0, 1-\kappa)$

Belgioioso, F. Fabiani, F. Blanchini, and S. Grammatico (2018), "*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.

Is strict pseudocontractivity also sufficient in Banach spaces?

Literature review with timeline on the iteration $x(k+1) = (1-\theta)x(k) + \theta \mathsf{T}(x(k))$

2025 - D. Deplano, S. Grammatico, and M. Franceschelli (*sufficient*):

- Banach spaces with p-norms with $p \in (1, \infty)$
- T is κ-strictly pseudocontractive and $\theta^r \in \left(0, \frac{1-\kappa}{c_p}\right)$ with $r = \min\{p, 2\}$ and $c_p \geq 1$ is a constant.

D. Deplano, S. Grammatico, and M. Franceschelli (2025), "*On the convergence of the Krasnoselskij iteration for strictly pseudocontractive operators*", IEEE European Control Conference.

2025 - D. Deplano, S. Grammatico, and M. Franceschelli (*sufficient*):

- Banach spaces with ∞-norms
- T is (b, c)-enriched weakly contractive and $\theta \in \left(0, \frac{1}{b+1}\right)$.

D. Deplano, S. Grammatico, and M. Franceschelli (2025), "*Non-Euclidean Enriched Contraction Theory for Monotone Operators and Monotone Dynamical Systems*", under review, available in arXiv.

Outline

Definition

Given a real Hilbert space $(\mathbb{R}^n, \|\cdot\|)$, an operator $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ is called κ-strictly pseudocontractive if, for some $\kappa \in (0,1)$ and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, it satisfies

$$\|\mathsf{T}(\boldsymbol{x}) - \mathsf{T}(\boldsymbol{y})\|^2 \le \|\boldsymbol{x} - \boldsymbol{y}\|^2 + \kappa \|\boldsymbol{x} - \boldsymbol{y} - (\mathsf{T}(\boldsymbol{x}) - \mathsf{T}(\boldsymbol{y}))\|^2.$$

Definition

Given a real Banach space $(\mathbb{R}^n, \|\cdot\|_p)$, an operator $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ is called κ-strictly pseudocontractive if, for some $\kappa \in (0,1)$ and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, there exists a linear mapping $\mathsf{L}_{\mathrm{j}_r(\boldsymbol{x}-\boldsymbol{y})} \in \mathsf{J}_r(\boldsymbol{x} - \boldsymbol{y})$ with $r = \min\{p, 2\}$ such that

$$\mathsf{L}_{\mathrm{j}_r(\boldsymbol{x}-\boldsymbol{y})}(\mathsf{T}(\boldsymbol{x}) - \mathsf{T}(\boldsymbol{y})) \le \|\boldsymbol{x} - \boldsymbol{y}\|_p^r - \frac{1-\kappa}{r} \|\boldsymbol{x} - \boldsymbol{y} - (\mathsf{T}(\boldsymbol{x}) - \mathsf{T}(\boldsymbol{y}))\|_p^r,$$

where $\mathsf{L}_{\boldsymbol{z}}$ denotes a linear mapping $\mathsf{L}_{\boldsymbol{z}}(\boldsymbol{x}) = \boldsymbol{z}^\top \boldsymbol{x}$ and J_r is the generalized dual mapping containing all continuous linear mappings such that

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

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

Definition

Given a real Hilbert space $(\mathbb{R}^n, \|\cdot\|)$, an operator $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ is called κ-strictly pseudocontractive if, for some $\kappa \in (0,1)$ and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, it satisfies

$$\|\mathsf{T}(\boldsymbol{x}) - \mathsf{T}(\boldsymbol{y})\|^2 \le \|\boldsymbol{x} - \boldsymbol{y}\|^2 + \kappa\|\boldsymbol{x} - \boldsymbol{y} - (\mathsf{T}(\boldsymbol{x}) - \mathsf{T}(\boldsymbol{y}))\|^2.$$

Definition

Given a real Banach space $(\mathbb{R}^n, \|\cdot\|_p)$, an operator $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ is called κ-strictly pseudocontractive if, for some $\kappa \in (0,1)$ and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, there exists a linear mapping $\mathsf{L}_{\mathrm{j}_r(\boldsymbol{x}-\boldsymbol{y})} \in \mathsf{J}_r(\boldsymbol{x} - \boldsymbol{y})$ with $r = \min\{p, 2\}$ such that

$$\mathsf{L}_{\mathrm{j}_r(\boldsymbol{x}-\boldsymbol{y})}(\mathsf{T}(\boldsymbol{x}) - \mathsf{T}(\boldsymbol{y})) \le \|\boldsymbol{x} - \boldsymbol{y}\|_p^r - \frac{1-\kappa}{r}\|\boldsymbol{x} - \boldsymbol{y} - (\mathsf{T}(\boldsymbol{x}) - \mathsf{T}(\boldsymbol{y}))\|_p^r,$$

where $\mathsf{L}_{\boldsymbol{z}}$ denotes a linear mapping $\mathsf{L}_{\boldsymbol{z}}(\boldsymbol{x}) = \boldsymbol{z}^\top \boldsymbol{x}$ and J_r is the generalized dual mapping containing all continuous linear mappings such that

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

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

Main convergence result in Banach spaces equipped with a p-norm with $p \in (1, \infty)$

Theorem 2 and Lemma 5 in [R1]

Consider a real Banach space $(\mathbb{R}^n, \|\cdot\|_p)$ and a map $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ such that $\text{fix}(\mathsf{T}) \neq \varnothing$. Then, the Krasnoselskij iteration

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

converges if T is κ-strictly pseudocontractive for any $p \in (1, \infty)$ and

$$\theta^{r-1} < (1-\kappa)/c_p, \qquad \text{with} \qquad r = \min\{p, 2\},$$

where the best possible constant c_p is given by:

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

with t_p being the unique solution of $(p-2)t^{p-1} + (p-1)t^{p-2} = 1$.

[R1] D. Deplano, S. Grammatico, and M. Franceschelli (2025), "*On the convergence of the Krasnoselskij iteration for strictly pseudocontractive operators*", IEEE European Control Conference.

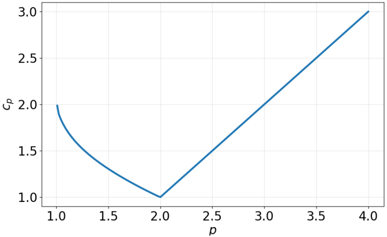
Brief discussion of the result

Good outcome - Correction of existing results in the literature:

- The results in the following two works actually hold only in Hilbert spaces:
  - C. E. Chidume and N. Shahzad (2010). "*Weak convergence theorems for a finite family of strict pseudocontractions*", Nonlinear Analysis: Theory, Methods & Applications.
  - D.R. Sahu and A. Petrusel (2011). "*Strong convergence of iterative methods by strictly pseudocontractive mappings in Banach spaces*", Nonlinear Analysis: Theory, Methods & Applications.

  The results in the following work actually hold only in Banach spaces with $[2,3]$-norms:
  - P. Cholamjiak and S. Suantai (2010). "*Weak convergence theorems for a countable family of strict pseudocontractions in Banach spaces*", Fixed Point Theory and Applications.

Bad outcome - The bound is still very mild, especially for $p \gg 2$.



An alternative property to strict pseudocontractivity

Definition

Given a real Banach space $(\mathbb{R}^n, \|\cdot\|)$, an operator $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ is called (b, c)-enriched weakly contractive if, for some $b \geq 0$, $c \in [0, b+1]$ and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, it satisfies

$$\|b(\boldsymbol{x} - \boldsymbol{y}) + \mathsf{T}(\boldsymbol{x}) - \mathsf{T}(\boldsymbol{y})\| \leq (b - c + 1)\|\boldsymbol{x} - \boldsymbol{y}\|.$$

A (b, c)-enriched weakly contractive operator is:

- Contractive if $0 = b < c < 1$;
- Weakly contractive (= nonexpansive) if $b = c = 0$;
- Enriched nonexpansive if $b > c = 0$.

Remark: We note that for $c \in (b, b+1]$, the coefficient on the right-hand side $(b - c + 1)$ is strictly less than one < 1, yielding a weak form of contractivity rather than nonexpansiveness.

Equivalency in Hilbert spaces

Proposition 1 in [R2]

Let $(\mathbb{R}^n, \|\cdot\|)$ be a real Hilbert space. An operator $T : \mathbb{R}^n \to \mathbb{R}^n$ is κ-strictly pseudocontractive if and only if it is $(b, 0)$-enriched weakly contractive with

$$b = \frac{k}{1-k}, \quad \text{or equivalently} \quad k = \frac{b}{b+1}.$$

[R2] D. Deplano, S. Grammatico, and M. Franceschelli (2025), "*Non-Euclidean Enriched Contraction Theory for Monotone Operators and Monotone Dynamical Systems*", under review, available in arXiv.

Enriched weak contractivity appears to be the perfect candidate property to generalize convergence results for the Krasnoselskij iteration in Banach spaces.

Equivalency in Hilbert spaces

Proposition 1 in [R2]

Let $(\mathbb{R}^n, \|\cdot\|)$ be a real Hilbert space. An operator $T : \mathbb{R}^n \to \mathbb{R}^n$ is κ-strictly pseudocontractive if and only if it is $(b, 0)$-enriched weakly contractive with

$$b = \frac{k}{1-k}, \quad \text{or equivalently} \quad k = \frac{b}{b+1}.$$

[R2] D. Deplano, S. Grammatico, and M. Franceschelli (2025), "*Non-Euclidean Enriched Contraction Theory for Monotone Operators and Monotone Dynamical Systems*", under review, available in arXiv.

Enriched weak contractivity appears to be the perfect candidate property to generalize convergence results for the Krasnoselskij iteration in Banach spaces.

Outline

Convergence result

Theorem 4 in [R2]

Consider an operator T that is (b,c)-enriched weakly contractive with $b>0$, $c\geq 0$. and w.r.t.

$$\|\boldsymbol{x}\|_{\infty,[\boldsymbol{\eta}]^{-1}} = \max_{i=1,\cdots,n} \frac{1}{\eta_i}|x_i|, \quad \text{for some} \quad \boldsymbol{\eta} \in \boldsymbol{\eta} \in \mathbb{R}^n_+,$$

If $c>0$, then the Krasnoselskij iteration converges to the unique fixed point $\operatorname{fix}(\mathsf{T}) = \{x^\star\}$ for

$$\theta \in \left(0, \frac{1}{b+1}\right].$$

Moreover, it holds that

$$\left\|\boldsymbol{x}(k+1) - \boldsymbol{x}^\star\right\| \leq (1-\theta c)\left\|\boldsymbol{x}(k) - \boldsymbol{x}^\star\right\|, \quad \forall k \in \mathbb{N}.$$

If $c = 0 \wedge \operatorname{fix}(\mathsf{T}) \neq \varnothing$, the iteration converges for $\theta \in (0, 1/(b+1))$ and fixed-point set is not a singleton.

[R2] D. Deplano, S. Grammatico, and M. Franceschelli (2025), "*Non-Euclidean Enriched Contraction Theory for Monotone Operators and Monotone Dynamical Systems*", under review, available in arXiv.

Frame Title

Theorem 1 in [R2]

Let $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ be a Lipschitz operator. For $b \geq 0$, $\boldsymbol{\eta} \in \mathbb{R}^n_+$, the following statements are equivalent:

(i) T is (b, c)-enriched weakly contractive w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$;

(ii) $|bI + D\mathsf{T}(\boldsymbol{x})|\boldsymbol{\eta} \leq (b - c + 1)\boldsymbol{\eta}$ for all $\boldsymbol{x} \in \mathbb{R}^n$.

Let $b^\star$ be the minimum b such that the above hold, then

$$0 \leq b^\star \leq \max\{0, \mathsf{diagL}(-\mathsf{T})\}, \quad \text{where} \quad \mathsf{diagL}(-\mathsf{T}) := \operatorname*{ess\,sup}_{\boldsymbol{x} \in \mathbb{R}^n} \max_{i \in \{1,\dots,n\}} (-\mathsf{T}(\boldsymbol{x}))_{ii}.$$

[R2] D. Deplano, S. Grammatico, and M. Franceschelli (2025), "*Non-Euclidean Enriched Contraction Theory for Monotone Operators and Monotone Dynamical Systems*", under review, available in arXiv.

Relationship with strongly monotone operators: definitions

Definition 12 and Equation (5) in [R3]

An operator $\mathsf{F} : \mathbb{R}^n \to \mathbb{R}^n$ is called c-strongly monotone w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$ if for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ it holds

$$\min_{i \in I_\infty([\boldsymbol{\eta}]^{-1}\boldsymbol{y})} \frac{(\mathsf{F}_i(\boldsymbol{x}) - \mathsf{F}(\boldsymbol{y}))(\boldsymbol{x}_i - \boldsymbol{y}_i)}{\eta_i^2} \geq c\|\boldsymbol{x} - \boldsymbol{y}\|_{\infty,[\boldsymbol{\eta}]^{-1}}.$$

where $I_\infty(\boldsymbol{v}) = \{i \in \{1, \dots, n\} \mid |\boldsymbol{v}_i| = \|\boldsymbol{v}\|_\infty\}$.

Lemma 14 in [R3]

A Lipschitz operator $\mathsf{F} : \mathbb{R}^n \to \mathbb{R}^n$ is c-strongly monotone w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$ if and only if

$$\lceil -D\mathsf{F}(\boldsymbol{x}) \rceil_{\mathsf{M}} \boldsymbol{\eta} \leq -c\boldsymbol{\eta} \text{ for almost every } \boldsymbol{x} \in \mathbb{R}^n,$$

where $\lceil M \rceil_{\mathsf{M}}$ is the Metzler majorant of the matrix M.

[R3] A. Davydov, S. Jafarpour, A.V. Proskurnikov, and F. Bullo (2024). "*Non-Euclidean monotone operator theory and applications*". Journal of Machine Learning.

Relationship with strongly monotone operators: equivalency of the two classes

Theorem 2 in [R2]

Let $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ be a Lipschitz operator. Consider the following statements:

(i) T is (b, c)-enriched weakly contractive w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$;

(ii) $\mathsf{F} := \mathsf{I} - \mathsf{T}$ is c-strongly monotone w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$.

Then, $(i) \Rightarrow (ii)$ holds for all $b \geq 0$ and $(i) \Leftarrow (ii)$ holds for $b \geq \mathrm{diagL}(-\mathsf{T})$.

[R2] D. Deplano, S. Grammatico, and M. Franceschelli (2025), "*Non-Euclidean Enriched Contraction Theory for Monotone Operators and Monotone Dynamical Systems*", under review, available in arXiv.

Relationship with strongly monotone operators: a numerical example

Consider the matrix

$$A = \frac{1}{2}\begin{bmatrix} -3 & 0 & 1 & -3 \\ 3 & -15 & -12 & -1 \\ 2 & -1 & -5 & -5 \\ -2 & 0 & -1 & -6 \end{bmatrix}.$$

and consider the Krasnoselskij iteration:

$$\boldsymbol{x}(k+1) = (1-\theta)\boldsymbol{x}(k) + \theta A\boldsymbol{x}(k).$$

The operator A : $\boldsymbol{x} \mapsto A\boldsymbol{x}$ is $(b,0)$-enriched weakly contractive w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$ with $\boldsymbol{\eta} = [0.09, 1, 0.22, 0.07]^\top$ and $b = 4$ because it holds $|bI + A|\boldsymbol{\eta} \leq (b+1)\boldsymbol{\eta}$.

The iteration converges for $\theta \in (0, \frac{1}{b+1})$, i.e., $\theta \in (0, 0.2)$ according to Th. 4 in [R2].

[R2] D. Deplano, S. Grammatico, and M. Franceschelli (2025), "*Non-Euclidean Enriched Contraction Theory for Monotone Operators and Monotone Dynamical Systems*", under review, available in arXiv.

Relationship with strongly monotone operators: a numerical example

The Krasnoselskij iteration can be re-written as the forward step method on the matrix $F = I - A$:

$$\boldsymbol{x}(k+1) = (1-\theta)\boldsymbol{x}(k) + \theta A\boldsymbol{x}(k) = \boldsymbol{x}(k) - \theta(I-A)\boldsymbol{x}(k) = \boldsymbol{x}(k) - \theta F\boldsymbol{x}(k).$$

where

$$F = \frac{1}{2}\begin{bmatrix} 5 & 0 & -1 & 3 \\ -3 & 17 & 12 & 1 \\ -2 & 1 & 7 & 5 \\ 2 & 0 & 1 & 8 \end{bmatrix}.$$

The operator $\mathsf{F} : \boldsymbol{x} \mapsto F\boldsymbol{x}$ is 0-strongly monotone w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$ with $\boldsymbol{\eta} = [0.09, 1, 0.22, 0.07]^\top$ because it holds $\lceil -D\mathsf{F}(\boldsymbol{x})\rceil_{\mathsf{M}}\boldsymbol{\eta} \le \mathbf{0}$.

The iteration converges for $\theta \in (0, \frac{1}{\mathsf{diagL(F)}})$, i.e., $\theta \in (0, 0.117)$ according to Th. 26(iii) in [R3].

[R3] A. Davydov, S. Jafarpour, A.V. Proskurnikov, and F. Bullo (2024). "*Non-Euclidean monotone operator theory and applications*". Journal of Machine Learning.

Application to zero-finding algorithms for monotone operators

Theorems 4 and 5 in [R2]

Let $\mathsf{F} : \mathbb{R}^n \to \mathbb{R}^n$ be a Lipschitz operator. If $\mathrm{zer}(\mathsf{F}) \neq \varnothing$ and F is c-strongly monotone with $c > 0$ and w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$, then the iteration ruled by

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

converges to an element of $\mathrm{zer}(\mathsf{F}) := \{\boldsymbol{x} \in \mathbb{R}^n : \mathsf{F}(\boldsymbol{x}) = \mathbf{0}\}$ for every

$$\theta \in \left(0, \frac{1}{b'+1}\right), \quad \text{with} \quad b' = \min\{b \geq 0 | (b-c+1)I - D\mathsf{F}(\boldsymbol{x})|\boldsymbol{\eta} \leq (b+1)\boldsymbol{\eta}, \boldsymbol{\eta} > 0\},$$

where $b' \leq \mathsf{diagL}(\mathsf{F}) - 1$. Moreover, the optimal convergence rate is achieved for $\theta = 1/(b^\star + 1)$ where

$$\begin{aligned}(b^\star, c^\star) = \underset{b\geq 0, c\geq 0, \boldsymbol{\eta}>0, \boldsymbol{x}\in\mathbb{R}^n}{\mathrm{argmin}} \quad & 1 - \frac{c}{b+1}, \\ \text{s.t.} \quad & |(b+1)I - D\mathsf{F}(x)|\boldsymbol{\eta} \leq (b-c+1)\boldsymbol{\eta}, \\ & c \leq b+1.\end{aligned}$$

Application to zero-finding algorithms for monotone operators

We consider operators $\mathsf{T} := \Phi \circ \mathsf{A}$ resulting from the composition of a nonlinear diagonal operator Φ with an affine operator A. This class of functions arises, for instance, in the context of training infinite-depth weight-tied neural network by solving fixed-point problems of the form

$$\boldsymbol{x}(k+1) = \mathsf{T}(\boldsymbol{x}(k)) := \Phi(A\boldsymbol{x}(k) + \boldsymbol{b}),$$

where $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ plays the role of a nonlinear activation function while A, $\boldsymbol{b}$ play the role of weight matrix and bias terms, ruling the continuous-time dynamics $\dot{\boldsymbol{x}}(t) = -\boldsymbol{x}(t) + \Phi(A\boldsymbol{x} + \boldsymbol{b})$.

Assumption

The nonlinear diagonal operator $\Phi(\boldsymbol{x}) = [\phi_1(x_1), \cdots, \phi_n(x_n)]^\top$ satisfies, for some $d_1 \le d_2$:

$$\frac{\phi_i(x) - \phi_i(x)}{x - y} \in [d_1, d_2], \qquad \forall x, y \in \mathbb{R}, x \neq y.$$

The most standard activation functions in machine learning satisfy these bounds, e.g., Leaky ReLu

$$\mathrm{LReLU}(x, \alpha) = \max\{\alpha x, x\}, \quad \text{with} \quad \alpha \in [0, 1]$$

satisfies it with $d_1 = \alpha$ and $d_2 = 1$.

Application to zero-finding algorithms for monotone operators

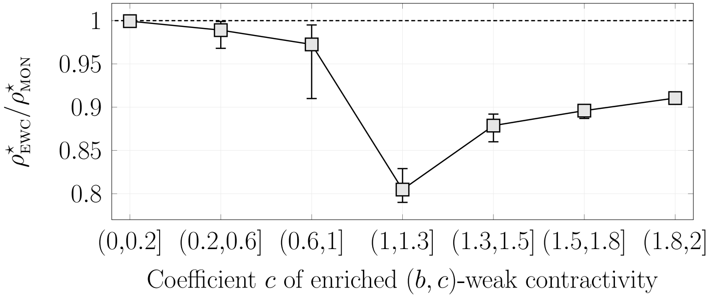


Figure: Display of the ratio $\rho^\star_{\mathrm{EWC}}/\rho^\star_{\mathrm{MON}}$ between the convergence rates averaged over 100 instances for different number of features $n \in [5, 200]$: values lower than 1 denote an improvement in the convergence rate obtained by exploiting (b, c)-enriched weak contractivity instead of c-strong monotonicity.

Relationship with strongly monotone operators: take-home message

1. **Enriched weak contractivity and strong monotonicity are equivalent classes.**

2. **Enriched weak contractivity is a more insightful property allowing to improve the convergence rate of the Krasnoselskij iteration.**

Relationship with strongly monotone operators: take-home message

1. **Enriched weak contractivity and strong monotonicity are equivalent classes.**

2. **Enriched weak contractivity is a more insightful property allowing to improve the convergence rate of the Krasnoselskij iteration.**

Relationship with monotone dynamical systems

Definition 1 and 3 in [R4]

A discrete-time dynamical system

$$\boldsymbol{x}(k+1) = \mathsf{T}(\boldsymbol{x}(k))$$

is called monotone if the operator $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ is order-preserving, i.e.,

$$\boldsymbol{x} \leq \boldsymbol{y} \Rightarrow \mathsf{T}(\boldsymbol{x}) \leq \mathsf{T}(\boldsymbol{y}), \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n.$$

Theorem 5 in [R4]

A Lipschitz operator $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ is order-preserving if and only if its Jacobian matrix is nonnegative almost everywhere, i.e.,

$$(D\mathsf{T}(\boldsymbol{x}))_{ij} \geq 0 \text{ for almost every } \boldsymbol{x} \in \mathbb{R}^n, i, j \in \{1, \ldots, n\}.$$

[R4] D. Deplano, M. Franceschelli, and A. Giua (2020), "*Nonlinear Perron–Frobenius approach for stability and consensus of discrete-time multi- agent systems*", Automatica.

Relationship with monotone dynamical systems

Definition 7 in [R2]

An operator $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ is called c-strictly $\boldsymbol{\eta}$-subhomogeneous if, for some $c \geq 0$ and for all $\boldsymbol{x} \in \mathbb{R}^n$, it satifies

$$\mathsf{T}(\boldsymbol{x} + \theta\boldsymbol{\eta}) \leq \mathsf{T}(\boldsymbol{x}) + \theta(1-c)\boldsymbol{\eta}, \qquad \forall \theta > 0.$$

Corollary 3 in [R2]

Consider a discrete-time dynamical system $\boldsymbol{x}(k+1) = \mathsf{T}(\boldsymbol{x}(k))$ and assume it is monotone. Then, the following statements are equivalent:

(i) T is (b,c)-enriched weakly contractive w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$;

(ii) T is c-strictly $\boldsymbol{\eta}$-subhomogeneous.

[R2] D. Deplano, S. Grammatico, and M. Franceschelli (2025), "*Non-Euclidean Enriched Contraction Theory for Monotone Operators and Monotone Dynamical Systems*", under review, available in arXiv.

Application to nonlinear consensus in monotone multi-agent systems

Theorem 6 in [R2]

Consider a discrete-time MAS with agents dynamics

$$x_i(k+1) = x_i(k) - \theta \sum_{i=1}^{n} a_{ij} f_{ij}\Big(x_i(k) - x_j(k)\Big),$$

If the local interaction rules f_{ij}, with $i = 1, \ldots, n$, are Lipschitz with constant $L \geq 0$ and satisfy:

1. $\partial f_{ij}/\partial x \geq 0$ for almost every $x \in \mathbb{R}$;

then the state trajectories globally, asymptotically converge to one of its equilibrium points, if any, for

$$\theta \in (0, [L \max_{i=1,\cdots,n} |\mathcal{N}_i|]^{-1}).$$

If it further holds that:

2. $f_{ij}(0) = 0$ and $f_{ij}(x) \neq 0$ a.e. in a neighborhood of 0;
3. the graph $\mathcal{G}$ has a globally reachable node;

then the MAS converges asymptotically to a consensus state.

Relationship with monotone dynamical systems: take-home message

1. **Enriched weak contractivity is equivalent to strict subhomogeneity for monotone dynamical systems.**

2. **Enriched weak contractivity allows to generalize known results about agreement in monotone multi-agent systems to non-continuously differentiable transition mappings.**

Relationship with monotone dynamical systems: take-home message

1. Enriched weak contractivity is equivalent to strict subhomogeneity for monotone dynamical systems.

2. Enriched weak contractivity allows to generalize known results about agreement in monotone multi-agent systems to non-continuously differentiable transition mappings.

Outline

Novel contributions presented in this talk

Contribution for Banach spaces equipped with a p-norms and $p \in (1, \infty)$

- Derivation of the strictest condition ensuring convergence of the Krasnoselskij iteration for κ-strictly pseudocontractive operators.

D. Deplano, S. Grammatico, and M. Franceschelli (2025), "*On the convergence of the Krasnoselskij iteration for strictly pseudocontractive operators*", IEEE European Control Conference.

Contributions for Banach spaces equipped with a ∞-norm

- Derivation of the strictest condition ensuring convergence of the Krasnoselskij iteration for (b, c)-enriched weakly contractive operators.
- Enriched weak contractivity and strong monotonicity are equivalent classes of operators, but the former allows to determine better convergence rates;
- The transition mapping of monotone dynamical systems is enriched weakly contractive if and only if T is subhomogeneous, ...

D. Deplano, S. Grammatico, and M. Franceschelli (2025), "*Non-Euclidean Enriched Contraction Theory for Monotone Operators and Monotone Dynamical Systems*", under review, available in arXiv.

Novel contributions presented in this talk

Contribution for Banach spaces equipped with a p-norms and $p \in (1, \infty)$

- Derivation of the strictest condition ensuring convergence of the Krasnoselskij iteration for κ-strictly pseudocontractive operators.

D. Deplano, S. Grammatico, and M. Franceschelli (2025), "*On the convergence of the Krasnoselskij iteration for strictly pseudocontractive operators*", IEEE European Control Conference.

Contributions for Banach spaces equipped with a ∞-norm

- Derivation of the strictest condition ensuring convergence of the Krasnoselskij iteration for (b, c)-enriched weakly contractive operators.
- Enriched weak contractivity and strong monotonicity are equivalent classes of operators, but the former allows to determine better convergence rates;
- The transition mapping of monotone dynamical systems is enriched weakly contractive if and only if T is subhomogeneous, ...

D. Deplano, S. Grammatico, and M. Franceschelli (2025), "*Non-Euclidean Enriched Contraction Theory for Monotone Operators and Monotone Dynamical Systems*", under review, available in arXiv.

Applications corroborating the theoretical findings

Zero finding algorithms for monotone operators

We derive sufficient conditions for the convergence of the forward step method applied to monotone operators and simulate it on linear operators and nonlinear diagonal operators.

Nonlinear consensus in monotone multi-agent systems

We derive sufficient conditions on the Lipschitz, nonlinear, heterogeneous, and asymmetric interaction rule between agents ensuring their convergence to a consensus state.

On the convergence of the Krasnoselskij iteration for strictly pseudocontractive operators

Thank you for your attention!

Diego Deplano

Email: diego.deplano@unica.it
Webpage: https://diegodeplano.github.io/

SCAN ME