Dynamical Analysis of Full-Range Cellular Neural Networks by Exploiting Differential Variational Inequalities

Guido De Sandre, Mauro Forti, Paolo Nistri, and Amedeo Premoli

***Abstract*—The paper considers the full-range (FR) model of cellular neural networks (CNNs) in the case where the neuron nonlinearities are ideal hard-comparator functions with two vertical straight segments. The dynamics of FR-CNNs, which is described by a differential inclusion, is rigorously analyzed by means of theoretical tools from set-valued analysis and differential inclusions. The fundamental property proved in the paper is that FR-CNNs are equivalent to a special class of differential inclusions termed *differential variational inequalities*. A sound foundation to the dynamics of FR-CNNs is then given by establishing the existence and uniqueness of the solution starting at a given point, and the existence of equilibrium points. Moreover, a fundamental result on trajectory convergence towards equilibrium points (complete stability) for reciprocal standard CNNs is extended to reciprocal FR-CNNs by using a generalized Lyapunov approach. As a consequence, it is shown that the study of the ideal case with vertical straight segments in the neuron nonlinearities is able to give a clear picture and analytic characterization of the salient features of motion, such as the sliding modes along the boundary of the hypercube defined by the hard-comparator nonlinearities. Finally, it is proved that the solutions of the ideal FR model are the uniform limit as the slope tends to infinity of the solutions of a model where the vertical segments in the nonlinearities are approximated by segments with finite slope.**

***Index Terms*—Cellular neural networks (CNNs), differential inclusions, full-range (FR) model, sliding modes, trajectory convergence, variational inequalities.**

NOTATION

$\mathbb{R}^n$	Real n-space.
A	$= [A_{ij}] \in \mathbb{R}^{n\times n}$, Square matrix.
A'	Transpose of A.
A^{-1}	Inverse of A.
x	$= (x_1,\ldots,x_n)' \in \mathbb{R}^n$, Column vector.
$\langle x, y\rangle$	$= \sum_{i=1}^n x_i y_i$, Scalar product of $x, y \in \mathbb{R}^n$.
$\|x\|_2$	$= (\sum_{i=1}^n x_i^2)^{1/2}$, Euclidean norm of $x \in \mathbb{R}^n$.
$B(0,R)$	$= \{y \in \mathbb{R}^n : \|y\|_2 < R\}$, n-dimensional open ball with center 0 and radius R.
$\mathrm{cl}(Q)$	Closure of set $Q \subset \mathbb{R}^n$.
$\mathrm{int}(Q)$	Interior of Q.
$\mathrm{bd}(Q)$	Boundary of Q.
$\mathrm{dist}(x,Q)$	$= \inf_{y\in Q} \|x-y\|_2$, Distance of $x \in \mathbb{R}^n$ from Q.
$\mathcal{T}_Q(x)$	Tangent cone to Q at x.
$\mathcal{N}_Q(x)$	Normal cone to Q at x.
$\mathcal{P}_Q(x)$	Projector of best approximation of x on Q.
$Q_1 \backslash Q_2$	Difference of sets $Q_1, Q_2 \subset \mathbb{R}^n$.
$m(Q)$	Element of Q with the smallest norm.
$Q_1 \pm Q_2$	$= \{x \in \mathbb{R}^n : x = x_1 \pm x_2, \text{with } x_1 \in Q_1, x_2 \in Q_2\}$
$\mathrm{span}\{v_i, i \in \mathcal{I}\}$	Linear subspace of $\mathbb{R}^n$ spanned by vectors $v_i \in \mathbb{R}^n, i \in \mathcal{I}$.
$V_1 \oplus V_2$	Direct sum of subspaces V_1 and V_2.
$\emptyset$	Empty set.
$\phi : \mathbb{R}^n \to \mathbb{R}$	(Conventional) single-valued function.
$\nabla\phi$	(Conventional) gradient of ϕ.
$f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$	Extended-valued function.
$\mathrm{graph}(f)$	Graph of f.
$\mathrm{epi}(f)$	Epigraph of f.
∂f	Generalized gradient of f.

Manuscript received September 25, 2004; revised November 29, 2006 and January 14, 2007. This paper was recommended by Associate Editor L. Trajkovic.

G. De Sandre is with STMicroelectronics, 20041 Agrate Brianza, Italy (e-mail: guido.de-sandre@st.com).

M. Forti and P. Nistri are with the Dipartimento di Ingegneria dell'Informazione, Università di Siena, 53100 Siena, Italy (e-mail: forti@dii.unisi.it; pnistri@dii.unisi.it).

A. Premoli is with the Dipartimento di Elettronica e Informazione, Politecnico di Milano, 20133 Milan, Italy.

Digital Object Identifier 10.1109/TCSI.2007.902607

I. INTRODUCTION

THE standard (S) cellular neural networks (S-CNNs) introduced by Chua and Yang [1] have been one of the most investigated neural paradigms for real time signal processing.

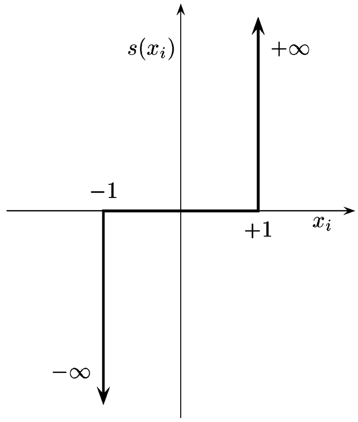


Fig. 1. Hard-comparator nonlinearity.

These networks have a number of advantages over other neural architectures, with the most important one being that they admit a simpler VLSI implementation because of the local and spatially invariant neuron interconnecting structure.

A number of additional advantages in the VLSI implementation of chips with a large number of cells have been achieved by the introduction of the so-called full-range (FR) model of CNNs [2]. In a FR-CNN each neuron is characterized by a self-loss hard-comparator nonlinearity $y = -s(x_i)$ where

$$s(x_i) = \begin{cases} (-\infty, 0], & x_i = -1 \\ 0, & x_i \in (-1, 1) \\ [0, +\infty), & x_i = 1 \end{cases}$$

and x_i is the voltage at the neuron capacitor (see Fig. 1). This nonlinearity is used to prevent the state variables x_i of FR-CNNs from exceeding the range $[-1, 1]$ and is different from that used in S-CNNs, which involves a standard piecewise-linear input–output function with unity gain inside the linearity region and ± 1 saturation levels outside that region [1].

It is worth pointing out that almost all recently manufactured chips [3]–[6] are based on the FR model. These include the well-known ACE4k [4] and ACE16k [6] chips, which implement the CNN Universal Machine [7], i.e., a general purpose vision device incorporating optical sensing and combining logic operations and spatial-temporal dynamics generated within a FR-CNN architecture. The same chips have been demonstrated especially suitable for the implementation of a real-time programmable bifurcation testbed to experimentally uncover new spatial-temporal patterns and complex dynamics [8].

The ideal nonlinearity $s(\cdot)$ exhibits two *vertical* straight segments for $x_i = \pm 1$, and thus it corresponds mathematically to a set-valued map. Therefore, FR-CNNs are described by a *differential inclusion* [9] unlike S-CNNs, where the dynamics obeys a (conventional) differential equation. The goal of this paper is to rigorously analyze FR-CNNs by using appropriate tools from set-valued analysis and differential inclusions. The fundamental property established is that the dynamics of FR-CNNs can be described by a special class of differential inclusions named *differential variational inequalities*. The paper then gives a sound theoretical foundation for the dynamics of FR-CNNs, by addressing the questions of the definition, existence, and uniqueness of the solution of FR-CNNs, and of the existence of at least one stationary solution (an equilibrium point). The analysis is valid in the general case where the neuron interconnection matrix T may be either symmetric or nonsymmetric. Moreover, under the hypothesis of a symmetric T, it is shown that FR-CNNs obey a gradient differential inclusion, and a fundamental result on trajectory convergence towards equilibrium points (complete stability) is obtained via a generalized Lyapunov approach. This extends to reciprocal FR-CNNs previous results on convergence for reciprocal S-CNNs [1].

To the authors' knowledge, FR-CNNs have been analyzed only by approximating the nonlinearity $s(\cdot)$ by a less-hard single-valued comparator function where the vertical segments in $s(\cdot)$ are replaced by segments with a very high but finite slope σ, which leads to FR-CNNs described by a differential equation. It would be more precise to use the name improved signal-range CNN model for the case with finite slope σ [3], since for this model the state variables exit the range $[-1, 1]$. For simplicity, we adopt the acronym FR-CNN also for the case with finite slope. In particular, basic results on the complete stability of FR-CNNs with finite-slope nonlinearities have been established in [3], [10]. Moreover, in [3] it is demonstrated that in the case of finite slope σ, FR-CNNs preserve most functionalities of the original S-CNN model [1]. In this paper, we thus aim to fill the gap in previous work by showing that it is possible to give a rigorous mathematical treatment and a proof of complete stability for the ideal FR-CNN model, i.e., the model obtained when the slope σ is infinity. As it usually happens in the framework of the theory of differential inclusions [9], [11], we will find that the analysis of the ideal case has also the following main advantages.

1) It is simpler and more direct than the approximation approach based on considering first the situation of σ finite and then studying the limit as σ tends to infinity;
2) It is able to give a clear picture and analytic characterization of the salient features of motion, such as the presence of sliding modes along the boundary of the hypercube $H = [-1, 1]^n = [-1, 1] \times [-1, 1] \times \cdots \times [-1, 1]$ defined by the hard-comparator nonlinearities (n is the number of neurons).

A different approach, which keeps exactly vertical straight segments in the ideal nonlinearity $s(\cdot)$, consists in an event-based scheduling of a sequence of affine systems describing the dynamics of the FR-CNN. As the state of the network is confined in the hypercube H, only a subset of the state variables is active (i.e., unconstrained), for any affine system in the sequence, while the remaining state variables are saturated (i.e., constrained to ± 1). The time-domain analysis implied by this approach requires the determination of the time instants when a state variable changes from being unconstrained to being constrained, or vice versa. In correspondence with these switching events, the system of differential equations governing the FR-CNN changes together with the subset of active state variables. We are able to reformulate this system within each inter-event time interval: at a given time instant, the order of the affine system equals the number of active state variables, and varies in the range $0, \ldots, n$. In this way, a solution to the FR-CNN equations is constructed by piecing together solutions of the different affine systems in the event-based sequence. An analogous method was applied in [12] to a class of modified Hopfield-type neural networks. This conventional approach allows the implementation of computer simulator algorithms by using standard numerical techniques [13]. However, it is our opinion that it does not easily allow the determination of general information about trajectory convergence and other global properties of FR-CNNs. Furthermore, such an approach suffers from the fundamental shortcomings pointed out in [14]. Specifically, there are difficulties in defining a solution in cases where a trajectory switches infinitely often between different affine systems during a finite time interval. This situation commonly occurs in switching dynamical systems. The present paper aims at formulating a unique equation system that incorporates the inequality constraints due to the hypercube boundary, is valid in the whole time interval of analysis and, hence, is independent of the number of constraints actually limiting the trajectory space. Moreover, by using tools from differential inclusions, the noted problems in the definition of a solution are resolved.

The outline of the paper is as follows. In the remaining part of this section we give the needed mathematical background for the analysis to be developed. Section II describes the model of FR-CNNs and discusses its analytic foundation. The main results on trajectory convergence of FR-CNNs are given in Section III. Finally, the conclusions drawn in the paper are summarized in Section IV.

A. Preliminaries

In this section, we present the needed definitions and results concerning set-valued maps and differential inclusions. We refer the reader to [9], [15] for a more thorough treatment.

1) Set-Valued Maps: A set-valued map $F : K \subset \mathbb{R}^n \multimap \mathbb{R}^n$ is a map that associates to each point $x \in K$ a set $F(x) \subset \mathbb{R}^n$. A set-valued map $F : K \multimap \mathbb{R}^n$ with nonempty values is said to be upper semicontinuous at $\tilde{x} \in K$ if for any open set U_y containing $F(\tilde{x})$ there exists a neighborhood U_x of $\tilde{x}$ such that $F(U_x) \subset U_y$. If K is closed, F has nonempty closed values and is bounded in a neighborhood of each point $x \in K$, then F is termed upper semicontinuous on K if and only if the graph of F

$$\mathrm{graph}(F) = \{(x, y) \in K \times \mathbb{R}^n : y \in F(x)\}$$

is closed.

2) Tangent and Normal Cones to Closed Convex Sets: Given a nonempty closed convex set $K \subset \mathbb{R}^n$, and a point $x \in K$, the tangent cone to K at x is defined as follows [16]:

$$\mathcal{T}_K(x) = \left\{ v \in \mathbb{R}^n : \liminf_{\rho \to 0^+} \frac{\mathrm{dist}(x + \rho v, K)}{\rho} = 0 \right\}$$

while the normal cone to K at x is

$$\mathcal{N}_K(x) = \{p \in \mathbb{R}^n : \langle p, v \rangle \leq 0 \qquad \forall v \in \mathcal{T}_K(x)\}.$$

We observe that both $\mathcal{T}_K(x)$ and $\mathcal{N}_K(x)$ are nonempty closed convex cones in $\mathbb{R}^n$. In particular, if $x \in \mathrm{int}(K)$, then $\mathcal{T}_K(x) = \mathbb{R}^n$ and so $\mathcal{N}_K(x) = \{0\}$.

Some tangent and normal cones are shown in Fig. 2 for a closed convex set $K \subset \mathbb{R}^2$. They are shown for points $x_{\mathrm{a}}, x_{\mathrm{b}}$, and x_{c} on $\mathrm{bd}(K)$, where the tangent line cannot be defined, and to a point x_{d} on $\mathrm{bd}(K)$ where the tangent line exists. For simplicity of representation, in the figure we have illustrated $x + \mathcal{T}_H(x)$ and $x + \mathcal{N}_H(x)$, i.e., the two cones translated at the related point. Clearly, the use of the normal and tangent cones is of interest mainly when they are referred to a point $x \in \mathrm{bd}(K)$ where the tangent line cannot be defined.

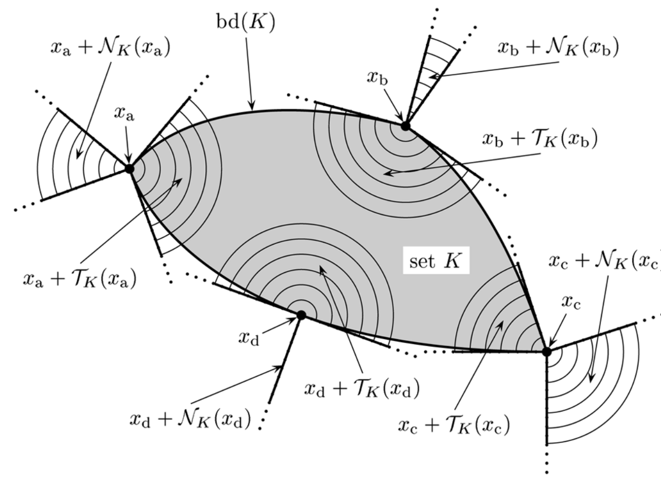


Fig. 2. Closed and convex set $K \subset \mathbb{R}^2$ with tangent and normal cones shown at some points on $\mathrm{bd}(K)$.

To any $x \in \mathbb{R}^n$ we can associate a unique point $\mathcal{P}_K(x) \in K$ satisfying

$$\|x - \mathcal{P}_K(x)\|_2 = \min_{y \in K} \|x - y\|_2 = \mathrm{dist}(x, K).$$

The operator $\mathcal{P}_K$ is called the projector of best approximation on K. In particular, we define by

$$m(K) \doteq \mathcal{P}_K(0)$$

the element $y \in K$ with the smallest norm.

3) Differential Calculus for Extended-Valued Functions: In this section, we consider extended-valued functions [9] which are allowed to take the value $+\infty$, i.e., functions $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$.

Given $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, we denote the epigraph of f by

$$\mathrm{epi}(f) = \{(x, r) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq r\}.$$

We recall that a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is convex if and only if its epigraph is convex [9, p. 22].

Definition 1 [9, p. 32]: Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be convex in $\mathbb{R}^n$. Suppose that f is finite at a point $x \in \mathbb{R}^n$. We define the generalized gradient of f at x to be the set

$$\partial f(x) = \{\xi \in \mathbb{R}^n : (\xi, -1) \in \mathcal{N}_{\mathrm{epi}(f)}(x, f(x))\}.$$

■

Fig. 3(a) shows the epigraph at a point $(x_0, f(x_0)) \in \mathbb{R}^2$ and the generalized gradient at $x_0 \in \mathbb{R}$ for a nonsmooth convex function $f : \mathbb{R} \to \mathbb{R}$. Fig. 3(b) depicts the epigraph and the generalized gradient in the case of a nonsmooth extended-valued function $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$, which is finite at $x_0 \in \mathbb{R}$ and assumes the value $+\infty$ in a left neighborhood of x_0.

If f is convex and differentiable at $x \in \mathbb{R}^n$, then $\partial f(x) = \{\nabla f(x)\}$ is single-valued. Moreover, $(\nabla f(x), -1)$ is the normal vector to $(1, \nabla f(x))$, while $(1, \nabla f(x))$ is the normal vector at the point $(x, f(x)) \in \mathbb{R}^{n+1}$ to the graph of f. Definition 1 thus generalizes to nonsmooth extended-valued functions the notion of the normal vector at a point of a graph. We observe that $\partial f(x)$ is in general a (possibly empty) closed

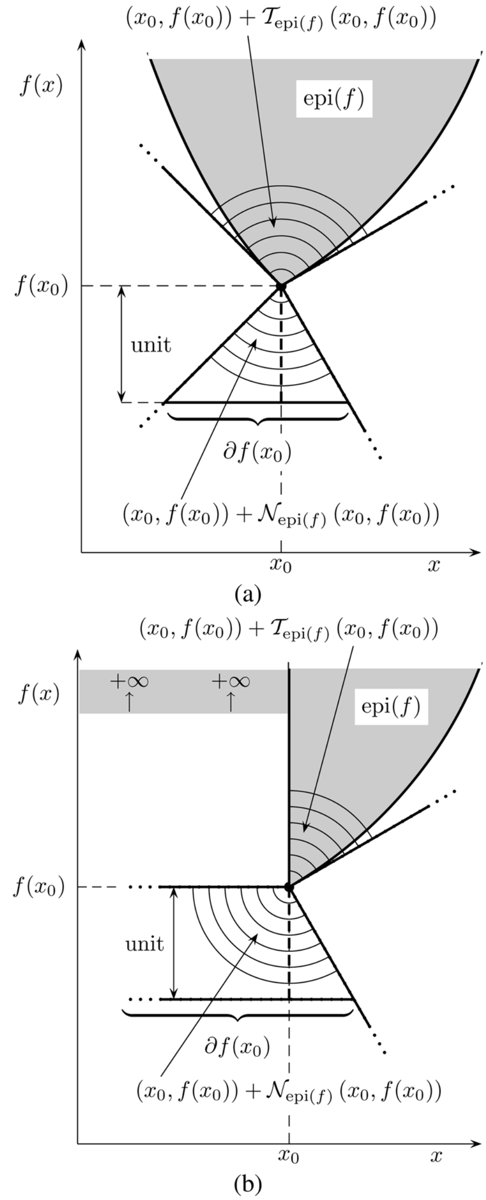


Fig. 3. Epigraph and generalized gradient of nonsmooth convex functions. (a) Conventional function $f : \mathbb{R} \to \mathbb{R}$. (b) Extended-valued function $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$.

convex set of $\mathbb{R}^n$. In particular, if f is a nonsmooth locally Lipschitz function then $\partial f(x)$ is a nonempty set and it coincides with the Clarke's generalized gradient. This concept has been already used in a number of engineering applications, see, e.g., [17], [18], and their references.

Let $\Psi_K : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be the indicator of the nonempty closed convex set $K \subset \mathbb{R}^n$, namely

$$\Psi_K(x) = \begin{cases} 0, & x \in K \\ +\infty, & x \notin K. \end{cases}$$

The function Ψ_K is convex in $\mathbb{R}^n$ and $\mathrm{epi}(\Psi_K)$ is a nonempty closed convex set in $\mathbb{R}^n \times \mathbb{R}$.

We recall the following results.

Property 1: For any point $x \in K$, we have

$$\partial \Psi_K(x) = \mathcal{N}_K(x). \qquad \blacksquare$$

Proof: See, e.g., [15, Prop. 2.4.12]. ■

In this paper, we will encounter extended-valued Lyapunov functions $W : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ of the form

$$W(x) = \phi(x) + \Psi_K(x) \tag{1}$$

where $\phi : \mathbb{R}^n \to \mathbb{R}$ is $C^1(\mathbb{R}^n)$ and $\Psi_K : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is the indicator of a nonempty closed convex set $K \subset \mathbb{R}^n$. The calculus of the time derivative of W along solutions $x(t)$ of a given differential inclusion is of key importance in the Lyapunov approach used in this paper. To this end, the following rule will play a crucial role.

Property 2 (Chain Rule): Consider the function W defined by (1). Suppose that $x : \mathbb{R} \to \mathbb{R}^n$ is differentiable at t and Lipschitz near t, and that $W(x(\cdot))$ is differentiable at t. Then

$$\frac{d}{dt} W(x(t)) = \langle \xi, \dot{x}(t) \rangle \qquad \forall \xi \in \partial W(x(t)) \tag{2}$$

where $\partial W(x(t)) = \nabla \phi(x(t)) + \mathcal{N}_K(x(t))$. ■

Proof: Since $x \to \phi(x)$ and $x \to \Psi_K(x)$ are regular functions at any $x \in K$ [15, Cor. p. 32, Prop, 2.4.12], we have that $\partial W(x) = \nabla \phi(x) + \partial \Psi_K(x)$ [15 , Th. 2.9.8] and, by Property 1, $\partial W(x) = \nabla \phi(x) + \mathcal{N}_K(x)$. Since $x(\cdot)$ is differentiable at t, and Lipschitz near t, then $x(\cdot)$ is strictly differentiable at t [15, Prop. 2.2.4]. Moreover, it can be proved that all assumptions of [15, Th. 2.9.9] are satisfied. In fact, Ψ_K is finite at any $x(t) \in K$, it is directionally Lipschitz at $x(t)$, and regular at this point. Therefore, the result directly follows by applying [15, Th. 2.9.9] . ■

4) Differential Variational Inequalities: Let $K \subset \mathbb{R}^n$ be a nonempty closed convex set and $G : K \to \mathbb{R}^n$ be a given function. Following [9, Ch. 5], a differential variational inequality is a differential inclusion of the form

$$\dot{x} \in G(x) - \mathcal{N}_K(x) \tag{3}$$

where $\mathcal{N}_K(x)$ is the normal cone to K at x. A solution $x(t)$ to (3), $t \in [t_1, t_2]$, with $t_1 < t_2$, is an absolutely continuous function such that

$$\begin{cases} x(t) \in K & \forall t \in [t_1, t_2] \\ \dot{x}(t) \in G(x(t)) - \mathcal{N}_K(x(t)), & \text{for a.a. } t \in [t_1, t_2] \end{cases} \tag{4}$$

where to simplify the notation we have used a.a. as an abbreviation of the term "almost all." If in addition

$$\dot{x}(t) = m(G(x(t)) - \mathcal{N}_K(x(t))) \tag{5}$$

for a.a. $t \in [t_1, t_2]$, then $x(t)$ is said to be a *slow* solution of (3).

The following lemma summarizes results on differential variational inequalities established in [9, Ch. 5], which are useful for the analysis in this paper.

Lemma 1: Suppose that K is a nonempty compact convex set and $G : K \to \mathbb{R}^n$ is continuous on K. Then the following are true.

1) For any $x_0 \in K$ there is at least a slow solution to the differential variational inequality

$$\begin{cases} x(0) = x_0, \quad x(t) \in K & \forall t \in [0, +\infty) \\ \dot{x}(t) \in G(x(t)) - \mathcal{N}_K(x(t)), & \text{for a.a. } \ t \in [0, +\infty) \end{cases}$$

which is also a solution of the projected differential equation

$$\begin{cases} x(0) = x_0, & x(t) \in K \quad \forall t \in [0, +\infty) \\ \dot{x}(t) = \mathcal{P}_{\mathcal{T}_K(x(t))} G(x(t)), & \text{for a.a. } t \in [0, +\infty) \end{cases} \tag{6}$$

where $\mathcal{T}_K(x(t))$ is the tangent cone to K at $x(t)$ and $\mathcal{P}_{\mathcal{T}_K(x(t))} G(x(t))$ is the projection of $G(x(t))$ on $\mathcal{T}_K(x(t))$.

2) There exists at least a solution $e \in K$ to the inclusion $0 \in G(e) - \mathcal{N}_K(e)$. ■

The above concept of a differential variational inequality has been recently generalized in [19]. We refer the reader to that paper also for a discussion on the importance of variational inequalities in modeling several physical systems that may also involve inequalities and discontinuous nonlinearities.

II. Theoretical Foundation of FR Model

The goal of this section is to give a theoretical foundation to the dynamics of FR-CNNs by exploiting tools from set-valued analysis and differential inclusions. In Section II-A we introduce the FR model of CNNs. In Section II-B we discuss the definition, existence and uniqueness of the solution of such model. Finally, in Section II-C we present results that are useful for the understanding of FR-CNN dynamics, and we compare the solutions of the ideal FR-CNN model with those of a modified model where the vertical segments in the nonlinearities are approximated by segments with finite slope.

A. FR Model of CNNs

In this paper, we consider CNNs whose dynamics is described by the differential inclusion

$$\dot{x} \in F_{\text{fr}}(x) = Tx + I - S(x) \tag{6a}$$

where $x \in H \subset \mathbb{R}^n$ is the vector of neuron state variables, $F_{\text{fr}} : H \multimap \mathbb{R}^n$ is the set-valued vector field defining (6a), and $H = [-1,1]^n = [-1,1] \times [-1,1] \times \cdots \times [-1,1]$, is a hypercube in $\mathbb{R}^n$. Furthermore, $T \in \mathbb{R}^{n \times n}$ is the constant neuron interconnection matrix, $I \in \mathbb{R}^n$ is the constant input, and $S(x) = (s(x_1), \ldots, s(x_n))' : H \multimap \mathbb{R}^n$ is a set-valued map where $s(x_i) : [-1,1] \multimap \mathbb{R}$ is defined as

$$s(x_i) = \begin{cases} (-\infty, 0], & x_i = -1 \\ 0, & x_i \in (-1, 1) \\ [0, +\infty), & x_i = 1. \end{cases} \tag{7}$$

Neural network (6a) corresponds to the so-called FR model of CNNs. This model significantly differs from that of S-CNNs originally introduced by Chua and Yang [1]. The main difference is that in (6a) each neuron has a self-loss nonlinearity which is represented by a hard-comparator function $s(\cdot)$ as in (7), while a standard piecewise-linear input–output neuron nonlinearity with unity gain in the linear region, and ± 1 saturation levels, is employed for S-CNNs. FR-CNNs were introduced in [3] with the aim of accelerating the operation of S-CNNs by reducing the allowable range of variation of the neuron state variables x_i. In fact, the hard-comparator nonlinearity $s(\cdot)$ is seen to prevent x_i from leaving the interval $[-1, 1]$. FR-CNNs have displayed a number of advantages in the VLSI implementation with respect to S-CNNs, including smaller power consumption, higher cell densities, and increased processing speed. Almost all chips manufactured recently exploit an array dynamics which can be modeled by (6a), see [2]–[6].

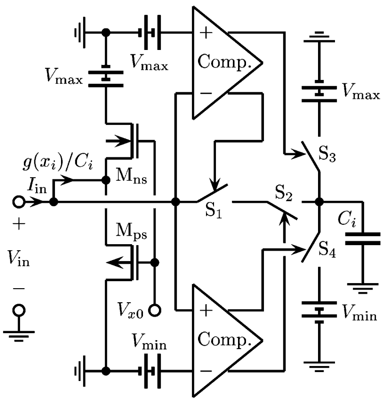


Fig. 4. Circuit implementing a FR-CNN neuron.

It is also worthy of note that the ACE4k chip described in [4] features a nonlinear neuron conductance that closely approximates the ideal hard-comparator nonlinearity $s(x_i)$, with x_i denoting the voltage at the neuron capacitor C_i. In fact, as discussed in [4, Sec. 3.5], the neuron conductance can be considered an open circuit in the range $x_i \in (-1, 1)$ (current $g(x_i)$ is zero except for leakage through C_i or the switches S_3, S_4, and diode-like clamps M_{ps} and M_{ns}), while the voltage remains clamped at a constant value $x_i = 1$ or $x_i = -1$ respectively for any positive or negative current $g(x_i)$, as shown in Fig. 4. An almost ideal voltage clamp is obtained by switching between two circuit topologies. The topology change is implemented by switches $\text{S}_1, \text{S}_2, \text{S}_3$, and S_4. When the voltage x_i reaches the limiting levels $x_i = 1$ or $x_i = -1, C_i$ is disconnected from any input current and connected to the constant supply voltage (high or low, according to x_i). Two clamping MOSFETs absorb any current entering the input node, without affecting the neuron voltage x_i. As the voltage x_i stays constant notwithstanding the nonidealities of the switches, the clamp can be considered ideal. Since $x_i = 1$ the neuron can sink any positive current, the current $g(1)$ through the neuron conductance may actually assume all positive values, i.e., $g(1)$ is a set-valued map. This is expressed mathematically by the inclusion $g(1) \in [0, +\infty)$. An analogous situation occurs for $x_i = -1$. This discussion justifies from a more physical viewpoint why the dynamics of FR-CNNs is naturally described by a differential inclusion as in (6a).

B. Existence and Uniqueness of Solution

Since (6a) is described by a differential inclusion, we consider what is meant by a solution of a Cauchy problem associated with (6a). Moreover, the issue of the existence and uniqueness of the solution for such a Cauchy problem, and the existence of at least one equilibrium point for (6a), must be addressed.

According to the theory of differential inclusions [9], a solution of (6a) on $[0, \tilde{t}], \tilde{t} > 0$, with initial condition $x(0) = x_0 \in H$, is an absolutely continuous function $x(t)$ on $[0, \tilde{t}]$ such that $x(t) \in H$ for $t \in [0, \tilde{t}]$ and for a.a. $t \in [0, \tilde{t}]$ we have $\dot{x}(t) \in Tx(t) + I - S(x(t))$.

In particular, an equilibrium point $e \in H$ is a stationary solution of (6a), hence it satisfies the inclusion $0 \in F_{\rm fr}(e) = Te + I - S(e)$.

To address the existence of solutions to a Cauchy problem associated with (6a), and the existence of stationary solutions, we begin by establishing a relationship between (6a) and a class of differential variational inequalities [see Section I-A4)]. To this end, note that

$$S(x) = \mathcal{N}_H(x)$$

for any $x \in H$, i.e., $S(x)$ coincides with the normal cone to H at x. Hence, (6a) can be rewritten as

$$\dot{x} \in Tx + I - S(x) = Tx + I - \mathcal{N}_H(x)$$

for $x \in H$. This means that the stated Cauchy problem associated with (6a) is equivalent to the following differential variational inequality

$$\begin{cases} x(0) = x_0 \in H, \quad x(t) \in H & \forall t \in [0, \tilde{t}] \\ \dot{x}(t) \in Tx(t) + I - \mathcal{N}_H(x(t)), & \text{for a.a. } t \in [0, \tilde{t}]. \end{cases} \tag{8}$$

The next property holds.

Property 3: For any $x_0 \in H$, there is at least a solution $x(t)$ of (6a) with initial condition $x(0) = x_0$, which is defined for $t \in [0, +\infty)$. Moreover, there is at least one equilibrium point $e \in H$ of (6a). ■

Proof: A Cauchy problem associated with (6a) is equivalent to the differential variational inequality (8). Hence, the result is an immediate consequence of Lemma 1. ■

Let us now consider the issue of uniqueness of the solution for a Cauchy problem associated with (6a). It is noted that the differential variational inequality (8) is equivalent to the projected differential equation $\dot{x}(t) = \mathcal{P}_{\mathcal{T}_H(x(t))}(Tx(t) + I)$ (cf. (6) of Lemma 1), whose right-hand side is a discontinuous function of the state $x(t)$ on $\mathrm{bd}(H)$. In general situations a discontinuous differential equation may possess multiple solutions to a given Cauchy problem [20]. Nevertheless, in the next property we show that since the affine vector field $Tx + I$ defining (6a) is Lipschitz, the FR model enjoys the physically important property of uniqueness of the solution starting at a given initial condition.

Property 4: For any $x_0 \in H$, there is a unique solution $x(t), t \in [0, +\infty)$, of (6a) with initial condition $x(0) = x_0$, which coincides with the slow solution

$$\begin{aligned} \dot{x}(t) &= m(Tx(t) + I - \mathcal{N}_H(x(t))) \\ &= \mathcal{P}_{\mathcal{T}_H(x(t))}(Tx(t) + I) \end{aligned} \tag{9}$$

for a.a. $t \in [0, +\infty)$. ■

Proof: We begin by establishing the uniqueness of the solution starting at a given point. Pick $x_0 \in H$. Let $x(t)$ and $y(t)$ be two solutions of (6a) that are defined on a common interval $[0, \tau]$, for some $\tau > 0$, and are such that $x(0) = y(0) = x_0$. Since $x(t)$ and $y(t)$ are absolutely continuous in $[0, \tau]$, then they are differentiable for a.a. $t \in [0, \tau]$ and we have

$$\begin{aligned} \frac{d}{dt}\left(\frac{1}{2}\|x(t) - y(t)\|_2^2\right) &= \langle x(t) - y(t), \dot{x}(t) - \dot{y}(t)\rangle \\ &= \langle x(t) - y(t), Tx(t) - Ty(t)\rangle \\ &\quad - \langle x(t) - y(t), v_x(t) - v_y(t)\rangle \end{aligned}$$

for some $v_x(t) \in S(x(t))$ and $v_y(t) \in S(y(t))$.

Since $Tx + I$ is a Lipschitz vector field on $\mathbb{R}^n$, we obtain

$$\begin{aligned} \frac{d}{dt}\left(\frac{1}{2}\|x(t) - y(t)\|_2^2\right) &\le \lambda \|x(t) - y(t)\|_2^2 \\ &\quad - \langle x(t) - y(t), v_x(t) - v_y(t)\rangle \end{aligned}$$

where $\lambda > 0$ is the Lipschitz constant.

Now, from (7) it follows that S is a maximal monotone operator and thus (see [9, Prop. 1, p. 159]), $\langle x(t) - y(t), w_x(t) - w_y(t)\rangle = \sum_{i=1}^{n}(x_i(t) - y_i(t))(w_{x,i}(t) - w_{y,i}(t)) \ge 0$ for any $w_x(t) \in S(x(t))$ and $w_y(t) \in S(y(t))$. Therefore

$$\frac{d}{dt}\left(\frac{1}{2}\|x(t) - y(t)\|_2^2\right) \le \lambda \|x(t) - y(t)\|_2^2. \tag{10}$$

Let $d(t) = \|x(t) - y(t)\|_2^2$. By applying Gronwall's inequality to $d(t)$, we obtain on the basis of (10)

$$d(t) \le d(0)e^{2\lambda t}, \quad t \in [0, \tau].$$

Therefore, since $d(0) = \|x(0) - y(0)\|_2 = 0$, we conclude that $d(t) = \|x(t) - y(t)\|_2 = 0$, for any $t \in [0, \tau]$. Since this argument holds for any choice of $\tau > 0$, it follows that there is a unique solution $x(t)$ of (6a) for $t \in [0, +\infty)$, starting at x_0 at $t = 0$.

Finally, the fact that $x(t)$ is the slow solution to (6a), and a solution to the projected differential equation $\dot{x}(t) = \mathcal{P}_{\mathcal{T}_H(x(t))}(Tx(t) + I)$, follows from (1) of Lemma 1. ■

Remark:

- When treating differential systems with discontinuous right-hand side, or differential inclusions, the concept of uniqueness of the solution has to be linked with the time direction of the solution evaluation, i.e., we have to distinguish between forward uniqueness and backward uniqueness (or right-side and left-side uniqueness, as in [14]). According to Property 4, the slow solution to (6a) possesses the former, but not necessarily the latter, as shown in the next example.

■

Example 1: Let us consider the second-order FR-CNN

$$\dot{x} \in \begin{bmatrix} 0.2 & -1 \\ 1 & 0.2 \end{bmatrix} x - \mathcal{N}_H(x)$$

where $x = (x_1, x_2)' \in H = [-1, 1] \times [-1, 1]$. It is easily seen that there are no equilibrium points on $\mathrm{bd(H)}$ and that there is an unstable focus at the origin. The constraint imposed by the vector field $F_{\rm fr}$ on $\mathrm{bd}(H)$ forces any nonstationary solution to converge to a unique limit cycle. As shown in Fig. 5, the solutions starting, respectively, at point $P_a = (0.4, 0)'$ and $P_b = (0.986, -0.373)'$ at $t = 0$, become coincident at the "confluence point" $P_c = (-0.104, 1)'$, at the time $t_c = 4.608$. Note that since the vector field never vanishes along the solution starting at point P_a, there is convergence of the solution to the limit cycle in finite time. Therefore, there is not backward uniqueness for the solutions starting at point P_c at t_c. An analogous conclusion holds for any point belonging to the intersection of the cycle and $\mathrm{bd}(H)$. The lack of backward uniqueness, however, does not pose any practical problem since we are interested in the solution only for forward values of time. ■

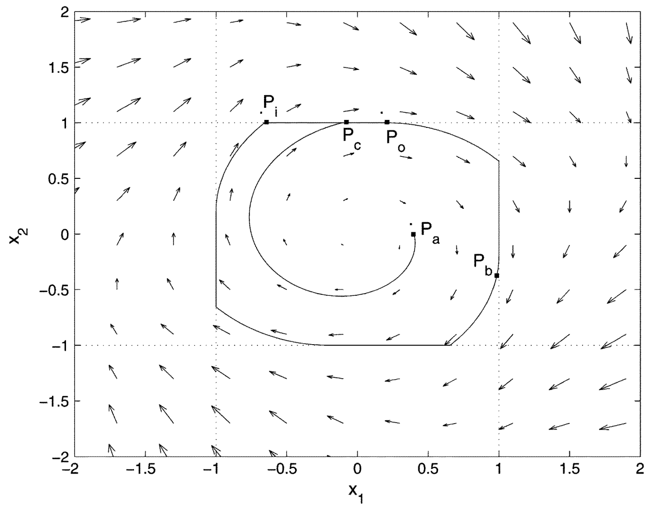


Fig. 5. Trajectories of the second-order FR-CNN in Example 1. The P_iP_o-segment corresponds to one of the four sliding modes of the trajectories.

C. Approximation of Solutions

In model (6a), the ideal hard-comparator function $s(\cdot)$ may assume arbitrarily large values at $x_i = \pm 1$. This would cause problems in the actual realization of this nonlinearity. In the next property we give an analytic expression for the values assumed by $s(\cdot)$, along each solution of (6a), and show that these values are bounded by an easily computable quantity.

Property 5: For any $x_0 \in H$, let $x(t), t \in [0,+\infty)$, be the solution of (6a) with initial condition $x(0) = x_0$. Then, the measurable function $\gamma : [0,+\infty) \to \mathbb{R}^n$ defined as

$$\gamma(t) = \mathcal{P}_{\mathcal{N}_H(x(t))}(Tx(t) + I)$$

is such that $\gamma(t) \in S(x(t)) = \mathcal{N}_H(x(t))$ and

$$\dot{x}(t) = \mathcal{P}_{\mathcal{T}_H(x(t))}(Tx(t) + I) = Tx(t) + I - \gamma(t)$$

for a.a. $t \in [0,+\infty)$, i.e., $\gamma(t)$ corresponds to the values assumed by nonlinearity S for a.a. $t \in [0,+\infty)$. Moreover, we have

$$|\gamma_i(t)| < \mu$$

for $i = 1,\ldots,n$ and a.a. $t \in [0,+\infty)$, where μ is any positive quantity satisfying

$$\mu > \max_{y \in \mathrm{bd}(H)} \|Ty + I\|_2. \tag{11}$$

■

Proof: Note that $Tx(t) + I = \mathcal{P}_{\mathcal{T}_H(x(t))}(Tx(t) + I) + \mathcal{P}_{\mathcal{N}_H(x(t))}(Tx(t) + I)$, hence from Property 4 we have

$$\dot{x}(t) = \mathcal{P}_{\mathcal{T}_H(x(t))}(Tx(t) + I) = Tx(t) + I - \gamma(t)$$

for a.a. $t \in [0,+\infty)$, where

$$\gamma(t) = \mathcal{P}_{\mathcal{N}_H(x(t))}(Tx(t) + I)$$

is a measurable function from $[0,+\infty)$ to $\mathbb{R}^n$, such that $\gamma(t) \in S(x(t)) = \mathcal{N}_H(x(t))$, for a.a. $t \in [0,+\infty)$.

Then, considering that H is a compact set, we obtain

$$\begin{aligned}|\gamma_i(t)| \le \|\gamma(t)\|_2 &= \|\mathcal{P}_{\mathcal{N}_H(x(t))}(Tx(t) + I)\|_2 \\ &\le \|(Tx(t) + I)\|_2 \le \max_{y \in \mathrm{bd}(H)} \|Ty + I\|_2 < \mu\end{aligned}$$

for $i = 1,\ldots,n$ and for a.a. $t \in [0,+\infty)$. ■

Observe that the upper bound μ in (11) is independent of the initial condition x_0, i.e., it is valid for any solution $x(t), t \in [0,+\infty)$, of (6a). This result is useful for the electronic implementation of the nonlinearity $s(\cdot)$ [4]. In fact, the next property shows that the dynamics of (6a) remains unchanged if we replace the unbounded nonlinearity $s(\cdot)$ in (7) with the bounded function $\hat{s} : [-1,1] \multimap \mathbb{R}$ defined as

$$\hat{s}(x_i) = \begin{cases} [-\mu, 0], & x_i = -1 \\ 0, & x_i \in (-1,1) \\ [0,\mu], & x_i = 1. \end{cases} \tag{12}$$

Property 6: For any $x_0 \in H$, let $x(t), t \in [0,+\infty)$, be the solution of (6a) with initial condition $x(0) = x_0$. Then, $x(t)$ is the solution of the differential inclusion

$$\dot{x} \in Tx + I - \hat{S}(x) \tag{13}$$

satisfying $x(0) = x_0$, where $\hat{S}(x) = (\hat{s}(x_1),\ldots,\hat{s}(x_n))' : H \multimap \mathbb{R}^n$ and $\hat{s}$ is defined in (12). Conversely, if $x(t), t \in [0,+\infty)$, is the solution of (13) with $x(0) = x_0 \in H$, then $x(t)$ is the solution of (6a) satisfying $x(0) = x_0$. ■

Proof: Since the upper bound μ in (11) is independent of the solution $x(t), t \in [0,+\infty)$, of (6a) satisfying $x(0) = x_0$, whenever $x_0 \in H$, then we obtain

$$\dot{x}(t) \in Tx(t) + I - \hat{S}(x(t))$$

for a.a. $t \in [0,+\infty)$, i.e., $x(t)$ is also the solution of (13) such that $x(0) = x_0$. The converse is obvious. Observe that the uniqueness of the solution to any Cauchy problem associated with (13) easily follows from Property 4 and the definition of μ in Property 5. ■

Finally, for any $\epsilon > 0$, let us consider the Lipschitz continuous single-valued map $\hat{S}_\epsilon(x) = (\hat{s}_\epsilon(x_1),\ldots,\hat{s}_\epsilon(x_n))' : H \to \mathbb{R}^n$, where $\hat{s}_\epsilon : [-1,1] \to \mathbb{R}$ is defined as

$$\hat{s}_\epsilon(x_i) = \begin{cases} -\mu + \frac{\mu}{\epsilon}(x_i + 1), & -1 \le x_i \le -1 + \epsilon \\ 0, & -1 + \epsilon \le x_i \le 1 - \epsilon \\ \mu + \frac{\mu}{\epsilon}(x_i - 1), & 1 - \epsilon \le x_i \le 1. \end{cases}$$

Note the vertical segments of $\hat{s}$ at $x_i = \pm 1$ have been approximated in $\hat{s}_\epsilon$ by segments with a very-high finite slope μ/ϵ (ϵ small). In the next property, it is shown that each solution of the ideal model (6a) can be obtained as the limit as the slope $\sigma = \mu/\epsilon \to +\infty$, of solutions of a modified model involving nonlinearity $\hat{s}_\epsilon$.

Property 7: For any $j = 1,2,\ldots$, and any $x_0 \in H$, let $x^j(t)$ be the solution of

$$\dot{x} = Tx + I - \hat{S}_\epsilon(x) \tag{14}$$

with initial condition $x(0) = x_0 \in H$, where $\epsilon = 1/j$. Then, $x^j(t) \in H$ for $t \in [0, +\infty)$ and $x^j(t) \to x(t)$ as $j \to +\infty$, uniformly on any compact interval $[0, \tau], \tau > 0$, where $x(t)$ is the solution of (6a) satisfying $x(0) = x_0$. ■

Proof: The validity of the property, when $x_0 \in \text{int}(H)$, follows from a general result [9, Th. 3, p. 98]. Nevertheless for the reader convenience we outline the proof in the specific case of system (14). First of all observe that if $|x_i^j(t)| = 1$, for some $t \geq 0$ and $i \in \{1, \ldots, n\}$, then on the basis of (11) we have $\dot{x}_i^j(t) \neq 0$ and $\text{sgn}\dot{x}_i^j(t) = -\text{sgn}x_i^j(t)$. Therefore, $x^j(t) \in H$ for any $t \in [0, +\infty)$. Furthermore, since $\{\hat{S}(x), x \in H\} = [-\mu, \mu]^n$, and $\hat{S}_{1/j}(x) \in [-\mu, \mu]^n$ for any $j = 1, 2, \ldots$, and any $x \in H$, we have that $\dot{x}^j(t) \in K = \{Tx + I + [-\mu, \mu]^n : x \in H\}$, for any $j = 1, 2, \ldots$, and $t \in [0, +\infty)$, where K is a compact subset of $\mathbb{R}^n$. In conclusion, $\{x^j(t)\}, j = 1, 2, \ldots$, is a sequence of C^1 functions which are equibounded and equicontinuous on $[0, +\infty)$. Thus, by [9, Th. 4, p. 13], there exists a subsequence $\{x^{j(k)}\}, k = 1, 2, \ldots$, of $\{x^j(t)\}$, which converges to an absolutely continuous function $x(t)$, uniformly on the compact subsets of $[0, +\infty)$, and $\dot{x}^j(t)$ converges to $\dot{x}(t)$ weakly in $L^1([0, +\infty), \mathbb{R}^n)$. From the result in [9, Th. p. 60], we have that

$$\dot{x}(t) \in Tx(t) + I - \hat{S}(x(t))$$

for a.a. $t \in [0, +\infty)$ and $x(0) = x_0$. Taking into account Property 6, we conclude that $x(t), t \in [0, +\infty)$, is a solution of (6a) such that $x(0) = x_0$. Due to the property of uniqueness of the solution of (6a) starting at any given point (Property 4), we conclude that the sequence $\{x^j(t)\}$ converges to $x(t)$, uniformly on the compact subsets of $[0, +\infty)$. ■

Remarks:

1) By the result in [9, Th. 3, p. 98], the same conclusion as in Property 7 holds true for any sequence of locally Lipschitz maps, $F_{\epsilon_j} : H \to \mathbb{R}^n, j = 1, 2, \ldots$, where $\epsilon_j \to 0$ as $j \to +\infty$, approximating the set-valued vector field F_{fr} defining (6a) in the sense of the graph, namely $\text{graph}(F_{\epsilon_j}) \subset \text{graph}(F_{\text{fr}}) + B(0, \epsilon_j)$. The existence of such approximating sequences is guaranteed by [9, Th. 1, p. 84].
2) Property 4 and Property 5 admit a simple geometrical interpretation. Consider the case $n = 2$. We first note some properties of the vector field F_{fr} defining (6a). If $x \in \text{int}(H)$, then $\mathcal{N}_H(x) = \{0\}$ and $\mathcal{T}_H(x) = \mathbb{R}^2$, i.e., $F_{\text{fr}}(x) = Tx + I$ is single-valued and, obviously, (6a) reduces to an affine system of differential equations, in a small neighborhood of x. Consider a point $x = (x_1, 1)'$ on an edge of $\text{bd}(H)$, with $-1 < x_1 < 1$. Fig. 6(a) depicts the cones $\mathcal{T}_H(x)$ and $\mathcal{N}_H(x)$, the affine vector field $Tx + I$, and its projections on $\mathcal{T}_H(x)$ and $\mathcal{N}_H(x)$. The figure also depicts vectors belonging to the set-valued map $F_{\text{fr}}(x)$ (dashed), and the vector with minimum norm $\mathcal{P}_{\mathcal{T}_H(x)}(Tx + I) = m(F_{\text{fr}}(x))$. On a vertex x of $\text{bd}(H)$, we have an analogous situation [Fig. 7(a)].
On the basis of Fig. 6(a) we can give the following interpretation, which is valid for a.a. t, to the velocity $\dot{x}(t)$ along a solution $x(t)$ of (6a). According to Property 4, on an edge point as in Fig. 6(b) we have $\dot{x}(t) = \mathcal{P}_{\mathcal{T}_H(x(t))}(Tx(t) + I) = m(F_{\text{fr}}(x(t)))$, i.e., $\dot{x}(t)$ is the projection on $\mathcal{T}_H(x(t))$ of the affine vector field $Tx(t) + I$ shown in Fig. 6(b). This corresponds to a sliding of $x(t)$ along the surface $\text{bd}(H)$. ■
On the other hand, the projection on $\mathcal{N}_H(x(t))$ of $Tx(t) + I$ gives the value $\gamma(t)$ assumed by the nonlinearity S at point $x(t)$, see Property 5. A similar interpretation can be given for a vertex of $\text{bd}(H)$ [Fig. 7(b)]. ■

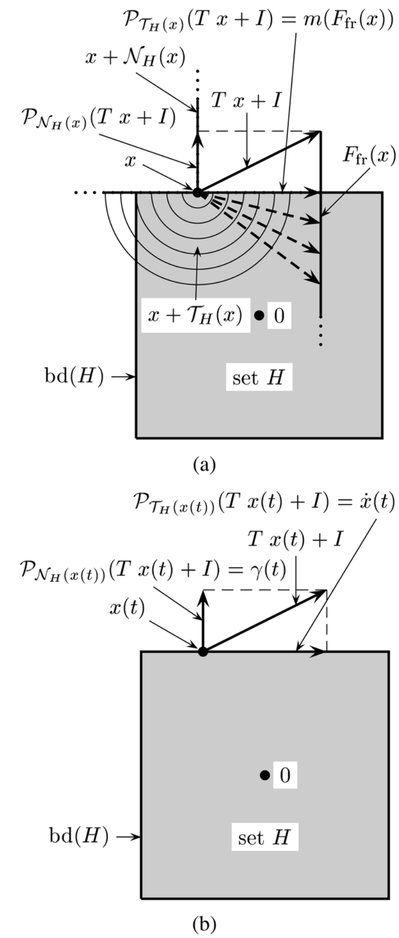


Fig. 6. Representation of a second-order FR-CNN. (a) Tangent and normal cones at a point x on an edge of $\text{bd}(H)$, decomposition of the affine vector field $Tx + I$ on these cones, set-valued vector field $F_{\text{fr}}(x)$, and vector with minimal norm $m(F_{\text{fr}}(x))$. (b) Velocity vector $\dot{x}(t)$ and value $\gamma(t)$ assumed by the nonlinearity S at point $x = x(t)$, on an edge of $\text{bd}(H)$.

III. Trajectory Convergence

In the previous section we have established dynamical properties of (6a) in the general case where the neuron interconnection matrix T may be symmetric or nonsymmetric. In Section III-A, under the hypothesis of symmetry for T, we address the convergence of the trajectories of (6a) towards equilibrium points by exploiting a Lyapunov approach. In Section III-B, we obtain further results on trajectory convergence towards minima of the Lyapunov function.

A. Convergence for Symmetric Interconnection Matrices

Let us recall some standard definitions. Property 3 showed that there is at least an equilibrium point of (6a). Let

$$E = \{x \in H : 0 \in Tx + I - S(x)\} \neq \emptyset$$

be the set of equilibrium points of (6a).

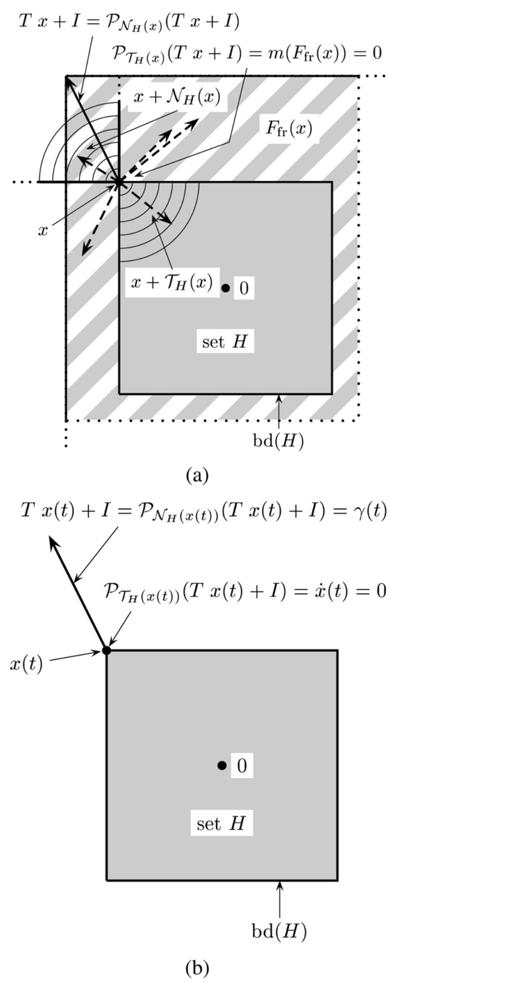


Fig. 7. Representation of a second-order FR-CNN. (a) Tangent and normal cones at a vertex x of $\text{bd}(H)$, decomposition of the affine vector field $Tx+I$ on these cones, set-valued vector field $F_{\text{fr}}(x)$, and vector with minimal norm $m(F_{\text{fr}}(x))$, in a case where x is an equilibrium point of (6a). (b) Velocity vector $\dot{x}(t)$ and value $\gamma(t)$ assumed by the nonlinearity S at point $x=x(t)$, on a vertex $x=x(t)$ of $\text{bd}(H)$.

Definition 2 ([21]): A neural network (6a) is said to be quasi-convergent if and only if, for any $x_0 \in H$, the solution $x(t)$ of (6a) with initial condition $x(0)=x_0$ satisfies

$$\lim_{t\to+\infty} \text{dist}(x(t),E)=0. \qquad \blacksquare$$

Equivalently, (6a) is quasi-convergent if and only if the ω-limit set of any trajectory $x(t)$ of (6a) is contained within E.

Definition 3 ([21]): A neural network (6a) is said to be convergent (or completely stable) if and only if, for any $x_0 \in H$, the solution $x(t)$ of (6a) with initial condition $x(0)=x_0$ satisfies

$$\lim_{t\to+\infty} x(t)=\tilde{e}$$

for some equilibrium point $\tilde{e}\in E$. ■

Equivalently, (6a) is convergent if and only if the ω-limit set of any trajectory $x(t)$ of (6a) is a singleton.

Suppose that the matrix T is symmetric, i.e., $T'=T$. To address trajectory convergence of (6a) we introduce the (candidate) extended-valued Lyapunov function $V:\mathbb{R}^n\to\mathbb{R}\cup\{+\infty\}$ given by

$$V(x)=-\frac{1}{2}x'Tx-x'I+\Psi_H(x) \qquad (15)$$

where $\Psi_H(x)$ is the indicator of H.

We then observe:

Property 8: Suppose that the neuron interconnection matrix T is symmetric. Then, (6a) is described by the gradient differential inclusion

$$\dot{x}\in F_{\text{fr}}(x)=Tx+I-S(x)=Tx+I-\mathcal{N}_H(x)=-\partial V(x)$$

for any $x\in H$. ■

Proof: Since H is a nonempty compact convex set, we have $\partial\Psi_H(x)=\mathcal{N}_H(x)$, for any $x\in H$ (cf. Property 1). The result in the property follows by noting that for a symmetric T, we have $\partial[(1/2)(x'Tx)]=Tx$ and $\partial(x'I)=I$. ■

Property 9: Suppose that the neuron interconnection matrix T is symmetric. Let $x(t), t\in[0,+\infty)$, be the solution of (6a) starting at $x(0)=x_0\in H$. Then, $V(x(\cdot))$ is differentiable for a.a. $t\in[0,+\infty)$, and

$$\begin{aligned}\frac{d}{dt}V(x(t)) &= -\|\dot{x}(t)\|_2^2\\ &= -\|m(Tx(t)+I-\mathcal{N}_H(x(t)))\|_2^2\le 0. \qquad (16)\end{aligned}$$

■

Proof: Since $x(t), t\in[0,+\infty)$, is absolutely continuous and $\|\dot{x}(t)\|_2$ is bounded on $[0,+\infty)$ by $\mu>0$ as given in (11), then $x(\cdot)$ is differentiable for a.a. $t\in[0,+\infty)$ and is Lipschitz near t. Furthermore, as in the Proof of Property 2 it turns out that V is regular, since $\Psi_H(x)$ and $\phi(x)=-(1/2)x'Tx-x'I$ are regular functions. Therefore, $W(x(\cdot))$ is differentiable for a.a. $t\in[0,+\infty)$, and Property 2 ensures that for a.a. $t\in[0,+\infty)$

$$\frac{d}{dt}V(x(t))=\langle\xi,\dot{x}(t)\rangle \qquad \forall\xi\in\partial V(x(t)). \qquad (17)$$

From Property 8 we obtain $\dot{x}(t)\in-\partial V(x(t))$ for a.a. $t\in[0,+\infty)$. Since the scalar product $\langle\xi,\dot{x}(t)\rangle$ is independent of the choice of ξ in $\partial V(x(t))$, we can choose $\xi=-\dot{x}(t)$ and obtain $dV(x(t))/dt=-\|\dot{x}(t)\|_2^2$. Finally, by recalling that $x(t)$ is the slow solution of (6a), and using (9), we have $dV(x(t))/dt=-\|m(Tx(t)+I-\mathcal{N}_H(x(t)))\|_2^2$. ■

Remarks:

1) Property 9 simply states that the function V is nonincreasing along the trajectories of (6a). Moreover it is strictly decreasing along nonstationary trajectories.
2) There is a simple geometric interpretation that explains why the scalar product in the chain rule (17) is independent of the choice of $\xi\in\partial V(x(t))$. Indeed, as can be seen from Fig. 6(a), in the case where a trajectory of (6a) passes at an edge point $x(t)\in\text{bd}(H)$, the set of vectors $F_{\text{fr}}(x(t))=-\partial V(x(t))$ has the same component in the direction of $\dot{x}(t)$. ■

On the basis of Properties 8 and 9, we are in a position to prove the main result on trajectory convergence of (6a).

Theorem 1: If the neuron interconnection matrix T is symmetric, then (6a) is quasi-convergent. If, in addition, the equilibrium points of (6a) are isolated then (6a) is convergent. ∎

Proof: Pick $x_0 \in H$, and let $x(t), t \in [0,+\infty)$, be the trajectory of (6a) with $x(0) = x_0 \in H$. Since $\Psi_H(x) = 0$ for $x \in H$, it follows that V is bounded from below on the compact set H. This implies that $V(x(t))$ is bounded from below for $t \in [0,+\infty)$.

Assume to the contrary, that $x(t)$ does not converge to E as $t \to +\infty$. Then there exists $\beta = 2\alpha > 0$ such that for any $T \geq 0$ there is a $t_* > T$ with $\mathrm{dist}(x(t_*), E) \geq 2\alpha$. We conclude that there exist $\beta = 2\alpha > 0$ and a strictly increasing sequence $\{t_n\}$, such that $t_n \to +\infty$ as $n \to +\infty$, and $\mathrm{dist}(x(t_n), E) \geq 2\alpha$ for all n.

Let us consider the compact set

$$K_\alpha = \mathrm{cl}\left(H \backslash \left(E + \frac{\alpha}{p}B(0,1)\right)\right)$$

with $p > 1$ chosen such that $K_\alpha \cap H \neq \emptyset$. We wish to evaluate $dV(x(t))/dt$ for $x(t) \in K_\alpha$. To this end, consider the map $\nu(x) : H \to \mathbb{R}$ defined by

$$\nu(x) = \|m(Tx + I - \mathcal{N}_H(x))\|_2.$$

Since the set-valued map $x \multimap Tx + I - \mathcal{N}_H(x)$ has nonempty closed convex values for $x \in H$ and is upper semicontinuous on H, then $\nu(x)$ is lower semicontinuous on H. Hence, its minimum exists in any compact set contained in H [16, Lemma 9.3.1, p. 361]. In particular, since $K_\alpha \cap E = \emptyset$, we have $\xi \neq 0$ for any $\xi \in Tx + I - \mathcal{N}_H(x)$ and any $x \in K_\alpha$. We obtain

$$\min_{x \in K_\alpha} \nu(x) = \nu_\alpha > 0.$$

Therefore, from (16) we have

$$\frac{d}{dt}V(x(t)) \leq -\nu_\alpha^2 < 0 \tag{18}$$

for a.a. $t \geq 0$ such that $x(t) \in K_\alpha$.

On the basis of (18), it is easy to reach a contradiction to the fact that $V(x(t))$ is bounded from below for $t \in [0,+\infty)$. Indeed, since for all n we have $\mathrm{dist}(x(t_n), E) \geq 2\alpha > \alpha(2p-1)/p$, it follows that $x(t) \in K_\alpha$ for any $t \in (t_n, t_n + \delta)$, and for all n, provided $0 < \delta \leq 2\alpha(p-1)/(p\mu_\alpha)$, where

$$\sup_{x \in K_\alpha} \|m(Tx + I - \mathcal{N}_H(x))\|_2 = \sup_{x \in K_\alpha} \nu(x) \leq \mu_\alpha = \max_{y \in K_\alpha} \|Ty + I\|_2.$$

On any of the time intervals $(t_n, t_n + \delta)$ the function V undergoes a negative jump which is lower than $-\delta\nu_\alpha^2$. We reach a contradiction to the fact that $V(x(t))$ is bounded from below for all $t \geq 0$ since by Property 9, V is nonincreasing along $x(t)$ for all $t \geq 0$. This contradiction implies that $\lim_{t \to +\infty} \mathrm{dist}(x(t), E) = 0$, hence that (6a) is quasi-convergent.

Suppose that the equilibrium points of (6a) are isolated. The ω-limit set of each trajectory $x(t), t \geq 0$, of (6a) is a nonempty closed connected set. Hence, the fact that $x(t) \to E$ as $t \to +\infty$, immediately implies that $x(t) \to \tilde{e}$ as $t \to +\infty$, for some $\tilde{e} \in E$. ∎

Remarks:

1) Theorem 1 extends to the FR model the classical result on complete stability for symmetric S-CNNs established in [1]. Moreover, Theorem 1 can also be thought of as the extension of previous results on complete stability of FR-CNNs obtained in the case where the slope is finite [3], [10].
2) Next we give a simple condition ensuring that the equilibrium points of (6a) are isolated (cf. Theorem 1).

Property 10: If matrix T and all its principal submatrices are nonsingular then the equilibrium points of (6a) are isolated. ∎

Proof: The set of equilibrium points of (6a) is given by $E = \{x \in H : 0 \in Tx + I - S(x)\} = \{x \in H : Tx + I \in \mathcal{N}_H(x)\}$. Since in the region $\mathrm{int}(H)$ we have $\mathcal{N}_H(x) = 0$, then (6a) has at most the equilibrium point $x = -T^{-1}I$ in $\mathrm{int}(H)$. Now, consider a generic partial saturation region $\{x \in H : |x_i| = 1, i \in \mathcal{I}; |x_i| < 1, i \notin \mathcal{I}\}$, where $\mathcal{I} \subseteq \{1,\ldots,n\}, \mathcal{I} \neq \emptyset$. If $x \in H$ is an equilibrium point of (6a) in such a region then x satisfies the algebraic equation $Tx + I = \gamma$, for some $\gamma \in \mathcal{N}_H(x)$. Equivalently we have

$$\sum_{i \notin \mathcal{I}} x_i T\eta_i + \sum_{i \in \mathcal{I}} \xi_i T\eta_i + I = \gamma \tag{19}$$

where $\xi_i = x_i \in \{-1,1\}$ for $i \in \mathcal{I}$ and $\eta_i \in \mathbb{R}^n, i = 1,\ldots,n$, are the canonical basis vectors of $\mathbb{R}^n$. Since $\mathcal{N}_H(x) = \{p \in \mathbb{R}^n : p = \sum_{i \in \mathcal{I}} \lambda_i \xi_i \eta_i, \lambda_i \geq 0\} \neq \{0\}$, we observe that (19) is equivalent to

$$\sum_{i \notin \mathcal{I}} x_i T\eta_i + \sum_{i \in \mathcal{I}} \xi_i T\eta_i - \sum_{i \in \mathcal{I}} \lambda_i \xi_i \eta_i = -I.$$

This in turn is equivalent to the systems of equations

$$\sum_{i \notin \mathcal{I}} T_{\ell i} x_i = -I_\ell - \sum_{i \in \mathcal{I}} \xi_i T_{\ell i} \eta_i \quad \ell \notin \mathcal{I}$$
$$\lambda_j \xi_i = \sum_{i \notin \mathcal{I}} T_{ji} x_i + I_j + \sum_{i \in \mathcal{I}} \xi_i T_{ji} \quad j \in \mathcal{I}$$

where $[T_{ij}]_{i,j=1}^n = T$ and $[I_j]_{j=1}^n = I$. Since by assumption $\det[T_{\ell i}]_{\ell, i \notin \mathcal{I}} \neq 0$, the first equation is uniquely solvable with respect to $\hat{x} = [x_i]_{i \notin \mathcal{I}}$. By substituting $\hat{x}$ in the second system, we obtain a unique vector $\hat{\lambda} = [\lambda_i]_{i \in \mathcal{I}}$ solving the system. In conclusion, there is at most one equilibrium point $\hat{x}$ in the considered region, with a corresponding $\hat{\lambda}$, such that $\hat{x}_i \in (-1,1)$ for any $i \notin \mathcal{I}$ and $\hat{\lambda} - i \geq 0$ for any $i \in \mathcal{I}$.

B. Further Results on Trajectory Convergence

When the neuron interconnection matrix T is symmetric, (6a) admits the Lyapunov function $V : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, given in (15), which can be rewritten as follows:

$$V(x) = \phi(x) + \Psi_H(x)$$

where $\phi : \mathbb{R}^n \to \mathbb{R}$ is given by

$$\phi(x) = -\frac{1}{2}x'Tx - x'I$$

and $\Psi_H : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is the indicator of H.

Observe that $V(x) = \phi(x)$ for any $x \in H$, while $V(x) = +\infty$ for $x \notin H$ by the definition of Ψ_H. Thus, it follows that the constrained minima of ϕ in H coincide with the unconstrained minima of V on $\mathbb{R}^n$. Moreover, if $x(t) \in H, t \geq 0$, is a trajectory of (6a), we have $V(x(t)) = \phi(x(t))$ for $t \geq 0$. As a consequence of Property 9, the function $\phi(x(t))$ is monotonically nonincreasing along the trajectories $x(t)$ of (6a), while along nonstationary trajectories $\phi(x(t))$ is strictly decreasing. This means that the neural network (6a) is suitable for finding constrained minima of ϕ in H. This is intuitively clear when realizing that Ψ_H plays the role of a barrier function, which prevents the trajectories of (6a) from leaving the hypercube H, see (15).

We consider this issue in detail. Denote by M the set of constrained minima of ϕ in H, and by M_s the set of strictly constrained minima of ϕ in H. Note that $M_s \subset M \subset C$, where C is the set of critical points of ϕ in H [15, Prop. 2.4.11]. By definition we have

$$C = \{x \in H : 0 \in -\nabla\phi(x) - \mathcal{N}_H(x)\}.$$

Moreover, the set of equilibrium points of (6a) is

$$E = \{x \in H : 0 \in Tx + I - S(x)\}.$$

Since for any $x \in H$ we have $-\nabla\phi(x) = Tx + I$, and $S(x) = \mathcal{N}_H(x)$, it follows that

$$C = E. \tag{20}$$

We obtain the following result from Theorem 1.

Corollary 1: Suppose that the neuron interconnection matrix T is symmetric. Then each solution $x(t)$ of (6a) converges to the set C of critical points of ϕ in H, as $t \to +\infty$. ■

Example 2: Consider the second-order symmetric FR-CNN

$$\dot{x} \in \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix} x - \mathcal{N}_H(x)$$

where $x = (x_1, x_2)' \in H = [-1, 1] \times [-1, 1]$. A straightforward computation shows that we have $C = E = \{(0,0)', (1,-3/5)', (-1,3/5)', (1,1)', (-1,-1)'\}$. We observe that: $E_1 = (0,0)'$ is an unstable and completely repelling equilibrium point; $E_2 = (1,-3/5)'$ and $E_3 = (-1,3/5)'$ are unstable saddle-type equilibrium points belonging to edges of $\mathrm{bd}(H)$; $E_4 = (-1,-1)'$ and $E_5 = (1,1)'$ are asymptotically stable equilibrium points on vertexes of $\mathrm{bd}(H)$. It follows that $M = M_s = \{E_4, E_5\}$.

Fig. 8 displays two trajectories of the neural network starting at points A and B, and converging towards the equilibria E_4 and E_5, respectively, which are strict local minima of ϕ in H. The figure also shows a trajectory starting near to E_1, and converging to E_2, which is a saddle point for ϕ constrained to H. ■

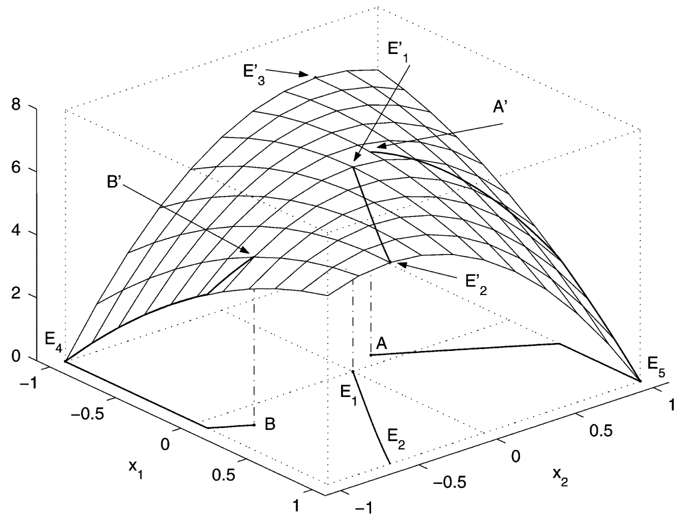


Fig. 8. Trajectories of a second-order convergent FR-CNN as in Example 2, and graph of the Lyapunov function $V(x) = \phi(x)$ for $x \in H$. For simplicity, we have added to ϕ a constant value, so that the minimum of ϕ in H is equal to 0 at both points E_4 and E_5.

Example 2 shows that in the general case there are trajectories of (6a) that converge to C, but not to M. However, it can be argued that the equilibrium points in the $C\backslash M$ subset are unstable and hence unobservable for the dynamics of (6a). This is proved in the next result.

Theorem 2: Suppose that the neuron interconnection matrix T is symmetric. Then M coincides with the set of stable equilibrium points of (6a), while M_s coincides with the set of asymptotically stable equilibrium points of (6a). ■

Proof: Suppose that $\hat{x} \in M$, or $\hat{x} \in M_s$. If $\hat{x} \in \mathrm{int}(H)$ then the result of the theorem is trivial, since (6a) reduces in a sufficiently small neighborhood of $\hat{x}$ to an affine system of differential equations which is the negative gradient of ϕ.

Suppose that $\hat{x} \in \mathrm{bd}(H)$. Since $\hat{x}$ is a critical point of ϕ in H, from (20) it is also an equilibrium point of (6a). We have

$$0 = T\hat{x} + I - \hat{\gamma}$$

where $\hat{\gamma} = -\nabla\phi(\hat{x}) = T\hat{x} + I \in \mathcal{N}_H(\hat{x})$.

Consider now the equations describing the dynamics of (6a)

$$\dot{x} \in Tx + I - \gamma$$

where $\gamma \in \mathcal{N}_H(x)$. The change of variables $y = x - \hat{x}$ yields

$$\dot{y} \in Ty + \hat{\gamma} - \gamma \tag{21}$$

where $\hat{\gamma} \in \mathcal{N}_H(y + \hat{x})$. Note that $\hat{x}$ has been translated into the equilibrium point $\hat{y} = 0$ of (21).

The next property, whose proof is given in Appendix I, holds.

Property 11: If $\hat{x} \in M$ then for sufficiently small $r > 0$ we have

$$\langle y, Ty + \hat{\gamma} - \gamma \rangle \leq 0 \tag{22}$$

for any $y \in B(0,r) \cap (H - \hat{x})$. If $\hat{x} \in M_s$ then the strict inequality holds in (22) for any $y \in B(0,r) \cap (H - \hat{x})\backslash\{0\}$. ■

From (22) it immediately follows that the squared distance $\|y(t)\|_2^2$ is nonincreasing along any trajectory $y(t)$ of (21), for any t such that $y(t)$ belongs to a sufficiently small neighborhood $B(0,r) \cap (H - \hat{x})$ of $\hat{y} = 0$, for $\hat{x} \in M$. Moreover, $\|y(t)\|_2^2$ is strictly decreasing when $y(t) \in B(0,r) \cap (H - \hat{x}) \backslash \{0\}$, if $\hat{x} \in M_s$. This implies that the equilibrium point $\hat{y} = 0$ of (21) is stable if $\hat{x} \in M$, and that it is asymptotically stable for $\hat{x} \in M_s$. Hence, this is true also of the equilibrium point $x = \hat{x}$ of (6a).

In general the set E of equilibrium points of (6a) is not connected. A connected component of E is a maximal connected set of equilibria. Assume now that $\hat{x} \in E$ is a stable equilibrium point of (6a). We want to prove that $\hat{x} \in M$. To this end, let $\hat{E}$ be the connected component of E containing $\hat{x}$, and assume that $\hat{x}$ is not a local minimum of ϕ in H. Thus, for any $n = 1, 2, \ldots$, there is a point $y_n \in B(\hat{x}, 1/n) \cap H$ such that $\phi(y_n) < \phi(\hat{x})$. Since ϕ is constant on $\hat{E}$, we have that $y_n \notin \hat{E}$ and $y_n \to \hat{x}$ as $n \to \infty$. Let U be a neighborhood of $\hat{x}$ such that $U \cap E = U \cap \hat{E}$. Observe that such a neighborhood always exists, since the number of connected components of E is finite for a quadratic function ϕ constrained to the hypercube H. Let $x_n(t), t \geq 0$, be the trajectory of (6a) such that $x_n(0) = y_n$. Since $\lim_{t \to +\infty} \mathrm{dist}(x_n(t), E) = 0$, and $V(x_n(t)) = \phi(x_n(t))$ is monotonically nonincreasing along $x_n(t)$, we have that $x_n(t)$ eventually leaves U for any n sufficiently large, contradicting the assumption that $\hat{x}$ is a stable equilibrium point of (6a).

Finally, assume that $\hat{x}$ is an asymptotically stable equilibrium point of (6a). Then there exists a neighborhood U of $\hat{x}$ such that $E \cap U \cap H = \{\hat{x}\}$, and for any $x_0 \in U \cap H$, the trajectory $x(t)$ of (6a) with $x(0) = x_0$ is such that $\lim_{t \to +\infty} x(t) = \hat{x}$. Assume to the contrary that $\hat{x}$ is not a strictly constrained minimum of ϕ in H. Then, there exists a sequence $y_n, n = 1, 2, \ldots$, with $y_n \in U \backslash \{\hat{x}\}$ and $y_n \to \hat{x}$ as $n \to +\infty$, such that $\phi(y_n) \leq \phi(\hat{x})$. Let $x_n(t), t \geq 0$, be the trajectory of (6a) such that $x_n(0) = y_n$. From Property 9, $V(x_n(t)) = \phi(x_n(t))$ is nonincreasing for $t \geq 0$, moreover $dV(x_n(t))/dt < 0$ for a.a. t such that $x_n(t) \in U \cap H \backslash \{\hat{x}\}$. This observation, together with the property $\phi(x_n(0)) = \phi(y_n) \leq \phi(\hat{x})$, for $n = 1, 2, \ldots$, yields a contradiction to the fact that $x_n(t) \to \hat{x}$ as $t \to +\infty$. ■

IV. Conclusion

By using theoretical tools from set-valued analysis and differential variational inequalities, the paper has developed a method to give a rigorous analytic foundation to the FR-CNN model where the neuron self-loss nonlinearities are ideal hard-comparator functions with two vertical straight segments. In particular, a fundamental result on complete stability for reciprocal S-CNNs has been extended to reciprocal FR-CNNs by using a Lyapunov approach generalized to a class of set-valued maps.

The method employed has advantages with respect to using event-based approaches for numerical simulation. First of all, it has allowed the incorporation of the constraints representing the nonlinearities in a unique equation system valid for the whole time domain. Moreover, it has permitted to overcome the problems encountered in trying to define a solution to the FR-CNN equations by piecing together solutions of the different affine systems among which a trajectory switches.

The analysis of the ideal case with vertical segments has proven simpler and more direct than the approach based on considering the case where the slope of the nonlinearities is finite and then studying the limit as the slope tends to infinity. Furthermore, the analysis has supplied a clear portrait of the salient dynamical features of FR-CNNs, including the presence of sliding modes on the boundary of the hypercube defined by the hard-comparator functions. This in turn has yielded simple expressions for the velocity vector and the time derivative of the Lyapunov function (symmetric case), which have a clear geometrical interpretation and are easy enough for use in engineering applications. It is worth noting that these results have been previously accepted without a formal mathematical proof, or have been justified only on the basis of heuristic arguments.

Appendix I
Proof of Property 11

We begin by noting that for any $y \in B(0,r) \cap (H - \hat{x})$ we have that $\langle y, \gamma \rangle = 0$. In fact, if y belongs to the interior of H, then $\gamma \in \mathcal{N}_H(y + \hat{x}) = \{0\}$. If $y \in \mathrm{bd}(H)$, it can be verified that y is orthogonal to the set $\mathcal{N}_H(y + \hat{x})$, whence we have $\langle y, \gamma \rangle = 0$. From the above observation it follows that

$$\langle y, Ty + \hat{\gamma} - \gamma \rangle = \langle y, Ty + \hat{\gamma} \rangle$$

for any $y \in B(0,r) \cap (H - \hat{x})$.

Now, let $\mathcal{J} = \{i \in \{1, \ldots, n\} : \hat{\gamma}_i \neq 0\}$, where $\hat{\gamma} = (\hat{\gamma}_1, \ldots, \hat{\gamma}_n)'$. If $\mathcal{J} = \emptyset$, i.e.: $\hat{\gamma} = 0$, then the result follows from Property 12 in Appendix II. Assume that $\mathcal{J} \neq \emptyset$. Without loss of generality and to simplify notation, suppose that $\mathcal{J} = \{1, \ldots, m\}$ for some $m \geq 1$, and $\hat{\gamma}_i > 0$ for all $i \in \mathcal{J}$. If for some index $j \in \mathcal{J}$, we have $\hat{\gamma}_j < 0$, we can repeat an analogous argument.

Let $\hat{Y} = \mathrm{span}\{\eta_i, i \in \mathcal{J}\}$, where $\eta_i, i = 1, \ldots, n$, is the canonical basis of $\mathbb{R}^n$, and let $\hat{Y}^\perp$ be the subspace of $\mathbb{R}^n$ orthogonal to $\hat{Y}$. Any vector $y \in B(0,r) \cap (H - \hat{x})$ can be decomposed as

$$y = \hat{y} + \hat{y}^\perp \in \hat{Y} \oplus \hat{Y}^\perp.$$

First suppose that $y = \hat{y} \in \hat{Y}, \hat{y} \neq 0$. We have

$$\langle \hat{y}, T\hat{y} + \hat{\gamma} \rangle \leq \lambda_{\max} \sum_{i=1}^{m} \hat{y}_i^2 + \sum_{i=1}^{m} \hat{y}_i \hat{\gamma}_i$$

where $\lambda_{\max}$ is the maximum eigenvalue of the symmetric matrix T. Since $\hat{y} \in H - \hat{x} \subset \mathcal{T}_H(\hat{x})$, and $\hat{\gamma} \in \mathcal{N}_H(\hat{x})$, it follows that $\hat{y}_i \hat{\gamma}_i \leq 0$, for any $i \in \mathcal{J}$. If $\lambda_{\max} \leq 0$, we immediately get the result in the property. Suppose that $\lambda_{\max} > 0$. Let us denote by $\mathcal{J}_- = \{i \in \mathcal{J} : \hat{y}_i < 0\}$, and assume that $\mathcal{J}_- = \{1, \ldots, \hat{m}\}$ for some $\hat{m} \geq 1$. We obtain

$$\langle \hat{y}, T\hat{y} + \hat{\gamma} \rangle \leq \lambda_{\max} \sum_{i=1}^{\hat{m}} \hat{y}_i^2 + \sum_{i=1}^{\hat{m}} \hat{y}_i \hat{\gamma}_i = \sum_{i=1}^{\hat{m}} |\hat{y}_i| (\lambda_{\max} |\hat{y}_i| - \hat{\gamma}_i).$$

If $\|\hat{y}\|_2$ is sufficiently small we have that

$$\lambda_{\max} |\hat{y}_i| - \hat{\gamma}_i < -\frac{\hat{\alpha}}{2} < 0$$

where $\hat{\alpha} = \min\{\hat{\gamma}_i, i = 1, \ldots, m\} > 0$. Hence

$$\langle \hat{y}, T\hat{y} + \hat{\gamma} \rangle \leq -\frac{\hat{\alpha}}{2} \sum_{i=1}^{\hat{m}} |\hat{y}_i| \leq -\frac{\hat{\alpha}}{2} \|\hat{y}\|_2.$$

Suppose next that $y = \hat{y}^\perp \in \hat{Y}^\perp$. Since we have $\langle \hat{y}^\perp, \hat{\gamma} \rangle = 0$, from Property 12 in Appendix II we obtain

$$\langle \hat{y}^\perp, T\hat{y}^\perp + \hat{\gamma} \rangle = \langle \hat{y}^\perp, T\hat{y}^\perp \rangle \leq 0$$

and if in addition $\hat{x} \in M_s$, we have that

$$\langle \hat{y}^\perp, T\hat{y}^\perp + \hat{\gamma} \rangle = \langle \hat{y}^\perp, T\hat{y}^\perp \rangle < 0$$

for $\hat{y}^\perp \neq 0$.

Finally, let $y = \hat{y} + \hat{y}^\perp \in \hat{Y} \oplus \hat{Y}^\perp$. If $r > 0$ is chosen sufficiently small, we obtain that

$$\begin{aligned} &\langle y, Ty + \hat{\gamma} \rangle \\ &= \hat{y}'T\hat{y} + \hat{y}'\hat{\gamma} + 2\hat{y}'T\hat{y}^\perp + (\hat{y}^\perp)'T\hat{y}^\perp \\ &\leq -\frac{\hat{\alpha}}{2}\|\hat{y}\|_2 + 2\|\hat{y}\|_2\|T\|_2\|\hat{y}^\perp\|_2 + (\hat{y}^\perp)'T\hat{y}^\perp \\ &= \|\hat{y}\|_2 \left(-\frac{\hat{\alpha}}{2} + 2\|T\|_2\|\hat{y}^\perp\|_2 \right) + (\hat{y}^\perp)'T\hat{y}^\perp \end{aligned}$$

where $\|T\|_2 = |\lambda_M|$ and λ_M is the eigenvalue with maximum absolute value of T. For sufficiently small $\|\hat{y}^\perp\|_2$ we have

$$-\frac{\hat{\alpha}}{2} + 2\|T\|_2\|\hat{y}^\perp\|_2 < -\frac{\hat{\alpha}}{4} < 0.$$

In conclusion, there exists $r > 0$ such that for any $y \in B(0, r) \cap (H - \hat{x})$ we obtain

$$\langle y, Ty + \hat{\gamma} \rangle \leq -\frac{\hat{\alpha}}{4}\|\hat{y}\|_2 + (\hat{y}^\perp)'T\hat{y}^\perp \leq 0.$$

Moreover, the strict inequality holds for $\hat{y} \neq 0$, when $\hat{x} \in M_s$. This completes the proof of the property. ■

Appendix II

Property 12: If $\hat{x} \in M$, then for sufficiently small $r > 0$ we have

$$0 \leq \phi(x) - \phi(\hat{x}) = -\frac{1}{2}(x - \hat{x})'T(x - \hat{x}) - (x - \hat{x})'\hat{\gamma} \quad (23)$$

for any $x \in B(\hat{x}, r) \cap H$, or equivalently

$$0 \leq -\frac{1}{2}y'Ty - y'\hat{\gamma} \quad (24)$$

for any $y \in B(0, r) \cap (H - \hat{x})$.

If $\hat{x} \in M_s$, then the strict inequality holds in (23) for any $x \in B(\hat{x}, r) \cap H \backslash \{\hat{x}\}$ and in (24) for any $y \in B(0, r) \cap (H - \hat{x}) \backslash \{0\}$. ■

Proof: For sufficiently small $r > 0$ we have for $\hat{x} \in M$

$$0 \leq \phi(x) - \phi(\hat{x}) = -\frac{1}{2}x'Tx - x'I + \frac{1}{2}\hat{x}'T\hat{x} + \hat{x}'I. \quad (25)$$

By the symmetry of T we have that

$$\begin{aligned} -\frac{1}{2}(x - \hat{x})'T(x - \hat{x}) - (x - \hat{x})'\hat{\gamma} &= -\frac{1}{2}x'Tx + x'T\hat{x} \\ &\quad -\frac{1}{2}\hat{x}'T\hat{x} - x'\hat{\gamma} + \hat{x}'\hat{\gamma}. \end{aligned}$$

Since $0 = T\hat{x} + I - \hat{\gamma}$, we thus obtain

$$\begin{aligned} &-\frac{1}{2}(x - \hat{x})'T(x - \hat{x}) - (x - \hat{x})'\hat{\gamma} \\ &\qquad = -\frac{1}{2}x'Tx - x'I + \frac{1}{2}\hat{x}'T\hat{x} + \hat{x}'I. \end{aligned}$$

By comparing this expression with (25) we obtain (23). The change of variables $y = x - \hat{x}$ yields (24).

The proof is the same in the case $\hat{x} \in M_s$. ■

Acknowledgment

The authors would like to thank the anonymous reviewers and Associate Editor for the insightful and constructive comments. They also wish to thank Prof. M. Gilli, Politecnico di Torino, Turin, Italy, for the helpful discussions on standard cellular neural networks (S-CNNs) and full-range CNNs.

References

[1] L. O. Chua and L. Yang, “Cellular neural networks: Theory,” *IEEE Trans. Circuits Syst.*, vol. 35, no. 10, pp. 1257–1272, Oct. 1988.
[2] A. Rodríguez-Vázquez, S. Espejo, R. Domínguez-Castro, J. L. Huertas, and E. Sánchez-Sinencio, “Current-mode techniques for the implementation of continuous- and discrete-time cellular neural networks,” *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 40, no. 3, pp. 132–146, Mar. 1993.
[3] S. Espejo, R. Carmona, R. Domínguez-Castro, and A. Rodríguez-Vázquez, “A VLSI oriented continuous time CNN model,” *Int. J. Circuit Theory Appl.*, vol. 24, pp. 341–356, 1996.
[4] G. Linán, S. Espejo, R. Domínguez-Castro, and A. Rodríguez-Vázquez, “ACE4k: An analog I/O 64 × 64 visual microprocessor chip with 7-bit analog accuracy,” *Int. J. Circuit Theory Appl.*, vol. 30, pp. 89–116, 2002.
[5] G. L. Cembrano, A. Rodríguez-Vázquez, S. E. Meana, and R. Domínguez-Castro, “ACE16k: A 128 × 128 focal plane analog processor with digital I/O,” *Int. J. Neural Syst.*, vol. 13, no. 6, pp. 427–434, Dec. 2003.
[6] A. Rodríguez-Vázquez, G. Liñán-Cembrano, L. Carranza, E. Roca-Moreno, R. Carmona-Galán, F. Jiménez-Garrido, R. Domínguez-Castro, and S. E. Meana, “ACE16k: The third generation of mixed-signal SIMD-CNN ACE chips towards VSoCs,” *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 5, pp. 851–863, May 2004.
[7] T. Roska and L. O. Chua, “The CNN universal machine: An analogic array computer,” *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 40, no. 3, pp. 163–173, Mar. 1993.
[8] I. Petrás, T. Roska, and L. O. Chua, “New spatial-temporal patterns and the first programmable on-chip bifurcation test bed,” *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 50, pp. 619–633, May 2003.
[9] J. P. Aubin and A. Cellina, *Differential Inclusions*. Berlin, Germany: Springer-Verlag, 1984.
[10] F. Corinto and M. Gilli, “Comparison between the dynamic behaviour of Chua-Yang and full-range cellular neural networks,” *Int. J. Circuit Theory Appl.*, vol. 31, pp. 423–441, Sept./Oct. 2003.
[11] V. I. Utkin, *Sliding Modes and Their Application in Variable Structure Systems*. Moscow, U.S.S.R.: MIR Publishers, 1978.
[12] J. Li, A. N. Michel, and W. Porod, “Analysis and synthesis of a class of neural networks: Linear systems operating on a closed hypercube,” *IEEE Trans. Circuits Syst.*, vol. 36, no. 11, pp. 1405–1422, Nov. 1989.
[13] G. De Sandre and A. Premoli, “Piecewise exponential approximation for fast time-domain simulation of 2-D cellular neural networks,” *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 51, pp. 400–405, Aug. 2004.
[14] A. F. Filippov, “Differential equations with discontinuous right-hand side,” *Transl. American Math. Soc.*, vol. 42, pp. 199–231, 1964.
[15] F. H. Clarke, *Optimization and Non-Smooth Analysis*. New York: Wiley, 1983.

[16] J. P. Aubin and H. Frankowska, *Set-Valued Analysis*. Boston, MA: Birkauser, 1990.
[17] B. E. Paden and S. S. Sastry, "Calculus for computing Filippov's differential inclusion with application to the variable structure control of robot manipulator," *IEEE Trans. Circuits Syst.*, vol. 34, no. 1, pp. 73–82, Jan. 1987.
[18] M. Forti, P. Nistri, and M. Quincampoix, "Generalized neural network for nonsmooth nonlinear programming problems," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 9, pp. 1741–1754, Sep. 2004.
[19] J. Pang and D. Stewart, "Differential variational inequalities," *Math. Progr. Ser. A* 2006 [Online]. Available: http://www.rpi.edu/~pangj/bib.html, submitted.
[20] A. F. Filippov, *Differential Equations with Discontinuous Right-Hand Side. Mathematics and its Applications (Soviet Series)*. Boston, MA: Kluwer, 1988.
[21] M. Hirsch, "Convergent activation dynamics in continuous time networks," *Neural Netw.*, vol. 2, pp. 331–349, 1989.

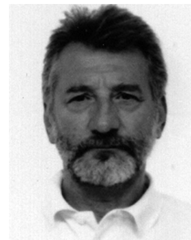
Guido De Sandre received the engineering degree (*cum laude*) in electronic engineering from the Politecnico di Milano, Milano, Italy, in 1995.

From 1997 he attended Ph.D. studies at the Department of Electronics and Information (DEI) of Politecnico di Milano. Since 1998, he has been with STMicroelectronics, within the Central Research and Development organization, where he was engaged in the test and design of flash memory test vehicles. His research interests are in the field of nonlinear circuits and cellular neural networks dynamics, but also in the area of analog and mixed-mode design, flash memories design and test in the context of both stand-alone and embedded applications.

Mauro Forti received the degree in electronics engineering from the University of Florence, Florence, Italy, in 1988.

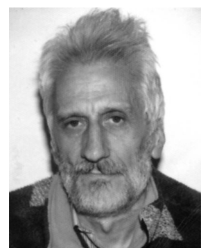
From 1991 to 1998, he was an Assistant Professor in applied mathematics and network theory in the Electronic Engineering Department, University of Florence. In 1998, he joined the Department of Information Engineering of the University of Siena, Siena, Italy, where he is currently Professor of electrical engineering. His main research interests are in the field of nonlinear circuits and systems, with emphasis on the qualitative analysis and stability of circuits modeling artificial neural networks. His research activity also includes aspects of electromagnetic compatibility.

Dr. Forti has served as an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: FUNDAMENTAL THEORY AND APPLICATIONS from 2001 to 2003. Since 2001, he has been serving as an Associate Editor of the IEEE TRANSACTIONS ON NEURAL NETWORKS.



Paolo Nistri was born in Rignano sull'Arno, Florence (Italy), in 1948. He received the doctoral degree in mathematics from the University of Florence, Florence, Italy, in 1972.

From 1974 to 1979 he was an Assistant Professor at the University of Calabria, Cosenza. In 1979, he moved to the Engineering Faculty of the University of Florence, where, from 1982 to 1998, he has been Associate Professor. Since 1998, he has been at the Engineering Faculty at the University of Siena, Siena, Italy, where he is Full Professor of Mathematical Analysis. His main scientific interests are in the field of dynamical systems, mathematical control theory, differential inclusions and topological methods. He is author of around 100 scientific publications, and he has been organizer and coordinator of several international scientific activities. He is an associate editor of international mathematical journals and he acts as referee of many of them.



Amedeo Premoli was born in Crema, Italy in 1942. He received the Laurea Degree (*summa cum laude)* in electronic engineering from Politecnico di Torino, Torino, Italy, in 1965.

From 1966 to 1986, he was a member of Istituto Elettrotecnico Nazionale Galileo Ferraris, Torino, Italy. From 1986 to 1995, he was Full Professor of Elettrotecnica (Basic Circuit Theory) at Dipartimento di Elettrotecnica, Elettronica ed Informatica (Dept. of Electrical and Computer Sciences), Università di Trieste, Trieste, Italy. He was Visiting Professor at EPFL of Lausanne, Switzerland, in September 1984, June 1987, and October 1990. Since 1995, he has been Full Professor of Elettrotecnica (Basic Circuit Theory) at Dipartimento di Elettronica e Informazione (Dept. of Electronic and Computer Sciences), Politecnico di Milano, Milano, Italy. His main research interests are focused on circuit and system theory and applications, including, initially, the analysis and design of distributed circuits, approximation of transfer functions, design of electrical active filters and multilayer coating optical filters, more recently, the numerical processing of measurements, applications of optimization techniques, electrical and electrothermal simulation and dynamics of electronic circuits and cellular neural networks. He is author or coauthor of about 140 scientific papers, among which about 80 in international journals, and a handbook on design of electrical filters.

Dr. Premoli was Associate Editor of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: FUNDAMENTAL THEORY AND APPLICATIONS from 1993 to 1995.