Non-Euclidean Enriched Contraction Theory for Monotone Operators and Monotone Dynamical Systems

Diego Deplano, Sergio Grammatico, Mauro Franceschelli

***Abstract*— We adopt an operator-theoretic perspective to analyze a class of nonlinear fixed-point iterations and discrete-time dynamical systems. Specifically, we study the Krasnoselskij iteration—at the heart of countless algorithmic schemes and underpinning the stability analysis of numerous dynamical systems—by focusing on a non-Euclidean vector space equipped with the diagonally weighted supremum norm.**

By extending the state of the art, we introduce the notion of *enriched weak contractivity*, which (i) is characterized by a simple, verifiable condition for Lipschitz operators, and (ii) yields explicit bounds on the admissible step size for the Krasnoselskij iteration. Our results relate the notion of weak contractivity with that of monotonicity of operators and dynamical systems by showing its use without loss of generality, thus yielding larger step sizes and improved convergence speed for broad classes of dynamical systems.

The results are applied to the design of zero-finding algorithms for monotone operators and nonlinear consensus dynamics in monotone multi-agent dynamical systems.

I. Introduction

We consider the Krasnoselskij iteration:

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

where $\theta \in (0,1)$ is the step size, and $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ is an operator. The Krasnoselskij iteration is a fundamental tool in fixed point theory, and its importance in system theory arises from its role in analyzing and designing iterative algorithms with guaranteed convergence to equilibrium points or invariant sets. Variations of the Krasnoselskij fixed-point iteration have been adopted to design distributed algorithms for computing fixed points in networks [1]–[5], splitting methods in distributed convex optimization [6]–[10], aggregative game theory [11]–[13], monotone dynamical systems [14]–[19], and monotone operator theory [20]–[22]. From a mathematical perspective, the convergence of the Krasnoselskij iteration is a fixed-point problem [23] on the operator T, or equivalently, a zero-finding problem for the operator $\mathsf{Id} - \mathsf{T}$, where Id denotes the identity operator. For instance, consensus in nonlinear multi-agent systems is equivalent to finding a collective state in the kernel of the nonlinear Laplacian operator [24]–[28].

In many of the above-mentioned applications, monotonicity plays a central role in guaranteeing convergence and stability, which enables the use of iterative schemes, including the Krasnoselskij iteration in (1), to compute fixed points efficiently [20]–[22]. On the other hand, in dynamical systems theory, monotonicity refers to systems whose trajectories preserve a partial order over time, which facilitates the design of distributed control laws in large-scale networked systems [14]–[19]. This dual relevance of monotonicity – both as an algebraic property of operators and as a dynamical property of systems – has motivated for decades the study of fixed point iterations like (1) in contexts where the operator T either demonstrates some form of monotonicity or induces monotonic behavior in the trajectories generated by the iteration [7], [13], [17], [29]–[32].

The aim of this paper is to introduce *enriched contraction theory* as a general framework encompassing both of these monotonicity notions. Toward this direction, we focus on the study of the Krasnoselskij iteration in non-Euclidean vector spaces equipped with a diagonally weighted supremum norm. Understanding its convergence under more general conditions would not only extend classical results but also open up new possibilities for applications in systems, control, learning, and optimization.

A. Literature review

One of the first convergence results for Krasnoselskij iterations dates back to 1955 and is due to Krasnoselskij [33], [34, 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 [23, Definition 1.8]. More than ten years later, Edelstein in [35] extended this result to $\theta \in (0,1)$ and *strictly convex* spaces [23, 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 [36, Theorem 1], see also [34, Theorem 6.4.3]. By limiting their analysis to Hilbert spaces, Marino and Xu in [37] 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 Krasnoselskij iteration, given $\theta < 1-\kappa$ [38, Theorem 1]. Marino and Xu in [37,

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@unica.it, mauro.franceschelli@unica.it.

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

Section 3] also posed the currently open question: "*Is strict pseudocontractivity also sufficient 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 [39]–[45].

A thorough answer to this question is given in [46] in the special case of real vector non-Euclidean spaces of finite dimension $n \in \mathbb{N}$ equipped with a p-norm such that $p \in (1, \infty)$: The Krasnoselskij iteration converges if $\theta^{r-1} < (1-\kappa)/c_p$ where $r = \min\{p, 2\}$ and $c_p \geq 1$ is a constant that depends on p, whose best (smallest) value has been found by Xu in [47, Corollary 2], and proved in [46, Lemmas 4-5]. Moreover, it has been shown that, for non-uniformly convex spaces with $p \in \{1, \infty\}$, strict pseudocontractivity is not sufficient for the convergence of the Krasnoselskij iteration. A natural open question arises: "*What is the most appropriate generalization of strict pseudocontractivity in non-Euclidean vector spaces that are not uniformly convex?*"

B. Main contributions

The first main contribution of this manuscript is to introduce the property of *enriched weak contractivity*. We show that this property serves as a natural generalization of strict pseudocontractivity in Banach spaces, while they are equivalent properties in Hilbert spaces (Proposition 1). Subsequently, by focusing on vector spaces equipped with a diagonally weighted norm, we prove the following original results:

- We provide a necessary and sufficient condition for enriched weak contractivity of Lipschitz operators (Theorem 1);
- We provide a bound on the maximum allowable step size ensuring the convergence of the Krasnoselskij iteration in (1) under the assumption that T is enriched weakly contractve (Theorem 4).

Regarding the relationship with monotone operators, we prove the following original results:

- A Lipschitz operator T is enriched weakly contractive if and only if the residual operator $\mathsf{Id} - \mathsf{T}$ is (strongly) monotone, where Id denotes the identity operator (Theorem 2);
- The Krasnoselskij iteration in (1) converges for larger allowable range of step sizes and improved contraction factor compared with those obtained by the state of the art on monotone operators [48] (Section III-C).

Next, on the relationship with monotone dynamical systems, we prove the following original results:

- For monotone dynamical systems $\boldsymbol{x}(k+1) = \mathsf{T}(\boldsymbol{x}(k))$, the operator T is enriched weakly contractive if and only if T is (strictly) subhomogeneous (Corollary 3);
- For monotone dynamical systems whose dynamics is ruled by the Krasnoselskij iteration in (1), an easy-to-verify bound on the maximum step size is derived (Corollary 4).

Concluding the manuscript, we apply the above theoretical results on two main applications along with numerical simulations corroborating the technical findings:

- Zero finding algorithms for monotone operators – We derive sufficient conditions for the convergence of the *forward step method* applied to monotone operators (Theorem 5) and simulate it on linear operators and nonlinear diagonal operators (Sections IV-A-IV-B).
- Nonlinear consensus in monotone multi-agent systems – We derive sufficient conditions on the nonlinear local interaction rule between agents ensuring their convergence to a consensus state (Theorem 6).

Structure of the paper. Section II provides the necessary background on enriched weak contractions, along with necessary and sufficient conditions for its validity and a comparison with other important operator-theoretic properties. Section III provides sufficient conditions for the convergence of the Krasnoselskij iteration on enriched weakly contractive operators in vector spaces equipped with a diagonally weighted norm. Sections IV-V discuss, respectively, the application of the theoretical results to design algorithms for computing zeros of monotone operators and distributed algorithms for reaching consensus in order-preserving multi-agent systems.

II. Enriched Weak Contractivity

A. Notation and preliminaries

The sets of real and integer numbers are denoted by $\mathbb{R}$ and $\mathbb{Z}$, while their restrictions to nonnegative and positive values are denoted by $\mathbb{R}_{\geq 0}$, $\mathbb{N}$ and $\mathbb{R}_+$, $\mathbb{N}_+$, respectively. Scalars $s \in \mathbb{R}$ are denoted by lowercase letters, while vectors $\boldsymbol{v} \in \mathbb{R}^n$ by boldface bold letters. The vectors of zeros and ones of dimension n are denoted by $\mathbf{0}_n$ and $\mathbf{1}_n$, respectively, and the subscript n is omitted if clear from the context. Matrices $M \in \mathbb{R}^{n \times n}$ are denoted by uppercase letters, and $\mathbb{S}^n_{\succ 0}$ denotes the set of positive definite symmetric matrices, i.e., such that $M^\top = M$ and $\boldsymbol{x}^\top M \boldsymbol{x} > 0$ for all $\boldsymbol{x} \in \mathbb{R}^n$. A matrix is said to be *Metzler* if its off-diagonal entries m_{ij} (with $i \neq j$) are nonnegative. The nonnegative majorant $|M|$ and the Metzler majorant $\lceil M \rceil_{\mathrm{M}}$ of a matrix are defined entry-wise by

$$(|M|)_{ij} := |m_{ij}|, \qquad (\lceil M \rceil_{\mathrm{M}})_{ij} := \begin{cases} a_{ii} & \text{if } i = j, \\ |m_{ij}| & \text{if } i \neq j. \end{cases} \quad (2)$$

Given a vector $\boldsymbol{v} \in \mathbb{R}^n$, the diagonal matrix whose entries are those of the vector $\boldsymbol{v}$ is denoted by $[\boldsymbol{v}]$, namely

$$([\boldsymbol{v}])_{ij} := \begin{cases} \boldsymbol{v}_i & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \quad (3)$$

Operators $\mathsf{T} : \mathcal{X}_1 \to \mathcal{X}_2$ 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 $\mathsf{Id} : \boldsymbol{x} \mapsto I\boldsymbol{x}$, but in general operators may be nonlinear. When $\mathcal{X}_2 \equiv \mathbb{R}$ sometimes block lowercase letters are used instead, e.g., $\mathsf{t} : \mathcal{X} \to \mathbb{R}$. Given a self-operator $\mathsf{T} : \mathcal{X} \to \mathcal{X}$, $\mathrm{fix}(\mathsf{T}) = \{\boldsymbol{x} \in \mathcal{X} : \mathsf{T}(\boldsymbol{x}) = \boldsymbol{x}\}$ denotes the set of fixed points and $\mathrm{zer}(\mathsf{T}) = \{\boldsymbol{x} \in \mathcal{X} : \mathsf{T}(\boldsymbol{x}) = \mathbf{0}\}$ denotes the set of zeros.

A *normed 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 will focus on the real vector space $\mathcal{X} = \mathbb{R}^n$, thus making any space $(\mathbb{R}^n, \|\cdot\|)$ a *Banach space* since every finite-dimensional normed vector space is complete (cfr. [23, Def. 1.5 and Rem. 2 on page 7] and [49, Theorem 5.33]). Hilbert spaces are Banach spaces where the inner product is well defined, i.e., $\langle \boldsymbol{x}, \boldsymbol{x}\rangle = \|\boldsymbol{x}\|^2$ for any $\boldsymbol{x} \in \mathbb{R}^n$. We will be specifically interested in real vector Banach spaces $(\mathbb{R}^n, \|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}})$ equipped with diagonally weighted ℓ_∞ norms defined by positive vectors $\boldsymbol{\eta} \in \mathbb{R}^n_+$ as follows

$$\|\boldsymbol{x}\|_{\infty,[\boldsymbol{\eta}]^{-1}} = \left\|[\boldsymbol{\eta}]^{-1}\boldsymbol{x}\right\|_\infty = \max_{i=1,\cdots,n} \frac{1}{\eta_i}|x_i|,$$

but we will also refer to Hilbert spaces $(\mathbb{R}^n, \|\cdot\|_{2,P})$ equipped with weighted ℓ_2 norms defined by positive definite symmetric matrices $P \in \mathbb{S}^n_{\succ 0}$ as follows

$$\|\boldsymbol{x}\|_{2,P} = \|P\boldsymbol{x}\|_2 = \sqrt{\boldsymbol{x}^\top P^2 \boldsymbol{x}}.$$

We now introduce the basic notions of contractivity and weak contractivity for general Banach spaces.

Definition 1. *[50] Given a real Banach space $(\mathbb{R}^n, \|\cdot\|)$, an operator $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ is called ℓ-Lipschitz if, for some $\ell \geq 0$ and for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, it satisfies*

$$\|\mathsf{T}(\boldsymbol{x}) - \mathsf{T}(\boldsymbol{y})\| \leq \ell \|\boldsymbol{x} - \boldsymbol{y}\|. \tag{4}$$

If we let $\mathsf{Lip}(\mathsf{T})$ be the minimum (or infimum) constant $\ell \geq 0$ which satisfies (4)*, i.e.,*

$$\mathsf{Lip}(\mathsf{T}) := \sup_{\boldsymbol{x}\neq\boldsymbol{y}} \frac{\|\mathsf{T}(\boldsymbol{x}) - \mathsf{T}(\boldsymbol{y})\|}{\|\boldsymbol{x} - \boldsymbol{y}\|}$$

then:

- *T is called ℓ-contractive if $\mathsf{Lip}(\mathsf{T}) \in (0, 1)$;*
- *T is called weakly contractive if $\mathsf{Lip}(\mathsf{T}) = 1$.*

Remark 1. *For a Lipschitz operator T, the Jacobian matrix $D\mathsf{T}(\boldsymbol{x})$ exists for almost every $\boldsymbol{x} \in \mathbb{R}^n$ by Rademacher's theorem. Moreover, it holds that*

$$\mathsf{Lip}(\mathsf{T}) = \operatorname*{ess\,sup}_{\boldsymbol{x}\in\mathbb{R}^n} \|D\mathsf{T}(\boldsymbol{x})\|,$$

where the essential supremum ignores the points in the set of Lebesgue measure zero where $D\mathsf{T}(\boldsymbol{x})$ does not exist.

B. The proposed definition of enriched weak contractivity

Next, we propose a generalization of the weak contractivity property called *enriched weak contractivity*.

Definition 2. *Given a real Banach space $\mathcal{B} = (\mathbb{R}^n, \|\cdot\|)$, 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*

$$\|b(\boldsymbol{x} - \boldsymbol{y}) + \mathsf{T}(\boldsymbol{x}) - \mathsf{T}(\boldsymbol{y})\| \leq (b - c + 1)\|\boldsymbol{x} - \boldsymbol{y}\|. \tag{5}$$

We note that for $c \in (b, b+1]$, the coefficient on the right-hand side $(b - c + 1)$ is strictly less than one, thus resulting in a weak form of contractivity. Enriched weak contractivity generalizes the notion of *enriched nonexpansiveness* proposed by Berinde in [50], which is a special case by letting $c = 0$, thus ensuring the coefficient on the right-hand side $b + 1$ is greater than or equal to one. Let us collect some useful special cases in the following remark.

Remark 2. *A (b, c)-enriched weakly contractive operator is:*

- *Contractive if $0 = b < c < 1$;*
- *Weakly contractive if $b = c = 0$;*
- *Enriched nonexpansive if $b > c = 0$ [50, Eq. (16)].*

C. Relationship with pseudocontractivity

Another important generalization of weakly contractive operators is that of strictly pseudocontractive operators [38]. We formally define this property for Hilbert spaces, as its general form for Banach spaces involves more advanced mathematical tools [46].

Definition 3. *[38, Definition 4] 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*

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

Notably, we show next that in the case of Hilbert spaces equipped with an inner product $\langle\cdot,\cdot\rangle$, κ-strict pseudocontractivity is equivalent to $(b, 0)$-enriched weak-contractivity with a specific relation between the parameters κ and b. On the other hand, for general Banach spaces strict pseudocontractivity and enriched weak contractivity are independent properties.

Proposition 1. *Let $\mathcal{H} = (\mathbb{R}^n, \|\cdot\|)$ be a real Hilbert space. An operator $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ is κ-strictly pseudocontractive if and only if it is $(b, 0)$-enriched weakly contractive with*

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

Proof: The following transformations hold in both directions, where the first inequality is the definition of enriched weak contractivity and the last is the definition of strict pseudocontractivity:

$$\begin{aligned}
\|b(\boldsymbol{x}-\boldsymbol{y})+\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y})\| &\leq (b-c+1)\|\boldsymbol{x}-\boldsymbol{y}\|,\\
\|b(\boldsymbol{x}-\boldsymbol{y})+\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y})\| &\leq (b+1)\|\boldsymbol{x}-\boldsymbol{y}\|,\\
\|b(\boldsymbol{x}-\boldsymbol{y})+\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y})\|^2 &\leq (b^2+2b+1)\|\boldsymbol{x}-\boldsymbol{y}\|^2,\\
\|\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y})\|^2 &\leq (2b+1)\|\boldsymbol{x}-\boldsymbol{y}\|^2\\
&\quad -2b\langle\boldsymbol{x}-\boldsymbol{y},\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y})\rangle,\\
\|\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y})\|^2 &\leq (2b+1)\|\boldsymbol{x}-\boldsymbol{y}\|^2\\
&\quad -2b\langle\boldsymbol{x}-\boldsymbol{y},\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y})\rangle\\
&\quad +b\|\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y})\|^2-b\|\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y})\|^2\\
(b+1)\|\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y})\|^2 &\leq (b+1)\|\boldsymbol{x}-\boldsymbol{y}\|^2\\
&\quad +b\|\boldsymbol{x}-\boldsymbol{y}-(\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y}))\|^2,\\
\|\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y})\|^2 &\leq \|\boldsymbol{x}-\boldsymbol{y}\|^2\\
&\quad +\frac{b}{b+1}\|\boldsymbol{x}-\boldsymbol{y}-(\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y}))\|^2.
\end{aligned}$$

We conclude that T is $(b,0)$-enriched weakly contractive if and only if it is κ-strictly pseudocontractive with $\kappa = b/(b+1)$. ■

D. Enriched weak contractivity of Lipschitz operators

The following theorem provides a necessary and sufficient condition for enriched weak contractivity of Lipschitz operators.

Theorem 1. *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

$$0\leq b^\star\leq\max\{0,\mathrm{diagL}(-\mathsf{T})\},$$

where

$$\mathrm{diagL}(-\mathsf{T}):=\operatorname*{ess\,sup}_{\boldsymbol{x}\in\mathbb{R}^n}\max_{i\in\{1,\ldots,n\}}(-\mathsf{T}(\boldsymbol{x}))_{ii}. \tag{6}$$

Proof: According to the definition of enriched weak contractivity in Definition 2, condition (i) means that the operator $b\mathsf{Id}+\mathsf{T}$ is Lipschitz with constant (see also Remark 1)

$$\mathsf{Lip}(\mathsf{T})=\operatorname*{ess\,sup}_{\boldsymbol{x}\in\mathbb{R}^n}\|D\mathsf{T}(\boldsymbol{x})\|=b-c+1;$$

then, condition (i) is equivalent to

$$\|D\mathsf{T}(\boldsymbol{x})\|_{\infty,[\boldsymbol{\eta}]^{-1}}\leq b-c+1,\ \text{ for almost every } \boldsymbol{x}\in\mathbb{R}^n. \tag{7}$$

Thus, the equivalence $(i)\Leftrightarrow(ii)$ follows from

$$\begin{aligned}\|bI+D\mathsf{T}(\boldsymbol{x})\|_{\infty,[\boldsymbol{\eta}]^{-1}}&\leq b-c+1\\ \left\|[\boldsymbol{\eta}]^{-1}(bI+D\mathsf{T}(\boldsymbol{x}))[\boldsymbol{\eta}]\right\|_{\infty}&\leq b-c+1\\ \left|[\boldsymbol{\eta}]^{-1}(bI+D\mathsf{T}(\boldsymbol{x}))[\boldsymbol{\eta}]\right|\mathbf{1}&\leq(b-c+1)\mathbf{1},\\ [\boldsymbol{\eta}]^{-1}|bI+D\mathsf{T}(\boldsymbol{x})|[\boldsymbol{\eta}]\mathbf{1}&\leq(b-c+1)\mathbf{1},\\ |bI+D\mathsf{T}(\boldsymbol{x})|\boldsymbol{\eta}&\leq(b-c+1)\boldsymbol{\eta}.\end{aligned}$$

For the last statement, let $b',c'>0$ be values of b,c such that (ii) holds, and consider the i-th row of (ii), namely

$$|b'+(D\mathsf{T}(\boldsymbol{x}))_{ii}|\eta_i+\sum_{j\neq i}|(D\mathsf{T}(\boldsymbol{x}))_{ij}|\eta_j\leq(b'-c'+1)\eta_i.$$

Note that if $b'\geq\mathrm{diagL}(-\mathsf{T})$ the argument of the first absolute value is nonnegative, namely,

$$b'+(D\mathsf{T}(\boldsymbol{x}))_{ii}\geq b'-\mathrm{diagL}(-\mathsf{T})\geq 0,$$

thus implying that condition (ii) becomes independent of b':

$$(D\mathsf{T}(\boldsymbol{x}))_{ii}\eta_i+\sum_{j\neq i}|(D\mathsf{T}(\boldsymbol{x}))_{ij}|\eta_j\leq(1-c)\eta_i.$$

This means that if (i)-(ii) hold for any $b'\geq\mathrm{diagL}(-\mathsf{T})$, then they hold also for any other $b\geq\mathrm{diagL}(-\mathsf{T})$, and therefore $b^*\leq\mathrm{diagL}(-\mathsf{T})$. The proof is completed by considering the complementary case $b^\star\leq b'<\mathrm{diagL}(-\mathsf{T})$. ■

Corollary 1. *Let $\mathsf{T}:\mathbb{R}^n\to\mathbb{R}^n$ be a Lipschitz operator. For $\boldsymbol{\eta}\in\mathbb{R}^n_+$, the following statements are equivalent:*
(i) T is weakly contractive w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$;
(ii) $|D\mathsf{T}(\boldsymbol{x})|\boldsymbol{\eta}\leq\boldsymbol{\eta}$ for almost every $\boldsymbol{x}\in\mathbb{R}^n$.

Corollary 2. *Let $\mathsf{T}:\mathbb{R}^n\to\mathbb{R}^n$ be a Lipschitz operator. For $\boldsymbol{\eta}\in\mathbb{R}^n_+$, the following statements are equivalent:*
(i) T is contractive w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$;
(ii) $|D\mathsf{T}(\boldsymbol{x})|\boldsymbol{\eta}<\boldsymbol{\eta}$ for almost every $\boldsymbol{x}\in\mathbb{R}^n$.

E. Relationship with monotone operators

We now compare the enriched weak contractivity property to that of monotonicity in Banach spaces of the type $(\mathbb{R}^n,\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}})$. In the view of [48], we report the definition of monotonicity in such spaces and the necessary and sufficient condition for Lipschitz operators.

Definition 4. *[48, Definition 12 and Equation (5)] 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*

$$\min_{i\in I_\infty([\boldsymbol{\eta}]^{-1}\boldsymbol{y})}\frac{(\mathsf{F}_i(\boldsymbol{x})-\mathsf{F}(\boldsymbol{y}))(\boldsymbol{x}_i-\boldsymbol{y}_i)}{\eta_i^2}\geq c\|\boldsymbol{x}-\boldsymbol{y}\|_{\infty,[\boldsymbol{\eta}]^{-1}}. \tag{8}$$

where $I_\infty(\boldsymbol{v})=\{i\in\{1,\ldots,n\}\mid|\boldsymbol{v}_i|=\|\boldsymbol{v}\|_\infty\}$. If (8) holds, then for $c=0$, the operator F is called monotone.

Proposition 2. *[48, Lemma 14] 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*

$$\lceil -D\mathsf{F}(\boldsymbol{x})\rceil_M\boldsymbol{\eta}\leq -c\boldsymbol{\eta}\ \text{for almost every } \boldsymbol{x}\in\mathbb{R}^n, \tag{9}$$

where $\lceil M\rceil_M$ is the Metzler majorant of the matrix M as in (2).

Notably, we show that the class of strongly monotone operators is exactly the same class of enriched weakly contractive operators. However, as discussed in more detail later, enriched weak contractivity is a more insightful property due to the extra parameter b.

Theorem 2. *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) $\mathsf{F}:=\mathsf{Id}-\mathsf{T}$ is c-strongly monotone w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$.
Then, the following hold:
(a) $(i)\Rightarrow(ii)$ holds for all $b\geq 0$;
(b) $(i)\Leftarrow(ii)$ holds for $b\geq\mathrm{diagL}(-\mathsf{T})$.

Proof: We start by recalling the necessary and sufficient condition for (ii) enriched weak contractivity in Theorem 1:

$$|bI+D\mathsf{T}(\boldsymbol{x})|\boldsymbol{\eta}\leq(b-c+1)\boldsymbol{\eta}\ \text{for almost every } \boldsymbol{x}\in\mathbb{R}^n,$$

and the necessary and sufficient condition for (ii) strong monotonicity of F in Proposition 2:

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

We prove statement (a) by the following steps:

$$\begin{aligned}\lceil -D\mathsf{F}(\boldsymbol{x})\rceil_{\mathrm{M}}\boldsymbol{\eta}&=\lceil D\mathsf{T}(\boldsymbol{x})-I\rceil_{\mathrm{M}}\boldsymbol{\eta}\\ &=\lceil D\mathsf{T}(\boldsymbol{x})\rceil_{\mathrm{M}}\boldsymbol{\eta}-\boldsymbol{\eta}\\ &=\lceil D\mathsf{T}(\boldsymbol{x})\rceil_{\mathrm{M}}\boldsymbol{\eta}-\boldsymbol{\eta}+b\boldsymbol{\eta}-b\boldsymbol{\eta}\\ &=\lceil bI+D\mathsf{T}(\boldsymbol{x})\rceil_{\mathrm{M}}\boldsymbol{\eta}-(b+1)\boldsymbol{\eta}\\ &\leq|bI+D\mathsf{T}(\boldsymbol{x})|\boldsymbol{\eta}-(b+1)\boldsymbol{\eta}\leq -c\boldsymbol{\eta}.\end{aligned}$$

Next, we prove statement (b) as follows:

$$\begin{aligned}|bI+D\mathsf{T}(\boldsymbol{x})|\boldsymbol{\eta}&=|(b+1)I-D\mathsf{F}(\boldsymbol{x})|\boldsymbol{\eta}\\ &\overset{(i)}{=}\lceil(b+1)I-D\mathsf{F}(\boldsymbol{x})\rceil_{\mathrm{M}}\boldsymbol{\eta}\\ &=(b+1)\boldsymbol{\eta}+\lceil -D\mathsf{F}(\boldsymbol{x})\rceil_{\mathrm{M}}\boldsymbol{\eta}\\ &\leq(b-c+1)\boldsymbol{\eta}\end{aligned}$$

where (i) holds for $b\geq\mathrm{diagL}(\mathsf{F})-1=\mathrm{diagL}(-\mathsf{T})$ as in (6). ■

F. Relationship with monotone systems

In this section, we also compare enriched weak contractivity with *order-preservation*, which makes discrete-time dynamical system $\boldsymbol{x}(k+1) = \mathsf{T}(\boldsymbol{x}(k))$ monotone [16], [51] in the sense of Kamke–Muller [52], [53]. Real vector spaces of finite-dimension can be equipped with the natural order relation $\leq$, yielding ordered vector spaces whose positive cone is the nonnegative orthant $\mathbb{R}^n_{\geq 0} = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x} \geq 0\}$. If between any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n_{\geq 0}$ for which there exists an order relation, the operator $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ is such that this relation is preserved for their images $\mathsf{T}(\boldsymbol{x})$ and $\mathsf{T}(\boldsymbol{y})$, then T is said to be *order-preserving*.

Definition 5. *[25, Definition 3] An operator $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ is called order-preserving if it holds*

$$\boldsymbol{x} \leq \boldsymbol{y} \Rightarrow \mathsf{T}(\boldsymbol{x}) \leq \mathsf{T}(\boldsymbol{y}), \qquad \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n. \tag{10}$$

Definition 6. *[25, Definition 1] 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.*

Proposition 3. *[16, Theorem 5] 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.,*

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

Monotone dynamical systems are of interest because they enjoy the enriched weak contractivity property w.r.t. a diagonally weighted ℓ_∞ if and only if the strict subhomogeneity property holds, which is defined next.

Definition 7. *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*

$$\mathsf{T}(\boldsymbol{x} + \theta\boldsymbol{\eta}) \leq \mathsf{T}(\boldsymbol{x}) + \theta(1-c)\boldsymbol{\eta}, \qquad \forall \theta > 0. \tag{11}$$

If (11) *holds with the equality sign, then T is called $\boldsymbol{\eta}$-homogeneous. If* (11) *holds for $c = 0$, then T is called $\boldsymbol{\eta}$-(sub)homogeneous.*

Proposition 4. *A Lipschitz operator $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ is c-strictly $\boldsymbol{\eta}$-subhomogeneous if and only if its Jacobian matrix satisfies*

$$D\mathsf{T}(\boldsymbol{x})\boldsymbol{\eta} \leq (1-c)\boldsymbol{\eta}, \text{ for almost every } \boldsymbol{x} \in \mathbb{R}^n.$$

For c-strictly $\boldsymbol{\eta}$-homogeneous operators, the necessary and sufficient condition is the above with the strict equality.

Proof: The necessity of the condition is proven via the definition of the directional derivative:

$$\begin{aligned} D\mathsf{T}(\boldsymbol{x})\boldsymbol{\eta} &= \lim_{\theta \to 0^+} \frac{\mathsf{T}(\boldsymbol{x} + \theta\boldsymbol{\eta}) - \mathsf{T}(\boldsymbol{x})}{\theta} \\ &\leq \lim_{\theta \to 0^+} \frac{\mathsf{T}(\boldsymbol{x}) + \theta(1-c)\boldsymbol{\eta} - \mathsf{T}(\boldsymbol{x})}{\theta} = (1-c)\boldsymbol{\eta}. \end{aligned}$$

The sufficiency of the condition is proven by the Newton–Leibnitz formula for vector-valued functions:

$$\begin{aligned} f(\boldsymbol{x} + \theta\boldsymbol{\eta}) - f(\boldsymbol{x}) &= \theta \int_0^1 Df(\boldsymbol{x} + s\theta\boldsymbol{\eta})\boldsymbol{\eta} ds \\ f(\boldsymbol{x} + \theta\boldsymbol{\eta}) - f(\boldsymbol{x}) &\leq \theta \int_0^1 (1-c)\boldsymbol{\eta} ds \\ f(\boldsymbol{x} + \theta\boldsymbol{\eta}) - f(\boldsymbol{x}) &\leq \theta(1-c)\boldsymbol{\eta}. \end{aligned}$$

■

This allows us to prove the following technical result.

Theorem 3. *Let $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$ be a Lipschitz operator such that its Jacobian matrix $D\mathsf{T}(\boldsymbol{x})$ is Metzler. If $b \geq$* $\mathrm{diagL}(-\mathsf{T})$*, 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.*

Proof: We prove $(i) \Rightarrow (ii)$ by

$$\begin{aligned} D\mathsf{T}(\boldsymbol{x})\boldsymbol{\eta} &= D\mathsf{T}(\boldsymbol{x})\boldsymbol{\eta} + b\boldsymbol{\eta} - b\boldsymbol{\eta} \\ &= b\boldsymbol{\eta} + D\mathsf{T}(\boldsymbol{x})\boldsymbol{\eta} - b\boldsymbol{\eta} \\ &\overset{(\star)}{=} |bI + D\mathsf{T}(\boldsymbol{x})|\boldsymbol{\eta} - b\boldsymbol{\eta} \\ &\leq (b-c+1)\boldsymbol{\eta} - b\boldsymbol{\eta} \\ &\leq (1-c)\boldsymbol{\eta}. \end{aligned}$$

where $(\star)$ holds by assumption $b \geq \mathrm{diagL}(-\mathsf{T})$.

We prove $(i) \Leftarrow (ii)$ by

$$\begin{aligned} |bI + D\mathsf{T}(\boldsymbol{x})|\boldsymbol{\eta} &\overset{(\star)}{=} b\boldsymbol{\eta} + D\mathsf{T}(\boldsymbol{x})\boldsymbol{\eta} \\ &\leq b\boldsymbol{\eta} + (1-c)\boldsymbol{\eta} \\ &= (b-c+1)\boldsymbol{\eta} \end{aligned}$$

where $(\star)$ holds by assumption $b \geq \mathrm{diagL}(-\mathsf{T})$.

■

Corollary 3. *Consider a discrete-time dynamical system $\boldsymbol{x}(k+1) = \mathsf{T}(\boldsymbol{x}(k))$ that 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.*

III. The Krasnoselskij iteration in a Banach space

A. Main convergence result

Consider the Krasnoselskij iteration,

$$\boldsymbol{x}(k+1) = \mathsf{T}_\theta(\boldsymbol{x}(k)) = (1-\theta)\boldsymbol{x}(k) + \theta\mathsf{T}(\boldsymbol{x}(k)), \tag{12}$$

with step size $\theta \in (0,1)$. The next lemma provides an upper bound on the maximum value of θ guaranteeing the convergence of the iteration, provided that the operator T is enriched weakly contractive as in Definition 2.

Lemma 1. *Consider a Banach space $(\mathbb{R}^n, \|\cdot\|)$ and an operator $\mathsf{T} : \mathbb{R}^n \to \mathbb{R}^n$. The following statements are equivalent:*

- T *is (b,c)-enriched weakly contractive w.r.t. $\|\cdot\|$;*
- T_θ *in* (12) *is Lipschitz for $\theta = 1/(b+1)$ w.r.t. $\|\cdot\|$ and with constant $\ell = 1 - \frac{c}{b+1}$.*

Proof: The following transformations hold in both directions, where the first inequality is the definition of enriched weak contractivity and the last is the definition of Lipschitz continuity for T_θ with $\theta = 1/(b+1)$:

$$\begin{aligned}
\|b(\boldsymbol{x}-\boldsymbol{y})+\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y})\| &\le (b-c+1)\|\boldsymbol{x}-\boldsymbol{y}\|,\\
\frac{1}{b+1}\|b(\boldsymbol{x}-\boldsymbol{y})+\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y})\| &\le \frac{b-c+1}{b+1}\|\boldsymbol{x}-\boldsymbol{y}\|,\\
\left\|\frac{b(\boldsymbol{x}-\boldsymbol{y})+\mathsf{T}(\boldsymbol{x})-\mathsf{T}(\boldsymbol{y})}{b+1}\right\| &\le \left(1-\frac{c}{b+1}\right)\|\boldsymbol{x}-\boldsymbol{y}\|,\\
\left\|\frac{b\boldsymbol{x}+\mathsf{T}(\boldsymbol{x})}{b+1}-\frac{b\boldsymbol{y}+\mathsf{T}(\boldsymbol{y})}{b+1}\right\| &\le \left(1-\frac{c}{b+1}\right)\|\boldsymbol{x}-\boldsymbol{y}\|,\\
\left\|\mathsf{T}_{\frac{1}{b+1}}(\boldsymbol{x})-\mathsf{T}_{\frac{1}{b+1}}(\boldsymbol{y})\right\| &\le \left(1-\frac{c}{b+1}\right)\|\boldsymbol{x}-\boldsymbol{y}\|.
\end{aligned}$$

■

We now state and prove the main result of the paper.

Theorem 4. *Consider an operator T that is (b,c)-enriched weakly contractive w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$ for some $\boldsymbol{\eta} \in \mathbb{R}^n_+$, $b > 0$, $c \ge 0$. Then the following hold:*

(a) If $c = 0$ and $\mathrm{fix}(\mathsf{T}) \neq \emptyset$, the iteration in (12) converges for $\theta \in \left(0, \frac{1}{b+1}\right)$ to some fixed point $\bar{\boldsymbol{x}} \in \mathrm{fix}(\mathsf{T})$;

(b) If $c > 0$, the iteration in (12) converges to the unique fixed point $\mathrm{fix}(\mathsf{T}) = \{x^\star\}$ for $\theta \in \left(0, \frac{1}{b+1}\right]$. Moreover, it holds that

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

Proof: The iteration can be rewritten as follows

$$\begin{aligned}
\mathsf{T}_\theta &= (1-\theta)\mathsf{Id}+\theta\mathsf{T} = (1-\theta)\mathsf{Id}+\theta\left((b+1)\mathsf{T}_{\frac{1}{b+1}}-b\mathsf{Id}\right)\\
&= (1-\theta(b+1))\mathsf{Id}+\theta(b+1)\mathsf{T}_{\frac{1}{b+1}}\\
&= (1-\alpha)\mathsf{Id}+\alpha\mathsf{T}_{\frac{1}{b+1}}, \quad \text{with} \quad \alpha=\theta(b+1).
\end{aligned}$$

Due to Lemma 1, the operator $\mathsf{T}_{\frac{1}{b+1}}$ ruling the iteration in (12) is either weakly contractive (if $c = 0$) or contractive (if $c = 0$) with contraction factor $1 - c/(b+1)$ or . If $c = 0$, by the known result of Ishikawa [36] [48, Lemma 11], the iteration converges to a fixed point (if there exists one) for any $\alpha \in (0,1)$, i.e., $\theta \in (0, 1/(b+1))$, thus proving (a). If $c > 0$, by the known Banach fixed point theorem, the iteration converges to the unique fixed point and the convergence rate is given by

$$\begin{aligned}
&\|D\mathsf{T}_\theta(\boldsymbol{x})\|_{\infty,[\boldsymbol{\eta}]^{-1}} = \|(1-\theta)I+\theta D\mathsf{T}(\boldsymbol{x})\|_{\infty,[\boldsymbol{\eta}]^{-1}}\\
&= \max_i\left\{|1-\theta+\theta(D\mathsf{T}(\boldsymbol{x}))_{ii}|+\theta\sum_{j\neq i}|(D\mathsf{T}(\boldsymbol{x}))_{ij}|\frac{\eta_j}{\eta_i}\right\}\\
&= \max_i\left\{|1-\theta+\theta(D\mathsf{T}(\boldsymbol{x}))_{ii}+\alpha-\alpha|+\theta\sum_{j\neq i}|(D\mathsf{T}(\boldsymbol{x}))_{ij}|\frac{\eta_j}{\eta_i}\right\}\\
&\overset{(i)}{\le} 1-\alpha+\theta\max_i\left\{|((D\mathsf{T}(\boldsymbol{x}))_{ii}+b)|+\sum_{j\neq i}|(D\mathsf{T}(\boldsymbol{x}))_{ij}|\frac{\eta_j}{\eta_i}\right\}\\
&\overset{(ii)}{=} 1-\alpha+\theta(b+1)\left\|D\mathsf{T}_{\frac{1}{b+1}}\right\|_{\infty,[\boldsymbol{\eta}]^{-1}}\\
&\overset{(iii)}{\le} 1-\alpha+\theta(b+1)\left(1-\frac{c}{b+1}\right)\overset{(iv)}{=}1-\frac{\alpha c}{b+1}=1-\theta c.
\end{aligned}$$

We note that (i) holds because since $1-\alpha$ with $\alpha \in (0,1)$ is always positive, it can be moved outside the absolute value by the triangle inequality; (i) uses the identity $\alpha - \theta = \theta b$, thus allowing to take θ as a common factor; (ii) holds by definition of the diagonally weighted ℓ_∞ norm and by the definition of $\mathsf{T}_{\frac{1}{b+1}} = (b \cdot \mathsf{Id} + \mathsf{T})/(b+1)$; (iii) holds by Lemma 1 which ensures that the Lipschitz constant of $\mathsf{T}_{\frac{1}{b+1}}$ is $1 - c/(b+1)$; (iv) holds by the identity $\alpha = \theta(b+1)$. ■

B. The special case of linear operators: Comparison with strict pseudocontractive operators

We start our discussion by recalling a recent result by Belgioioso *et al.* [38] in the case of linear operators and Hilbert spaces $(\mathbb{R}^n, \|\cdot\|_{2,P})$. Under these assumptions, the iteration converges if and only if the map T is k-strictly pseudocontractive which, in Hilbert spaces, is equivalent to being $(b,0)$-enriched weakly contractive, by Proposition 1. Let us formally report this result in view of Proposition 1.

Proposition 5. *[38, Theorem 1] Consider the iteration*

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

For $b \ge 0$, the following statements are equivalent:

- *$\exists P \in \mathbb{S}^n_{\succ 0} : \mathsf{A} : \boldsymbol{x} \mapsto A\boldsymbol{x}$ is $(b,0)$-enriched weakly contractive w.r.t. $\|\cdot\|_{2,P}$;*
- *$\boldsymbol{x}(k)$ converges to $\mathrm{fix}(\mathsf{A})$ for $\theta \in \left(0, \frac{1}{b+1}\right)$.*

On the contrary, we show that enriched weak contractivity w.r.t. a diagonally weighted ℓ_∞ norm is only sufficient for the convergence of the iteration, but not necessary. We do so by providing a counter-example for a weakly contractive matrix, i.e., a $(0,0)$-enriched weakly contractive matrix.

Theorem 1 and Corollary 1 provide a method to verify (enriched) weak contractivity of an operator w.r.t. a diagonal weighted ℓ_∞. In the case of linear operators $\mathsf{A} : \boldsymbol{x} \mapsto A\boldsymbol{x}$, weak contractivity w.r.t. a diagonally weighted ℓ_∞ norm reduces to a linear feasibility program (LP) of the kind $|A|\boldsymbol{\eta} \le \boldsymbol{\eta}$; on the other hand, weak contractivity w.r.t. a weighted ℓ_2 is a semidefinite feasibility program (SDP) of the kind $A^\top P A \preccurlyeq P$ [38, Lemma 3]. An important consideration due to [38, Lemma 3] is that the SDP feasibility is a necessary and sufficient condition for (at most marginal) stability of the system $\boldsymbol{x}(k+1) = A\boldsymbol{x}(k)$, i.e., all eigenvalues of A are within the unit disk and those on the boundary are semi-simple. On the contrary, the next example shows that the LP feasibility is not necessary, while it is sufficient as proved in the following Proposition 6.

Example 1. *Consider the matrix:*

$$A = \begin{bmatrix} 1 & 1 \\ 0 & a \end{bmatrix}, \quad \textit{with} \quad a \in (0,1).$$

The eigenvalues $\{a, 1\}$ of A are semi-simple and within the unit disk, thus the iteration $\boldsymbol{x}(k+1) = A\boldsymbol{x}(k)$ converges. However, the operator $\mathsf{A} : \boldsymbol{x} \to A\boldsymbol{x}$ is not weakly contractive w.r.t. any diagonally weighted ℓ_∞ norm because the LP *program $|A|\boldsymbol{\eta} \leq \boldsymbol{\eta}$ is not feasible for $\boldsymbol{\eta} \in \mathbb{R}^n_+$:*

$$\begin{cases} A\boldsymbol{\eta} \leq \boldsymbol{\eta} \\ \boldsymbol{\eta} > \mathbf{0} \end{cases} \Leftrightarrow \begin{cases} \eta_1 + \eta_2 \leq \eta_1 \\ a\eta_2 \leq \eta_2 \\ \eta_1 > 0 \\ \eta_2 > 0 \end{cases} \Leftrightarrow \begin{cases} \eta_2 \leq 0 \\ a \leq 1 \\ \eta_1 > 0 \\ \eta_2 > 0 \end{cases}.$$

The above example shows that weak contractivity w.r.t. a diagonally weighted ℓ_∞ norm is not necessary for (at least marginal) stability of a linear time-invariant system. In the next proposition, we show that it is sufficient.

Proposition 6. *For a linear operator $\mathsf{A} : \boldsymbol{x} \mapsto A\boldsymbol{x}$, weak contractivity w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$ implies that:*

($\star$) *All eigenvalues of A are in the unit disk and those on the boundary of the unit disk are semi-simple.*

Vice versa, the above condition does not imply weak contractivity of A w.r.t. the norm $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$.

Proof: We first note that weak contractivity is equivalent to $(0,0)$-enriched weak contractivity, which holds if and only if (Corollary 2)

$$|A|\boldsymbol{\eta} \leq \boldsymbol{\eta} \tag{13}$$

When $\mathsf{T} = \mathsf{A}$ is a linear operator $\mathsf{A} : \boldsymbol{x} \to A\boldsymbol{x}$, then $D\mathsf{T}(\boldsymbol{x}) = D\mathsf{A}(\boldsymbol{x}) = A$ for all $\boldsymbol{x} \in \mathbb{R}^n$. Consider the similarity transformation given by $N = [\boldsymbol{\eta}]$, i.e., $A_N = N^{-1}|A|N$. By construction, matrices $|A|$ and A_N have the same eigenvalues and are both nonnegative. Exploiting (13), it can be show that A_N is row-substochastic,

$$A_N\mathbf{1} = N^{-1}|A|N\mathbf{1} = N^{-1}|A|\boldsymbol{\eta} \leq N^{-1}\boldsymbol{\eta} = \mathbf{1},$$

and, in turn, the eigenvalues of A are in the unit disk:

$$\rho(A) \leq \rho(|A|) = \rho(A_N) \leq \max(A_N\mathbf{1}) \leq \max(\mathbf{1}) \leq 1,$$

For the sake of contradiction, assume there is an eigenvalue with magnitude 1 that is not semi-simple, i.e., its algebraic multiplicity is strictly greater than its geometric multiplicity. This implies that the system $\boldsymbol{x}(k) = A\boldsymbol{x}(k-1)$ is unstable, i.e., at least one component of $\boldsymbol{x}(k)$ grows unbounded as $k \to \infty$. This contradicts the assumption that the linear operator A is weakly contractive because it implies that $\boldsymbol{x}(k)$ must remain bounded. Instead, since weak contractivity of A is equivalent to $\|A\|_{\infty,[\boldsymbol{\eta}]^{-1}} \leq 1$ as in (7), it holds that:

$$\begin{aligned} \lim_{k\to\infty} \|\boldsymbol{x}(k)\|_{\infty,[\boldsymbol{\eta}]^{-1}} &\leq \lim_{k\to\infty} \left\|A^k\boldsymbol{x}(0)\right\|_{\infty,[\boldsymbol{\eta}]^{-1}} \\ &= \lim_{k\to\infty} \|A\|^k_{\infty,[\boldsymbol{\eta}]^{-1}} \|\boldsymbol{x}(0)\|_{\infty,[\boldsymbol{\eta}]^{-1}} \\ &= \|\boldsymbol{x}(0)\|_{\infty,[\boldsymbol{\eta}]^{-1}}. \end{aligned}$$

This proves that weak contractivity of A implies ($\star$). The contrary does not hold by Example 1, thus completing the proof. ■

C. Comparison with monotone operators and monotone dynamical systems

By Theorem 2, we know that F is c-strongly monotone, then $\mathsf{T} = \mathsf{Id} - \mathsf{F}$ is surely (b,c)-enriched weakly contractive for any $b \geq \mathrm{diagL}(-\mathsf{T})$. Thus, the value $\mathrm{diagL}(-\mathsf{T})$ works as an upper bound to minimum value $b^\star$ for which the T is enriched weakly contractive. This is coherent with Theorem 1. This means that the information provided by the monotonicity property is somehow less than that provided by enriched weak contractivity, in the sense that monotonicity of F allows to find an upper bound on the constant of enriched weak contractivity of T, but lower values may be still valid. We provide an example illustrating this case.

Example 2. *Consider the matrix*

$$A = \frac{1}{2}\begin{bmatrix} -3 & 0 & 1 & -3 \\ 3 & -15 & -12 & -1 \\ 2 & -1 & -5 & -5 \\ -2 & 0 & -1 & -6 \end{bmatrix} \tag{14}$$

and the Krasnoselskij iteration

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

According to Theorem 1, the operator $\mathsf{A} : \boldsymbol{x} \mapsto A\boldsymbol{x}$ is $(b,0)$-enriched weakly contractive w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$ if and only $|bI + A|\boldsymbol{\eta} \leq (b+1)\boldsymbol{\eta}$. It can be verified that this holds with the choices of $\boldsymbol{\eta} = [0.09, 1, 0.22, 0.07]^\top$ and $b = 4$, which satisfies the upper bound $\mathrm{diagL}(-A) = 15/2$ given by Theorem 1. Therefore, by Theorem 4, the iteration converges for $\theta \in (0, \frac{1}{b+1})$, i.e., $\theta \in (0, 0.2)$.

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

$$\begin{aligned} \boldsymbol{x}(k+1) &= (1-\theta)\boldsymbol{x}(k) + \theta A\boldsymbol{x}(k) \\ &= \boldsymbol{x}(k) - \theta(I-A)\boldsymbol{x}(k) = \boldsymbol{x}(k) - \theta F\boldsymbol{x}(k), \end{aligned}$$

where

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

According to [48, Theorem 26(iii)], if the operator $\mathsf{F} : \boldsymbol{x} \to F\boldsymbol{x}$ is monotone as in Definition 4 (with $c = 0$), the iteration converges for $\theta \in \left(0, \frac{1}{\mathrm{diagL}(\mathsf{F})}\right)$ where $\mathrm{diagL}(\mathsf{F}) = 1 + \mathrm{diagL}(-A)$. We note that this range of feasible values of θ corresponds to the one obtained from the upper bound given in Theorem 1. The operator F is monotone if and only if there exists $\boldsymbol{\eta} = [\eta_1, \cdots, \eta_n] \in \mathbb{R}^n_+$ such that (cfr. Proposition 2)

$$-F_{ii}\eta_i + \sum_{j\neq i} |F_{ij}|\eta_j \leq 0.$$

The above holds for the same $\boldsymbol{\eta} = [0.09, 1, 0.22, 0.07]^\top$ and, in turn, the iteration converges for $\theta \in \big(0, 0.117\big)$.

To make the results easily verifiable, we did not consider the coefficient c of enriched weak contractivity and that of strong monotonicity. A more precise characterization of A is that it is (b,c)-enriched weak contractivity for $b \approx 4.006$ and $c \approx 0.0227$, while a more precise characterization of F is that it is c-strongly monotone with the same $c \approx 0.0227$.

We conclude the section with an easy-to-verify bound on the step size for monotone dynamical systems by combining the results of Theorem 3 and Theorem 4.

Corollary 4. *Consider a monotone dynamical system with dynamics* $\boldsymbol{x}(k+1) = (1-\theta)\boldsymbol{x}(k) + \theta \mathsf{T}(\boldsymbol{x}(k))$ *where* T *is a Lipschitz operator whose Jacobian matrix is Metzler and let* $L = \mathrm{diagL}(-\mathsf{T})$. *Then, the following statements hold:*

- *If* T *is* $\boldsymbol{\eta}$*-subhomogeneous and* $\mathrm{fix}(\mathsf{T}) \neq \emptyset$, *then the iteration in* (12) *converges to some fixed point* $\bar{\boldsymbol{x}} \in \mathrm{fix}(\mathsf{T})$ *for* $\theta \in \left(0, \frac{1}{1+L}\right)$.
- *If* T *is strictly subhomogeneous, the iteration in* (12) *converges to the unique fixed point* $\mathrm{fix}(\mathsf{T}) = \{x^\star\}$ *for* $\theta \in \left(0, \frac{1}{1+L}\right]$ *with convergence rate optimized at* $\theta = 1/(1+L)$.

IV. Application to Zero-Finding Algorithms for Monotone Operators

The problem of finding a zero of an operator (e.g., the zero of the Laplacian operator in consensus problems [25]–[28]) is usually translated into the problem of finding a fixed point of a suitably defined operator. According to Theorem 2, the problem of finding a zero of a (strongly) monotone operator F is equivalent to the problem of finding a fixed point of the enriched weakly contractive operator $\mathsf{T} = \mathsf{Id} - \mathsf{F}$, namely:

$$\mathsf{F}(\boldsymbol{x}) = \mathbf{0} \quad \Leftrightarrow \quad \mathsf{T}(\boldsymbol{x}) = \boldsymbol{x}.$$

The standard "*forward step method*" to find a zero of the monotone operator F consists in iterating the "*forward operator*"

$$\mathsf{S}_{\theta\mathsf{F}} := \mathsf{Id} - \theta\mathsf{F}, \tag{15}$$

Theorem 5. *Let* $\mathsf{F} : \mathbb{R}^n \to \mathbb{R}^n$ *be a Lipschitz operator. If* $\mathrm{zer}(\mathsf{F}) \neq \emptyset$ *and* F *is strongly monotone w.r.t.* $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$, *then the iteration ruled by* $\boldsymbol{x}(k+1) = \mathsf{S}_{\theta\mathsf{F}}(\boldsymbol{x}(k))$ *with* $\mathsf{S}_{\theta\mathsf{F}}$ *as in* (15) *converges to an element of* $\mathrm{zer}(\mathsf{F})$ *for every*

$$\theta \in \left(0, \frac{1}{b^\star + 1}\right),$$

where $b^\star \leq \mathrm{diagL}(\mathsf{F}) - 1$ *is given by*

$$b^\star = \min\{b \geq 0 | (b+1)I - D\mathsf{F}(\boldsymbol{x}) | \boldsymbol{\eta} \leq (b+1)\boldsymbol{\eta}, \boldsymbol{\eta} > 0\}.$$

Proof: The forward step method corresponds to the Krasnoselskij iteration with operator $\mathsf{T} := \mathsf{Id} - \mathsf{F}$:

$$\mathsf{S}_{\theta\mathsf{F}} = \mathsf{Id} - \theta\mathsf{F} = (1-\theta)\mathsf{Id} + \theta(\mathsf{Id} - \mathsf{F}) = (1-\theta)\mathsf{Id} + \theta\mathsf{T}.$$

By Theorem 2, F is c-strongly monotone if and only if T is (b,c)-enriched weakly contractive for some $b \geq 0$ which, in turn, is equivalent to $|bI + D\mathsf{T}(\boldsymbol{x})|\boldsymbol{\eta} \leq (b-c+1)\boldsymbol{\eta}$ for almost every $\boldsymbol{x} \in \mathbb{R}^n$ by Theorem 1. The smallest admissible value $b^\star$ of b is obtained by selecting $c = 0$. Theorem 4 ensures that the iteration converges to an element of $\mathrm{fix}(\mathsf{T})$ for any $0 < \theta < 1/(b+1) \leq 1/(b^\star+1)$. The proof is completed by exploiting Theorem 1 which yields $b^\star \leq \mathrm{diagL}(-\mathsf{T}) = \mathrm{diagL}(\mathsf{F} - \mathsf{Id}) = \mathrm{diagL}(\mathsf{F}) - 1$. ∎

We now provide some numerical simulations to illustrate that Theorem 5 represents a generalization of [48, Theorem 26] since it allows for larger step sizes and, in turn, improved convergence rate.

A. Affine operators

We first consider the simple case where $\mathsf{T} := \mathsf{A}$ is an affine operator $\mathsf{A} : \boldsymbol{x} \mapsto A\boldsymbol{x} + \boldsymbol{b}$ and fixed-point problems of the form

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

Consider now A and $\boldsymbol{b}$ as follows:

$$A = \begin{bmatrix} -1.07 & -0.17 & -0.53 & -0.33 \\ 0.07 & 0.42 & -0.07 & 0.15 \\ -0.13 & -0.10 & -0.06 & -0.30 \\ 0.04 & 0.05 & -0.21 & 0.40 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

In this case, there is a unique fixed point given by

$$\boldsymbol{x}^\star \approx \begin{bmatrix} 0.04 & -2.14 & -0.25 & -1.76 \end{bmatrix}^\top.$$

It is possible to exploit either enriched weak contractivity or strong monotonicity to determine the range of possible values for the Krasnoselskij iteration w.r.t. $\|\cdot\|_\infty$:

- *Enriched weak contractivity*: Theorem 1 ensures convergence of the Krasnoselskij iteration for $0 \leq \theta \leq \frac{1}{b^\star+1} \leq 0.645$, because the lowest value of b such that A is $(b,0)$-enriched weakly contractive is $b \geq b^\star = 0.55$;
- *Monotonicity*: Theorem 26(iii) in [48] ensures convergence of the Krasnoselskij iteration for $0 \leq \theta \leq \frac{1}{\mathrm{diagL}(\mathsf{Id}-\mathsf{A})} = 0.48$, because $\mathsf{Id} - \mathsf{A}$ is monotone.

The above derivations show that exploiting enriched weak contractivity – instead of monotonicity – allows to enlarge the range of admissible step sizes ensuring the convergence of the Krasnoselskij iteration, which is coherent with the relationship $b^\star \leq \max\{0, \mathrm{diagL}(-\mathsf{A})\} = 1.07$ provided in Theorem 1.

We now discuss how to chose θ to guarantee the best (smallest) convergence rate. According to Theorem 4, the convergence rate depends on both parameters b and c determining (b,c)-enriched weakly contractive of the operator A and, in particular, it is upper bounded by

$$1 - \frac{c}{b+1}. \tag{16}$$

In general, there may be different pairs of b, c for which A is enriched weakly contractive. Therefore, we solve a minimization problem whose objective function is the bound in (16) to find the best possible pair (b,c) such that enriched weak contractivity is guaranteed:

$$\begin{aligned} (b^\star, c^\star) = \underset{b \geq 0, c \geq 0}{\mathrm{argmin}}\ & 1 - \frac{c}{b+1}, \\ \text{s.t.} \quad & |bI + A|\mathbf{1} \leq (b - c + 1)\mathbf{1}, \\ & c \leq b + 1. \end{aligned}$$

In our example, the solution to the above problem is $b^\star = 0.695$, $c^\star = 0.29$, ensuring that the convergence rate is not greater than 0.83 when $\theta = 1/(b^\star+1) = 0.59$. Instead, according to [48], selecting $\theta = \frac{1}{\mathrm{diagL}(\mathsf{Id}-\mathsf{A})} = 0.48$ ensures a worse factor, namely 0.86. Fig. 1(left) shows the convergence of the distance between the iteration $\boldsymbol{x}(k)$ and the fixed point $\boldsymbol{x}^\star$ when the forward step method in (15) 0.48 is applied with the two different choices of θ, corroborating the theoretical results by revealing an improved convergence rate.

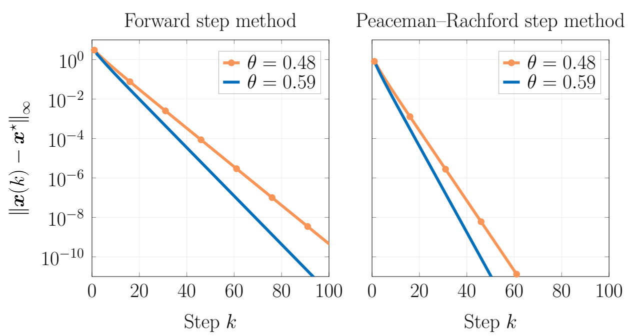


Fig. 1. Residual versus number of iterations for the (left) forward-step method and (right) Peaceman-Rachford Splitting.

We also compare the performance of the forward-backward and the Peaceman–Rachford step methods with the same choices of the step size, as their convergence can be guaranteed by retracing the results of [48] in light of Theorem 5. The results for the forward-backward step method are comparable with those of the forward step method and thus are omitted, while the results for the Peaceman-Rachford step method are shown in Fig. 1(right).

B. Composition of nonlinear diagonal operators with affine operators

In this section, we consider the special class of "*nonlinear diagonal*" operators $\Phi : \boldsymbol{x} \mapsto [\phi_1(x_1), \cdots, \phi_n(x_n)]^\top$ where $\phi_i : \mathbb{R} \to \mathbb{R}$ are assumed to be Lipschitz. We consider operators $\mathsf{T} := \Phi \circ \mathsf{A}$ resulting from the composition of a nonlinear diagonal operator with an affine operator. This class of functions arises, for instance, in the context of training infinite-depth weight-tied neural network [20], [22], [54] by solving fixed-point problems of the form

$$\boldsymbol{x}(k+1) = \mathsf{T}(\boldsymbol{x}(k)) := \Phi(A\boldsymbol{x}(k) + \boldsymbol{b}), \tag{17}$$

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 $\dot{\boldsymbol{x}}(t) = -\boldsymbol{x}(t) + \Phi(A\boldsymbol{x} + \boldsymbol{b})$. We remark that, in the context of neural network, the vector $\boldsymbol{b}$ is usually the result of two separate terms, namely the term related to the inputs $U\boldsymbol{u}$ where $U \in \mathbb{R}^{n\times m}$ are the input-injection weights and $\boldsymbol{u} \in \mathbb{R}^m$ are the inputs, and the constant term related to the biases $\boldsymbol{p} \in \mathbb{R}^n$, yielding $\boldsymbol{b} = U\boldsymbol{u} + \boldsymbol{p}$.

Let us make the following technical assumption.

Assumption 1. *The nonlinear diagonal operator $\Phi(\boldsymbol{x}) = [\phi_1(x_1), \cdots, \phi_n(x_n)]^\top$ satisfies, for some $d_1 \le d_2$, the following condition*

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

Popular activation functions used in machine learning satisfy these bounds, such as the Leaky ReLu function

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

satisfies Assumption 1 with $d_1 = \alpha$ and $d_2 = 1$. This function has been introduced by Kaiming He *et al.* in [55] as an attempt to fix the "*dying ReLU*" problem, i.e., the occurrence of dead neurons when using the ReLu function, which corresponds to the Leaky ReLu with $\alpha = 0$.

To test the performance resulting from the exploitation of enriched weak contractivity versus strong monotonicity, we randomly generate a matrix $M \in \mathbb{R}^{n\times n}$ and a vector $\boldsymbol{b} \in \mathbb{R}^n$ following a Gaussian distribution with mean $\mu = 0$ and variance $\sigma^2 = 1/n$. To ensure that T is enriched weakly contractive with respect to the diagonally weighted norm $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$, we first derive a sufficient condition:

$$\begin{aligned}|bI + D\mathsf{T}(\boldsymbol{x})|\boldsymbol{\eta} &= |bI + D\Phi(A\boldsymbol{x}(k) + \boldsymbol{b})|\boldsymbol{\eta}\\ &\le \max\{|bI + d_1A|\boldsymbol{\eta}, |bI + d_2A|\boldsymbol{\eta}\}\\ &\le (b - c + 1)\boldsymbol{\eta},\end{aligned}$$

Then, we compute A by projecting M such that the above condition is satisfied and the diagonal elements of A are free:

$$\begin{aligned}\min_{b\ge c-1, A\in\mathbb{R}^{n\times n}} \quad & \|(A - M) - \mathrm{diag}(A - M)\|_F,\\ \text{s.t.} \quad & |bI + d_1A|\boldsymbol{\eta} \le (b - c + 1)\boldsymbol{\eta}\\ & |bI + d_2A|\boldsymbol{\eta} \le (b - c + 1)\boldsymbol{\eta}\end{aligned}$$

for a given, fixed coefficient $c \ge 0$ and a vector $\boldsymbol{\eta} \in \mathbb{R}^n_+$, where $\|\cdot\|_F$ denotes the Frobenius norm and $\mathrm{diag}(\cdot)$ returns a diagonal matrix whose elements are those in the diagonal of the input matrix. For instance, one matrix of dimension $n = 5$ is

$$A = \frac{1}{2}\begin{bmatrix} 0.278 & 0.111 & -0.280 & -0.134 & -0.098\\ -0.189 & -0.739 & -0.207 & -0.105 & -0.066\\ -0.408 & -0.355 & -0.203 & 0.301 & 0.039\\ 0.252 & 0.246 & -0.225 & -0.537 & -0.046\\ 0.144 & 0.253 & 0.288 & 0.225 & -0.395\end{bmatrix},$$

for which the operator T is (b, c)-enriched weakly contractive with $b = 0.537$, $c = 0.324$ and for $\boldsymbol{\eta} = [2.673,\ 1.181,\ 2.215,\ 1.261,\ 1.498]^\top$.

In the experiments, we selected different choices for the parameter $c = \{0.2, 0.6, 1, 1.25, 1.5, 1.75, 2\}$ and we considered the Leaky ReLu activation function with $\alpha = 0.1$. Also, we considered matrices $A \in \mathbb{R}^{n\times n}$ with growing size $n \in [5, 200]$ and, for each size n and each parameter c, we run 10 different experiments, computing the optimal step size θ both exploiting our Theorem 4 and Theorem 26 in [48]. In particular, we denote by $\theta^\star_{\mathrm{EWC}}$ the optimal step size obtained by minimizing the convergence rate $\rho^\star_{\mathrm{EWC}}$ provided by Theorem 4, which exploits enriched weak contractivity, for any possible weighted norm $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$:

$$\rho^\star_{\mathrm{EWC}} = 1 - c^\star_{\mathrm{EWC}}\theta^\star_{\mathrm{EWC}}, \qquad \theta^\star_{\mathrm{EWC}} = \frac{1}{b^\star_{\mathrm{EWC}} + 1},$$

where $b^\star_{\mathrm{EWC}}, c^\star_{\mathrm{EWC}}, , \boldsymbol{\eta}^\star_{\mathrm{EWC}}$ are solutions of

$$\begin{aligned}\operatorname*{argmin}_{b\ge 0, c\ge 0, \boldsymbol{\eta}>\mathbf{0}} \quad & 1 - \frac{c}{b+1},\\ \text{s.t.} \quad & |bI + d_1A|\boldsymbol{\eta} \le (b - c + 1)\boldsymbol{\eta}\\ & |bI + d_2A|\boldsymbol{\eta} \le (b - c + 1)\boldsymbol{\eta}\\ & c \le b + 1.\end{aligned}$$

On the other hand, we denote by $\theta^\star_{\mathrm{MON}}$ the optimal step size obtained by minimizing the convergence rate $\rho^\star_{\mathrm{MON}}$ provided by [48, Theorem 26(ii)], which exploits monotonicity, for any possible weighted norm $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$, namely

$$\rho^\star_{\text{MON}} = 1 - c^\star_{\text{MON}}\theta^\star_{\text{MON}}, \qquad \theta^\star_{\text{MON}} = \frac{1}{1 - \min\limits_{i\in\{1,\cdots n\}}\{\alpha\cdot a_{ii}, a_{ii}\}},$$

where $c^\star_{\text{MON}}, \boldsymbol{\eta}^\star_{\text{MON}}$ are the solutions of (cfr. [56, Theorem 21(ii)])

$$\begin{aligned} \operatorname*{argmin}_{c\geq 0,\boldsymbol{\eta}>\mathbf{0}} \ & 1 - c\theta^\star_{\text{MON}}, \\ \text{s.t.} \quad & \lceil d_1 A\rceil_{\text{M}}\boldsymbol{\eta} \leq (1-c)\boldsymbol{\eta} \\ & \lceil d_2 A\rceil_{\text{M}}\boldsymbol{\eta} \leq (1-c)\boldsymbol{\eta} \\ & c \leq 1/\theta^\star_{\text{MON}} \end{aligned}$$

The results of the experiments are given in Fig 2, where we report the ratio between the two convergence rates $\rho^\star_{\text{MON}}$ and $\rho^\star_{\text{EWC}}$. Ratio values less than 1 denote an improvement in the convergence rate bound thanks to the exploitation of enriched weak contractivity. For small values of $c \approx 0$, we notice that the ratio is very close to 1 because the parameter b does not influence significantly the convergence rate bound as it is close to 1. Instead, for larger values of $c \approx 1$ we observe that the highest improvement in the convergence rate bound is reached. Finally, for large values of $c \gg 1$ this improvement is incrementally lost because in the limit of $c \to \infty$ it holds that $b \to \infty$, which implies that the convergence rate bound becomes again very close to 1.

V. Application to Nonlinear Consensus in Monotone Multi-Agent Systems

In this section, we consider multi-agent systems (MASs) composed of $n \in \mathbb{N}$ agents modeled as dynamical systems with scalar state $x_i \in \mathbb{R}$, seeking consensus via discrete-time, distributed information exchange:

$$x_i(k+1) = x_i(k) - \theta\sum_{i=1}^{n} a_{ij} f_{ij}\Big(x_i(k) - x_j(k)\Big), \tag{18}$$

where $f_{ij}:\mathbb{R}\to\mathbb{R}$ are Lipschitz nonlinear functions. We limit our attention to monotone discrete-time MASs in the sense of Definition 6, i.e., systems whose dynamics is ruled by an order-preserving operator. For the dynamics in (18), order-preservation is guaranteed by $\partial f_{ij}/\partial x \geq 0$ a.e. according to

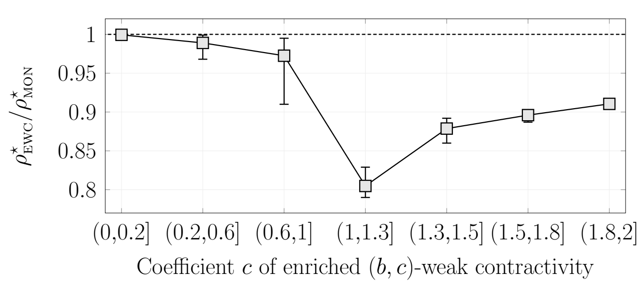


Fig. 2. Display of the ratio $\rho^\star_{\text{EWC}}/\rho^\star_{\text{MON}}$ between the convergence rates averaged over 100 instances for different number of features $n \in [5, 200]$: values lower than 1 denote an improvement in the convergence rate obtained by exploiting (b,c)-enriched weak contractivity instead of c-strong monotonicity.

Proposition 3. Monotone MASs are also called *cooperative* especially in the continuous-time framework [14], [57]–[60].

Coefficients $a_{ij} \in \{0,1\}$ are such that $a_{ij} = 1$ if node j can transmit information to node i, $a_{ij} = 0$ if not, and $a_{ii} = 0$. If $a_{ij} = 1$, then agent j is said to be a *neighbor* of agent i, and we denote its set of neighbors by $\mathcal{N}_i = \{j \in \mathcal{V} : (i,j) \in \mathcal{E}\}$. The pattern of interconnections among the agents is given by a graph $\mathcal{G} = (\mathcal{V},\mathcal{E})$ where $\mathcal{V} = \{1,\ldots,n\}$ is the set of nodes representing the agents and $\mathcal{E} \subseteq \mathcal{V}\times\mathcal{V}$ is a set of directed edges. A directed edge $(i,j) \in \mathcal{E}$ exists if agent i is influenced by agent j, i.e., $a_{ij} = 1$. The matrix $A = \{a_{ij}\} \in \mathbb{R}^{n\times n}$ is called the adjacency matrix. By defining

$$\mathsf{L}(\boldsymbol{x}) := \text{diag}(\boldsymbol{x})A - A\text{diag}(\boldsymbol{x}), \qquad g(A) := A \mapsto A\mathbf{1},$$

we can rewrite the dynamics of the agents' network as follows

$$\begin{aligned} \boldsymbol{x}(k+1) &= \boldsymbol{x}(k) - \theta(g\circ f\circ \mathsf{L})(\boldsymbol{x}(k)) \\ &= (1-\theta)\boldsymbol{x}(k) + \theta(\mathsf{Id} - g\circ f\circ\mathsf{L})(\boldsymbol{x}(k)) \end{aligned} \tag{19}$$

where $f = [\cdots, f_{ij}, \cdots] : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ and where $\circ$ denotes the composition operator.

The main result of this section is Theorem 6, which provides sufficient conditions on the local interaction rules f_{ij} such that the operator $\mathsf{T} := \mathsf{Id} - g\circ f\circ\mathsf{L}$ in (18) is enriched weakly contractive. Due to Theorem 4, enriched weak contractivity implies stability of the system and convergence of its state trajectories toward an equilibrium point, while accounting for heterogeneous local interaction rules. This result is particularly interesting from a control perspective when addressing the problem of steering a MAS toward specific equilibrium points, by relying on partial and relative information, without the intervention of a central controller, as in formation control [61] or distributed optimization [62]. Moreover, we provide some extra sufficient conditions ensuring that the equilibrium point set coincides with the so-called consensus set $\mathcal{C} = \{\alpha\mathbf{1} : \alpha \in \mathbb{R}\}$. The proposed sufficient condition is graph theoretical and based on the graph $\mathcal{G}$ describing the pattern of interconnections among the agents: It requires that there exists a globally reachable node in $\mathcal{G}$ and that consensus states are equilibrium points.

Theorem 6. *Consider a discrete-time MAS with agents dynamics as in* (18). *If the local interaction rules $f_{ij} : \mathcal{X} \to \mathbb{R}$, with $i = 1,\ldots,n$, are Lipschitz with constant $L \geq 0$ and satisfy:*

(i) $\partial f_{ij}/\partial x \geq 0$ for almost every $x \in \mathbb{R}$;

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

$$\theta \in \left(0, \frac{1}{L\max\limits_{i=1,\cdots,n}|\mathcal{N}_i|}\right). \tag{20}$$

If it further holds that:

(ii) $f_{ij}(0) = 0$ and $f_{ij}(x) \neq 0$ a.e. in a neighborhood of 0;

(iii) the graph $\mathcal{G}$ has a globally reachable node;

then the MAS converges asymptotically to a consensus state.

Proof: The MAS state update can be written as the Krasnoselskij iteration (19) on the operator $\mathsf{T} := \mathsf{Id} - f\circ\mathsf{L}$, which is $(b, 0)$-enriched weakly contractive for

$$b \geq \mathrm{diagL}(-\mathsf{T}) = \mathrm{diagL}(g \circ f \circ \mathsf{L}) - 1 = L \max_{i=1,\cdots,n} |\mathcal{N}_i| - 1$$

w.r.t. $\|\cdot\|_\infty$ due to Theorem 3, as it is is $\mathbf{1}$-subhomogeneous (i.e., $\mathsf{T}(\boldsymbol{x} + \theta\mathbf{1}) = \mathsf{T}(\boldsymbol{x}) + \theta\mathbf{1}$) and its Jacobian matrix is Metzler by construction and by condition (i). When the system has at least one equilibrium point, one can thus exploit the result in Corollary 4 to establish that, for any initial condition and for all θ as in (20), the state trajectories of the MAS converge to one of its equilibrium points, completing the first part of the proof.

While condition (ii) implies that all consensus states are equilibrium points, condition (iii) implies that there are no other equilibrium points other than the consensus states. The graph $\mathcal{G}$ is aperiodic for any θ as in (20), which ensures that the diagonal elements of the Jacobian matrix are strictly positive and, in turn, the presence of a self-loop at each node; moreover, the graph contains a globally reachable node due to condition (iii).

By means of the definition of directional derivative, for every equilibrium point $\boldsymbol{x}_e$, it holds

$$\begin{aligned} D\mathsf{T}^{\pm}(c\mathbf{1})\boldsymbol{x}_e &= \lim_{h\to 0^\pm} \frac{\mathsf{T}(c\mathbf{1} + h\boldsymbol{x}_e) - \mathsf{T}(c\mathbf{1})}{h} \\ &= \lim_{h\to 0^\pm} \frac{c\mathbf{1} + h\boldsymbol{v} - c\mathbf{1}}{h} = \lim_{h\to 0^\pm} \frac{h\boldsymbol{x}_e}{h} = \boldsymbol{x}_e. \end{aligned} \tag{21}$$

By replacing $\boldsymbol{x}_e = \mathbf{1}$, (21) implies that the left/right Jacobian matrices $D\mathsf{T}^{\pm}$ are row-stochastic at any consensus point $c\mathbf{1}$ with $c \in \mathbb{R}$. By exploiting [63, Theorem 5.1], we conclude that $D\mathsf{T}^{\pm}$ have a simple unitary eigenvalue $\lambda = 1$ with corresponding eigenvector equal to $\mathbf{1}$, unique up to a scaling factor.

Since the Krasnoselskij iteration is nonexpansive by Lemma 1, then the set of equilibrium points is either empty or closed and convex by [64, Theorem 1]. Now, if there is an equilibrium point $x_e \notin \mathcal{C}$ that is not a consensus point, then all points $c\mathbf{1} + hx_e$ with $h \in [0, 1]$ and $c \in \mathbb{R}$ are also equilibrium points by construction, and thus

$$\begin{aligned} D\mathsf{T}^{\pm}(c\mathbf{1})x_e &= \lim_{h\to 0^\pm} \frac{\mathsf{T}(c\mathbf{1} + hx_e) - \mathsf{T}(c\mathbf{1})}{h} \\ &= \lim_{h\to 0^\pm} \frac{c\mathbf{1} + hx_e - c\mathbf{1}}{h} = \lim_{h\to 0^\pm} \frac{hx_e}{h} = x_e, \end{aligned}$$

This means that x_e is a second eigenvector (other than $v = \mathbf{1}$) of the unitary eigenvalue $\lambda = 1$ of $D\mathsf{T}^{\pm}$, which is a contradiction with respect to [63, Theorem 5.1]. This proves that there do not exist equilibrium points x_e outside the consensus set, completing the second part of the proof. ■

Theorem 6 generalizes [16, Theorem 6] for multi-agent systems of the form (18) by accommodating non-continuously differentiable functions. This framework enables analysis of heterogeneous, asymmetric interactions such as:

$$f_{ij}(x) = \max(\alpha x, x), \qquad f_{ij}(x) = \min(\alpha x, x),$$

with $\alpha > 0$ ruling the asymmetry of the interaction. Notably, neighboring agents may also exhibit distinct interaction rules ($f_{ij} \neq f_{ji}$). We remark that the above result can be generalized to deal with the special case $\alpha = 0$, which leads to unilateral interactions in the sense of [14], by considering more general notions of graphs, e.g., bicolored interaction graphs [14, Definition 10].

VI. Conclusions and Future Work

The proposed *enriched weak contractivity* property emerges as a powerful tool for generalizing foundational results in the theory of monotone operators and monotone dynamical systems, specifically by: (i) enabling the convergence of fixed-point iterations for monotone operators with larger step sizes in the Krasnoselskij iteration, thereby yielding improved bounds on convergence rates; and (ii) facilitating agreement in monotone multi-agent networks by accommodating Lipschitz, nonlinear, heterogeneous, and asymmetric interactions among agents.

Future work will investigate the role of enriched weak contractivity in other non-Euclidean spaces, e.g., those equipped with diagonally weighted ℓ_1 norms. In fact, the iteration of operators may converge to a fixed point despite not being contractive with respect to the ℓ_∞ norm (Example 1) or the ℓ_2 norm [65, Example 9], while it is still possible that they are contractive with respect to an ℓ_1 norm. An interesting direction is the emergence of consensus in nonlinear multi-agent systems in the case of antagonistic interactions [66]. Enriched weak contractivity might be the right tool to generalize known results on monotone mulit-agent networks with cooperative interaction to the case of possibly antagonistic interactions.

References

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

[2] 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, 2020.

[3] 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, 2022.

[4] 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.

[5] 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.

[6] 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, 2016.

[7] 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, 2019.

[8] 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, 2019.

[9] 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.

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

[11] 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, 2015.

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

[13] D. Gadjov and L. Pavel, "A passivity-based approach to nash equilibrium seeking over networks," *IEEE Transactions on Automatic Control*, vol. 64, no. 3, 2018.

[14] S. Manfredi and D. Angeli, "Necessary and sufficient conditions for consensus in nonlinear monotone networks with unilateral interactions," *Automatica*, vol. 77, 2017.

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

[16] 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, 2023.

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

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

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

[20] E. Winston and J. Z. Kolter, "Monotone operator equilibrium networks," *Advances in neural information processing systems*, vol. 33, 2020.

[21] C. Pabbaraju, E. Winston, and J. Z. Kolter, "Estimating lipschitz constants of monotone deep equilibrium models," in *International conference on learning representations*, 2021.

[22] S. Jafarpour, A. Davydov, A. Proskurnikov, and F. Bullo, "Robust implicit networks via non-euclidean contractions," *Advances in Neural Information Processing Systems*, vol. 34, 2021.

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

[24] 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.

[25] D. Deplano, M. Franceschelli, and A. Giua, "A nonlinear perron–frobenius approach for stability and consensus of discrete-time multiagent systems," *Automatica*, vol. 118, 2020.

[26] J. C. Bronski and L. DeVille, "Spectral theory for dynamics on graphs containing attractive and repulsive interactions," *SIAM Journal on Applied Mathematics*, vol. 74, no. 1, 2014.

[27] L. DeVries, A. Sims, and M. D. Kutzer, "Kernel design and distributed, self-triggered control for coordination of autonomous multiagent configurations," *Robotica*, vol. 36, no. 7, 2018.

[28] J. Qu, Z. Ji, Y. Liu, and C. Lin, "The controllability and structural controllability of laplacian dynamics," *International Journal of Control*, vol. 98, no. 5, 2025.

[29] H. Daniels and M. Velikova, "Monotone and partially monotone neural networks," *IEEE Transactions on Neural Networks*, vol. 21, no. 6, 2010.

[30] P. Yi and L. Pavel, "Distributed generalized nash equilibria computation of monotone games via double-layer preconditioned proximal-point algorithms," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 1, 2018.

[31] P. Tabuada and B. Gharesifard, "Universal approximation power of deep residual neural networks through the lens of control," *IEEE Transactions on Automatic Control*, vol. 68, no. 5, 2022.

[32] D. Mikulincer and D. Reichman, "Size and depth of monotone neural networks: Interpolation and approximation," *IEEE Transactions on Neural Networks and Learning Systems*, 2024.

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

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

[35] M. Edelstein, "A remark on a theorem of M.A. Krasnoselski," *Amer. Math. Monthly*, vol. 73, 1966.

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

[37] 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, 2007.

[38] 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, 2018.

[39] 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, 2009.

[40] ——, "Strong convergence theorems for strict pseudo-contractions in q-uniformly smooth Banach spaces," *Nonlinear Analysis: Theory, Methods* & *Applications*, vol. 70, no. 9, 2009.

[41] 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, 2010.

[42] 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, 2010.

[43] 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, 2010.

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

[45] ——, "Strong convergence for a countable family of strict pseudo-contractions in q-uniformly smooth Banach spaces," *Computers* & *Mathematics with Applications*, vol. 62, no. 2, 2011.

[46] D. Deplano, S. Grammatico, and M. Franceschelli, "On the convergence of the krasnoselskij iteration for strictly pseudocontractive operators," Available: https://diegodeplano.github.io/files/ECC25-kSpC-withapp.pdf.

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

[48] A. Davydov, S. Jafarpour, A. V. Proskurnikov, and F. Bullo, "Non-euclidean monotone operator theory and applications," *Journal of Machine Learning Research*, vol. 25, no. 307, 2024.

[49] J. K. Hunter and B. Nachtergaele, *Applied analysis*. World Scientific Publishing Company, 2001.

[50] V. Berinde and M. Păcurar, "Fixed points theorems for unsaturated and saturated classes of contractive mappings in Banach spaces," *Symmetry*, vol. 13, no. 4, 2021.

[51] D. Angeli and E. D. Sontag, "Monotone control systems," *IEEE Transactions on automatic control*, vol. 48, no. 10, 2003.

[52] M. Müller, "Über das fundamentaltheorem in der theorie der gewöhnlichen differentialgleichungen," *Mathematische Zeitschrift*, vol. 26, no. 1, 1927.

[53] E. Kamke, "Zur theorie der systeme gewöhnlicher differentialgleichungen. ii.," *Acta Mathematica 58*, 1932.

[54] G. Li, M. Müller, B. Ghanem, and V. Koltun, "Training graph neural networks with 1000 layers," in *International conference on machine learning*, PMLR, 2021.

[55] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE international conference on computer vision*, 2015.

[56] A. Davydov, A. V. Proskurnikov, and F. Bullo, "Non-euclidean contraction analysis of continuous-time neural networks," *IEEE Transactions on Automatic Control*, 2024.

[57] P. De Leenheer and D. Aeyels, "Stability properties of equilibria of classes of cooperative systems," *IEEE Transactions on Automatic Control*, vol. 46, no. 12, 2001.

[58] M. W. Hirsch and H. L. Smith, "Competitive and cooperative systems: A mini-review," in *Positive systems*, Springer, 2003.

[59] H. L. Smith, *Monotone Dynamical Systems: An Introduction to the Theory of Competitive and Cooperative Systems: An Introduction to the Theory of Competitive and Cooperative Systems*, 41. American Mathematical Soc., 2008.

[60] V. S. Bokharaie, O. Mason, and M. Verwoerd, "D-stability and delay-independent stability of homogeneous cooperative systems," *IEEE Transactions on automatic control*, vol. 55, no. 12, 2010.

[61] K.-K. Oh, M.-C. Park, and H.-S. Ahn, “A survey of multi-agent formation control,” *Automatica*, vol. 53, 2015.

[62] T. Yang, X. Yi, J. Wu, Y. Yuan, D. Wu, Z. Meng, Y. Hong, H. Wang, Z. Lin, and K. H. Johansson, “A survey of distributed optimization,” *Annual Reviews in Control*, vol. 47, 2019.

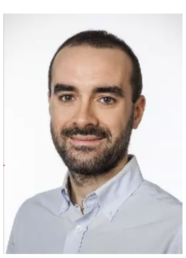
[63] F. Bullo, *Lectures on Network Systems*. Version 0.96, 2018.

[64] P. J. S. Ferreira, “The existence and uniqueness of the minimum norm solution to certain linear and nonlinear problems,” *Signal processing*, vol. 55, no. 1, 1996.

[65] F. Bullo, P. Cisneros-Velarde, A. Davydov, and S. Jafarpour, “From contraction theory to fixed point algorithms on riemannian and non-euclidean spaces,” in *2021 60th IEEE Conference on Decision and Control (CDC)*, IEEE, 2021.

[66] C. Altafini, “Consensus problems on networks with antagonistic interactions,” *IEEE transactions on automatic control*, vol. 58, no. 4, 2012.

Diego Deplano (S’17–M’21) is Assistant Professor (RTD-A) at the Department of Electrical and Electronic Engineering, University of Cagliari, Italy. He received the B.S., M.S., and PhD degrees in Electronic Engineering “cum laude” from the University of Cagliari, Italy, respectively in ’15, ’17, and ’21. He was awarded the best PhD Thesis defended in the area of Systems and Control Engineering at an Italian University by SIDRA. He spent visiting periods at the Nanyang Technological University (NTU), Singapore, at the Centre National de la Recherche Scientifique (CNRS), Grenoble, France, and at the University of Toronto (UofT), Toronto, Canada. He has lectured seminars at KTH in Stockholm, Sweden, and at Peking University in Beijing, China. He is a member of the Conference Editorial Board (CEB) for the IEEE Control Systems Society (CSS) and serves as Associate Editor for the IEEE Conference on Decision and Control (CDC) and the IEEE American Control Conference (ACC) since 2025. He also serves as an Associate Editor for the IEEE Conference on Automation Science and Engineering (CASE) since 2024. His research interests include open multi-agent systems, distributed optimization and learning, dynamic consensus problems, monotone theory, contraction theory.



Sergio Grammatico (M’16–SM’19) received the bachelor’s degree in computer engineering and the master’s and Ph.D. degrees in automatic control from the University of Pisa, Italy, in 2008, 2009, and 2013, respectively. He is currently an Associate Professor with the Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands. From 2012 to 2017, he held research positions with the University of California Santa Barbara, CA, USA, with ETH Zürich, Zürich, Switzerland, and with TU Eindhoven, The Netherlands. Dr. Grammatico was the recipient of the 2021 Roberto Tempo Best Paper Award. He is currently an Associate Editor for IEEE Transactions on Automatic Control and IFAC Automatica. His research interests include multi-agent systems, dynamic game theory and optimal control.

Mauro Franceschelli (M’11–SM’22) is Associate Professor at the Department of Electrical and Electronic Engineering, University of Cagliari, Italy. He received the Laurea degree in Electronic Engineering “cum laude” in 2007 and the PhD degree in 2011 from the University of Cagliari. He spent visiting periods at the Georgia Institute of Technology (GaTech), and the University of California at Santa Barbara (UCSB), USA. In 2013 he received a fellowship from the National Natural Science Foundation of China (NSFC), at Xidian University, Xi’an, China. In 2015 he was awarded a position of Assistant Professor (RTD-A) funded by the Italian Ministry of Education, University and Research (MIUR) under the 2014 call “Scientific Independence of Young Researchers” (SIR). He is currently an Associate Editor for the IEEE Transactions on Automatic Control (TAC) and IFAC Nonlinear Analysis: Hybrid Systems (NAHS). He is a member of the Conference Editorial Board (CEB) for the IEEE Control Systems Society (CSS) since 2019. He serves as Associate Editor for the IEEE Conference on Automation Science and Engineering (CASE) since 2015. He was an Associate Editor for the IEEE Transactions on Automation Science and Engineering (T-ASE) 2021-2023 and Guest Associate Editor for the IEEE Open Journal of Control Systems (OJ-CSYS) in 2024. He is IEEE Senior Member and Officer (vice-Chair) for the IEEE Italy Section Chapter of the IEEE Control Systems Society since 2025. His research interests include consensus problems, open multi-agent systems, multi-robot systems, networks of energy systems, resilient networks, distributed optimization and learning.