Entropy-Based Stability of Fractional Self-Organizing Maps with Different Time Scales

By C. A Peña Fernández

Abstract- The behavior of self-organizing neural maps, which develop through a combination of long and short-term memory, involves different time scales. Such a neural network's activity is characterized by a neural activity equation representing the fast phenomenon and a synaptic information efficiency equation representing the slow part of the neural system. The work reported here proposes a new method to analyze the dynamics of self-organizing maps based on the flow-invariance principle, considering the performance of the system's different time scales. In this approach, the equilibrium point is determined based on the estimate for the entropy at each iteration of the learning rule, which is generally sufficient to analyze existence and uniqueness. In this sense, the viewpoint reported here proves the existence and uniqueness of the equilibrium point on a fractional approach by using a Lyapunov method extension for Caputo derivatives when $0 < \alpha < 1$. Furthermore, the global exponential stability of the equilibrium point is proven with a strict Lyapunov function for the flow of the system with different time scales and some numerical simulations.

ENTROPYBASEDSTABILITYOFFRACTIONALSELFORGANIZINGMAPSWITHDIFFERENTTIMESCALES

Strictly as per the compliance and regulations of:

Entropy-Based Stability of Fractional Self-Organizing Maps with Different Time Scales

C. A Peña Fernández

Abstract- The behavior of self-organizing neural maps, which develop through a combination of long and short-term memory, involves different time scales. Such a neural network's activity is characterized by a neural activity equation representing the fast phenomenon and a synaptic information efficiency equation representing the slow part of the neural system. The work reported here proposes a new method to analyze the dynamics of self-organizing maps based on the flow-invariance principle, considering the performance of the system's different time scales. In this approach, the equilibrium point is determined based on the estimate for the entropy at each iteration of the learning rule, which is generally sufficient to analyze existence and uniqueness. In this sense, the viewpoint reported here proves the existence and uniqueness of the equilibrium point on a fractional approach by using a Lyapunov method extension for Caputo derivatives when $0 < \alpha < 1$. Furthermore, the global exponential stability of the equilibrium point is proven with a strict Lyapunov function for the flow of the system with different time scales and some numerical simulations.

Highlights

- Self-organizing neural maps develop through long and short-term memory and have fast and slow activity equations.
- Lyapunov method extended for Caputo derivatives helps estimate and examine entropy behavior during learning.
- Self-organizing neural maps based on competitive differential equations lack entropy-based synaptic efficiency.
- The final stage of training is not affected by external stimuli.
- Pattern learning has lower entropy-rate without external stimuli during learning process.

I. Introduction

Self-organizing neural maps, also known as competitive neural networks, are an important class of neural networks. These networks focus on two key aspects: the ability to store desired patterns as stable equilibrium points and the mutual interference between neuron and learning dynamics. This research examines how cortical cognitive maps work by using a self-organizing map with differential equations for neural activity levels, short-term memory (STM), and synaptic information efficiency, long-term memory (LTM).

Author: Electrical Engineering Department and Technological Innovation Center of Federal Institute of Bahia, R. Emídio dos Santos, s/n, Salvador, Bahia, Brazil. e-mail: cesarfernandez@ifba.edu.br

STM and LTM models are usually based on classical Grossberg's approach or Amari's model for primitive neural competition [1, 2]. These models often involve mutually inhibitory neurons with fixed synaptic connections [3, 4, 1]. Researchers have studied competitive neural systems using flow invariance theory and singular perturbation theory on large-scale networks. These networks have two types of state variables that describe the slow unsupervised dynamics of synapses and the fast dynamics of neural activity. The fast dynamics are usually represented by the goal of the self-organizing map, such as clustering or recognition. One example is the Willshaw-Malsburg model [1, 5] of topographic formation, which uses solutions of equations of synaptic self-organization coupled with the field equation of neural excitations to improve the understanding of the dynamics of cortical cognitive maps [4, 6]. However, the design using classical competitive differential equations is not broad enough because the synaptic efficiency is not dependent on entropy. Therefore, this paper extends these approaches to incorporate systems where external stimuli can modify the synaptic information efficiency. This is done by estimating the entropy of information transfer between neurons. In other words, this research focuses on LTM in terms of how information transfer evolves. Fractional differential equations are used to model synaptic efficiency based on entropy estimation, which accurately improves the models based on STM and LTM approaches [7, 8, 9, 10, 11].

The study proposes an alternate model for information transfer between neurons based on entropy. It applies McMillan-Shannon's approach to fractional competitive differential equations to determine the mathematical conditions for when STM and the estimation of entropy related to LTM have bounded trajectories. The study uses an alternative version of Lyapunov functions to examine exponential stability in fractional-order systems [7]. This proposal is more comprehensive than previous studies, such as those conducted by [3, 4, 1, 8]. It presents a strict Lyapunov function for the neural multi-time scale system, which demonstrates the existence and uniqueness of the equilibrium point. Additionally, this proposal provides some conditions for global exponential stability based on singular perturbation theory and variable entropy-dependent synaptic efficiency.

The paper is structured as follows: In Section 2, the mathematical background related to self-organizing maps modeled by Caputo derivatives is presented. This allows for the inclusion of fractional order in the differential equations associated with STM and LTM dynamics. Section 3 analyzes the equilibrium point from the perspective of synaptic efficiency, and covers the existence and uniqueness of the equilibrium point. Section 4 presents numerical simulations that provide findings about SOM's equilibrium point and entropy when exposed to external stimulus. Finally, Section 5 offers closing remarks and final comments.

II. Mathematical Background

The synaptic information efficiency (SIE) can be defined according to the existence of an ergodic process related to binned input and output spike

[4] M. London, A. Schreibman, M. Häusser, M. E. Larkum, I. Segev, The information efficacy of a synapse, Nature Neuroscience 5 (4) (2002) 332–340. doi:10.1038/nn826.

trains whose length is N bins [4]. In computational experiments, N is associated with iterations during the learning procedure on time domain. In this way, by assuming m_{ij} the synaptic efficiency between i-th and the j-th neuron then let S_{in}, S_{out} the binned input and output spike trains, respectively, both defined by $\{\sigma_1, \ldots, \sigma_N\}$, where $\sigma_i \in \{0,1\}$ represents i-th bin, and assume that this string is the realization of a stationary and ergodic stochastic process. By using the Shannon-MacMillan-Brieman's theorem [12, 13] $-\frac{1}{N} \log_2 p(\sigma_1, \ldots, \sigma_N) \to H$, where H is the entropy rate of X (events set) and $p(\sigma, \ldots, \sigma_N)$ is the probability of obtaining the string $(\sigma_1, \ldots, \sigma_N)$ as a realization of X. In that sense, the estimate for the entropy at N-th iteration will be represented by $\hat{H}_N = -\frac{1}{N} \log_2 \hat{p}(\sigma_1, \ldots, \sigma_N)$. So, for i-th neuron,

$$\text{SIE}_i = \hat{H}_N(\sigma_1, \ldots, \sigma_N)_i - \hat{H}_N(\sigma_1, \ldots, \sigma_N | S_{in})_i$$

$$= \frac{1}{N} \log_2 \hat{p}_i(\sigma_1, \ldots, \sigma_N | S_{in}) - \frac{1}{N} \log_2 \hat{p}_i(\sigma_1, \ldots, \sigma_N),$$

where $\hat{H}_N(\sigma_1, \ldots, \sigma_N | S_{in})$ is the estimated output spike train entropy given the input spike train S_{in}, also known as conditional entropy (see [4]). In this way, for N iterations,

$$m_{ij} = \frac{1}{N} \log_2 \beta_{ij} \triangleq \frac{1}{N} \log_2 \frac{\hat{p}_i(\sigma_1, \ldots, \sigma_N | S_{in}) \hat{p}_j(\sigma_1, \ldots, \sigma_N)}{\hat{p}_j(\sigma_1, \ldots, \sigma_N | S_{in}) \hat{p}_i(\sigma_1, \ldots, \sigma_N)}. \tag{1}$$

Nevertheless, on the time domain, β_{ij} can be represented as $\beta_{ij}(t)$, for $t \geq 0$, since at N-th iteration there is some time related to the learning procedure. Sometimes, to guarantee coherence with symbolic representation for N iterations, the notation β_{ij} is retained.

Since the nature of the transfer process related to synaptic information becomes more precise using fractional differential equations, the Caputo definition for the fractional derivative is more suitable because it incorporates initial conditions and its integer order derivatives. Although there are several definitions regarding the fractional derivative of order $\alpha \geq 0$, in the time domain the general network equations describing the temporal evolution of the STM and LTM states for the j-th neuron of M neurons are

$$\frac{\varepsilon}{\Gamma(1-\alpha)} \int_0^t \frac{x_j^{(m+1)}(\lambda)}{(t-\lambda)^\alpha} d\lambda = -a_j x_j + \sum_{i=1}^M D_{ij} f(x_i) + \frac{B_j}{N} \sum_{i=1}^P \log_2 \beta_{ij} y_i$$

$$\frac{1}{N\Gamma(1-\alpha)} \int_0^t \frac{[\log_2 \beta_{ij}(\lambda)]^{(m+1)}}{(t-\lambda)^\alpha} d\lambda = -\frac{1}{N} \log_2 \beta_{ij} + y_i f(x_j),$$

where α is the Caputo's fractional order defined by $\alpha = m + \gamma$, $m \in \mathbb{Z}^+$, $0 < \gamma \leq 1$; $\Gamma(\cdot)$ is the Gamma function, x_j is the current activity level (STM),

a_j is the time constant of the neuron, B_j is the contribution of the external stimulus term, $f(x_i)$ is the neuron's output, y_i is the external stimulus, and ε is the fast time-scale associated with the STM state. D_{ij} represents a synaptic connection parameter between the i-th neuron and j-th neuron.

According to the definition of m_{ij} in (1) and unlike [1, 9, 6], the self-organizing map is implicitly modeled by a network of sources emitting input signals with a prescribed probability distribution and external stimulus $\boldsymbol{y} \triangleq [y_i]$. By using the dynamic transform $w_j = \langle \boldsymbol{y}, \log_2 \boldsymbol{\beta}_j \rangle$ the model gets as follows:

$$
\frac{\varepsilon}{\Gamma(1-\alpha)} \int_0^t \frac{x_j^{(m+1)}(\lambda)}{(t-\lambda)^\alpha} d\lambda = -a_j x_j + \sum_{i=1}^M D_{ij} f(x_i) + \frac{B_j}{N} w_j \tag{2}
$$

$$
\frac{1}{N\Gamma(1-\alpha)} \int_0^t \frac{w_j^{(m+1)}(\lambda)}{(t-\lambda)^\alpha} d\lambda = -\frac{1}{N} w_j + \|\boldsymbol{y}\|^2 f(x_j), \tag{3}
$$

where the external stimuli are assumed to be normalized vectors of unit magnitude $\|\boldsymbol{y}\|^2 = 1$.

As each string is the realization of a stationary and ergodic stochastic process, it will be assumed a stochastic column matrix for each iteration of $\boldsymbol{\beta}_j$, such that there exists $P \triangleq [\bar{p}_{ij}] \in \mathbb{R}^{M \times M}$, with $\sum_{i=1}^M \bar{p}_{ik} = 1$ and $\boldsymbol{\beta}_j^{\sigma+1} = P\boldsymbol{\beta}_j^\sigma$ for $(\sigma+1)$-th iteration. In this way, it can be noted that $\dot{\boldsymbol{\beta}}_j = \lim_{\Delta t \to 0} \frac{1}{\Delta^k t} (P-I)^k \boldsymbol{\beta}_j$, such that the sum of elements in each column of $(P-I)^k \triangleq [p_{ij}]$ is equal to zero, i.e.,

$$
\beta_{ij}^{(k)} = \lim_{\Delta t \to 0} \frac{1}{\Delta^k t} \sum_{s=1}^M p_{is} \beta_{sj}, \qquad \sum_{i=1}^M p_{ik} = 0, \quad \text{for } \forall \, k. \tag{4}
$$

III. EQUILIBRIUM AND GLOBAL ASYMPTOTIC STABILITY

The existence and uniqueness of the equilibrium point are given based on flow-invariance while the global exponential stability will be based on a strict Lyapunov function for fractional-order approaches. It is well-known that the flow-invariance theory provides a qualitative viewpoint about the dynamics of a system. To begin with, it will be defined the following theorem before presenting the main results of this paper.

Theorem 1. Consider the system of fractional differential equations:

$$
\frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{x_i^{(m+1)}(\lambda)}{(t-\lambda)^\alpha} d\lambda = -a_i x_i + \sum_{j=1}^M D_{ij} f(x_j) + \frac{B_i}{N} w_i, \tag{5}
$$

$$\frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{w_i^{(m+1)}(\lambda)}{(t-\lambda)^\alpha} d\lambda = -\frac{1}{N} w_i + f(x_i),, \tag{6}$$

for $i = 1, \ldots, M$, where $a_i > 0$ for all $i = 1, \ldots, M$, and f is locally Lipschitz and bounded, that is, there exists a constant $C > 0$ such that $-C \leq f(x) \leq C$ for all $x \in \mathbb{R}$. Then for any $\varepsilon > 0$ and for any initial condition $\{x(0), w(0)\} \in \mathbb{R}^{2M}$ there exists a $T \geq 0$ such that

$$w_i(t) \in [-C - \varepsilon, C + \varepsilon], \qquad x_i(t) \in [-l_i - \varepsilon, l_i + \varepsilon],$$

for all $i = 1, \ldots, M$, with equilibrium point $e = [\bar{x}_i \quad \bar{w}_i] = [\bar{x}_i \quad \langle \boldsymbol{y}, \log_2 \bar{\boldsymbol{\beta}}_i \rangle]$, where $\forall \bar{\beta}_{ij} \in \bar{\boldsymbol{\beta}}_i$ satisfies

$$\int_0^t \int_0^\lambda \frac{y_j}{\bar{\beta}_{ij}^2 \ln 2} \sum_{r=1}^M \sum_{q=1}^M \sum_{s=1}^M p_{is} p_{sq} p_{ir} \bar{\beta}_{qj} \bar{\beta}_{rj} \cdot d\lambda^2 = 0 \tag{7}$$

$$\int_0^t \frac{y_j}{\bar{\beta}_{ij}^2 \ln 2} \sum_{r=1}^M \sum_{q=1}^M \sum_{s=1}^M p_{is} p_{sq} p_{ir} \bar{\beta}_{qj} \bar{\beta}_{rj} \cdot d\lambda = 0 \tag{8}$$

$$S_m \sum_{s=1}^M p_{is} \bar{\beta}_{sj} y_j = \frac{y_j}{\bar{\beta}_{ij} \ln 2} \sum_{r=1}^M \sum_{q=1}^M \sum_{s=1}^M p_{is} p_{sq} p_{ir} \bar{\beta}_{qj} \bar{\beta}_{rj}, \tag{9}$$

for $m = 0, 1$ and 2, respectively, and $S_m \in \mathbb{Z}^+$ for all $t \geq T$.

Proof. Since f is Lipschitz, system (5)-(6) has local existence and uniqueness of solutions. Furthermore, since f is uniformly bounded, there exist constants $K_1, \ldots, K_5$ such that

$$\left| \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{x_i^{(m+1)}(\lambda)}{(t-\lambda)^\alpha} d\lambda \right| \leq K_1 + K_2 \|x_i(t)\| + K_3 \|w_i(t)\|$$

$$\left| \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{w_i^{(m+1)}(\lambda)}{(t-\lambda)^\alpha} d\lambda \right| \leq K_4 + K_5 \|w_i(t)\|,$$

thus all solutions are defined globally (for all $t \geq 0$).

Given $\varepsilon > 0$, let $\delta_i > 0$ be defined as

$$\delta_i = \begin{cases} \min\left\{ \frac{a_i \varepsilon}{2|B_i|}, \varepsilon \right\} & B_i \neq 0 \\ \varepsilon & B_i = 0 \end{cases},$$

such that $-|B_i|\delta_i + a_i\varepsilon \geq a_i\varepsilon/2$, for all $i = 1, \ldots, M$. Then for $t \geq 0$ and for $w_i(t) \leq -C - \delta_i$ the following inequality holds:

$$\frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{w_i^{(m+1)}(\lambda)}{(t-\lambda)^\alpha} d\lambda \geq -\frac{1}{N}(-C-\delta_i) + f(x_i) = \frac{1}{N}\delta_i +$$

$$\left[f(x_i) + \frac{1}{N}C \right] \geq \delta_i > 0. \tag{10}$$

Similarly, for $t \geq 0$ and for $w_i(t) \geq C + \delta_i$, the following inequality holds:

$$\frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{w_i^{(m+1)}(\lambda)}{(t-\lambda)^\alpha} d\lambda \leq -\frac{1}{N}(C+\delta_i) + f(x_i) = -\frac{1}{N}\delta_i +$$

$$\left[f(x_i) - \frac{1}{N}C \right] \leq -\delta_i < 0. \tag{11}$$

Since $\alpha \leq 1$ then $\Gamma(1-\alpha) > 0$ and $(t-\lambda)^\alpha > 0$. In that sense, from (10)-(11), both inequalities are guaranteed if and only if $w_i^{(m+1)}(\lambda) > 0$ and $w_i^{(m+1)}(\lambda) < 0$, respectively.

By mathematical induction, it can be noted that

$$w_i^{(m+1)} = \sum_{j=1}^M \left[\sum_{k=0}^m \frac{A_k}{\beta_{ij}^{m+1-k}} (\dot{\beta}_{ij})^{m+1-k} y_j^{(k)} + \frac{S_k}{\beta_{ij}^{m+1-k}} \beta_{ij}^{(k+1)} y_j^{(m-k)} \right.$$

$$\left. + \frac{d^{m-2}}{d\lambda^{m-2}} \frac{\dot{\beta}_{ij}\ddot{\beta}_{ij}y_j}{\beta_{ij}^2 \ln 2} \right),$$

where $A_k, B_k \in \mathbb{Z}^+$. So, for $w_i(t) \leq -C - \delta_i$ and $w_i(t) \geq C + \delta_i$,

$$\sum_{j=1}^M \left[\sum_{k=0}^m \frac{A_k}{\beta_{ij}^{m+1-k}} (\dot{\beta}_{ij})^{m+1-k} y_j^{(k)} + \frac{S_k}{\beta_{ij}^{m+1-k}} \beta_{ij}^{(k+1)} y_j^{(m-k)} \right] >$$

$$- \frac{d^{m-2}}{d\lambda^{m-2}} \frac{\dot{\beta}_{ij}\ddot{\beta}_{ij}y_j}{\beta_{ij}^2 \ln 2} \right), \tag{12}$$

$$\sum_{j=1}^M \left[\sum_{k=0}^m \frac{A_k}{\beta_{ij}^{m+1-k}} (\dot{\beta}_{ij})^{m+1-k} y_j^{(k)} + \frac{S_k}{\beta_{ij}^{m+1-k}} \beta_{ij}^{(k+1)} y_j^{(m-k)} \right] <$$

$$- \frac{d^{m-2}}{d\lambda^{m-2}} \frac{\dot{\beta}_{ij}\ddot{\beta}_{ij}y_j}{\beta_{ij}^2 \ln 2} \right), \tag{13}$$

respectively.

By using (4) in (12), it can be obtained

$$
\lim_{\Delta t \to 0} \sum_{j=1}^{M} \left[\sum_{k=0}^{m} \frac{A_k}{\Delta^{m-k-2} t \beta_{ij}^{m+1-k}} \sum_{s=1}^{M} p_{is} \beta_{sj} y_j^{(k)} \right.
$$

$$
+ \frac{S_k}{\Delta^{k-2} t \beta_{ij}^{m+1-k}} \sum_{s=1}^{M} p_{is} \beta_{sj} y_j^{(m-k)} \Bigg] >
$$

$$
- \lim_{\Delta t \to 0} \frac{d^{m-2}}{d\lambda^{m-2}} \left(\frac{y_j}{\beta_{ij}^2 \ln 2} \sum_{r=1}^{M} \sum_{q=1}^{M} \sum_{s=1}^{M} p_{is} p_{sq} \beta_{qj} p_{ir} \beta_{rj} \right).
$$

For $k < m$, the evaluation of limits above yields,

$$
\infty > - \frac{d^{m-2}}{d\lambda^{m-2}} \left(\frac{y_j}{\beta_{ij}^2 \ln 2} \sum_{r=1}^{M} \sum_{q=1}^{M} \sum_{s=1}^{M} p_{is} p_{sq} \beta_{qj} p_{ir} \beta_{rj} \right), \tag{14}
$$

which represents an obvious condition. For $k = m$,

$$
\lim_{\Delta t \to 0} \sum_{j=1}^{M} \left[\frac{A_m \Delta^2 t}{\beta_{ij}} \sum_{s=1}^{M} p_{is} \beta_{sj} y_j^{(m)} + \frac{S_m}{\Delta^{m-2} t \beta_{ij}} \sum_{s=0}^{M} p_{is} \beta_{sj} y_j \right]
$$

$$
> - \lim_{\Delta t \to 0} \frac{d^{m-2}}{d\lambda^{m-2}} \left(\frac{y_j}{\beta_{ij}^2 \ln 2} \sum_{r=1}^{M} \sum_{q=1}^{M} \sum_{s=1}^{M} p_{is} p_{sq} \beta_{qj} p_{ir} \beta_{rj} \right).
$$

Therefore, by evaluating the limits for $m > 2$, the inequality (14) is obtained again. However, for $0 \le m \le 2$ (i.e., $m = 0, 1$ and 2),

$$
\int_0^t \int_0^\lambda \left(\frac{y_j}{\beta_{ij}^2 \ln 2} \sum_{r=1}^{M} \sum_{q=1}^{M} \sum_{s=1}^{M} p_{is} p_{sq} p_{ir} \beta_{qj} \beta_{rj} \right) d\lambda^2 > 0, \tag{15}
$$

$$
\int_0^t \left(\frac{y_j}{\beta_{ij}^2 \ln 2} \sum_{r=1}^{M} \sum_{q=1}^{M} \sum_{s=1}^{M} p_{is} p_{sq} p_{ir} \beta_{qj} \beta_{rj} \right) d\lambda > 0, \tag{16}
$$

$$
S_m \sum_{s=1}^{M} p_{is} \beta_{sj} y_j < \frac{y_j}{\beta_{ij} \ln 2} \sum_{r=1}^{M} \sum_{q=1}^{M} \sum_{s=1}^{M} p_{is} p_{sq} p_{ir} \beta_{qj} \beta_{rj}, \tag{17}
$$

respectively. Since the operator $d^{m-2}/d\lambda^{m-2}$ becomes an integral for $m < 2$.

Notes

So, for $w_i(t) \le -C - \delta_i$, the inequalities (15)-(17) guarantee $w^{(m+1)} > 0$. The same treatment applies to (13), it is only to change the less than symbol in (15)-(17) by greater than symbol. Therefore, for any $i \in \{1, \ldots, M\}$ there exists a $T_i > 0$ such that

$$w_i(t) \in [-C - \delta_i, C + \delta_i] \subseteq [-C - \varepsilon, C + \varepsilon], \qquad (18)$$

for all $t \ge T_i$. So, the equilibrium point $\bar{w}_i = \langle \boldsymbol{y}, \log_2 \bar{\boldsymbol{\beta}}_i \rangle \in [-C - \varepsilon, C + \varepsilon]$, where $\forall \bar{\beta}_{ij} \in \bar{\boldsymbol{\beta}}_i$ satisfies

$$\int_0^t \int_0^\lambda \left. \frac{y_j}{\bar{\beta}_{ij}^2 \ln 2} \sum_{r=1}^M \sum_{q=1}^M \sum_{s=1}^M p_{is} p_{sq} p_{ir} \bar{\beta}_{qj} \bar{\beta}_{rj} \right) d\lambda^2 = 0, \quad (m = 0)$$

$$\int_0^t \left. \frac{y_j}{\bar{\beta}_{ij}^2 \ln 2} \sum_{r=1}^M \sum_{q=1}^M \sum_{s=1}^M p_{is} p_{sq} p_{ir} \bar{\beta}_{qj} \bar{\beta}_{rj} \right) d\lambda = 0, \quad (m = 1)$$

$$S_m \sum_{s=1}^M p_{is} \bar{\beta}_{sj} y_j = \frac{y_j}{\bar{\beta}_{ij} \ln 2} \sum_{r=1}^M \sum_{q=1}^M \sum_{s=1}^M p_{is} p_{sq} p_{ir} \bar{\beta}_{qj} \bar{\beta}_{rj} \quad (m = 2).$$

Defining $T_S = \max T_i$ then $w_i(t) \in [-C - \varepsilon, C + \varepsilon]$ holds for all $i \in \{1, \ldots, M\}$ and for all $t \ge T_S$.

Now, let $t \ge T_S$. For $x_i(t) \le -l_i - \varepsilon$, (5) and (18) imply that

$$\frac{1}{\Gamma(1 - \alpha)} \int_0^t \frac{x_i^{(m+1)}(\lambda)}{(t - \lambda)^\alpha} d\lambda \ge a_i(l_i + \varepsilon) + \sum_{j=1}^M D_{ij} f(x_j)$$

$$+ \frac{B_i}{N}(-C - \delta_i) \ge a_i l_i + a_i \varepsilon - C \left(\sum_{j=1}^M |D_{ij}| + \frac{|B_i|}{N} \right) - |B_i| \delta_i.$$

By defining $l_i = \frac{C}{a_i} \left(\sum_{j=1}^M |D_{ij}| + \frac{1}{N} |B_i| \right) > 0$, for $i = 1, \ldots, M$ then

$$\frac{1}{\Gamma(1 - \alpha)} \int_0^t \frac{x_i^{(m+1)}(\lambda)}{(t - \lambda)^\alpha} d\lambda \ge -|B_i| \delta_i + a_i \varepsilon \ge \frac{a_i \varepsilon}{2} > 0.$$

Similarly, for $t \ge T_S$ and for $x_j(t) \ge l_i + \varepsilon$, (5) and (18) imply that

$$\frac{1}{\Gamma(1 - \alpha)} \int_0^t \frac{x_i^{(m+1)}(\lambda)}{(t - \lambda)^\alpha} d\lambda \le |B_i| \delta_i - a_i \varepsilon \le -\frac{a_i \varepsilon}{2} < 0.$$

Since $\alpha \le 1$ then $\Gamma(1 - \alpha) > 0$ and $(t - \lambda)^{\alpha} > 0$. In that sense, both before inequalities are guaranteed if and only if $x_i^{(m+1)}(\lambda) > 0$ and $x_i^{(m+1)}(\lambda) < 0$, respectively. Therefore, for any $i \in \{1, \ldots, M\}$ there exists a $T_i > 0$ such that

$$x_i(t) \in [-l_i - \varepsilon, l_i + \varepsilon]$$

for all $t \ge T_i$. Defining $T_X = \max T_i$ then $x_i(t) \in [-l_i - \varepsilon, l_i + \varepsilon]$ holds for all $i \in \{1, \ldots, M\}$ and for all $t \ge T_X$. □

The system (5)-(6) is dissipative in $\mathbb{R}^{2M}$ and therefore, it has a compact global attractor

$$\mathcal{A} \subseteq D = \prod_{i=1}^{M} [-l_i, l_i] \times \prod_{i=1}^{M} [-C, C].$$

It follows from the proof of Theorem 1 that the set D is flow invariant under (5)-(6). In other words, D is positively invariant set of (5)-(6), i.e., any solution starting in D at $t = 0$ remains in D for all $t \ge 0$. Furthermore, from the proof of Theorem 1 the set H that contains D can be contracted to a point, and since D is flow-invariant with respect to (5)-(6) then by the Brower fixed point theorem implies that there exists a point $e \in D$ is an equilibrium point of (5)-(6).

Theorem 2. Suppose that $f(x)$ is $\mathcal{C}^1$ with $|\dot{f}(x)| < k$ for all x and

$$a_i > k \left(\sum_{j=1}^{M} |D_{ij}| + |B_i| \right), \qquad i = 1, \ldots, M, \qquad (19)$$

then the equilibrium e is unique.

Proof. At the equilibrium point, $f(x_i) = \frac{1}{N} w_i$ from (6). Substituting these expressions in (5), it is obtained

$$0 = -a_i x_i + \sum_{j=1}^{M} D_{ij} f(x_j) + B_i f(x_i), \quad \text{for } i = 1, \ldots, M.$$

Since $a_i > 0$, x_i can be expressed as

$$x_i = \frac{1}{a_i} \left(\sum_{j=1}^{M} D_{ij} f(x_j) + B_i f(x_i) \right) = G_i(x_1, \ldots, x_M).$$

The inequality (19) implies that

$$|G_i(x_1', \ldots, x_M') - G_i(x_1'', \ldots, x_M'')| < k|(x_1', \ldots, x_M') - (x_1'', \ldots, x_M'')|,$$

i.e., G is a contracting map in $\mathbb{R}^M$. Consequently, there exists a unique fixed point of G. The x_i-coordinates of this fixed point uniquely determine the w_i-coordinates via $f(x_i) = \frac{1}{N}w_i$. Therefore, the equilibrium point is unique. $\square$

Now, let $e = [\bar{x}_i \ \bar{w}_i] = [\bar{x}_i \ \langle \boldsymbol{y}, \log_2 \bar{\boldsymbol{\beta}}_i \rangle]$ be the equilibrium point of (5)-(6) and introduce the change of variables $\phi_i = x_i - \bar{x}_i$, $\varphi_i = w_i - \bar{w}_i$ which shifts e to the origin. Specifically, if $f_i(\phi_i) = f(\phi_i + \bar{x}_i) - \frac{1}{N}\bar{w}_i$, then $f_i(0) = 0$ and (5)-(6) may be rewritten as

$$\frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{\phi_i^{(m+1)}(\lambda)}{(t-\lambda)^\alpha} d\lambda = -a_i\phi_i + \sum_{j=1}^M D_{ij}f(\phi_j) + \frac{B_i}{N}\varphi_i, \qquad (20)$$

$$\frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{\varphi_i^{(m+1)}(\lambda)}{(t-\lambda)^\alpha} d\lambda = -\frac{1}{N}\varphi_i + f(\phi_i) \ \text{ for } i = 1,\dots,M. \qquad (21)$$

By assuming that there exists a Lyapunov function $V(t,\phi_i,\varphi_i)$ and class-K functions γ_i (for $i = 1,2,3$) satisfying

$$\gamma_1(|\phi_i|,|\varphi_i|) \le V(t,\phi_i,\varphi_i) \le \gamma_1(|\phi_i|,|\varphi_i|),$$

$$\frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{V^{(m+1)}(\lambda,\phi_i,\varphi_i)(\lambda)}{(t-\lambda)^\alpha} d\lambda \le -\gamma_3(|\phi_i|,|\varphi_i|),$$

then the system (20)-(21) becomes asymptotically stable at equilibrium point e [7].

Lemma 1. *[see [7]] Let $r(t) \in \mathbb{R}$ be a continuous and derivable function. Then, for any time $t \ge t_0$*

$$\frac{1}{2\Gamma(1-\alpha)} \int_0^t \frac{[r^2(\lambda)]^{(m+1)}}{(t-\lambda)^\alpha} d\lambda \le \frac{r(t)}{2\Gamma(1-\alpha)} \int_0^t \frac{r(\lambda)^{(m+1)}(\lambda)}{(t-\lambda)^\alpha} d\lambda.$$
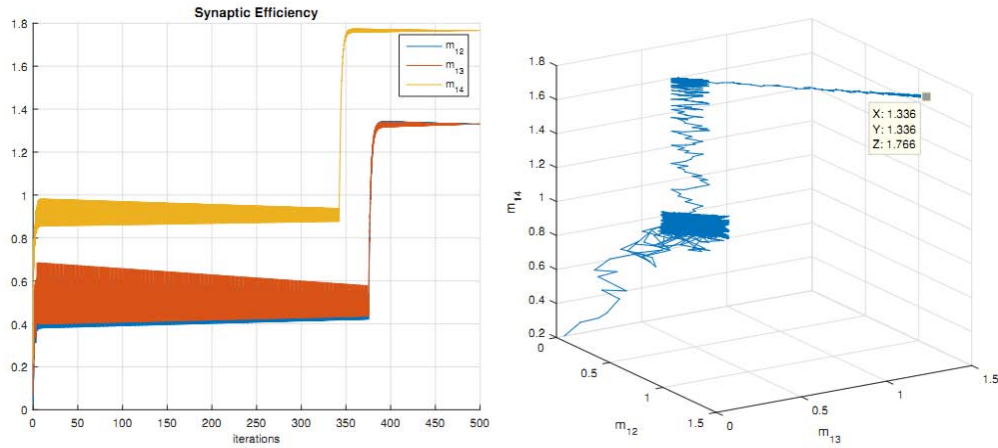
Notes

Notes

Figure 1: Evolution of synaptic efficiency m_{ij} (for $i = 1, j = 2, 3, 4$) toward equilibrium point $(1.336, 1.336, 1.768)$ after $N = 500$ iterations when each neuron receives an external stimulus based on neighborhood's winner.
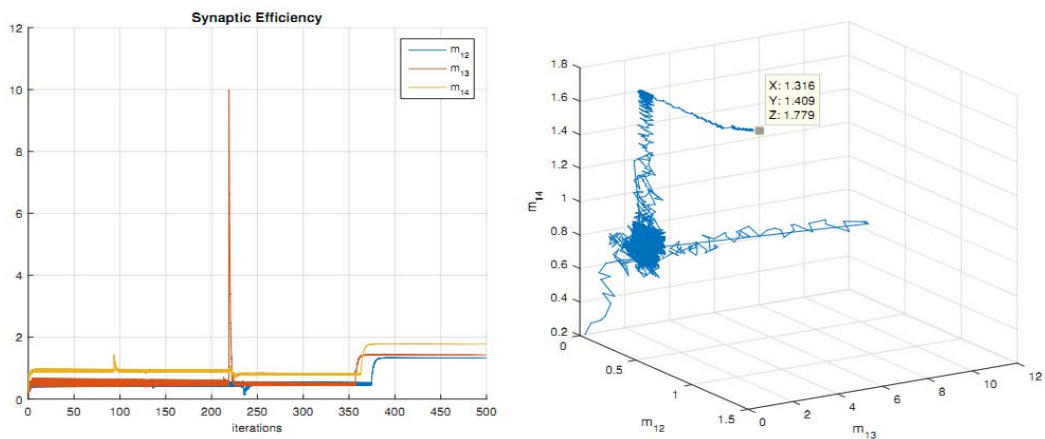

Figure 2: Evolution of synaptic efficiency m_{ij} (for $i = 1, j = 2, 3, 4$) toward equilibrium point $(1.316, 1.409, 1.779)$ after $N = 500$ iterations when each neuron updates its weights without external stimulus.

Theorem 3. Suppose that $f(x)$ is $\mathcal{C}^1$ with $|\dot{f}(x)| \leq k$ for all x and $a_i > 0$.

If

$$\max_i \left\{ \frac{1}{2} \left(\frac{|B_i|}{a_i} + k \right) + \sum_{j=1}^{M} \frac{1}{2} k \left(\frac{|D_{ij}|}{a_i} + \frac{|D_{ji}|}{a_j} \right) \right\} < 1 \qquad (22)$$

then e is a global attractor for the system (20)-(21). Moreover, all solutions of (20)-(21) converge to e exponentially fast as $t \to \infty$.

Proof. The global convergence will be proved by using a Lyapunov function for (20)-(21). Let

$$V(\phi_i, \varphi_i) = \frac{1}{2} \sum_{i=1}^{M} \left(\frac{\phi_i^2}{a_i} + \varphi_i^2 \right),$$

then the fractional derivative for $V(\phi_i, \varphi_i)$ is defined by

$$\frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{V^{(m+1)}(\phi_i, \varphi_i)}{(t-\lambda)^\xi} d\lambda =$$

$$\frac{1}{2\Gamma(1-\alpha)} \sum_{i=1}^{M} \left(\frac{1}{a_i} \int_0^t \frac{[\phi_i^2(\lambda)]^{(m+1)}}{(t-\lambda)^\alpha} d\lambda + \int_0^t \frac{[\varphi_i^2(\lambda)]^{(m+1)}}{(t-\lambda)^\alpha} d\lambda \right),$$

and by using the Lemma 1 together with (20)-(21) yields

$$\frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{V^{(m+1)}(\phi_i, \varphi_i)}{(t-\lambda)^\alpha} d\lambda \leq -\sum_{i=1}^{M} \left(\phi_i^2 + \varphi_i^2 \right) \tag{23}$$

$$+ \sum_{i=1}^{M} \sum_{j=1}^{M} D_{ij} \frac{\phi_i}{a_i} f_j(\phi_j) + \sum_{i=1}^{M} \varphi_i \left(\frac{\phi_i}{a_i} B_i + f_i(\phi_i) \right).$$

Since $f_i(0) = 0$ and $|\dot{f}_i(x)| = |\dot{f}(x + \bar{x}_i)| < k$, it is possible to have $|f_i(\phi_i)| < k|\phi_i|$. Consequently, with this last fact and the Minkowski inequality, (23) can be replaced by the inequality

$$\frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{V^{(m+1)}(\phi_i, \varphi_i)}{(t-\lambda)^\alpha} d\lambda < -\sum_{i=1}^{M} \left(\phi_i^2 + \varphi_i^2 \right)$$

$$+ \sum_{i=1}^{M} \sum_{j=1}^{M} |D_{ij}| k \frac{1}{a_i} |\phi_i| |\phi_j| + \sum_{i=1}^{M} \left(\frac{|B_i|}{a_i} + k \right) |\varphi_i| |\phi_i|.$$

The right-hand side of this inequality is given by the quadratic form with the matrix $-Q$ where Q has the following structure:

$$Q = \begin{bmatrix} 1 - \frac{1}{2}k \left(\frac{|D_{11}|}{a_1} + \frac{|D_{11}|}{a_1} \right) & -\frac{1}{2} \left(\frac{|B_1|}{a_1} + k \right) & -\frac{1}{2}k \left(\frac{|D_{12}|}{a_1} + \frac{|D_{21}|}{a_2} \right) & 0 & \cdots \\ -\frac{1}{2} \left(\frac{|B_1|}{a_1} + k \right) & 1 & 0 & 0 & \cdots \\ -\frac{1}{2}k \left(\frac{|D_{21}|}{a_2} + \frac{|D_{12}|}{a_1} \right) & 0 & 1 - \frac{1}{2}k \left(\frac{|D_{22}|}{a_2} + \frac{|D_{22}|}{a_2} \right) & -\frac{1}{2} \left(\frac{|B_2|}{a_2} + k \right) & \cdots \\ 0 & 0 & -\frac{1}{2} \left(\frac{|B_2|}{a_2} + k \right) & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

According to Gerschgorin's theorem applied to $-Q$, there are M disks centered at $z = -1$ (in $\mathbb{C}$-plane) with radius $\frac{1}{2}\left(\frac{|B_i|}{a_i} + k\right)$. In the same way, there are M disks centered at $z = \frac{1}{2}k\left(\frac{|D_{ii}|}{a_i} + \frac{|D_{ii}|}{a_i}\right) - 1$ and radius $\frac{1}{2}\left(\frac{|B_i|}{a_i} + k\right) + \sum_{j=1}^{M}\frac{1}{2}k\left(\frac{|D_{ij}|}{a_i} + \frac{|D_{ji}|}{a_j}\right)$. If condition (22) is valid then all eigenvalues q of $-Q$ satisfy $q < 1$ or

$$\frac{1}{2}k\left(\frac{|D_{ii}|}{a_i} + \frac{|D_{ii}|}{a_i}\right) - 2 < q < \frac{1}{2}k\left(\frac{|D_{ii}|}{a_i} + \frac{|D_{ii}|}{a_i}\right).$$

Since $\frac{1}{2}k\left(\frac{|D_{ii}|}{a_i} + \frac{|D_{ii}|}{a_i}\right) > 0$ for $\forall i$ then $-Q$ is positive definite and Q is negative definite. Therefore, let $\xi > 0$ the smallest eigenvalue of $-Q$ such that

$$\frac{1}{\Gamma(1-\alpha)}\int_0^t \frac{V^{(m+1)}(\phi_i, \varphi_i)}{(t-\lambda)^\alpha}d\lambda < -\xi\sum_{i=1}^{M}\left(\phi_i^2 + \varphi_i^2\right),$$

and consequently, V is a strict Lyapunov function for (20)-(21). Moreover, $2V \le \sum_{i=1}^{M}(\phi_i^2 + \varphi_i^2)$ thus

$$\frac{1}{\Gamma(1-\alpha)}\int_0^t \frac{V^{(m+1)}(\phi_i, \varphi_i)}{(t-\lambda)^\alpha}d\lambda < -2\xi V.$$

This result implies that V converges to zero exponentially fast, and the solutions (ϕ_i, φ_i) converge to the origin also exponentially fast, i.e., solutions (x, w) in system (5)-(6) converges exponentially fast to equilibrium $e = [\bar{x} \ \langle \boldsymbol{y}, \log_2 \bar{\boldsymbol{\beta}}_1 \rangle \dots \langle \boldsymbol{y}, \log_2 \bar{\boldsymbol{\beta}}_M \rangle]$.
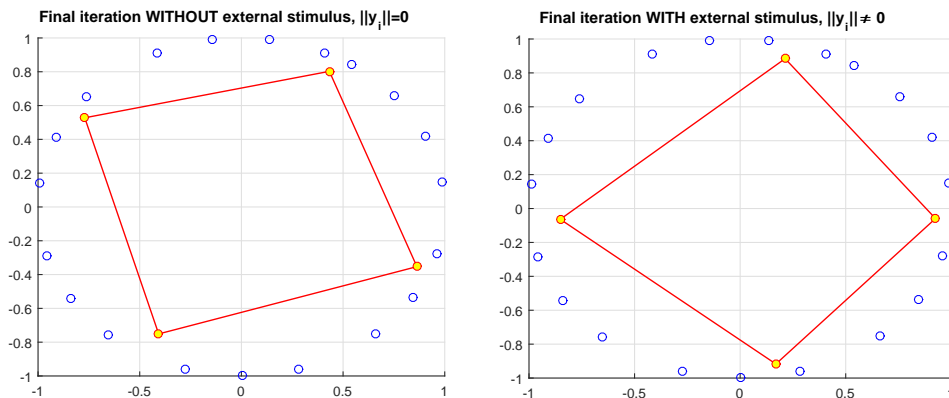

Figure 3: Comparing last iteration ($N = 500$ iterations) related to SOM for circle pattern composed by 20 points.

IV. Accessing Numerical Simulations

To validate the proposed model, a self-organizing map with four neurons ($M = 4$) is utilized in the following example. The aim is to find the best updating of synaptic weights to adjust a circle shape that consists of 20

patterns and has a radius of 0.5. This will be achieved using unsupervised Hebb's learning rule with $N = 500$ and a learning rate of 0.05. Let's assume that each neuron in a neural network has an activation value of $a_j = 0.5$. Two parameters, D_{ij} and B_j, were set randomly for each neuron within the range of 0 to 1. Additionally, α was set to 0.5, m to 0, and ε to 10^{-4}. To analyze the behavior of the equilibrium point in this model, the algorithm ran for 500 iterations and observed how the equilibrium point behaves. Theorem 1 was used to determine the stability of this point. Our results show that the model is locally stable, which means that it is capable of reaching an equilibrium point and maintaining it over time.

In order to illustrate this point, Fig. 1 illustrates the equilibrium point $(1.336, 1.336, 1.768)$ after 500 iterations with no external stimulus, meaning that $\|y_i\| = 0$ and $m_{ij} = \frac{1}{N} \log_2 \beta_{ij}$ for $i = 1$ and $j = 2, 3, 4$. This means that all sequences $(\sigma_1, \ldots, \sigma_{500})$ for events set $X = \emptyset$ between two neurons have a probability of one, and the entropy H_i $(i = 1, 2, 3, 4)$ approaches zero, which is the ideal situation.

The following text describes the behavior of neurons when an external stimulus is applied. So, let's compare two figures - Fig. 1 and Fig. 2. The former shows the neurons' behavior when no external stimuli are applied, while the latter shows their behavior when stimuli are applied. The amplitude of the impulses generated by the stimuli is set at 0.001, and the stimuli are defined by one impulse over the threshold. The sequences $(\sigma_1, \ldots, \sigma_{500})$ associated with each neuron are considered as a result of updating synaptic weights according to the training rule at each iteration. In a neural network, when a neuron is declared a winner and its weights are updated, the weights of the neighboring neuron also need to be updated. Therefore, for each iteration in the process, every neuron must have a term called σ_i in its sequence. This term should be equal to one when the neuron is updated and zero when it doesn't receive any updates in its weights. Figure 3 illustrates that neurons form a circular pattern consisting of 20 points; however, the final position of neurons changes due to external stimuli during the last iteration.

For both cases in Figures 1 and 2, the equilibrium point $(\bar{\beta}_{12}, \bar{\beta}_{13}, \bar{\beta}_{14})$ is represented by $(1.336, 1.336, 1.768)$ and $(1.316, 1.409, 1.779)$. By assuming only the synaptic efficiencies m_{12}, m_{13}, m_{14}, condition (7) becomes the following three conditions:

$$\frac{p_{11}}{\ln 2} \sum_{s=1}^{M} p_{1s} p_{s1} \int_0^t \int_0^\lambda y_4 \cdot d\lambda^2 = 0,$$

$$\frac{p_{11}}{\ln 2} \sum_{s=1}^{M} p_{1s} p_{s1} \int_0^t \int_0^\lambda y_3 \cdot d\lambda^2 = 0,$$

$$\frac{p_{11}}{\ln 2} \sum_{s=1}^{M} p_{1s} p_{s1} \int_0^t \int_0^\lambda y_2 \cdot d\lambda^2 = 0,$$

i.e., the external stimulus has no contribution in the final stage of training for neurons connected to neuron 1.

On another side, it can be observed from Figs. 1 and 2 that the evolution of m_{1j} $(j = 2, 3, 4)$ ensures the condition $SIE_1 > SIE_j$, more specifically,

$$H_i(\sigma_1, \ldots, \sigma_{500}|S_{in}) - H_j(\sigma_1, \ldots, \sigma_{500}|S_{in})$$
$$> H_i(\sigma_1, \ldots, \sigma_{500}) - H_j(\sigma_1, \ldots, \sigma_{500}).$$

This result indicates that the motion of neurons without an external stimulus is less disorganized than the motion of neurons with an external stimulus, as expected.

V. Final Remarks

The following paper establishes the global exponential stability of SOMs. To this end, this proposal has used fractional order derivatives to describe cognitive cortical maps that result from self-organization, with both LTM and STM approaches. It applied McMillan-Shannon's approach to fractional competitive differential equations, which helps to understand the entropy behavior of SOM in response to external stimuli. The proposal not only proves the existence and uniqueness of the equilibrium point but also shows how fractional order derivative operators can improve our understanding of entropy associated with synaptic efficiency. Synaptic efficiency was modeled by using sequences of updating impulses at each iteration, all of them as a result of applying external stimuli, not only for the winner neuron but also for neighbor neurons. In this way, the concept of conditional entropy plays an important role in the proposal as it shows that it must be greater than entropy to ensure convergence to the equilibrium point.

In future studies, it is hoped to investigate the correlation between the development of information transfer and the division of fast dynamics into distinct time scales. This will involve integrating the Dynamic Confined Space of Velocities criterion (DCSV) [14] into restricted self-organizing maps, which are extensively utilized to address constrained multi-objective optimization issues (CMOPs) [15].

References Références Referencias

1. A. Meyer-Baese, S. Pilyugin, Y. Chen, Global exponential stability of competitive neural networks with different time scales, IEEE Transactions on Neural Networks 14 (3) (2003) 716–719. doi: 10.1109/tnn.2003.810594.
2. S.-I. Amari, Competitive and Cooperative Aspects in Dynamics of Neural Excitation and Self-Organization, Springer Berlin Heidelberg, 1982, pp. 1–28. doi: 10.1007/978-3-642-46466-91.

3. M. Lemmon, B. Vijaya Kumar, Emulating the dynamics for a class of laterally inhibited neural networks, Neural Networks 2 (3) (1989) 193–214. doi: 10.1016/0893-6080(89)90004-x.

4. M. London, A. Schreibman, M. Häusser, M. E. Larkum, I. Segev, The information efficacy of a synapse, Nature Neuroscience 5 (4) (2002) 332–340. doi:10.1038/nn826.

5. D. J. Willshaw, V. D. Malsburg, How patterned neural connections can be setup by self-organization, Proceedings of the Royal Society of London. Series B. Biological Sciences 194 (1117) (1976) 431–445. doi: 10.1098/rspb.1976.0087.

6. S. Najafian, E. Koch, K. L. Teh, J. Jin, H. Rahimi-Nasrabadi, Q. Zaidi, J. Kremkow, J.-M. Alonso, A theory of cortical map formation in the visual brain, Nature Communications 13 (1) (Apr. 2022). doi: 10.1038/s41467-022-29433-y.

7. N. Aguila-Camacho, M. A. Duarte-Mermoud, J. A. Gallegos, Lyapunov functions for fractional order systems, Communications in Nonlinear Science and Numerical Simulation 19 (9) (2014) 2951–2957. doi: 10.1016/j.cnsns.2014.01.022.

8. S. Zhang, Y. Yu, J. Yu, Lmi conditions for global stability of fractional order neural networks, IEEE Transactions on Neural Networks and Learning Systems 28 (10) (2017) 2423–2433. doi: 10.1109/tnnls.2016.2574842.

9. J. Jia, X. Huang, Y. Li, J. Cao, A. Alsaedi, Global stabilization of fractional order memristor-based neural networks with time delay, IEEE Transactions on Neural Networks and Learning Systems 31 (3) (2020) 997–1009. doi: 10.1109/tnnls.2019. 2915353.

10. A. Wu, Z. Zeng, Global mittag–leffler stabilization of fractional-order memristive neural networks, IEEE Transactions on Neural Networks and Learning Systems 28 (1) (2017) 206–217. doi: 10.1109/tnnls. 2015.2506738.

11. J. Chen, Z. Zeng, P. Jiang, Global mittag-leffler stability and synchronization of memristor-based fractional-order neural networks, Neural Networks 51 (2014) 1–8. doi: 10.1016/j.neunet.2013.11.016.

12. T. M. Cover, J. A. Thomas, Elements of Information Theory, Wiley-Interscience, 2006.

13. G. N. Saridis, Entropy In Control Engineering, World Scientific Publishing Company, 2001.

14. C. A. Peña Fernjández, An ergodic selection method for kinematic configurations in autonomous, flexible mobile systems, Journal of Intelligent amp; Robotic Systems 109 (1) (Aug. 2023). doi: 10.1007/s10846-023-01933-z.

15. C. He, M. Li, C. Zhang, H. Chen, P. Zhong, Z. Li, J. Li, A self-organizing map approach for constrained multi-objective optimization problems, Complex amp; Intelligent Systems 8 (6) (2022) 5355–5375. doi: 10.1007/s40747-022-00761-2.