Stability of Nonexpansive Monotone Systems and Application to Recurrent Neural Networks

Diego Deplano, *Member, IEEE*, Mauro Franceschelli, *Senior, IEEE*, Alessandro Giua, *Fellow, IEEE*

***Abstract*— This letter shows that trajectories of continuous-time monotone systems (in the sense of Kamke-Muller) converge to equilibrium points if their vector field is continuously differentiable and if they are nonexpansive w.r.t. a diagonally weighted infinity norm. Differently from the current literature trend, the system is not required to be contractive but merely nonexpansive, thus allowing for multiple equilibrium points. Easy-to-check conditions on the vector field to verify that the system is both monotone and nonexpansive are provided. This is done by showing that nonexpansiveness is implied by subhomogeneity of the system, a generalization of the translation invariance property. We apply the results in the context of RNNs, thus providing sufficient conditions for convergence of the state trajectories of nonexpansive monotone neural networks that are not contractive.**

***Index Terms*— Monotone Systems, Type-K Monotone, Subhomogeneous, Nonexpansive, Neural Networks.**

I. Introduction

Dynamical systems whose trajectories preserve a partial order have sparked considerable interest in numerous fields: such systems are usually called *monotone* [1], in the sense of Kamke-Muller [2], [3] and are such that any pair of ordered initial conditions give rise to ordered solutions. Monotonicity appears naturally in real-world phenomena and engineering applications, including chemical reactions [4], [5], biological systems [6], flow networks [7], [8], phase-coupled oscillators [9], [10], opinion dynamics [11], mechanical systems [12], and so on. Within the systems and control community, many authors are currently interested in monotone systems. Among them, Manfredi and Angeli have studied the case of monotone networks with unilateral interactions [13]. Como and Lovisari have considered monotone dynamical flow networks [7], [8], a topic of interest for Coogan and Arcak as well [14]. In particular, Coogan has recently presented a tutorial paper on mixed monotonicity, which extends the usual notion of monotonicity [6]. Also worth mentioning is the line of research on eventually monotone systems pursued by Altafini and Mauroy [11], [15], as well as the framework of differentially positive systems drawn up by Forni and Sepulchre [12]. For insights on new advances and applications of monotone systems, we refer interested readers to the work of Smith [1].

Contraction theory is becoming a popular framework [16]–[18], providing powerful tools for establishing stability properties of nonlinear dynamical systems. In general, a dynamical system is contractive if every two trajectories converge to one another, thus resulting in systems with a unique equilibrium (for time-invariant systems). On the other hand, convergence of trajectories toward equilibrium points is still possible when the system is not contractive but only nonexpansive, that is every two trajectories do not diverge from one another. It is clear that the class of nonexpansive systems is broader than contractive systems, naturally allowing for multiple equilibrium points. While classical approaches mostly focus on contraction with respect to the Euclidean ℓ_2-norm, recent works have shown that the stability of monotone systems can be studied for contractive systems [19] and nonexpansive systems [10], [20], [21] with respect to non-Euclidean norms. For instance, it is known that for a monotone system satisfying the translation invariance or the conservation law, nonexpansiveness naturally arises with respect to the supremum norm [22, Lemma 2.7.2] or the taxicab norm [22, Proposition 2.8.1], respectively.

We have recently shown in [10] that smooth continuous-time dynamical systems, i.e., systems with a continuously differentiable vector field, which are also monotone, satisfy a stricter notion of monotonicity called *type-K monotonicity* recently exploited in [10], [20], [23] also in the context of multi-agent systems. The main feature of type-K monotonicity is that, when considering discrete-time systems, it prevents periodic state trajectories with periods exceeding one, while simple monotonicity cannot. The **main contribution** of this manuscript is leveraging type-K monotonicity to prove that:

- Trajectories of smooth monotone systems that are nonexpansive w.r.t. a diagonally weighted sup-norm converge toward equilibrium points, if any exist (Propositions 1-2);
- Smooth monotone systems are nonexpansive if and only if they are subhomogeneous (Theorem 1);
- Necessary and sufficient conditions for monotonicity and subhomogeneity are given in terms of the Jacobian matrix of their vector field (Lemmas 1-2).

We also apply our novel results to the convergence analysis of recurrent neural networks (RNN), with a focus on Hopfield and firing-rate dynamics. In particular, we prove that:

- Monotonicity and subhomogeneity of these neural networks ensure convergence of their state trajectories **even if their dynamics are not contractive** (Theorem 2).

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

Structure of the paper. Section II introduces the notation and preliminaries on monotone and nonexpansive systems. Section III contains our main results and a tutorial example. In Section IV the results are applied to the analysis of nonexpansive RNNs. In Section V we give concluding remarks.

II. Notation and Preliminaries

The set of real and integer numbers are denoted by $\mathbb{R}$ and $\mathbb{Z}$, while their restriction to nonnegative values are denoted with $\mathbb{R}_{\geq 0}$, $\mathbb{N}$, respectively. Matrices $M \in \mathbb{R}^{n\times n}$ are denoted by uppercase letters, vectors $\boldsymbol{v} \in \mathbb{R}^n$ by boldface bold letters, scalars $s \in \mathbb{R}$ by lowercase letters, while sets and spaces $\mathcal{S}$ are denoted by uppercase calligraphic letters. Elements of a matrix M or a vector $\boldsymbol{v}$ are denoted by m_{ij} and v_i, where i and j denotes the corresponding row and column. A matrix M is Metzler if its off-diagonal elements $m_{ij} \geq 0$ with $i \neq j$ are nonnegative. The vectors of zeros and ones of dimension n are denoted by $\mathbf{0}_n$ and $\mathbf{1}_n$, respectively. A diagonal matrix is written as $[\boldsymbol{v}]$ with diagonal elements $v_1, \dots, v_n$. $I_n = [\mathbf{1}_n]$ is the identity matrix of dimension n. The element-wise product between vectors or matrices of appropriate dimensions is denoted by the symbol $\odot$. We denote by $\|\cdot\|$ the vector norm in $\mathbb{R}^n$ and corresponding induced norms on $\mathbb{R}^{n\times n}$. We will be specifically interested in the diagonally weighted sup-norm, defined by a positive vector $\boldsymbol{\eta} \in \mathbb{R}^n_+$ as follows

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

Any weighted norm is equivalent to the standard sup-norm:

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

A. Dynamical systems

We consider continuous-time autonomous *dynamical systems* $\dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t))$, with $\boldsymbol{x}(t) \in \mathcal{X}$ denoting the *state* of the system at time $t \in \mathbb{R}$ and $\mathcal{X} \subseteq \mathbb{R}^n$ denotes the *state space*.

Assumption 1: *The vector field $f : \mathcal{X} \to \mathbb{R}^n$ is continuously differentiable and the state space $\mathcal{X} \subseteq \mathbb{R}^n$ is convex.*

Under Assumption 1, the Jacobian of the vector field f is denoted by $Df(x)$. A dynamical system can be described in terms of its flow $\varphi(t, \boldsymbol{x}_0)$ denoting the state at time t as

$$\boldsymbol{x}(t) = \varphi(t, \boldsymbol{x}_0), \quad \forall t \geq 0, \quad \text{with} \quad \boldsymbol{x}(0) = \boldsymbol{x}_0.$$

The sequence of all values taken by the state vector is called the *trajectory* of the system. A point $\boldsymbol{x}_e \in \mathcal{X}$ is called an *equilibrium point* if $f(\boldsymbol{x}_e) = 0$, and the set of equilibrium points is denoted by $\mathcal{F}(f) = \{\boldsymbol{x}_e \in \mathcal{X} : f(\boldsymbol{x}_e) = \mathbf{0}\}$. A trajectory starting at $\boldsymbol{x}_0$ is said to converge asymptotically toward an equilibrium point $\boldsymbol{x}_e$ if $\lim_{t\to\infty} \varphi(t, \boldsymbol{x}_0) = \boldsymbol{x}_e$. We conclude by defining the properties of nonexpansiveness and contractivity for dynamical systems.

Definition 1 (Nonexpansiveness and contractivity): *Let $\|\cdot\|$ be a norm in $\mathbb{R}^n$. A system on $\mathcal{X} \subseteq \mathbb{R}^n$ is nonexpansive if for all $\boldsymbol{x}_0, \boldsymbol{y}_0 \in \mathcal{X}$ it holds*

$$\|\varphi(t, \boldsymbol{x}_0) - \varphi(t, \boldsymbol{y}_0)\| \leq \|\boldsymbol{x}_0 - \boldsymbol{y}_0\|, \qquad t \geq 0.$$

If the inequality holds strictly, the system is contractive.

B. Convergence of nonexpansive monotone systems

Consider the Euclidean space $\mathbb{R}^n$ equipped with the standard partial order $\leq$. Dynamical systems in $(\mathcal{X}, \leq)$, with $\mathcal{X} \subseteq \mathbb{R}^n$, whose flow preserves such order w.r.t. initial conditions are referred to as "order-preserving" [20], [22] or "monotone" [4], [5]; this manuscript uses of the latter denomination. We formally define the monotonicity property in Definition 2, along with the special class termed "type-K monotonicity" in Definition 3, introduced by us in [10], [20], [23].

Definition 2 (Monotonicity): *A system on $\mathcal{X} \subseteq \mathbb{R}^n$ is "monotone" if for all $\boldsymbol{x}_0, \boldsymbol{y}_0 \in \mathcal{X}$ it holds:*

$$\boldsymbol{x}_0 \leq \boldsymbol{y}_0 \Rightarrow \varphi(t, \boldsymbol{x}_0) \leq \varphi(t, \boldsymbol{y}_0), \quad \forall t \geq 0.$$

Definition 3 (Type-K Monotonicity): *A system on $\mathcal{X} \subseteq \mathbb{R}^n$ is "type-K monotone" if it is monotone and if for all $\boldsymbol{x}_0, \boldsymbol{y}_0 \in \mathcal{X}$ and for all $i = 1, \dots, n$ it holds:*

$$\boldsymbol{x}_0 \leq \boldsymbol{y}_0 \ \wedge \ x_{0,i} < y_{0,i} \Rightarrow \varphi_i(t, \boldsymbol{x}_0) < \varphi_i(t, \boldsymbol{y}_0), \quad \forall t \geq 0.$$

We have recently shown in [10] that for continuous-time smooth dynamical systems, monotonicity and type-K monotonicity are equivalent properties, and they can be verified by the sign structure of the Jacobian of the vector field.

Lemma 1: *For a system $\dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t))$ under Assumption 1, the following statements are equivalent:*
(a) The system is monotone;
(b) The system is type-K monotone;
(c) The Jacobian $Df(\boldsymbol{x})$ is Metzler for all $\boldsymbol{x} \in \mathcal{X}$.

Proof: Under Assumption 1, $(a) \Leftrightarrow (b)$ holds by [10, Theorem 3] and $(a) \Leftrightarrow (c)$ holds by [10, Proposition 2]. Note that $(a) \Leftrightarrow (c)$ was originally proved in [4]. ■

A nice feature of type-K monotonicity is that it allows to prove convergence toward equilibrium points for systems that are nonexpansive w.r.t. the sup-norm $\|\cdot\|_\infty$ and admit at least one equilibrium point. This result, which we explicitly prove here for monotone systems with a continuously differentiable vector field, was also exploited in [10, Theorem 1] to prove convergence toward equilibrium points for K-subtopical systems – that is, systems that are type-K monotone and $\mathbf{1}$-subhomogeneous as eq. (1), both in continuous- and discrete-time.

Proposition 1: *Consider a system $\dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t))$ under Assumption 1 satisfying the following:*
- *the system is monotone and nonexpansive w.r.t. $\|\cdot\|_\infty$;*
- *the set of equilibrium points $\mathcal{F}(f) \neq \emptyset$ is not empty.*

Then all equilibrium points are stable and each trajectory converges asymptotically to one of them.

Proof: Since the system is monotone and continuously differentiable, then it is also type-K monotone by Lemma 1. Since the system is nonexpansive w.r.t. $\|\cdot\|_\infty$, by [22, Lemma 2.7.2] we know that the flow is such that

$$\varphi(t, \boldsymbol{x}_0 + \alpha\mathbf{1}) \leq \varphi(t, \boldsymbol{x}_0) + \alpha\mathbf{1}, \qquad \forall\alpha \geq 0, \ \forall t \geq 0. \quad (1)$$

As it will be formalized in Section III, systems satisfying the above property are called "$\mathbf{1}$-subhomogeneous", or "plus-subhomogeneous" (see [10, Definition 2]). Type-K monotonicity and plus-subhomogeneity of the systems ensure stability of equilibrium points [10, Lemma 4] and convergence of all trajectories to equilibrium points [10, Theorem 1]. ■

We remark that the "monotonicity of a dynamical system" is to be intended in the sense of Kamke-Muller [2], [3], which must not be confused with the "monotonicity of an operator" in functional analysis [24, Definition 12.1]. This clarification is important in this manuscript as the notion of monotonicity of an operator is used in the same context of our application, that is recurrent neural networks. For instance, in [25], [26] the notion of monotonicity of an operator is generalized to Banach spaces by using weak pairings as a substitute for inner products, which is then used to effciently compute equilibria of recurrent neural networks via fixed-point iterations.

Example 1: *We now provide two examples proving that the two notions of monotonicity are different. First, recall that a linear dynamical system, both in continuous-time* $\dot{\boldsymbol{x}}(t) = A\boldsymbol{x}(t)$ *and in discrete-time* $\boldsymbol{x}(k+1) = A\boldsymbol{x}(k)$*, is monotone in the sense of Definition 2 if* A *is nonnegative [10, Proposition 2 and Theorem 5]. Secondly, recall that a linear operator* $\mathcal{A} : \boldsymbol{x} \mapsto A\boldsymbol{x}$ *is monotone in the sense of [24, Definition 12.1] if and only if it is positive semidefinite, i.e.,* $\boldsymbol{x}^\top A\boldsymbol{x} \geq 0$ *for all* $\boldsymbol{x}$ *[24, Example 12.2]. Let us define:*

$$A_1 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

One can verify that: 1) a dynamical system ruled by A_1 *is monotone because* $A_1 \geq 0$*, but the operator* $\mathcal{A}_1$ *is not monotone since for* $\boldsymbol{x} = [-2,\ 2,\ 1]^\top$ *it holds* $\boldsymbol{x}^\top A_1 \boldsymbol{x} = -1$*; 2) a dynamical system ruled by* A_2 *is not monotone because* $A_2 \not\geq 0$*, but the operator* $\mathcal{A}_2$ *is monotone because it is symmetric with nonnegative eigenvalues.*

III. Novel Convergence Results

We start this section by generalizing the stability and convergence results in Proposition 1 to systems that are nonexpansive w.r.t. a weighted sup-norm.

Proposition 2: *Consider a system* $\dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t))$ *under Assumption 1 satisfying the following:*

- *the system is monotone nonexpansive w.r.t.* $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$*;*
- *the set of equilibrium points* $\mathcal{F}(f) \neq \emptyset$ *is not empty.*

Then all equilibrium points are stable and each trajectory converges asymptotically to one of them.

Proof. Consider the change of variable $\boldsymbol{z}(t) = [\boldsymbol{\eta}]^{-1}\boldsymbol{x}(t)$ with $\boldsymbol{\eta} \in \mathbb{R}^n_+$, yielding the system $\dot{\boldsymbol{z}}(t) = g(\boldsymbol{z}(t))$. Let $\varphi(t,\cdot)$ and $\phi(t,\cdot)$ denote the flows of the system in the original and new sets of coordinates, respectively. By Assumption 1, both vector fields are continuously differentiable and related by:

$$\phi(t,\boldsymbol{z}) = [\boldsymbol{\eta}]^{-1}\varphi(t,[\boldsymbol{\eta}]\boldsymbol{z}) = [\boldsymbol{\eta}]^{-1}\varphi(t,\boldsymbol{x}), \tag{2}$$

$\forall \boldsymbol{x},\boldsymbol{z} \in \mathbb{R}^n$ s.t. $\boldsymbol{z} = [\boldsymbol{\eta}]^{-1}\boldsymbol{x}$. This means that any trajectory $(\varphi(t,\boldsymbol{x}))_{t\geq 0}$ has the same behavior of the trajectory $(\phi(t,\boldsymbol{z}))_{t\geq 0}$. Thus, the proof reduces to show that Proposition 1 holds for the system in the new set of coordinates:

- $\mathcal{F}(g) \neq \emptyset$ if and only if $\mathcal{F}(f) \neq \emptyset$, indeed, for any $\boldsymbol{x}_e \in \mathcal{F}(f)$ then $\boldsymbol{z}_e = [\boldsymbol{\eta}]^{-1}\boldsymbol{x}_e \in \mathcal{F}(g)$ and vice versa.
- One system is monotone if and only if the other one is, due to Lemma 1 and the fact that $[\boldsymbol{\eta}]Dg(\boldsymbol{z}) = Df([\boldsymbol{\eta}]\boldsymbol{z})$.
- The original system is nonexpansive w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$ if and only if the system in the new set of coordinates is nonexpansive w.r.t. $\|\cdot\|_\infty$ due to eq. (2), which yields the following (bidirectional) chain of inequalities:

$$\begin{gathered}\|\phi(t,\boldsymbol{z})-\phi(t,\boldsymbol{v})\|_\infty \leq \|\boldsymbol{z}-\boldsymbol{v}\|_\infty, \\ \left\|[\boldsymbol{\eta}]^{-1}\varphi(t,\boldsymbol{x})-[\boldsymbol{\eta}]^{-1}\varphi(t,\boldsymbol{y})\right\|_\infty \leq \left\|[\boldsymbol{\eta}]^{-1}\boldsymbol{x}-[\boldsymbol{\eta}]^{-1}\boldsymbol{y}\right\|_\infty, \\ \|\varphi(t,\boldsymbol{x})-\varphi(t,\boldsymbol{y})\|_{\infty,[\boldsymbol{\eta}]^{-1}} \leq \|\boldsymbol{x}-\boldsymbol{y}\|_{\infty,[\boldsymbol{\eta}]^{-1}}.\end{gathered} \tag{3}$$

■

Remark 1 (Comparison with Theorem 21 in [21]): *Our Proposition 2 has the advantage of not requiring piecewise real analyticity of the vector field to prove convergence of the trajectories to the set of equilibrium points. Another advantage is the fact that Lemmas 1-2, together with Theorem 1, provide easy-to-check conditions to apply Proposition 2, which can be verified by looking at each row of the Jacobian matrix independently. For instance, this is particularly useful in the context of multi-agent systems where these conditions translate into properties of the local interaction rules between agents. On the other hand, [21, Theorem 21] applies to general systems, not necessarily monotone, which are nonexpansive w.r.t. a norm* $\|\cdot\|_{p,Q}$ *where* $p \in \{1,\infty\}$ *and* $Q \in \mathbb{R}^{n\times n}$ *is invertible.*

We are going to prove in Theorem 1 that monotone systems are nonexpansive w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$ if and only if they are $\boldsymbol{\eta}$-subhomogeneus, as defined next.

Definition 4 (Subhomogeneity): *A dynamical system on* $\mathcal{X} \in \mathbb{R}^n$ *is "*$\boldsymbol{\eta}$*-subhomogeneous", where* $\boldsymbol{\eta} \in \mathbb{R}^n_+$ *is a positive vector, if for all initial conditions* $\boldsymbol{x}_0 \in \mathcal{X}$ *it holds:*

$$\varphi(t,\boldsymbol{x}_0+\alpha\boldsymbol{\eta}) \leq \varphi(t,\boldsymbol{x}_0)+\alpha\boldsymbol{\eta}, \qquad \forall\alpha\geq 0,\ \forall t\geq 0.$$

The system is $\boldsymbol{\eta}$*-homogeneous if the equality holds for* $\alpha \in \mathbb{R}$*.*

Note that $\boldsymbol{\eta}$-subhomogeneity encompasses properties like plus-subhomogeneity, where $\boldsymbol{\eta} = \mathbf{1}$ [10], [22], and translation invariance, corresponding to $\boldsymbol{\eta}$-homogeneity [5], [18]. We now state the main result of this section, which (together with the following Lemma 2) provides an operative way to use the stability and convergence results in Proposition 2.

Theorem 1: *Consider a monotone system* $\dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t))$ *under Assumption 1. Then, it is* $\boldsymbol{\eta}$*-subhomogeneous if and only if it is nonexpansive w.r.t.* $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$*.*

Proof: Consider the same change of variable in the proof of Proposition 2, i.e., $\boldsymbol{z}(t) = [\boldsymbol{\eta}]^{-1}\boldsymbol{x}(t)$ with $\boldsymbol{\eta} \in \mathbb{R}^n_+$. We first prove that the system is $\boldsymbol{\eta}$-subhomogeneous if and only if the system in the new set of coordinates is $\mathbf{1}$-subhomogeneous by the following (bidirectional) chain of inequalities,

$$\begin{gathered}\phi(t,\boldsymbol{z}_0+\alpha\mathbf{1}) \leq \phi(t,\boldsymbol{z}_0)+\alpha\mathbf{1} \\ [\boldsymbol{\eta}]^{-1}\varphi(t,[\boldsymbol{\eta}]\boldsymbol{z}_0+\alpha[\boldsymbol{\eta}]\mathbf{1}) \leq [\boldsymbol{\eta}]^{-1}\varphi(t,[\boldsymbol{\eta}]\boldsymbol{z}_0)+\alpha\mathbf{1} \\ [\boldsymbol{\eta}]^{-1}\varphi(t,\boldsymbol{x}_0+\alpha\boldsymbol{\eta}) \leq [\boldsymbol{\eta}]^{-1}\varphi(t,\boldsymbol{x}_0)+\alpha\mathbf{1} \\ \varphi(t,\boldsymbol{x}_0+\alpha\boldsymbol{\eta}) \leq \varphi(t,\boldsymbol{x}_0)+\alpha\boldsymbol{\eta}\end{gathered}$$

Secondly, the system is monotone if and only if the system in the new set of coordinates is monotone, as already proven in the proof of Proposition 2. Under monotonicity, $\mathbf{1}$-subhomogeneity is equivalent to nonexpansiveness w.r.t. $\|\cdot\|_\infty$, because for any $t \geq 0$ Lemma 2.7.2 in [22] holds for the self-map $\phi^t(\boldsymbol{x}) := \phi(t, \boldsymbol{x}) : \mathcal{X} \to \mathcal{X}$. In turn, it is equivalent to nonexpansiveness w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$ of the system by eq. (3). ■

We also provide two equivalent necessary and sufficient conditions for $\boldsymbol{\eta}$-subhomogeneity for monotone systems.

Lemma 2: *For a monotone system $\dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t))$ under Assumption 1, the following statements are equivalent:*

(a) the system is $\boldsymbol{\eta}$-subhomogeneous;

(b) the vector field satisfies $f(\boldsymbol{x}+\alpha\boldsymbol{\eta}) \leq f(\boldsymbol{x})$, $\forall \boldsymbol{x} \in \mathcal{X}, \alpha \geq 0$;

(c) the Jacobian satisfies $Df(\boldsymbol{x})\boldsymbol{\eta} \leq \mathbf{0}$, $\forall \boldsymbol{x} \in \mathcal{X}$.

Proof: Under Assumption 1 we have that:

- $(a) \Rightarrow (b)$ is logically equivalent to its contrapositive $\neg(b) \Rightarrow \neg(a)$, where "$\neg$" denotes the "not" logical operation. If (b) does not hold, then there exist a point $\boldsymbol{x} \in \mathcal{X}$ and a component $i \in \{1, \dots, n\}$ such that
$$f_i(\boldsymbol{x} + \alpha\boldsymbol{\eta}) > f_i(\boldsymbol{x}).$$
By the continuous differentiability of the vector field (Assumption 1), there exists $T > 0$ such that the distance at time $t = T$ between the i-th components of the flows is greater than the initial distance at time $t = 0$, namely,
$$\varphi_i(T,\boldsymbol{x}+\alpha\boldsymbol{\eta})-\varphi_i(T,\boldsymbol{x}) > \underbrace{\varphi_i(0,\boldsymbol{x}+\alpha\boldsymbol{\eta})-\varphi_i(0,\boldsymbol{x})}_{\alpha\eta_i}. \quad (4)$$
Eq. (4) implies that the system is not $\boldsymbol{\eta}$-subhomogeneous, i.e., (a) does not hold.
- $(b) \Rightarrow (a)$ is proven by contradiction. Consider any point $\boldsymbol{x} \in \mathbb{R}^n$ and, for the sake of contradiction, assume that there exists a finite time $T > 0$ after which condition (a) does not hold and let T be the minimum such time, while up to T it holds by condition (b) and the continuous differentiability of the flow (a consequence of Assumption 1). Namely, there exists an arbitrary small $\varepsilon > 0$ such that
$$\varphi(t, \boldsymbol{x} + \alpha\boldsymbol{\eta}) \leq \varphi(t, \boldsymbol{x}) + \alpha\boldsymbol{\eta}, \quad t \in [0, T], \quad (5)$$
$$\varphi(t, \boldsymbol{x} + \alpha\boldsymbol{\eta}) \not\leq \varphi(t, \boldsymbol{x}) + \alpha\boldsymbol{\eta} \quad t \in (T, T+\varepsilon]. \quad (6)$$
Let $\boldsymbol{v} \geq 0$ be the nonnegative vector that fills the gap in the inequality at time T, i.e.,
$$\varphi(T, \boldsymbol{x} + \alpha\boldsymbol{\eta}) = \varphi(T, \boldsymbol{x}) + \alpha\boldsymbol{\eta} - \boldsymbol{v}. \quad (7)$$
We now find a contradiction to eq. (6). Let $t > T$, then:
$$\begin{aligned} \varphi(t,\boldsymbol{x}+\alpha\boldsymbol{\eta}) &\overset{(i)}{=} \varphi(t-T,\varphi(T,\boldsymbol{x}+\alpha\boldsymbol{\eta})) \\ &\overset{(ii)}{=} \varphi(t-T,\varphi(T,\boldsymbol{x})+\alpha\boldsymbol{\eta}-\boldsymbol{v}) \\ &\overset{(iii)}{\leq} \varphi(t-T,\varphi(T,\boldsymbol{x})+\alpha\boldsymbol{\eta}) \\ \exists\delta^\star>0: \quad &\overset{(iv)}{\leq} \varphi(t-T,\varphi(T,\boldsymbol{x}))+\alpha\boldsymbol{\eta}, \quad \forall t\in[T,T+\delta^\star) \\ &\overset{(v)}{\leq} \varphi(t,\boldsymbol{x})+\alpha\boldsymbol{\eta} \end{aligned}$$
where (i) and (v) hold by the group law, which applies to continuously differentiable flows (cfr. [27, Section 7.1]); (ii) holds by eq. (7); (iii) holds by monotonicity; (iv) holds by assumption (b), which implies $f(\varphi(T, \boldsymbol{x}) + \alpha\boldsymbol{\eta}) \leq f(\varphi(T, \boldsymbol{x}))$, and by the continuous differentiability of the flow, which implies that $\exists\delta^\star > 0$ such that $\varphi(\delta, \varphi(T, \boldsymbol{x}) + \alpha\boldsymbol{\eta}) \leq \varphi(\delta, \varphi(T, \boldsymbol{x})) + \alpha\boldsymbol{\eta}$ for all $\delta \in [0, \delta^\star]$. This contradicts the existence of $\varepsilon > 0$ in eq. (6), i.e., there does not exists a finite T such that eq. (6) holds. In turn, eq. (5) holds for all $T \geq 0$, i.e., condition (a) holds.
- $(b) \Rightarrow (c)$ is proven by the definition of the directional derivative,
$$\begin{aligned} Df(\boldsymbol{x})\boldsymbol{\eta} &= \lim_{\alpha\to0^+} \frac{f(\boldsymbol{x}+\alpha\boldsymbol{\eta}) - f(\boldsymbol{x})}{\alpha} \\ &\overset{(b)}{\leq} \lim_{\alpha\to0^+} \frac{f(\boldsymbol{x}) - f(\boldsymbol{x})}{\alpha} = \mathbf{0}. \end{aligned}$$
- $(c) \Rightarrow (b)$ is proven by the Newton-Leibnitz formula for vector-valued continuously-differentiable functions:
$$\begin{aligned} [\alpha\boldsymbol{\eta}]^{-1}(f(\boldsymbol{x}+\alpha\boldsymbol{\eta}) - f(\boldsymbol{x})) &= \int_0^1 Df(\boldsymbol{x}+s\alpha\boldsymbol{\eta})ds \\ [\alpha\boldsymbol{\eta}]^{-1}(f(\boldsymbol{x}+\alpha\boldsymbol{\eta}) - f(\boldsymbol{x})) &\leq \mathbf{0} \\ f(\boldsymbol{x}+\alpha\boldsymbol{\eta}) &\leq f(\boldsymbol{x}), \end{aligned}$$
where $[\alpha\boldsymbol{\eta}]^{-1}$ is a diagonal matrix with elements $1/\alpha\eta_i$. ■

Remark 2: *The results of Lemma 2 and Theorem 1 are compatible with known results in the literature. In particular, let the diagonally weighted logarithmic sup-norm of a matrix $M \in \mathbb{R}^{n\times n}$ be denoted by*
$$\mu_{\infty,[\boldsymbol{\eta}]^{-1}}(M) = \max_{i=1,\cdots,n} \Big(m_{ii} + \sum_{j=1,j\neq i}^{n} \frac{\eta_j}{\eta_i}|m_{ij}|\Big). \quad (8)$$
Statement (c) in Lemma 2 is equivalent to $\mu_{\infty,[\boldsymbol{\eta}]^{-1}}(Df(\boldsymbol{x})) \leq \mathbf{0}$, $\forall \boldsymbol{x} \in \mathcal{X}$ according to [18, Lemma 4.17] and, in turn, it is equivalent to nonexpansiveness w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$ according to [28, Theorem 29]. In other words, systems satisfying the conditions in [28, Theorem 29] are subhomogeneous. This yields the following open question: "Does Theorem 1 hold without Assumption 1?".

Let us discuss an example of a monotone and subhomogeneous system, inspired from [19, Example 4.3], that is not contractive w.r.t. any diagonally weighted norm but nonexpansive w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$ for some $\boldsymbol{\eta} > 0$, and whose trajectories converge according to Proposition 1.

Example 2: *Consider the class of dynamical systems on $\mathbb{R}^2$*
$$\begin{aligned} \dot{x}_1(t) &= -x_1(t) + \alpha x_2(t) - \gamma g(x_1) \\ \dot{x}_2(t) &= \beta x_1(t) - x_2(t) \end{aligned}$$
where $\alpha, \beta, \gamma \geq 0$ and $g : \mathbb{R} \mapsto \mathbb{R}_{\geq 0}$ is continuosly differentiable with $g(0) = 0$ and positive derivative $g'(x) \geq 0$ for all $x \in \mathbb{R}$. Thus, the system is monotone because the Jacobian of the vector field $f := [f_1, f_2]^\top$ is Metzler,
$$Df(x_1, x_2) = \begin{bmatrix} -1 - \gamma\frac{d}{dt}g(x_1) & \alpha \\ \beta & -1 \end{bmatrix}.$$

Subhomogeneity is verified for some vector $\boldsymbol{\eta} \in \mathbb{R}^n_+$ by solving the system of linear equations $Df(x_1, x_2)\boldsymbol{\eta} \leq 0$,

$$\begin{cases} -(1-\gamma\frac{d}{dt}g(x_1))\eta_1 + \alpha\eta_2 \leq 0 \\ \beta\eta_1 - \eta_2 \leq 0 \end{cases} \Rightarrow \quad \eta_2 \in [\beta\eta_1, \frac{1}{\alpha}\eta_1],$$

which is a set of admissible solutions that hold for all values of $\gamma \geq 0$. Thus, the system is $\boldsymbol{\eta}$-subhomogeneous if $\alpha\beta \leq 1$. Since the origin is an equilibrium point of the system, one can exploit Proposition 1 and Theorem 1 to prove the convergence of all trajectories toward some equilibrium point.

Now, consider the special case $\alpha = 0.5$, $\beta = 2$, and $\gamma = 0$, for which the system becomes linear with dynamics

$$\dot{\boldsymbol{x}}(t) = A\boldsymbol{x}(t), \quad A = \begin{bmatrix} -1 & 0.5 \\ 2 & -1 \end{bmatrix}, \quad \lambda_1 = 0, \lambda_2 = -2.$$

Since the matrix has a null eigenvalue it is singular, then the system admits infinitely many equilibrium points, which forms the span of the eigenvector $\boldsymbol{v}_1 = [1,\ 2]^\top$. This implies that the system cannot be contracting w.r.t. any diagonally weighted norm. Indeed, the so-called logarithmic norm [18, Section 2.4] is lower bounded by the greatest eigenvalue for any $p \in [1, \infty]$ and any $\boldsymbol{\eta} \in \mathbb{R}^n_+$, and thus it is surely nonnegative (see [29, Lemma 1]). This implies that the system is non-contracting according to [28, Theorem 29]. In contrast, Proposition 1 ensures the convergence of the system's trajectories despite the fact it is non-contracting but only nonexpansive.

IV. Stability of Nonexpansive Monotone RNNs

We consider two models of RNNs [29], [30], the Hopfield and the firing-rate models, with dynamics

$$\dot{\boldsymbol{x}}(t) = f_{\text{H}}(\boldsymbol{x}(t)) := -C\boldsymbol{x}(t) + A\Phi(\boldsymbol{x}(t)) + \boldsymbol{b}, \qquad (9)$$

$$\dot{\boldsymbol{x}}(t) = f_{\text{FR}}(\boldsymbol{x}(t)) := -C\boldsymbol{x}(t) + \Phi(A\boldsymbol{x}(t) + \boldsymbol{b}), \qquad (10)$$

where $C \in \mathbb{R}^{n\times n}$ is a positive diagonal matrix, $A \in \mathbb{R}^{n\times n}$ is an arbitrary matrix, $\boldsymbol{b} \in \mathbb{R}^n$ is a constant input, and $\Phi : \mathbb{R}^n \mapsto \mathbb{R}^n$ is an activation function satisfying Assumption 2.

Assumption 2: *Activation functions are diagonal, i.e., $\Phi(\boldsymbol{x}) = [\phi_1(x_1), \cdots, \phi_n(x_n)]^\top$ where each $\phi_i : \mathbb{R} \mapsto \mathbb{R}$ is continuously differentiable and globally Lipschitz, i.e., there exists finite $d_1 \leq d_2$ such that for all $i = 1, \ldots, n$ it holds*

$$\frac{d}{dx}\phi_i(x) \in [d_1, d_2], \qquad \forall x \in \mathbb{R},$$

and the Lipschitz constant is given by $\bar{d} = \max\{|d_1|, |d_2|\}$

We now study their convergence toward equilibrium points.

Theorem 2: *Consider Hopfield and firing-rate neural networks as in eqs. (9)-(10) with activation function satisfying Assumption 2. Let $A_\star = \min\{d_1A, d_2A\}$ and $A^\star = \max\{d_1A, d_2A\}$ satisfy the following conditions:*

a) $A_\star$ is Metzler (monotonicity);
b) $\exists\boldsymbol{\eta} \in \mathbb{R}^n_+ :\ (A^\star - C)\boldsymbol{\eta} \leq \mathbf{0}$ ($\boldsymbol{\eta}$-subhomogeneity).

Then, all their trajectories converge to some equilibrium point, if any exists.

Proof: Under Assumption 2, both Hopfield and firing-rate neural networks are monotone if and only if condition $a)$ holds. In such case, The Jacobian matrix computed at a generic point $\boldsymbol{x} \in \mathbb{R}$ is lower bounded by

$$Df_{\text{H}}(\boldsymbol{x}) = AD\Phi(\boldsymbol{x}) - C \geq \min\{d_1A, d_2A\} - C = A_\star - C$$
$$Df_{\text{FR}}(\boldsymbol{x}) = D\Phi(A\boldsymbol{x} + \boldsymbol{b})A - C \geq \min\{d_1A, d_2A\} - C = A_\star - C$$

where $A_\star - C$ is Metzler if and only if $A_\star$ is Metzler, since C is diagonal. By Lemma 1, the Jacobian is Metzler if and only if the smooth system is monotone. Secondly, we prove that both Hopfield and firing-rate neural networks are $\boldsymbol{\eta}$-subhomogeneous if there is $\boldsymbol{\eta} \in \mathbb{R}^n_+$ such that condition $b)$ holds. For both networks it holds:

$$\begin{aligned} Df_{\text{H}}(\boldsymbol{x})\boldsymbol{\eta} &= (AD\Phi(\boldsymbol{x}) - C)\boldsymbol{\eta} \\ &\leq (\max\{d_1A, d_2A\} - C)\boldsymbol{\eta} = (A^\star - C)\boldsymbol{\eta} \\ Df_{\text{FR}}(\boldsymbol{x})\boldsymbol{\eta} &= (D\Phi(A\boldsymbol{x} + \boldsymbol{b})A - C)\boldsymbol{\eta} \\ &\leq (\max\{d_1A, d_2A\} - C)\boldsymbol{\eta} = (A^\star - C)\boldsymbol{\eta}. \end{aligned}$$

Thus, if $(A^\star - C)\boldsymbol{\eta} \leq \mathbf{0}$ then both Jacobians are non-positive and, in turn, the system is subhomogeneous by Lemma 2.

We have proved that conditions $a)$ and $b)$ imply that both neural networks are monotone and $\boldsymbol{\eta}$-subhomogeneous. Thus Theorem 1 ensures that they are also nonexpansive w.r.t. $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$ and Proposition 1 ensures the convergence of all trajectories toward equilibrium points, if any exists. ■

A. Comparison with contractive neural networks

We compare our results with those provided in Section V of the recent work of Davydov, Proskurnikov, and Bullo [29], whose extended version with all proofs and some additional results is [30]. Let $\mu_{\infty,[\boldsymbol{\eta}]^{-1}}(\cdot)$ denote the diagonally weighted logarithmic sup-norm as in eq. (8). Then, Theorem 21 in [30] gives the following condition for contraction of Hopfield neural networks

$$\max\begin{Bmatrix} \mu_{\infty,[\boldsymbol{\eta}]^{-1}}(\bar{d}A - (\bar{d} - d_1)A \odot I - C) \\ \mu_{\infty,[\boldsymbol{\eta}]^{-1}}(\bar{d}A - (\bar{d} - d_2)A \odot I - C) \end{Bmatrix} < 0, \qquad (11)$$

where $\bar{d}$ is the Lipschitz constant of the activation functions (Assumption 2), I is the identity matrix, and $\odot$ denotes the element-wise product. Theorem 23 in [30] gives the following condition for contraction of firing-rate neural networks,

$$\max\begin{Bmatrix} \mu_{\infty,[\boldsymbol{\eta}]^{-1}}(d_1A - C) \\ \mu_{\infty,[\boldsymbol{\eta}]^{-1}}(d_2A - C) \end{Bmatrix} < 0. \qquad (12)$$

Under condition $a)$ of Theorem 2, which implies that d_1A and d_2A are Metzler, both conditions in eq. (11)-(12), reduce to

$$\mu_{\infty,[\boldsymbol{\eta}]^{-1}}(A^\star - C) < 0, \qquad (13)$$

which is a stricter than condition $b)$ in Theorem 2, indeed,

$$\mu_{\infty,[\boldsymbol{\eta}]^{-1}}(A^\star - C) < 0 \Leftrightarrow (A^\star - C)\boldsymbol{\eta} < \mathbf{0}.$$

Thus the class of neural networks identified by Theorem 2 includes networks that are not considered in [29], [30].

Example 3 (Nonexpansive RNNs): *Consider a Hopfield or a firing-rate RNN with dynamics ruled by*

$$C = I, \quad A = \begin{bmatrix} 0 & 0.5 \\ 2 & 0 \end{bmatrix}, \quad \phi(x) = \tanh(x),$$

where Assumption 2 is satisfied with $d_1 = 0$ and $d_2 = 1$, and thus $\overline{d} = 1$. Condition (13) reads as

$$\mu_{\infty,[\boldsymbol{\eta}]^{-1}}(A - C) = \max\left\{\frac{\eta_2}{2\eta_1}, \frac{2\eta_1}{\eta_2}\right\} - 1 < 0$$

which has no feasible solution. Thus, the system is not contracting w.r.t. to $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$, instead it is nonexpansive w.r.t. to the $\|\cdot\|_{\infty,[\boldsymbol{\eta}]^{-1}}$ for $\boldsymbol{\eta} = \boldsymbol{v}_1$ where $\boldsymbol{v}_1 = [1,\ 2]^\top$ is the eigenvector of A associated with the eigenvalue $\lambda_1 = 1$, because all conditions of Theorem 1 hold:

- *Assumption 1 is satisfied because the activation function is the continuously differentiable hyperbolic tangent;*
- *The system is monotone since the Jacobian $Df(\boldsymbol{x}_0) \geq -C$ is Metzler for any $\boldsymbol{x}_0 \in \mathbb{R}^n$ according to Lemma 1, because C is diagonal.*
- *The system is $\boldsymbol{v}_1$-subhomogeneous according to Lemma 2, because the Jacobian satisfies $Df(\boldsymbol{x}_0)\boldsymbol{v}_1 \leq (A - C)\boldsymbol{\eta} = A\boldsymbol{v}_1 - C\boldsymbol{v}_1 = \boldsymbol{v}_1 - \boldsymbol{v}_1 = \mathbf{0}$*

Thus, the neural network satisfies conditions of Theorem 2 and thus all trajectories converge to some equilibrium point.

Other examples of nonexpansive RNNs that are nonexpansive but not contracting can be found for any nonnegative matrix $A \geq 0$ and choosing:

1) $C = \lambda_{\text{MAX}} I$, where λ_{MAX} is the largest eigenvalue of A: the system is nonexpansive w.r.t. $\|\cdot\|_{\infty,\boldsymbol{v}^{-1}}$ where $\boldsymbol{v}$ is the eigenvector associated with λ_{MAX};
2) $C = \text{diag}(A\mathbf{1})$. In this case, the system is nonexpansive w.r.t. $\|\cdot\|_\infty$;
3) $C = \text{diag}((A\boldsymbol{\eta}))[\boldsymbol{\eta}]^{-1}$ for any $\boldsymbol{\eta} \geq 0$. In this case, the system is nonexpansive w.r.t. $\|\cdot\|_{\infty,\boldsymbol{\eta}^{-1}}$.

V. Conclusions

It has been shown that smooth monotone systems that are nonexpansive w.r.t. a diagonally weighted infinity norm exhibit aperiodic state trajectories that converge to one of the equilibrium points. Notably, this differs from prevailing trends in the literature by not requiring the system to be contractive, thus accommodating multiple equilibrium points. This nice behavior is ensured thanks to the fact that smooth monotone systems naturally enjoy a stricter notion of monotonicity called *type-K monotonicity* [10], [20], [23], which prevents periodic trajectories. These findings apply also to RNNs, allowing us to provide sufficient convergence conditions for nonexpansive monotone neural networks that lack contractive properties.

References

[1] H. L. Smith, "Monotone dynamical systems: Reflections on new advances & applications," *Discrete & Continuous Dynamical Systems*, vol. 37, no. 1, 2017.

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

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

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

[5] D. Angeli and E. D. Sontag, "Translation-invariant monotone systems, and a global convergence result for enzymatic futile cycles," *Nonlinear Analysis: Real World Applications*, vol. 9, no. 1, 2008.

[6] S. Coogan, "Mixed monotonicity for reachability and safety in dynamical systems," in *59th IEEE Conference on Decision and Control*, 2020.

[7] G. Como, E. Lovisari, and K. Savla, "Convexity and robustness of dynamic traffic assignment and freeway network control," *Transportation Research Part B: Methodological*, vol. 91, 2016.

[8] E. Lovisari, G. Como, and K. Savla, "Stability of monotone dynamical flow networks," in *53th IEEE Conference on Decision and Control*, 2014.

[9] A. Mauroy and R. Sepulchre, "Contraction of monotone phase-coupled oscillators," *Systems & Control Letters*, vol. 61, no. 11, 2012.

[10] D. Deplano, M. Franceschelli, and A. Giua, "Novel stability conditions for nonlinear monotone systems and consensus in multi-agent networks," *IEEE Transactions on Automatic Control*, 2023.

[11] C. Altafini and G. Lini, "Predictable dynamics of opinion forming for networks with antagonistic interactions," *IEEE Transactions on Automatic Control*, vol. 60, no. 2, 2014.

[12] F. Forni and R. Sepulchre, "Differentially positive systems," *IEEE Transactions on Automatic Control*, vol. 61, no. 2, 2015.

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

[14] S. Coogan and M. Arcak, "Stability of traffic flow networks with a polytree topology," *Automatica*, vol. 66, 2016.

[15] A. Sootla and A. Mauroy, "Operator-theoretic characterization of eventually monotone systems," *IEEE Control Systems Letters*, vol. 2, no. 3, 2018.

[16] S. Coogan, "A contractive approach to separable Lyapunov functions for monotone systems," *Automatica*, vol. 106, 2019.

[17] H. Tsukamoto, S.-J. Chung, and J.-J. E. Slotine, "Contraction theory for nonlinear stability analysis and learning-based control: A tutorial overview," *Annual Reviews in Control*, vol. 52, 2021.

[18] F. Bullo, *Contraction Theory for Dynamical Systems*, 1.1. Kindle Direct Publishing, 2023.

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

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

[21] S. Jafarpour, P. Cisneros-Velarde, and F. Bullo, "Weak and semi-contraction for network systems and diffusively coupled oscillators," *IEEE Transactions on Automatic Control*, vol. 67, no. 3, 2021.

[22] B. Lemmens and R. Nussbaum, *Nonlinear Perron-Frobenius Theory*. Cambridge University Press, 2012.

[23] 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*, 2018.

[24] R. T. Rockafellar and R. J.-B. Wets, *Variational analysis*. Springer Science & Business Media, 2009, vol. 317.

[25] A. Davydov, S. Jafarpour, A. V. Proskurnikov, and F. Bullo, "Non-euclidean monotone operator theory with applications to recurrent neural networks," in *2022 IEEE 61st Conference on Decision and Control*, IEEE, 2022.

[26] A. Davydov, S. Jafarpour, A. V. Proskurnikov, and F. Bullo, "Non-euclidean monotone operator theory and applications," *arXiv preprint arXiv:2303.11273*, 2023.

[27] M. W. Hirsch, S. Smale, and R. L. Devaney, *Differential equations, dynamical systems, and an introduction to chaos*. Academic press, 2012.

[28] A. Davydov, S. Jafarpour, and F. Bullo, "Non-euclidean contraction theory for robust nonlinear stability," *IEEE Transactions on Automatic Control*, vol. 67, no. 12, 2022.

[29] A. Davydov, A. V. Proskurnikov, and F. Bullo, "Non-euclidean contractivity of recurrent neural networks," in *IEEE American Control Conference (ACC)*, 2022.

[30] A. Davydov, A. V. Proskurnikov, and F. Bullo, "Non-euclidean contraction analysis of continuous-time neural networks," *arXiv preprint arXiv:2110.08298*, 2021.